ABSTRACT

Title of Document:	THE IMPACT OF PRELIMINARY MODEL SELECTION ON LATENT GROWTH MODEL PARAMETER ESTIMATES
	Hsiu-Fei Wang, Doctor of Philosophy, 2010
Directed By:	Professor Gregory R. Hancock, Department of Measurement, Statistics, and Evaluation

In behavioral and social sciences, model selection and parameter estimation are treated as two separate steps of data analysis. The second step, parameter estimation, is generally conducted on the assumption that the model selected in step one is a correct model, and thus it is performed using the same data set that was used in step one. This two-step process ignores the effects of model uncertainty on parameter estimation, and thus may ultimately lead to misleading or invalid inferences.

The problems arising from the use of the two-step process have been well investigated in the context of regression. In the case of latent growth modeling (LGM), however, there have been no such published studies. This present study was thus designed to investigate the possible problems arising from the use of this two-step process in LGM. The goals of this study were: (1) To examine the subsequent impact of preliminary model selection using information criteria on LGM parameter

estimates; (2) To assess the data splitting method as a possible way to mitigate the effects of model uncertainty.

Two Monte Carlo simulation studies were conducted to achieve these goals. Study 1 was conducted using the same data set for both model selection and parameter estimation,, to investigate the possible impact of preliminary model selection in terms of model selection accuracy, relative parameter biases, and coverage rate. Study 2 was conducted using different split-data sets for both model selection and parameter estimation, to assess the data splitting method as a possible way to mitigate the effects of model uncertainty.

The major finding of this study was that inference based on AIC or BIC model selection leads to additional bias in, and overestimates the sampling variability of, the parameter estimates. The results of simulation studies showed that the post-model-selection parameter estimator has larger relative parameter biases, larger relative variance biases, and smaller coverage rate of confidence interval, than those of the true-model-selection estimator. These post-model-selection problems due to model uncertainty, unfortunately, still existed when the data splitting method was applied.

THE IMPACT OF PRELIMINARY MODEL SELECTION ON LATENT GROWTH MODEL PARAMETER ESTIMATES

By

Hsiu-Fei Wang

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
2010

Advisory Committee:
Professor Gregory R. Hancock, Chair
Professor Robert G. Croninger
Professor Hong Jiao
Professor Robert W. Lissitz
Professor George Macready

Dedication

To Tom, whose wisdom and love have inspired me to be the best I can be, and whose patience and support have made this all possible and rewarding.

Acknowledgements

With deepest gratitude, I wish to thank every person who has inspired, encouraged, and illuminated me while conducting this study. I would especially like to acknowledge and express my appreciation to the following people for their magnificent support and help.

First , I would like to send a special and heartfelt thanks to my advisor, Dr. Hancock, who believed in me more than I did myself and who magically transformed the seemingly daunting task of doing a dissertation into an enjoyable challenge. His passion, style, and guidance have had a greater influence over this work, and me, than he knows. Likewise, I would like to thank my dissertation committee members, Dr. Croninger, Dr. Jiao, Dr. Lissitz, and Dr. Macready for their enormous help and constructive suggestions given to me in creating this study.

I would also like to thank Wen-Chyi, Meng-Ju, Jui-Chen, and Phillip for providing computational help and resources. I thank Dr. Harring for his instant response to my special request for installation of the necessary software in the computer lab for my simulation study. I thank Dr. Smith for his enormous patience and help in explaining and answering my questions in statistics. I thank Audrey, Charles, and Tom, who helped to correct my English and gave me such valuable feedback reading the first draft.

I thank Yu, Jimmy, Shu-Ling, Wei-Heng, and TH for their more than ten years of loving and genuine support, and for motivating me to higher achievements. I thank Jennifer for believing in me and inspiring me to do research. I thank Emilie for understanding my feelings and for all the long and interesting conversations we had until dawn. I thank Jim for his constant encouragement, and for his wake-up calls in the library. I thank my precious friends: Rachna, Christy, Long, Robin, Rachelle, Jaehwa, Ru, Ya-Fei, Fang, Bonita, Gabriel, Wendy, Jun, Kyle, Mu-Liang, Yu-Ru, Chin-Fang, Daphne, Lin, Sophie, Pepper, Hsin-Hua, Rex, Emiko, and Sasha, Sonia, Elizabeth, and Dennis, who are my great cheerleaders and incredible supporters, accompanying me through all the difficult times.

Finally, but most especially of all, I thank my family for your truly unconditional love and for being there whenever, wherever, and however I needed. I can never thank you enough.

Table of Contents

List of Tables

Chapter 1: Introduction

1.1 Problems of Model Selection

One of the main tasks of applied statisticians and data analysts is to construct and evaluate a statistical model that describes and summarizes the behavior of an object of study. In the model-building process, researchers begin by representing the observations in terms of random variables, then fitting a model to the data, and finally providing an estimate of the parameters. If the data-model fit is good, this statistical model is commonly seen as a convenient conceptual representation of the observed phenomenon and as an abstract mechanism generating the observed data.

In behavioral and social sciences, however, it may not be possible to specify the true model from the analysis of an observed finite data set because the true mechanism, which generated the collected data, might be very complex and difficult to recognize. It is increasingly common, therefore, for several candidate models to be considered and to be fitted to a collected data set at the same time. As such, data mining processes and model selection techniques are useful, and might be unavoidable, in deciding on an appropriate model to fit and explain the data.

Once a model has been selected, by whatever model selection criteria are deemed reasonable, estimations and inferences are made about model parameters using the same data set under the assumption that the selected model is the true model. In other words, although model selection, and parameter estimation and inference are treated as two separate stages of data analysis, they are typically performed using the same data set.

Unfortunately, this two-step practice results in at least three problems:

(1) The use of the same data set for both model selection and parameter estimation ignores model uncertainty, that is, that the selected model might be wrong (Chatfield, 1995).

(2) Because the estimation procedure depends upon the outcome of model selection, the properties of preliminary model estimators (e.g., the shape of the distribution) and related statistics (e.g., the estimates of mean squared prediction error and the value of R^2) might be different from those had the model been known *a priori* (Breiman, 1988; Hurvich & Tsai, 1990; Pötscher, 1991; Rencher & Pun, 1980). Consequently, bias may exist. For example, Bancroft (1944) derived a mathematical formula to show the bias of regression coefficient estimators that resulted from preliminary model selection.

(3) Because the use of a model selection procedure affects the asymptotic distribution of parameter estimators and related statistics, the validity of the subsequent inference procedures may be severely affected (Miller, 1984; Zhang, 1992).

To summarize, because a true model is seldom known in the behavioral and social sciences, data-driven model selection procedures are commonly used. When fitting a model to data, the choice of the model, and the subsequent parameter estimation and inference procedures, are often based on the same data set. As a result, problems emerge which may ultimately lead to misleading and invalid inferences.

1.2 Purpose of Research

The two-step process is used in many types of analysis, including regression and structural equation modeling (SEM). In the particular case of latent growth modeling (LGM), a specific type of SEM which will be addressed in this study, the two-step process is typically carried out as follows. (1) First, one evaluates a growth model in which growth parameters are treated as latent variables, and repeated measures are treated as multiple indicators of the latent variables in order to capture the trends of changes. Several plausible candidate models might be considered and fitted to a collected data set (e.g., linear growth models with and without covariance between intercept factor and slope factor, and quadratic growth models with and without relations among residuals over time). If models are nested, models are then compared by using differences in chi-square statistics under the assumption of multivariate normality. If models are not nested, then model fit indices are used. In this first step, the evaluation of data-model fit for competing models is of primary interest. Based on the data-model fit evaluations, a single appropriate model is selected. (2) The second step then is conducted by using the same data set to estimate and test the specific parameters (e.g., means of growth factors) within the selected model, and to make inferences.

This two-step process in LGM is similar to that used in regression and might be expected to be subject to similar problems. The problems arising from the use of the two-step process have been well investigated in the case of regression. In fact, the discussion in part 1.1 above (problems of model selection), was based almost entirely on literature regarding regression. In the case of LGM, however, many applied researchers use popular computer software to run series of computer trials and they then choose the best fitting model that result from these series of computer trials. They may not recognize that the two-step process is being used in these software..Therefore, they may not consider investigating the possible problems inherent in model selection. So, there have been no such published studies regarding the impact of preliminary model selection in LGM.

Although the two-step process in LGM might be vulnerable to the same criticism as that in regression, LGM is different from regression in terms of model selection in an important way. In regression, the candidate models usually contain different variables (i.e., predictors), whereas in LGM the candidate models usually contain identical variables, but with different arrangements among those variables (e.g., the candidate quadratic growth models might contain different covariances among latent growth factors). This difference between regression and LGM is significant enough so we cannot blindly apply what has been found in regression to LGM. This study is thus designed to investigate the possible problems arising from the use of the two-step process in LGM. The study has two goals:

(1) To examine the impact of preliminary model selection using information criteria on latent growth model parameter estimates, and

(2) To assess the data splitting method as a possible way to mitigate the effects of model uncertainty.

These goals were accomplished through two Monte Carlo simulation studies.

Chapter 2 of this dissertation reviews the current literature regarding the basic idea of model selection, including model selection methods, problems of model selection, and the possible ways to overcome or at least mitigate the potential effects of model selection. In addition, the literature of model fit and model selection in SEM, specifically in the context of LGM, are also examined.

Chapter 3 explicates the research design. The first section of Chapter 3 describes the Monte Carlo simulation study, including the populations from which the data are drawn, the manipulated factors, and the data generation procedures. The second section of Chapter 3 presents outcome measures and data analysis.

Chapter 4 presents the summary of the simulation results. The current investigation includes two Monte Carlo simulation studies. Results are looked at from three aspects: (1) model selection accuracy, (2) relative parameter biases, and (3) coverage rate.

Chapter 5 includes the discussion of the findings, the limitation of this study, and possible future research directions.

Chapter 2: Literature Review

The Number of Cases Without Negative Estimates of Variance When Fitting the Data to the Linear Model.This chapter will explore the basic idea of model selection, the problems existing in model selection, and spossible ways to overcome or mitigate these problems. In addition, the literature on model fit and model selection in SEM, specifically in the context of LGM, will be examined.

2.1 Model Selection

Truth in social sciences is usually complicated. In some cases, researchers try to model the phenomenon of interest in order to make an appropriate decision or prediction. The model building process generally consists of three main components: model specification, model fitting, and model selection. It is usually an iterative process. Take SEM as an example. During model specification, a researcher's hypotheses are expressed as structural equation models in the form of diagrams or series of equations. A model's variables and the directionalities of presumed relations among observed or latent variables are specified. Once a model is specified, researchers use computer programs to evaluate model fit and estimate the model parameters. If the researcher's initial model does not fit the data very well, it might be necessary to respecify the model with different relations among variables, or possibly different variables. During this iterative process, several candidate models might be considered. These candidate models might have nested or non-nested relations with each other, depending on the researcher's specifications. In the end, the best single model, assuming adequate fit, is selected.

Obviously, it may not be possible to find a model representing exact truth or full reality from the analysis of a finite amount of data. There is some uncertainty about how to decide on the appropriate specifications. Therefore, in practice, preliminary tests have been used as an aid in choosing an appropriate specification/model. In most situations, a researcher is forced to ask which model has the best fit for a given set of data, and usually has to settle for inferences based on a good approximating model. Thus, the critical issue is, "What is the best model to use?" The problem of choosing from among a limited range of alternative models using only the available data is known as the *model selection problem* (Burnham & Anderson, 2002).

2.1.1 Model Selection Methods

Various procedures can be used to select appropriate models. Null hypothesis testing has been viewed as a popular basis for model selection. In the particular context of regression, sequential testing has often been used, either forward or backward methods (Burnham & Anderson, 2002).

The second approach to model selection is the use of likelihood ratio. A likelihood ratio approach can be used to determine goodness of fit and leads to a chi-square test on the assumption of multivariate normality (Bollen, 1989). In cases where data violate the normality assumption, a robust model chi-square test statistic is used instead (Satorra & Bentler, 1994), and thus relevant chi-square difference tests have to be adjusted, as described by Satorra and Bentler (2001). These testing-based methods are employed when models are nested. The definition of "nested models", however, is slightly different in regression versus in SEM. In regression the candidate models usually contain different variables (i.e., predictors), whereas in SEM the candidate models may contain identical variables, but with different arrangements among those variables. In regression, Model I is nested within Model II if Model I's set of variables is a subset of the variables for Model II. In SEM, Model I is nested within Model II if Model I's set of parameters to be estimated is a subset of the parameters to be estimated for Model II.

Yet, another approach to model selection is the use of Information Criterion (IC) measures, such as the Akaike Information Criterion (AIC; Akaike, 1978) and the Bayesian Information Criterion (BIC; Schwarz, 1978), which may be used whether models are nested or non-nested. In their general form, information criterion indices are based on the log likelihood (Log L) of a fitted model, where each IC measure applies a different correction for the number of model parameters and/or sample size in order to balance goodness of fit and complexity. More complex models usually fit data better, but the additional parameters may not represent anything useful. The concept of parsimony is employed in these methods of model selection. That is, if many models fit data equally well, the simplest model is preferred.

The IC measures considered in this study are two most commonly used, AIC and BIC. The AIC is defined as

$$\text{AIC} = -2 \log L + 2t,$$

where L is the maximum likelihood for the model and t is the number of free model parameters. The BIC is defined as

$$\text{BIC} = -2 \log L + t \log(n),$$

where n is sample size.

Cross-validation has also been suggested as a model evaluation method (Cudeck & Browne, 2003; Shao, 1993). In this method, the data are divided into two partitions. The first partition is used for model fitting and the second is used for model validation. Then a new partition is selected, and this whole process is repeated many (e.g., hundreds of) times. Some criterion, such as minimum squared prediction error, is then chosen as an index for model selection. The disadvantage of cross-validation is that the data need to be split into two or more parts. This can be a serious problem when the sample size is small.

2.1.2 Problems of Model Selection

In the context of regression, much has been written concerning the impact of preliminary model selection when a data-dependent model selection procedure has been used (Hurvich & Tsai, 1990; Leeb, 2005; Miller, 1990; Rencher & Pun, 1980). In such situations, data are used both to select a parsimonious model and to estimate the model parameters and their precision. Possible problems of this data-driven model selection practice in regression are: (1) ignorance of the model selection uncertainty; (2) the properties of preliminary model estimators and related statistics might be different from those when the model is known *a priori*; (3) the validity of the inference procedures may be affected. Each of these problems will be elaborated upon below.

First, the use of the same data set for both model selection and inference prompts a concern for model selection uncertainty in that the best selected model might be wrong (Chatfield, 1995). According to Draper (1995) and Hodges (1987), there are typically three main sources of uncertainty in the process of model building: (a) uncertainty about the structure of the model; (b) uncertainty about estimates of the model parameters, assuming that the structure of the model is known; (c) unexplained random variation in observed variables, even when the structure of the model and the values of the model parameters are known. Uncertainty about model structure might result from different sources, such as model misspecification (e.g., omitting a variable or constraining a parameter by mistake) or choosing from among alternative models of quite different structures. Draper and Hodges also noted that ignoring the effects of uncertainty about model structure results in estimated sampling variances and covariances that are too low, and thus the achieved confidence interval coverage will be below the minimal value. Chatfield (1995) pointed out "Statisticians must stop pretending that model uncertainty does not exist and begin to find ways of coping with it" (p. 422).

Second, because the estimation procedure depends upon the outcome of model selection, the properties of preliminary model estimators and related statistics might be different from those when the model is known *a priori*. For example, in the case of regression Hurvich and Tsai (1990) concluded that the conditional coverage rates are much smaller than the nominal coverage rates, assuming the model was known in advance. Rencher and Pun (1980) demonstrated that a model selected by the best subset regression method tends to have an inflated value of R^2. Pötscher (1991) investigated the asymptotic properties of preliminary model estimators and derived the asymptotic distribution of parameter estimators and related statistics. His research showed that although the mean of the asymptotic distribution of parameter estimators is unaffected by model selection, the variance will increase due to the model selection uncertainty and the shape of the distribution may change. Bancroft (1944) was concerned with the bias of a regression coefficient which resulted from pretests for model selection. Consequently, he derived a mathematical formula to investigate the bias of a regression coefficient estimator when the model had not been determined *a priori*. Miller (1990) provided a technical discussion of model selection bias in the context of linear regression. He warned that *p*-values from subset selection software were lacking foundation, and large biases in regression coefficients were often caused by data-based model selection. Breiman (1988) also showed that models selected by the various data-driven methods can produce strongly biased estimates of mean squared prediction error.

Third, because the use of a model selection procedure affects the asymptotic distribution of parameter estimators and related statistics, the validity of the subsequent inference procedures may be severely affected. Miller (1984) showed that, if one starts with a model selected from the data, then regression estimators may be biased and standard hypothesis tests may not be valid. Also, Zhang (1992) investigated the impact of model selection on statistical inferences in linear regression. His results showed that although variable selection did not have much impact on the inferences for the error variance, the sizes of the nominal confidence sets tend to be inflated if they are derived based on the selected model. Leeb and Pötscher (2005) commented, “naïve use of inference procedures that do not take into account the model selection step can be highly misleading” (p. 22).

2.1.3 Ways to Overcome or Mitigate the Problems of Model Selection

In this section, three primary methods for overcoming or at least mitigating the non-trivial biases which result from data-dependent specification searches are described: computational methods, Bayesian model averaging approach (BMA), and data splitting.

Computational Methods

A variety of computational methods have been examined including resampling, bootstrapping, and jackknifing (e.g., Faraway, 1992; Hjorth, 1994). Faraway (1992) wrote a program to simulate the data-analytic actions in a regression analysis. He investigated model selection bias and tested bootstrapping, jackknifing, and sample-splitting for mitigating the problem. Faraway's simulation results suggested that bootstrapping and jackknifing can provide more realistic, although not perfect, estimates of the error and thus can reduce the bias resulting from preliminary model selection. The sample-splitting estimator has less bias but at the expense of additional variance. Breiman (1992) suggested that bootstrapping can give nearly unbiased estimates of the mean square prediction error in regression models selected by using data-driven selection procedures.

Although research to date has shown bootstrapping to be an appealing alternative for reducing the bias due to model uncertainty, three issues clearly remain to be addressed. First, in bootstrapping the original parent data set may not represent the population. In this case, if a data set from Model I happens to have characteristics which suggest Model II, then the bootstrap samples are also likely to favor Model II rather than the true Model I (Chatfield, 1995). Second, an inappropriate choice of the resampling algorithm in bootstrapping may lead to problematic results. For example, Freeman, Navidi, and Peters (1988) indicated that resampling which is conditional on the fitted model must be avoided; otherwise, bootstrap samples will not reflect the true extent of model uncertainty. Faraway (1992) also showed that a resampling algorithm, which was conditional on the model, seriously underestimated the variance of the quantities of interest. Third, while bootstrapping is asymptotically consistent, it does not provide general finite-sample guarantees. Freeman et al. (1988) found that the bootstrap method worked reasonably well for adjusting the bias due to variable selection in regression, when the ratio of observations to predictor variables was large. The method, however, began to break down when this ratio was small. Nevitt and Hancock (2001) stated that using the bootstrap method was unwise when sample size was less then or equal to 100 because the standard error bias and variability were highly inflated.

Bootstrapping is not considered in the current study for two reasons. First, one of the sample sizes used in this study will be n=100 in order to control the level of Type II error and power for choosing a correct model. The failure of bootstrapping with relatively small sample sizes (Ichikawa & Konishi, 1995; Nevitt & Hancock, 2001) suggests that bootstrapping may not be an appropriate method in this study. Second, bootstrapping may be conservative in its control over Type I error in model rejections at the expense of the power to reject a misspecified model (Nevitt & Hancock, 2001). In this study the likelihood to choose a correct model will be controlled in order to make possible the situations of overfitting and underfitting, thus allowing the effects of preliminary model selection in the two situations to be examined. Therefore, if a simulated parent data set from the true model happens by chance to have characteristics which suggest a misspecified model, then the bootstrap samples are also likely to tend to fit the wrong model rather than the true model.

Bayesian Model Averaging

Bayesian model averaging (BMA) is an alternative method designed to help account for the inherent uncertainty in model selection (Draper, 1995; Hoeting, Madigan, Raftery, & Volinsky, 1999; Raftery, 1996). Instead of choosing a single best model, BMA takes into account model uncertainty by averaging over a variety of plausible competing models. In conducting BMA, the appropriate prior probabilities of the models, and the prior distributions of the parameters given a model, are specified. The data are then used to calculate posterior probabilities for the different models. Afterwards, the posterior model probability for each competing model is considered a weight, and then an average of the model-specific point estimates for a parameter is calculated. The Bayesian point estimate of a parameter is its posterior mean calculated by employing each posterior model probability as a weight.

Although BMA avoids the need to select a single best model and instead helps account for the model uncertainty in the model selection process by mixing several models, there are difficulties in applying BMA. For example, the number of plausible competing models could be very large. In such cases, an arbitrary cut-off point may be used to reduce the number of models by discarding those with low posterior probability (Chatfield, 1995). Another difficulty associated with BMA is setting up an appropriate prior probability for each of the various competing models (Hjort & Claeskens, 2003). Moreover, because BMA does not lead to a single best model but instead averages over a range of entertained models, the description and the interpretation of the estimates across different models turn out to be difficult. Extra caution should be exercised when interpreting a parameter over a set of plausible models (Chatfield, 1995). Finally, competing models associated with BMA are assumed to have common parameters; this is not the case, however, in many SEM model comparisons. In LGM in particular, plausible candidate models usually contain the same variables but may vary in how the variables are connected by the parameters. Therefore, BMA is not considered in the current study.

Data Splitting

In addition to computational methods and BMA, another possible solution to the conditionality problem of using the same data set for both model identification and inference is to conduct model selection, and parameter estimation and inference, on separate sets of the data. Tukey (1980) stated, "Often, confirmation requires a new unexplored set of data" (p. 821). As for how to obtain these new data, however, different researchers have different opinions. Some researchers suggest that model validation needs to be carried out on a completely new set of data. For example, Anscombe (1967) stated that "the only real validation of a statistical analysis, or of any statistical inquiry, is confirmation by independent observation" (p. 6). It is not always possible, however, to collect more data. Some experiments or data collection processes are so costly that it is necessary to derive as much information out of the existing data as possible. Other researchers, for instance Hurvich and Tsai (1992), suggested that data splitting provides a possible substitute for a true replicate sample in model validation. With data splitting, one problem is deciding how to split the sample (Picard & Cook, 1984). Another problem is that fitting a model to just part of the data will result in a loss of efficiency.

What is clear is that using more than one data set, whenever possible, is a wise way to cope with model uncertainty. Miller (1984) and Hurvich and Tsai (1992) recommended data splitting as a possible way to reduce the bias resulting from preliminary model selection. Therefore, in the current study the data splitting technique is employed. The simulated data will be randomly separated into two parts. The first part will be used to choose an appropriate model. The second part will then be used to estimate the parameters and make inferences from the chosen model.

2.2 Approach for Assessing Individual Changes

Understanding how some aspect of an individual changes over time has long been an area of research interest in the social and behavioral sciences, including education (e.g., Bryk & Raudenbush, 1987; Goldstein, 2003), psychology (e.g., Meredith & Tisak, 1990; Willett & Sayer, 1994), and sociology (e.g., Duncan & Duncan, 1996; Patterson, 1993). Such research requires multiple measurements from the same individuals, taken at different times. Different methods can then be used to analyze the longitudinal data, helping researchers assess the within-individual changes, and explaining how the changes may differ across individuals (e.g., Collins & Sayer, 2001; Hancock & Lawrence, 2006).

Traditionally, there are several methods to characterize these changes over time. These include repeated-measures ANOVA, multivariate analysis of variance (MANOVA), univariate analysis of covariance (ANCOVA), multivariate analysis of covariance (MANCOVA), and auto-regressive and cross-lagged multiple regression techniques. These methods, however, either require stringent data assumptions, (e.g., the assumptions of sphericity in the repeated-measures ANOVA), or can only detect changes at the group level but not at the individual level.

Recently, growth curve modeling (GCM) has been developed to help overcome some of these limitations in the traditional approaches. GCM allows one to study a wide range of parameters of change, including linear and nonlinear effects (MacCallum, Kim, Malarkey, & Kiecolt-Glaser, 1997; Mehta & West, 2000; Muthén & Curran, 1997). It simultaneously focuses on their variances, covariances, and mean values over time, providing a more complete picture of changes at both the group level and the individual level (Rogosa & Willett, 1985).

GCM may be carried out either within the framework of multilevel linear modeling (MLM), or within the framework of SEM. MLM is a statistical technique that addresses clustered data (i.e., observations are nested within individuals). In MLM, a multilevel mixed (i.e., with fixed and random effects) regression model is used to study change. The level-1 submodel is specified to capture the trends of the within-individual changes; the level-2 submodel is used to capture the inter-individual differences in growth parameters (Goldstein, 2003; Raudenbush & Bryk, 2002; Singer & Willett, 2003).

SEM, on the other hand, is a statistical technique for testing and estimating hypothesized causal relationships among observed and latent variables (Hoyle, 1995). In SEM, the growth parameters are treated as latent variables. Repeated measures are treated as multiple indicators of the latent variables in order to capture the trends of changes (Meredith & Tisak, 1990; Willett & Sayer, 1994).

Under SEM, modeling changes may take one of several names, such as latent growth modeling (LGM), latent growth curve analysis, or latent trajectory models. Numerous reviews of LGM for modeling change have been published in recent years (Bollen & Curran, 2006; Duncan, Duncan, & Strycker, 2006; Hancock & Lawrence, 2006).

A latent growth model can be represented in matrix notation in terms of a data model, a covariance structure, and a mean structure. The data model is as follows:

$$\mathbf{y} = \boldsymbol{\tau} + \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon},$$

where $\mathbf{y}$ represents the vector of observation; $\boldsymbol{\tau}$ represents the vector of intercept; $\boldsymbol{\Lambda}$ represents the factor loadings; $\boldsymbol{\eta}$ represents the latent growth variable; $\boldsymbol{\varepsilon}$ represents the error term**.**

From the data model, one can derive a covariance structure (Σ) and a mean structure (μ). The covariance structure is:

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}\,\boldsymbol{\Phi}\,\boldsymbol{\Lambda}' + \boldsymbol{\Theta}_{\varepsilon},$$

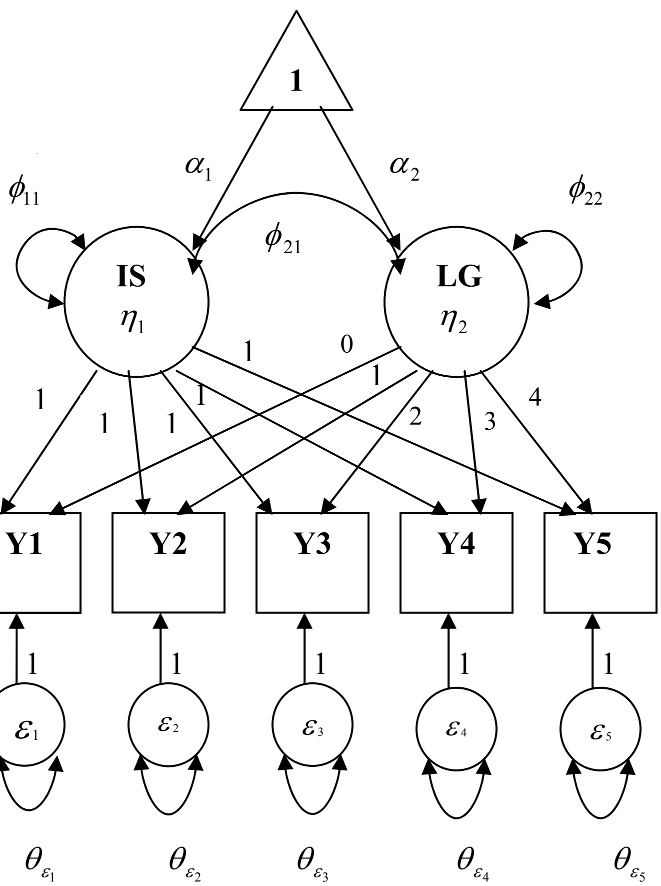
where $\boldsymbol{\Sigma}$ represents the population variances and covariances of the observed variables**;** $\boldsymbol{\Lambda}$ represents the factor loading**;** $\boldsymbol{\Phi}$ represents the factor variances and covariances; $\boldsymbol{\Theta}_{\varepsilon}$ represents the error variances and covariances. The mean structure is:

$$\mu = \boldsymbol{\tau} + \boldsymbol{\Lambda}\,\boldsymbol{\alpha},$$

where μ represents the population means of observed variables; $\boldsymbol{\tau}$ is the vector of intercepts; $\boldsymbol{\Lambda}$ is factor loadings; $\boldsymbol{\alpha}$ is the latent variable means. In LGM, the parameters of interest are contained in the matrices $\boldsymbol{\Lambda}$, $\boldsymbol{\Phi}$, and $\boldsymbol{\Theta}_{\varepsilon}$, and the vector $\boldsymbol{\alpha}$.

Figure 1 shows a complete path diagram for a typical linear LGM for five equally spaced time points.

Figure 1. Linear Latent Growth Model



This linear latent growth model contains the following elements:

(1) Two latent growth factors: Initial Status (IS)/Intercept and Linear Growth (LG)/ Slope. IS (η_1) represents the intercept of an individual's growth trajectory and

conveys the individual's true status at the initial measured time point. LG (η_2) represents the slope of an individual's growth trajectory and provides the individual's true rate of change per unit time.

(2) A pseudovariable, which assumes a constant value of 1 for all observations. The inclusion of a pseudovariable in the latent growth model allows for factor means to be estimated.

(3) Five outcome variables. *Y1-Y5* represent five continuous outcomes measured at equally spaced time points.

(4) Five error terms. ε_1 - ε_5 represent the degree of deviation between the observed outcome and the expected outcome from the latent growth model.

(5) Loadings and parameters to be estimated. The parameters and loadings are presented in the matrices $\mathbf{\Lambda}$, $\mathbf{\Phi}$, the vector $\boldsymbol{\alpha}$, and the matrix $\mathbf{\Theta}_{\boldsymbol{\varepsilon}}$. The details are described below.

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix},$$

$$\mathbf{\Phi} = \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix},$$

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix},$$

$$\mathbf{\Theta_\varepsilon} = \begin{bmatrix} \theta_{\varepsilon_1} & & & & \\ 0 & \theta_{\varepsilon_2} & & & \\ 0 & 0 & \theta_{\varepsilon_3} & & \\ 0 & 0 & 0 & \theta_{\varepsilon_4} & \\ 0 & 0 & 0 & 0 & \theta_{\varepsilon_5} \end{bmatrix}.$$

Elements of $\mathbf{\Lambda}$ represent the loadings. The first column is fixed to 1, reflecting the fact that each individual's intercept remains constant over the repeated measures; the second column reflects the hypothesis of linear growth with equal time intervals.

Elements of $\mathbf{\Phi}$ represent the variances and covariances of these aspects of change. In the case of linear LGM shown in Figure 1, the $\mathbf{\Phi}$ matrix contains the intercept variance (ϕ_{11}), the slope variance (ϕ_{22}), and the covariance of intercepts and slopes (ϕ_{21}). The intercept variance (ϕ_{11}) displays how much diversity exists initially for the ability of interest. The slope variance (ϕ_{22}) conveys the diversity in growth rates across individuals. The covariance of intercepts and slopes (ϕ_{21}) shows to what extent the rate of growth is related to the initial status of the ability.

Elements of the vector $\boldsymbol{\alpha}$ represent the mean of intercept (α_1) and the mean of slope (α_2). In contrast to the covariance structure, the mean structure contains information about change at the aggregate level. The mean of intercept (α_1) captures the average initial status and the mean of slope (α_2) represents the average growth rate across different times.

Elements of $\mathbf{\Theta}_{\varepsilon}$ represent variances and covariances of error terms, indicating the portion of the variance in the data not associated with the hypothesized latent curve.

As mentioned previously, growth may be modeled within either an MLM or LGM paradigm, each of which has advantages and disadvantages. MLM is better at incorporating levels of clustered data (Wu, West, & Taylor, 2009). MLM also handles more easily the case where people are measured at different time points (Mehta & West, 2000). LGM, on the other hand, has two advantages in modeling changes. First, in LGM, error variances and covariances may be estimated freely or specified to conform to a predetermined pattern, whereas in MLM error variances are constrained to remain equal over time. Second, LGM is able to model a latent outcome variable with multiple indicators at each time point (Hancock, Kuo, & Lawrence, 2001) and to include other measured or latent variables that can serve as correlates, predictors, or outcomes of the latent growth parameter (Meredith & Tisak, 1990; Muthén & Curran, 1997).

MLM and LGM approaches provide the same analsis results when modeling linear growth with homoscedastic residuals (Hox, 2000), whereas the empirical results of LGM and MLM may not necessarily be the same because more flexibility exists in LGM. In this study, the LGM approach will be used because LGM offers more flexibility in testing a nonlinear growth hypothesis and LGM has flexibility in the specification of the variances and covariances of the repeated measurements. By implementing LGM in this study, both linear and nonlinear growth models will be easily set up to examine how the information criteria measures perform to distinguish the two models in the context of model misspecification.

2.3 Model Fit in SEM

The principal considerations in SEM are evaluating model fit and estimating individual model parameters. The overall evaluation of the fit of a model in SEM is obviously important; concern over individual model parameters is pointless if a hypothesized model is not consistent with the data.

A common approach in SEM is to test the underlying structure of hypothesized models and to report some index of the goodness of fit of those models to the data. Goodness of fit is the empirical correspondence between a model's predictions and observed data. The concept of evaluating model fit in SEM can be described as follows: suppose the formal representation of the model is $\Sigma = \Sigma(\theta)$, where θ represents the parameters of the model, which are traditionally specified as freely estimated and/or fixed to specific values. The fitting process involves finding a set of parameter estimates $\hat{\theta}$ which minimize F, the maximum likelihood discrepancy function.

$$F = [\ln|\hat{\mathbf{\Sigma}}| + \mathrm{tr}(\mathbf{S}\hat{\mathbf{\Sigma}}^{-1}) - \ln|\mathbf{S}| - p] + (\mathbf{m} - \hat{\boldsymbol{\mu}})'\hat{\mathbf{\Sigma}}'^{-1}(\mathbf{m} - \hat{\boldsymbol{\mu}}),$$

where $\mathbf{S}$ is the observed covariance matrix, $\hat{\mathbf{\Sigma}}$ is the model implied covariance matrix based on optimum parameter estimates, p is the number of indicator variables, $\mathbf{m}$ is the vector of observed sample means of the indicator variables, and $\hat{\boldsymbol{\mu}}$ is the model-implied mean vector. $[\ln|\hat{\mathbf{\Sigma}}| + \mathrm{tr}(\mathbf{S}\hat{\mathbf{\Sigma}}^{-1}) - \ln|\mathbf{S}| - p]$ is the fit associated with the covariance structure portion of the model and $(\mathbf{m}-\hat{\boldsymbol{\mu}})'\hat{\Sigma}'^{-1}(\mathbf{m}-\hat{\boldsymbol{\mu}})$ is the fit associated with the mean structure portion of the model.

The estimates $\hat{\theta}$ minimizing F yield an implied covariance matrix ($\hat{\boldsymbol{\Sigma}}$) as similar as possible to the observed covariance matrix (**S**) of measured variables, and an implied mean vector ($\hat{\boldsymbol{\mu}}$) as similar as possible to the observed sample means vector (**m**) of the indicator variables. To the degree that $\hat{\boldsymbol{\Sigma}}$ resembles **S**, the minimized F will tend to be small, reflecting good fit. The situation when $\hat{\boldsymbol{\Sigma}} = \mathbf{S}$, $\hat{\boldsymbol{\mu}} = \boldsymbol{m}$, and $F = 0$ denotes perfect fit. If the match between the model's predictions and observed data is deemed adequate (by reaching or exceeding some benchmark) the model is said to show good fit (Preacher, 2006).

Specifically, the evaluation of goodness of fit for a latent growth model is generally carried out by assessing various global fit indices including (1) the chi-square statistic, (2) information criteria (e.g., AIC and BIC), and (3) data-model fit indices such as the root mean squared error of approximation (RMSEA; Browne & Cudeck, 1993), Nonnormed Fit Index (NNFI; Tucker & Lewis, 1973), and comparative fit index (CFI; Bentler, 1990). Generally, a good fit is indicated by the values of RMSEA less than 0.05 or NNFI and CFI greater than 0.95.

In the process of evaluating model fit in LGM, there are generally two main specification issues that need to be addressed (Kline, 2004). First, how many growth parameters need to be included in the model? That is, assuming there is change over time, is a latent linear growth factor sufficient to model change? Or is it necessary to also include a nonlinear factor? The second specification issue is related to the error variances and covariances. For example, are the errors independent over time? If not, what is the pattern of correlation between the errors of the repeated measures variable? Consequently, by taking into account the various combinations of possible growth factors and errors, several possible models may be considered at the same time. When a fit index is used to evaluate different plausible models, the index becomes a model selection criterion because the objective is to select the model that is optimal, given the data.

2.4 Model Selection in SEM

In SEM, it is common that in the model-building process several plausible candidate models might be considered and fitted to a single data set to see whether the fit can be improved. Fit indices are used to evaluate the plausible models, and then a single appropriate model is selected. Based on the selected model, parameter estimation and testing are conducted using the same data set.

It is obvious that model selection plays an important role in SEM. The literature regarding model selection in SEM, however, mostly focuses on evaluating model selection criteria/model fit indices themselves. For example, Coffman and Millsap (2006) showed that the global fit indices (e.g., the chi-square statistic, RMSEA, and CFI) which are generally used to evaluate the overall fit of a structural equation model can be misleading within the context of LGM. Therefore, they examined the usefulness of assessing individual fit in latent growth models and concluded that the evaluation of model fit at the level of the individual in LGM is an important addition to the assessment of the overall model. As another example, Wu et al. (2009) discussed the issues that arise in the evaluation of fit of latent curve growth modeling from the perspectives of both the SEM and MLM frameworks. First, they showed how the four sources of misfit in latent growth models - two related to the mean structure and two related to the covariance structure - can be reflected in fit indices from the SEM and MLM frameworks. Second, they stated that the availability and interpretation of measures of model fit depend on the type of longitudinal data (i.e., balanced on time with complete data, balanced on time with data missing at random, and unbalanced on time) being analyzed.

In addition, in SEM, some studies have examined whether an index reliably identifies the true model and consistently identifies a specific model across replications. For instance, Whittaker and Stapleton (2006) assessed the performance of eight cross-validation indices in terms of true model selection rate as well as consistency of model selection under different conditions including sample size, factor loading, model misspecification, and nonnormality. They suggested that the performance of the cross-validation indices tended to improve as factor loading and sample size increased but performed less well as nonnormality increased.

In SEM, the literature regarding model selection has mostly focused on evaluating the model selection criteria/model fit indices themselves. The effect of model selection, however, has not been widely investigated. This study, therefore, is designed to evaluate the performance of the information criteria used for model selection in LGM (a particular case of SEM) and to examine the impact of preliminary model selection on latent growth model parameter estimates. Also, data splitting is to be assessed as a possible way to mitigate the effects of model uncertainty. The details of methodology are elaborated in the following Chapter 3.

Chapter 3: Methodology

The research design is explicated in this chapter. The first section of this chapter describes the Monte Carlo simulation study, including the population from which the data will be drawn, the manipulated factors, the data generation procedures, and the data splitting procedure. The second section presents the outcome measures and data analysis.

3.1 Monte Carlo Simulation Study

To achieve the goals of this study, a series of Monte Carlo simulations are conducted in which many samples are drawn from populations with known values for the parameter estimates. This approach has the benefit of having a known growth pattern and known population values as a baseline for evaluating the ability to select a true model. Also, it takes into account the effects of sampling variability on parameter estimates and provides information on whether parameter estimates can be recovered.

3.1.1 Specifications in the Monte Carlo Study

In the Monte Carlo study, two kinds of variables need to be specified: the Monte Carlo variables and the population variables. The Monte Carlo variables include sample size and the number of replications. The population variables, which determine the generation of the sample data, include the number of latent factors in a model, the mean of the factors, the variance and covariance of the factors and errors, and the loadings. Among those variables, the number of replications, the number of latent factors in a model, the mean of the factors, the variance of intercept, the variance of slope, the covariance of the factors and errors, and the loadings are constant. Data-generating model, sample size, the variance of the quadratic factor, error variance, and model selection criteria are manipulated. The specifications are presented in Table 1 and explicated in 3.1.1.1 and 3.1.1.2.

Table 1. Specifications in the Monte Carlo Study

Data-Generating Model *	Linear LGM	Quadratic LGM
Parameters		
Mean intercept (α_1)	10	10
Mean linear slope (α_2)	25	25
Mean quadratic slope (α_3)	N/A	-0.1
Mean error	0	0
Intercept variance (ϕ_{11})	10	10
Linear slope variance (ϕ_{22})	2	2
Quadratic slope variance* (ϕ_{33})	N/A	0.01, 0.05, and 0.1
Error variance * (θ_ε)	2.5, 5, and 10	2.5, 5, and 10
Intercept/linear slope covariance (ϕ_{21})	-1.34164	-1.34164
Intercept/ quadratic slope covariance (ϕ_{31})	0	0
Linear/ quadratic slope covariance (ϕ_{23})	0	0
Loadings of intercept	[1,1,1,1,1]	[1,1,1,1,1]
Loadings of linear slope	[0,1,2,3,4]	[0,1,2,3,4]
Loadings of quadratic slope	N/A	[0,1,4,9,16]
Sample size *	100,200, 500, and 700	100,200, 500, and 700
Replications	1000	1000

* Manipulated factors in the Monte Carlo simulation study

3.1.1.1 Constant Factors

Means of Growth Factors

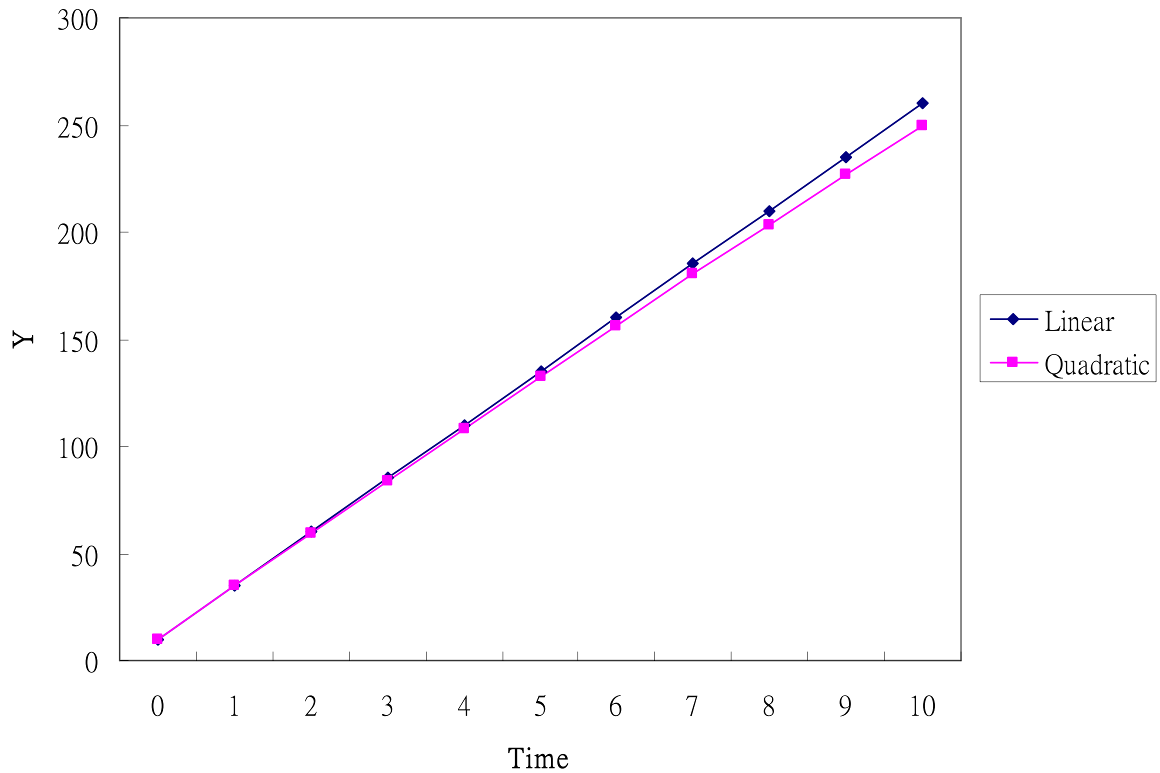
Data are generated by using the following population values for the latent growth factors with the assumption that variables are multivariate normally distributed. There are two latent factors in the linear growth model and three in the quadratic growth model. For both the linear and the quadratic growth models, the mean of the intercept factor (α_1) is 10 and the mean of the linear slope factor (α_2) is 25. The mean of the quadratic slope factor (α_3) is set at -0.1. Matrix representations for the means of growth factors in the linear and quadratic models used in this study are as follows:

$$\boldsymbol{\alpha}_{Linear} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 25 \end{bmatrix},$$

$$\boldsymbol{\alpha}_{Quadratic} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 25 \\ -0.1 \end{bmatrix} .$$

In mathematics, the difference between a straight line and a quadratic curve depends only on the magnitude of the quadratic variable coefficient. The greater the magnitude of the quadratic variable coefficient, the greater the discrepancy between a straight line and a quadratic curve. The pilot study showed that the magnitudes of the mean of the intercept factor and the mean of the slope factors did not significantly influence the probability of misspecifying between linear and quadratic models. When the magnitude of the mean of the quadratic slope factor increased, however, the probability of misspecifying between linear and quadratic growth models decreased, because the difference between the two models was more obvious. More specifically, the pilot study indicated that when the quadratic factor mean was of greater magnitude (e.g., -0.2, -0.3, -0.5, -1), the power to choose a correct model was greater than 0.99. In those cases, the effects of preliminary model selection could not be examined because a correct model was almost always chosen. Therefore, in this study relatively large values (10 and 25) for the means of the intercept and slope factors were chosen to contrast with the small quadratic factor mean (-0.1). The small magnitude of the mean of the quadratic slope factor makes the data generated by the quadratic growth model not dramatically different from those generated by the linear growth model and thus creates the possibility of model misspecification, allowing for examination of the effects of preliminary model selection on parameter estimates. Also, the negative value of the mean of the quadratic slope factor makes the quadratic growth trajectory reflect the general pattern of learning curves that display negative acceleration of changing rate, i.e., quick progress in learning during the initial stages followed by gradually slower improvement over time. Figure 2 shows that the trajectories of the linear model and the quadratic model are similar to each other.

Figure 2. Linear and Quadratic Growth Trajectories



Variance/Covariance of Growth Factors

The variance of the intercept factor is 10 and the variance of the slope growth factor is 2, reflecting a commonly seen variance ratio of 5:1 (Muthén & Muthén, 2002). The covariance between the intercept and slope growth factors is -1.34164, reflecting a moderate correlation of -0.3. This parameter value is based on Hancock and Lawrence's (2006) study. The negative intercept-slope covariance indicates that the object of interest at the initial time point will have a systematic negative relationship with the rate of change over time. This implies that individuals with lower initial status on the ability of interest tend to grow more whereas those with higher initial status on the ability of interest tend to grow less. This correlates with the general pattern of individual learning processes, first speeding up and then slowing down as the practically achievable level of improvement or the upper limit of measurements (e.g., a test ceiling) is reached.

Loadings

The loadings [1, 1, 1, 1, 1] are set for the intercept factor and [0, 1, 2, 3, 4] for the linear slope factor in both the linear and quadratic growth models. The loadings for the quadratic factor are [0, 1, 4, 9, 16] in the quadratic growth model. Matrix representations for the loadings in the linear model and the quadratic model used in this study are as follows:

$$\mathbf{\Lambda}_{Linear} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix},$$

$$\mathbf{\Lambda}_{Quadratic} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{bmatrix}.$$

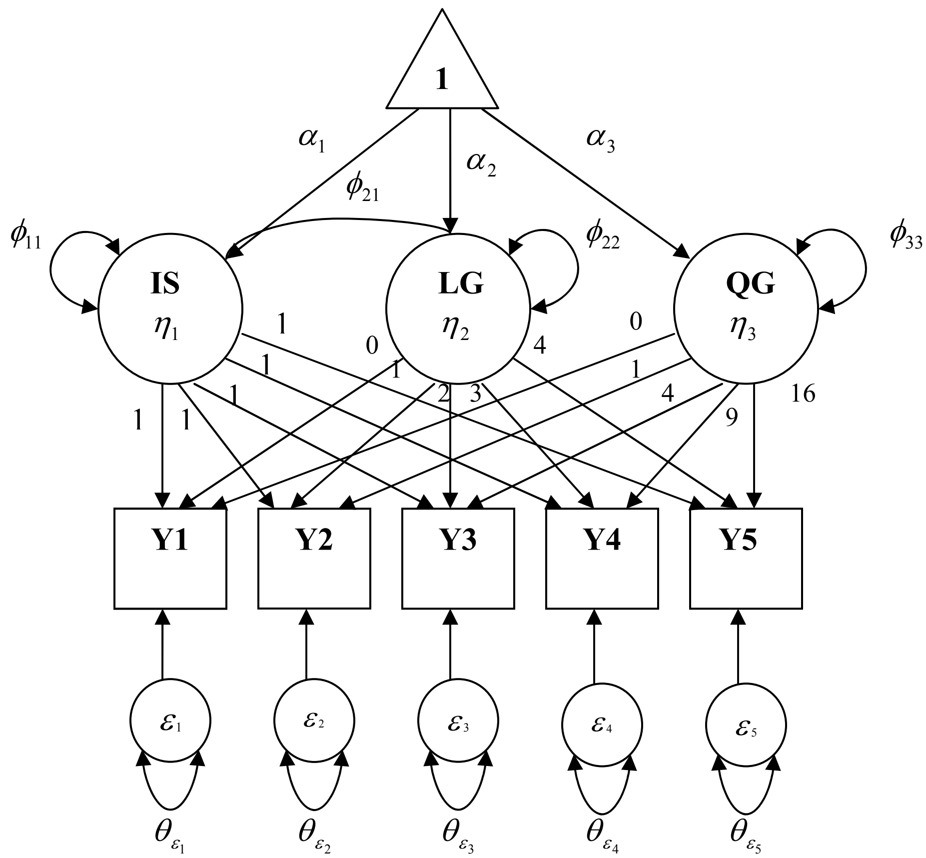
3.1.1.2 Manipulated Factors

Five factors will be manipulated in this study. They are the data-generating model, sample size, variance of quadratic factor, variance of error, and model selection criteria. The purpose of manipulating these factors is to control the level of type II error and power for choosing a correct model. In this current study, the likelihood of choosing a correct model is controlled in order to make possible the situations of overfitting and underfitting. For example, manipulating the magnitude of quadratic factor variance makes it possible to select a quadratic growth model even when data are generated from a linear growth model and vice versa. Consequently, the effects of preliminary model selection in underfitting and overfitting situations can be examined. A pilot study was conducted to choose parameter values that would manipulate how often the wrong model was selected. The range of parameter values to be used for a certain manipulated factor was decided by fixing the level of power and keeping the values of all other manipulated factors the same. The details of the pilot study will be described in the following sections discussing each of the manipulated factors.

Data-Generating Model

The first manipulated factor is the type of data-generating model. A linear growth model and a non-linear growth model are considered in this study. The linear growth model includes two latent growth factors: Initial Status (IS)/Intercept and Linear Growth (LG)/Slope, and assumes the trajectories examined are simple linear functions of time. The non-linear model includes these same two latent growth factors plus a third latent growth factor, Quadratic Growth (QG), which captures a quadratic trend over time. Figure 1, in section 2.2, illustrates the linear growth model used in this study with five repeated measures and with covariance and mean structure parameters. Figure 3 illustrates the non-linear growth model, in this case a quadratic model, which assumes that growth is governed by two factors, linear and quadratic.

Figure 3. Quadratic Latent Growth Model



These models were chosen for two reasons. First, the difference between these two models makes it possible to examine how each information criterion performs to distinguish the two models in the context of model misspecification. Second, these two models were chosen because they are nested. In this study, two types of misspecification will be examined: overfitting (i.e., the quadratic growth model is selected when the liner growth model is true) and underfitting (i.e., the linear growth model is selected when the quadratic growth model is true). Using two nested models makes it possible to determine in what circumstances overfitting or underfitting tends to occur, and to what extent model misspecification influences parameter estimation in overfitting and underfitting circumstances.

Sample Size

The second manipulated factor is the sample size. Four sample sizes are used: 100, 200, 500, and 700. These are numbers of observations typically seen in practice. The results of pilot analyses, shown in Figure 4, suggested that sample size ranged from 15 to 351 for power fixed at the values of 0.2, 0.5, and 0.8 as variance of the quadratic factor ranged from 0.01 to 0.1 and with error variance fixed at 5. The pilot study also showed, as illustrated in Figure 5, that sample size ranged from 14 to 699 for power fixed at the values of 0.2, 0.5, and 0.8 as error variance ranged from 1.3 to 10 and with variance of the quadratic factor fixed at 0.01. In this current study, however, sample size less than 100 is not used in order to ensure sufficient observations for data splitting.

Figure 4. Sample Size by Power and Variance of the Quadratic Factor

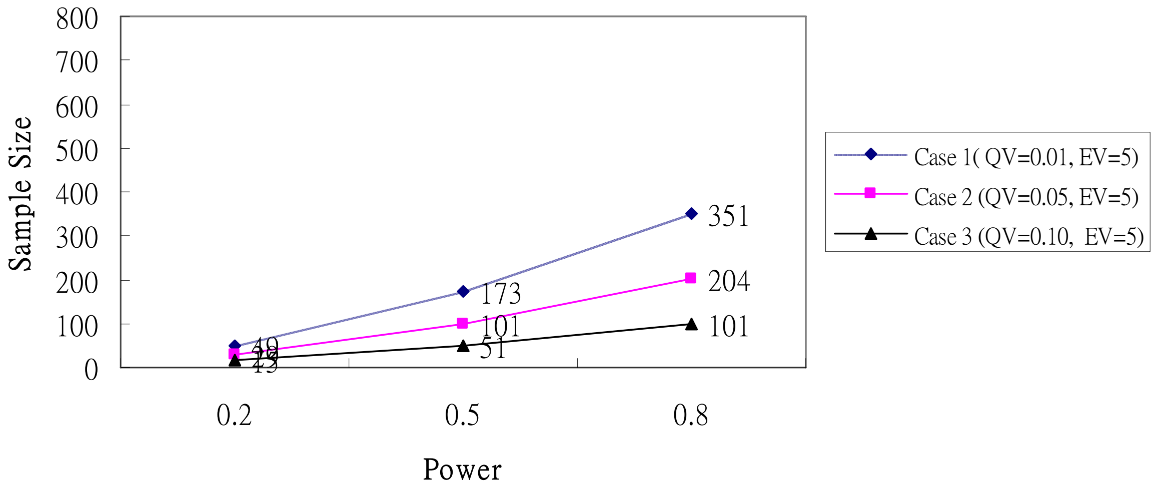
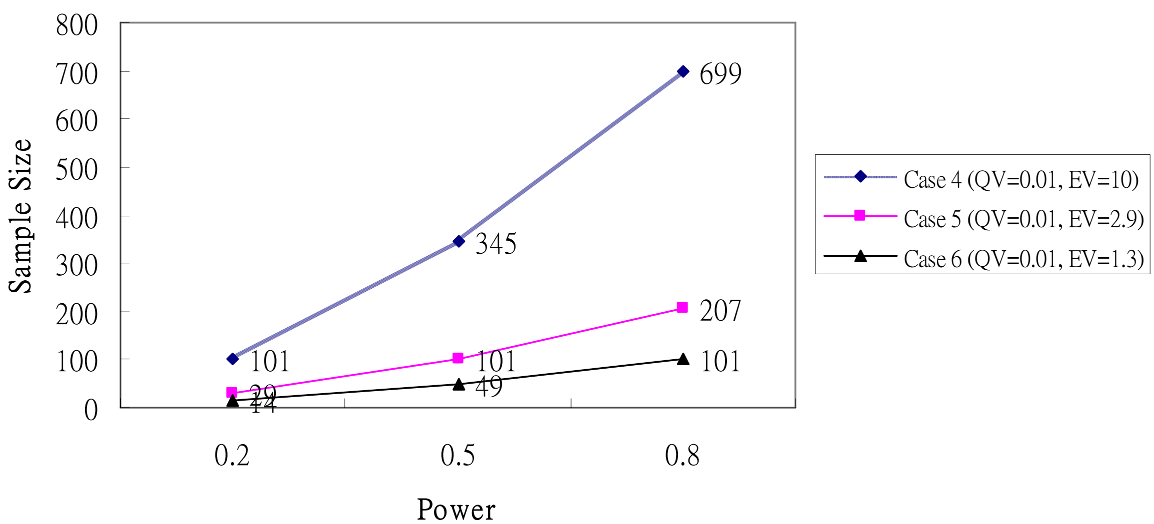


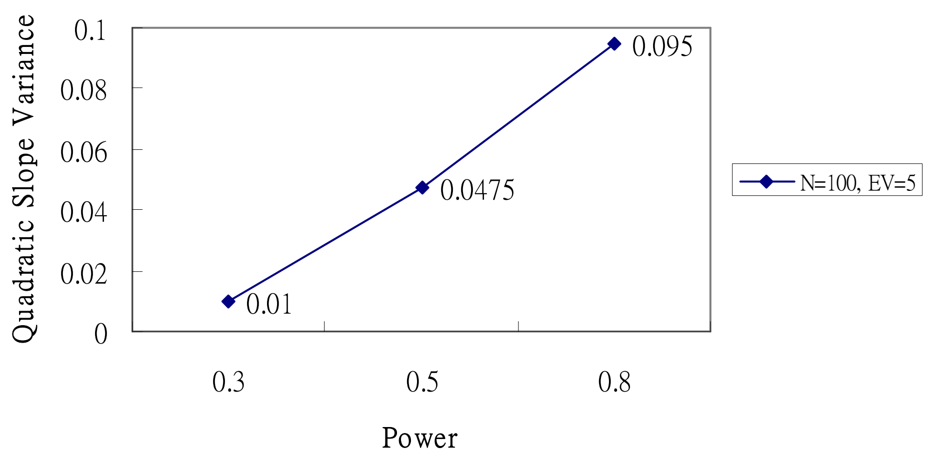
Figure 5. Sample Size by Power and Error Variance



Variance of the Quadratic Factor

The third manipulated factor is variance of the quadratic factor (ϕ_{33}). Three values are used: 0.01, 0.05, and 0.10. Pilot analysis suggested that the larger the variance of the quadratic factor, the higher the power to choose the quadratic model when it is true. Figure 6 presents the possible values for variance of the quadratic factor with sample size 100 and an error variance of 5.

Figure 6. Values for Variance of the Quadratic Factor by Power



The variance of the quadratic factor ranged from 0.01 to 0.095 for fixed power at the values of 0.3, 0.5, and 0.8 with sample size of 100 and an error variance of 5. Our pilot study also showed that with sample size larger than 100 the power is generally greater than 0.8 when variance of the quadratic factor is 0.01 to 0.095. Thus, conditions with variance of quadratic factor significantly greater than 0.095 are deemed not to create sufficient model misspecification for this current study to examine the effects of model uncertainty on parameter estimates and therefore are not analyzed further. The factor variance/covariance matrices for the two models, then, are:

$$\boldsymbol{\varphi}_{Linear} = \begin{bmatrix} \phi_{11} & \\ \phi_{21} & \phi_{22} \end{bmatrix} = \begin{bmatrix} 10 & \\ -1.34 & 2 \end{bmatrix},$$

$$\boldsymbol{\varphi}_{Quadratic} = \begin{bmatrix} \phi_{11} & & \\ \phi_{21} & \phi_{22} & \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}$$

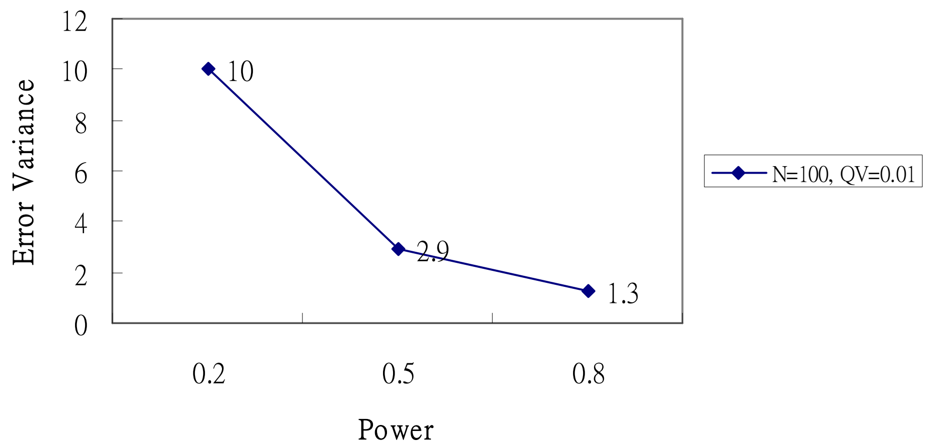
$$= \begin{bmatrix} 10 & & \\ -1.34 & 2 & \\ 0 & 0 & 0.01 \end{bmatrix} \text{ or } \begin{bmatrix} 10 & & \\ -1.34 & 2 & \\ 0 & 0 & 0.05 \end{bmatrix} \text{ or } \begin{bmatrix} 10 & & \\ -1.34 & 2 & \\ 0 & 0 & 0.1 \end{bmatrix}.$$

Error Variance

The fourth manipulated factor is the error variance. Three values are used: 2.5, 5, and 10. Pilot analysis suggested that the smaller the error variance, the greater the power to choose the quadratic model when it is true. The error variance cannot, however, be set to zero; otherwise, the variance and covariance matrix will be singular. Figure 7 presents the values of the error variance with sample size 100 and with the variance of the quadratic factor 0.01 at different levels of power.

The error variance ranged from 10 to 1.3 with power 0.2 to 0.8 respectively, with sample size 100 and the variance of the quadratic factor of 0.01. Our pilot study also showed that with sample size larger than 100 or variance of the quadratic factor greater than 0.01 power values are consistently higher than 0.9 when the error variance is 2.9. As such, only the range of 2.5 to 10 for the error variance is considered in this current study in order to create sufficient model misspecification.

Figure 7. Values of the Error Variance by Power



The error variance is varied from 2.5 to 5 to 10, reflecting 0.8, 0.67, and 0.5 growth curve reliability (R^2 values) at the initial time point for both the linear and quadratic growth models. The growth curve reliability (R^2 values) is defined as the proportion of total variance at a certain time point explained by the growth curve factors (Muthén & Muthén, 2002). In the present study, the R^2 value is computed by using the following formula,

$$R^2(Y_t) = (\phi_i + x_t^2\phi_s + x_t^4\phi_q + 2x_t\phi_{is} + 2x_t^2\phi_{iq} + 2x_t^3\phi_{sq})/(\phi_i + x_t^2\phi_s + x_t^4\phi_q + 2x_t\phi_{is} + 2x_t^2\phi_{iq} + 2x_t^3\phi_{sq} + \theta_t)$$

where ϕ_i is the intercept variance, x_t is the time score at time t, ϕ_s is the linear slope variance, ϕ_q is the quadratic slope variance, ϕ_{is} is the intercept/linear slope covariance, ϕ_{iq} is the intercept/quadratic slope covariance (set at zero in this study), ϕ_{sq} is the linear/quadratic slope covariance (set at zero in this study), and θ_t is the error variance for the outcome at time t. Here the x_t time scores are chosen as 0, 1, 2, 3, and 4. Muthén and Muthén (2002) reported that the R^2 value of the outcome variable ranged from 0.5 to 0.74. The R^2 value of the outcome variables in the current study ranges from 0.5 to 0.96 as presented in Table 2.

Table 2. Growth Curve Reliability (R^2) at Each Time Point

Model	Error Variance	Quadratic Slop Variance	Outcome Variable				
			Y1	Y2	Y3	Y4	Y5
Linear	EV=2.5		0.80	0.79	0.83	0.89	0.93
	EV=5		0.67	0.65	0.72	0.80	0.86
	EV=10		0.50	0.48	0.56	0.67	0.76
Quadratic	EV=2.5	QV=0.01	0.80	0.79	0.84	0.89	0.93
		QV=0.05	0.80	0.79	0.84	0.91	0.95
		QV=0.10	0.80	0.79	0.85	0.92	0.96
	EV=5	QV=0.01	0.67	0.65	0.72	0.81	0.87
		QV=0.05	0.67	0.65	0.73	0.83	0.90
		QV=0.10	0.67	0.65	0.74	0.85	0.92
	EV=10	QV=0.01	0.50	0.48	0.56	0.67	0.77
		QV=0.05	0.50	0.48	0.57	0.71	0.82
		QV=0.10	0.50	0.48	0.59	0.74	0.85

Although the error variances are manipulated, the specific error variances set in a model are equal over time. Also, the error covariances are fixed to zero to represent the hypothesis that errors are uncorrelated over time. In other words, homoscedastic and independent error variance is assumed. The error variance/covariance matrices for the two models are as follows:

$$\mathbf{\Theta}_{Linear\&Quadratic} = \begin{bmatrix} \theta_{\varepsilon_1} & & & & \\ 0 & \theta_{\varepsilon_2} & & & \\ 0 & 0 & \theta_{\varepsilon_3} & & \\ 0 & 0 & 0 & \theta_{\varepsilon_4} & \\ 0 & 0 & 0 & 0 & \theta_{\varepsilon_5} \end{bmatrix}$$

$$= \begin{bmatrix} 2.5 & & & & \\ 0 & 2.5 & & & \\ 0 & 0 & 2.5 & & \\ 0 & 0 & 0 & 2.5 & \\ 0 & 0 & 0 & 0 & 2.5 \end{bmatrix} \text{or} \begin{bmatrix} 5 & & & & \\ 0 & 5 & & & \\ 0 & 0 & 5 & & \\ 0 & 0 & 0 & 5 & \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \text{or} \begin{bmatrix} 10 & & & & \\ 0 & 10 & & & \\ 0 & 0 & 10 & & \\ 0 & 0 & 0 & 10 & \\ 0 & 0 & 0 & 0 & 10 \end{bmatrix}$$

Model Selection Criteria

The two model selection criteria considered in this study are AIC and BIC. For each cell of this study, these two measures are used to select a model with better goodness of fit. Based on the log likelihood of a fitted model, both AIC and BIC take into account the statistical goodness of fit. Each criterion, however, applies a different penalty. AIC penalizes the number of parameters to be estimated whereas BIC penalizes both the number of parameters and sample size. With different penalty terms, AIC and BIC might choose different models. In this study, the performance of both AIC and BIC will be evaluated.

In summary, by manipulating the five factors noted in Section 3.1.1.2 above, the level of Type II error for selecting a wrong model and the power for choosing a correct model were controlled. In this way, model misspecifications (i.e., overfitting and underfitting) were made possible. The values of manipulated factors were chosen based on the results of pilot analyses, using sample sizes commonly seen in practice. Manipulating these factors resulted in a total of 96 different conditions, as shown in Table 3.

Table 3. Factor Manipulation

Factors	Condition Specification	No. of Conditions	No. of Conditions
Data-generating model	Linear and quadratic growth models	Linear growth model	Quadratic growth model
Model selection criteria	AIC and BIC	2	2
Sample size	100, 200, 500, 700	4	4
Error variance	2.5, 5, 10	3	3
Variance of quadratic factor	0.01, 0.05, 0.1	--	3
Total		24	72

In the case of underfitting, data generated by a quadratic growth model were fitted to a linear model, and the power to choose a correct model (the quadratic model in this case) by AIC was evaluated. Table 4 shows the results. The power ranges from 0.203 to 0.999 under the different conditions created by factor manipulation.

Calculation of the power of the tests that compare the quadratic growth model (i.e., the full model) and the linear growth model (i.e., the reduced and the null model) in the case of underfitting is based on a method suggested by Saris and Satorra (1993) and Hancock (2006). As an example, in Table 4, with a sample size of 100, error variance of 2.5, and variance of quadratic growth factor of 0.01, the power to choose a quadratic growth model when it is true is 0.537. This result is obtained by the procedure described below.

Table 4. Power to Choose Quadratic Model by AIC in the Case of Underfitting

Sample Size	Error Variance	Variance of Quadratic Factor		
		QV=0.01	QV=0.05	QV=0.10
N=100	EV= 2.5	0.537	0.749	0.959
	EV= 5	0.322	0.460	0.733
	EV= 10	0.203	0.265	0.426
N=200	EV= 2.5	0.833	0.966	0.999
	EV= 5	0.550	0.750	0.960
	EV= 10	0.326	0.448	0.708
N=500	EV= 2.5	0.997	0.999	0.999
	EV= 5	0.916	0.989	0.999
	EV= 10	0.651	0.827	0.981
N=700	EV= 2.5	0.999	0.999	0.999
	EV= 5	0.978	0.999	0.999
	EV= 10	0.796	0.933	0.998

First, the quadratic growth model with population values described above was set as the true model and the model-implied covariance matrix for this specification was calculated. Second, the model-implied covariance matrix generated by the quadratic growth model was fitted into the linear growth model. The linear growth model yielded a model fit function value F, which was 0.05458 in this run. The estimated noncentrality parameter corresponding to the test of the reduced model was (N-1) * F, which was 5.40342 in this run. Third, the likelihood of choosing the quadratic growth model (which is the true model) was calculated. Using the value of the estimated noncentrality parameter (5.40342) and the tables for the noncentral χ^2, and taking into account that the number of degrees of freedom for the comparison of the quadratic and the linear models ($df_{diff} = df_L - df_Q$) is 4, and that the critical value is 8, the power of the model test would be 0.537. This result suggested that in this specific case of underfitting (i.e., when the quadratic growth model is the true model to generate the data but the data are fit into the linear growth model), there would be a 53.7 % chance of choosing the quadratic growth model and a 46.3 % chance of choosing the linear growth model by using AIC.

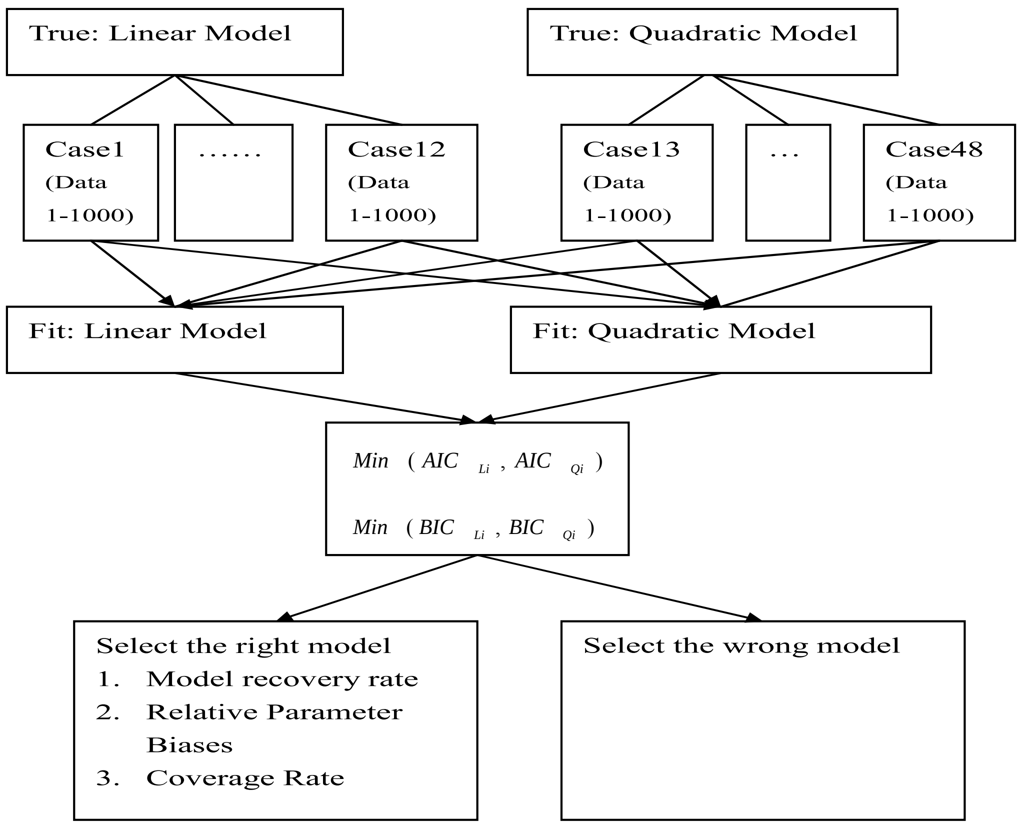
3.1.2 Monte Carlo Simulation Procedures

The current investigation includes two Monte Carlo simulation studies. Study 1 presents a simulation study designed to evaluate the impact of preliminary model selection on latent growth model parameter estimates when the choice of the model and the subsequent parameter estimation procedures are based on the same data set. Study 1 also explores the relative behavior of AIC and BIC on specific sets of data. Study 2 presents a simulation study to assess the method of data splitting to mitigate the effects of model uncertainty. The procedures of Monte Carlo simulation used in the two studies are described as follows.

Study 1. Conduct Model Selection and Parameter Estimation on the Same Data Set

The procedure for study 1 is displayed in Figure 8 and described below.

Figure 8. Procedure for Study 1



Step 1: Setting Conditions for Simulation

(1) Selecting a true model. First a linear growth model and then a quadratic growth model is considered as a true model in this study.
(2) Setting parameter values for that true model. The parameters are the elements of the means and the variance/covariance matrices. The values of the parameters used in this study are shown in Table 1.
(3) Setting sample size. With the consideration of power to retain the null model, 100, 200, 500, and 700 are implemented for linear and quadratic models.

For illustration purposes, Figure 9 presents the plots for the data generated under different manipulated conditions when sample size is fixed to 100.

Step 2: Running the Simulation

The simulation is run according to the specifications described above, with 1,000 replications. The details are described as follows:
(1) An R program is used to generate data for 48 different conditions. Table 5 summarizes the conditions of factor manipulation and data generation.

Table 5. Conditions of Data Generation

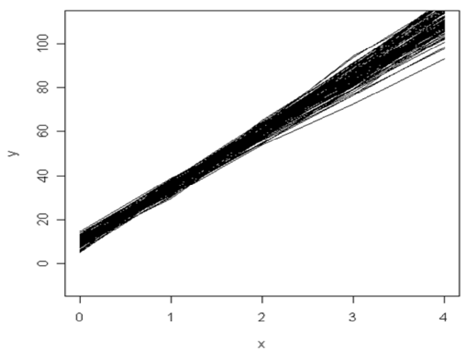
Factors	Condition Specification	No. of Conditions	No. of Conditions
		Linear growth model	Quadratic growth model
Sample size	100,200,500, 700	4	4
Error variance	2.5, 5, 10	3	3
Variance of quadratic factor	0.01, 0.05, 0.1	--	3
Total		12	36

(2) Both linear and quadratic models are applied to fit the entire sample. Structural parameter estimates and relevant fit indices are gathered for further data analysis. Start values for modeling simulated data are established using the parameter values set in the true model. Model estimations are calculated in all cases by maximum likelihood under the assumption of normality. The EQS program is used to analyze the data.

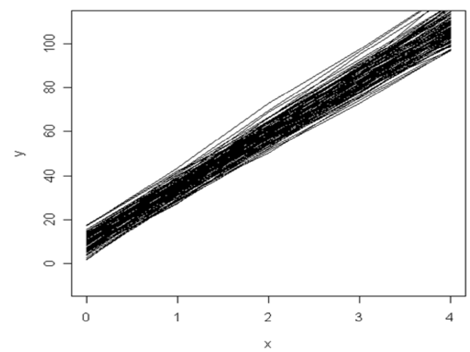
(3) The process is repeated 1,000 times per cell. The number of replications is set to be 1,000 to ensure sufficient reliability in the summary information. The maximum number of iterations to convergence for each model fitting and parameter estimating is set to 500. Any replication that fails to converge is discarded and replaced with another yielding a convergent result.

Figure 9. Plots of Each Condition

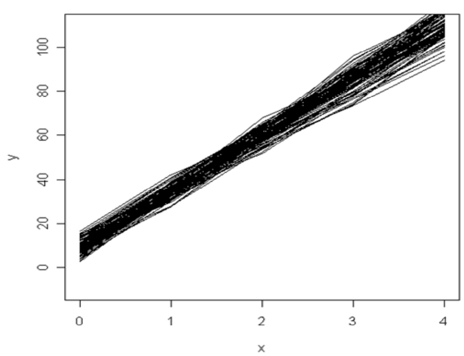
Linear LGM, N=100, Error Variance=2.5



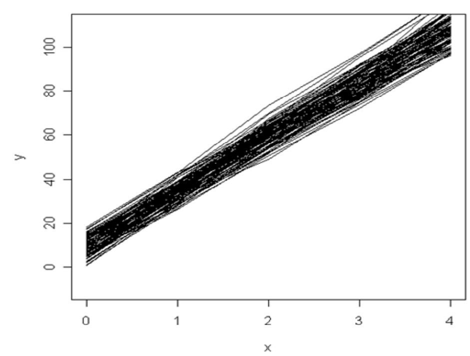
Quadratic LGM, N=100, QV=0.01, EV=2.5



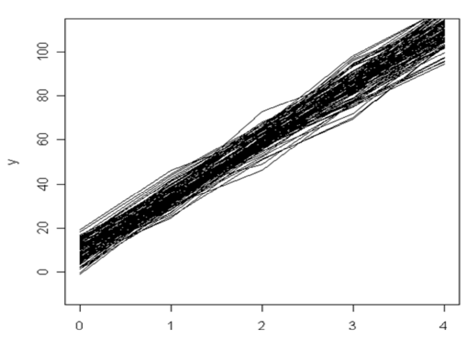
Error variance=5



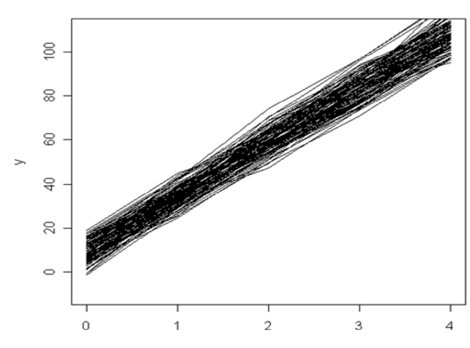
QV=0.01, EV=5



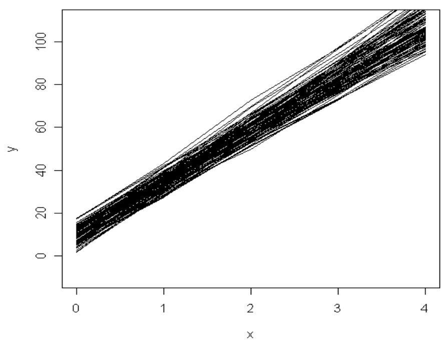
Error variance=10



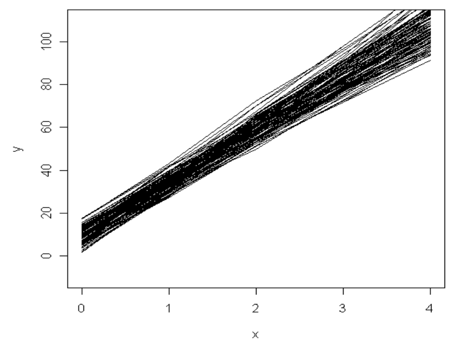
QV=0.01, EV=10



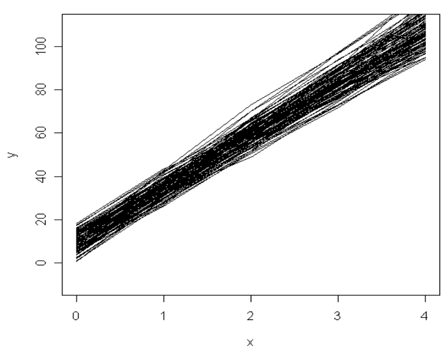
Quadratic LGM, N=100, QV=0.05, EV=2.5



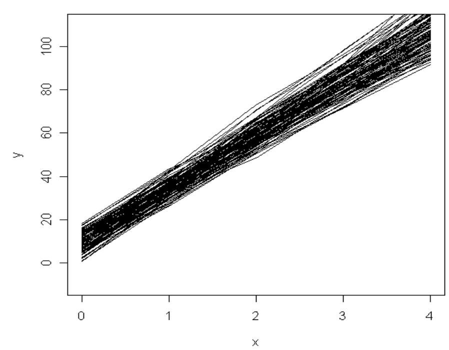
Quadratic LGM, N=100, QV=0.1, EV=2.5



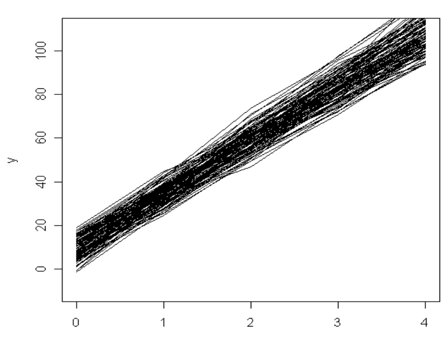
QV=0.05, EV=5



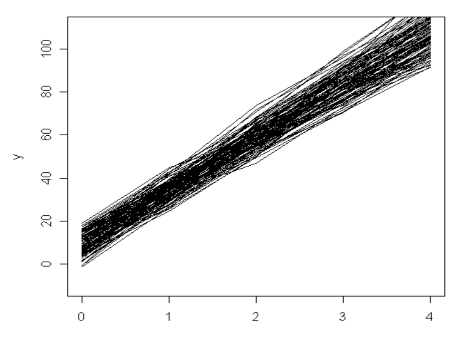
QV=0.1, EV=5



QV=0.05, EV=10



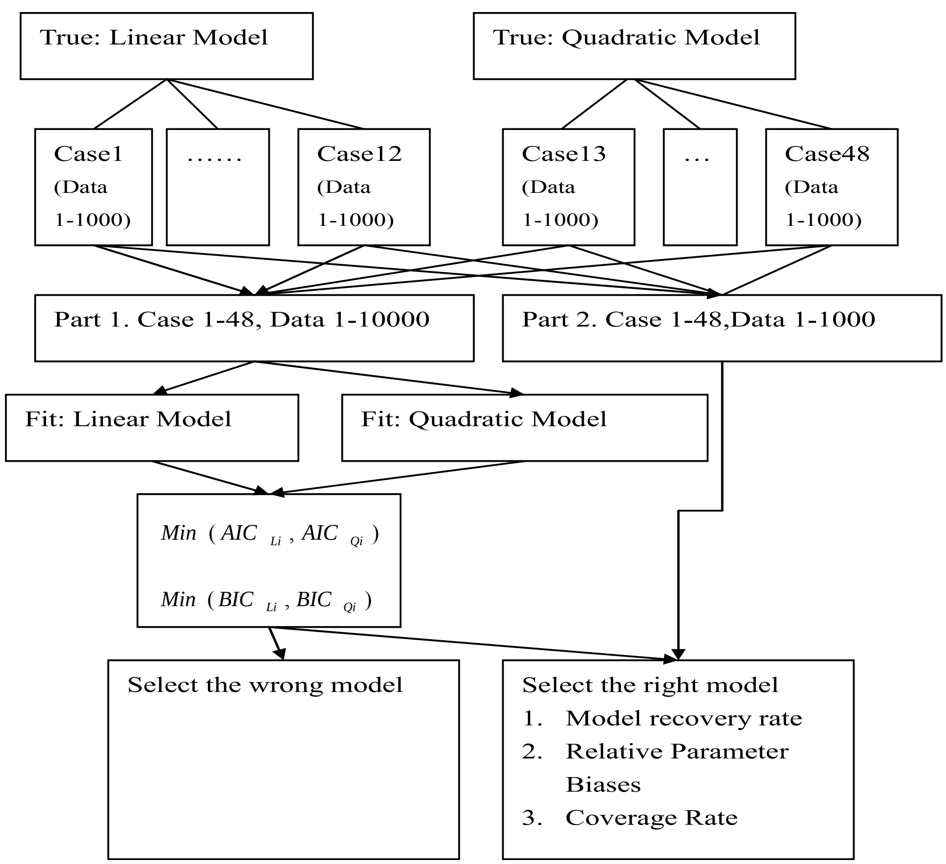
QV=0.1, EV=10



Study 2. Conduct Model Selection and Parameter Estimation on Different Data Sets

Hurvich and Tsai (1992) stated that data splitting provides a possible substitute for a true replicate sample in model validation and thus suggested a possible remedy based on data splitting to solve the problem resulting from the use of the same data set for both structural identification and inference. Therefore, in the current study the data splitting technique is employed as a possible way to mitigate the effects of model uncertainty. The procedure for study 2 is displayed in Figure 10 and described below.

Figure 10. Procedure for Study 2



Step1. Splitting Data by 50% vs. 50%

Each original data set generated from Study 1 is randomly separated into two parts by 50% vs. 50 % data splitting.

Step 2. Running Simulation

(1) The first part of the original data set is used to fit the linear and quadratic models. Relevant fit indices are determined and an appropriate model is selected based on model fit evaluation. The second part of the original data set is then used to estimate the parameters based on the model selected.

(2) The process is repeated 1,000 times per cell. The maximum number of iterations to convergence for each model fitting and parameter estimating is set to 500. Any replication that fails to converge during the run for model selection or the run for parameter estimation is discarded and replaced with another replication yielding a convergent result.

3.2 Outcome Measures and Data Analysis

Several outcome measures are gathered in the process of simulation for further data analysis. They are the relevant fit indices (i.e., AIC and BIC) and the structural parameter estimates for mean intercept, mean linear slope, mean quadratic slope, intercept variance, linear slope variance, quadratic variance, error variance, and covariance of the intercept and linear slopes.

The current study then examines the consequences of preliminary model selection by using these outcome measures. The data will be analyzed from different aspects including (1) model selection accuracy, (2) relative parameter biases, and (3) coverage rate.

3.2.1 Model Selection Accuracy

In this study, AIC and BIC are used to select a model from two plausible models; the lower value of a given information criterion indicates the better fitting model. Because each criterion applies different penalties, it is possible that each IC may point toward a different model as the better model. Therefore, this study examines whether AIC or BIC reliably identifies the true model and consistently identifies a specific model across replications under different conditions including sample size, underfitting, and overfitting. The performance of AIC and BIC is measured by the success rate of selecting correct models in the iterative process of model selection.

3.2.2 Relative Parameter Biases

Parameter biases are examined to assess the impact of preliminary model selection on latent growth model parameter estimates. Generally, in simulation studies, it is expected that the estimated parameter values are close to the population parameters. If the estimated parameter values significantly deviate from the population parameter values, the deviation might result from the preliminary model selection.

Relative parameter bias is calculated by using the following formula,

$$\text{Relative parameter bias} = (\bar{\hat{\theta}} - \theta) / \theta,$$

where $\bar{\hat{\theta}}$ is the parameter estimate average over the replications of the Monte Carlo study, and θ is the population value. Relative parameter bias will be examined for the following parameters: mean intercept, mean linear slope, mean quadratic slope, intercept variance, linear slope variance, quadratic slope variance, error variance, and covariance of intercept and linear slope.

3.2.3 Coverage Rate

The effect of the preliminary model selection on the coverage rate of confidence intervals for the growth model parameters will be examined. It is the goal in simulation studies that the analysis models are able to accurately recover the population parameters because the data are generated from the previously set true model. Usually, a coverage rate (i.e., the number of replications whose confidence intervals contain the true population parameter) is used to evaluate the ability of the analysis model to recover the population parameters. In the present study, coverage rates are calculated for each of the estimated parameters in the model.

For the interval coverage, both the unconditional and the conditional coverage probability will be examined. Unconditional coverage is defined as the coverage of confidence intervals without model selection. Conditional coverage is contingent upon selecting a correct model. To determine the coverage probability, the proportion of the confidence intervals that cover the “true” parameter is computed for each of the design conditions.

Chapter 4: Results

This chapter presents the summaries of the results from the simulation studies. The current investigation includes two Monte Carlo simulation studies. Study 1 conducts both model selection and parameter estimation using the same data set. Study 2 conducts both model selection and parameter estimation using different split-data sets. The simulation results for study 1 are presented in section 4.1 and for study 2 in section 4.2. Results in each section are looked at from three aspects: (1) model selection accuracy, (2) relative parameter biases, and (3) coverage rate.

4.1 Study 1: Conducting Model Selection and Parameter Estimation Using Same Data Set

4.1.1 Model Selection Accuracy

In this study, the performances of AIC and BIC are examined to see whether they reliably identify the true model and consistently identify a specific model across replications under different conditions including sample size, underfitting, and overfitting. The performances of AIC and BIC are evaluated from the aspects of model selection accuracy and model selection consistency by examining model recovery rate (i.e., the success rate of selecting correct models in the iterative process of model selection).

Table 6 displays the model recovery rates of AIC and BIC under each condition. For example, in case 1, for the 1,000 replications with data generated by the linear model, with sample size 100, and an error variance of 2.5, the AIC index correctly selected the linear model 949 times, the BIC index 998 times.

Table 6. Model Recovery Rate in Each Condition

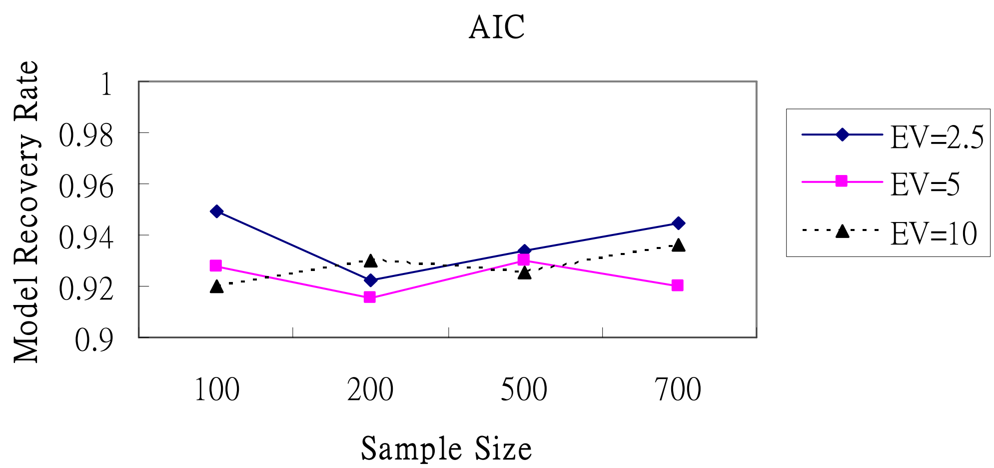
	True Model	Sample Size	Error Variance	Quadratic Variance	AIC	BIC
Case1	L	N=100	EV=2.5	--	0.949	0.998
Case2	L		EV=5	--	0.928	0.999
Case3	L		EV=10	--	0.920	1.000
Case4	L	N=200	EV=2.5	--	0.922	1.000
Case5	L		EV=5	--	0.915	1.000
Case6	L		EV=10	--	0.930	1.000
Case7	L	N=500	EV=2.5	--	0.934	1.000
Case8	L		EV=5	--	0.930	1.000
Case9	L		EV=10	--	0.925	1.000
Case10	L	N=700	EV=2.5	--	0.945	1.000
Case11	L		EV=5	--	0.920	0.999
Case12	L		EV=10	--	0.936	1.000
Case13	Q	N=100	EV=2.5	QV=0.01	0.490	0.057
Case14	Q			QV=0.05	0.746	0.180
Case15	Q			QV=0.10	0.958	0.599
Case16	Q	N=100	EV=5	QV=0.01	0.276	0.018
Case17	Q			QV=0.05	0.460	0.047
Case18	Q			QV=0.10	0.725	0.174
Case19	Q	N=100	EV=10	QV=0.01	0.182	0.003
Case20	Q			QV=0.05	0.240	0.014
Case21	Q			QV=0.10	0.437	0.028
Case22	Q	N=200	EV=2.5	QV=0.01	0.810	0.170
Case23	Q			QV=0.05	0.973	0.469
Case24	Q			QV=0.10	0.999	0.930
Case25	Q	N=200	EV=5	QV=0.01	0.524	0.036
Case26	Q			QV=0.05	0.731	0.105
Case27	Q			QV=0.10	0.967	0.468
Case28	Q	N=200	EV=10	QV=0.01	0.328	0.002
Case29	Q			QV=0.05	0.431	0.014
Case30	Q			QV=0.10	0.679	0.088
Case31	Q	N=500	EV=2.5	QV=0.01	0.999	0.697
Case32	Q			QV=0.05	1.000	0.959
Case33	Q			QV=0.10	1.000	1.000
Case34	Q	N=500	EV=5	QV=0.01	0.894	0.180
Case35	Q			QV=0.05	0.991	0.527
Case36	Q			QV=0.10	1.000	0.957
Case37	Q	N=500	EV=10	QV=0.01	0.643	0.030
Case38	Q			QV=0.05	0.829	0.070
Case39	Q			QV=0.10	0.975	0.403
Case40	Q	N=700	EV=2.5	QV=0.01	1.000	0.907
Case41	Q			QV=0.05	1.000	0.998
Case42	Q			QV=0.10	1.000	1.000
Case43	Q	N=700	EV=5	QV=0.01	0.982	0.340
Case44	Q			QV=0.05	0.999	0.749
Case45	Q			QV=0.10	1.000	0.998
Case46	Q	N=700	EV=10	QV=0.01	0.757	0.036
Case47	Q			QV=0.05	0.920	0.159
Case48	Q			QV=0.10	0.995	0.684

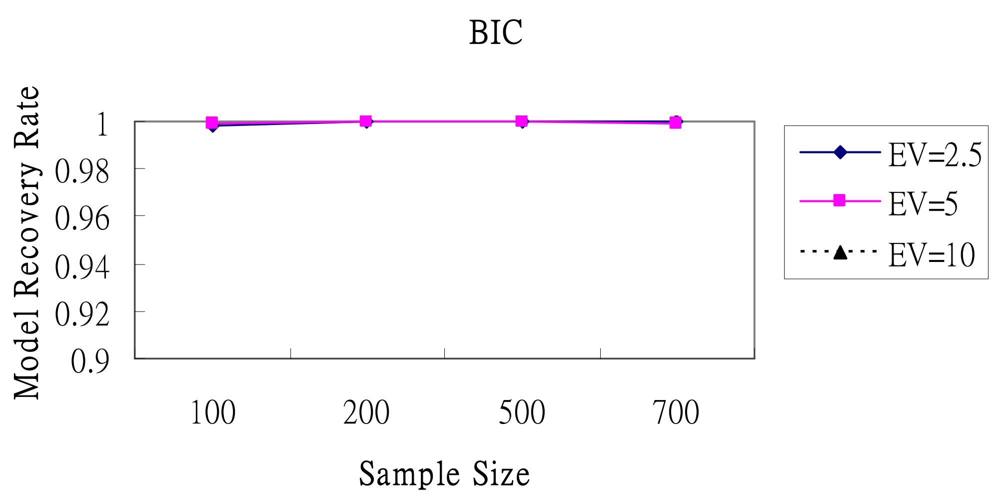
Overfitting

When the true model is the linear model (i.e., cases 1-12), three findings are noted. First, the model recovery rate of AIC ranges between 0.915 and 0.949 and the model recovery rate of BIC ranges between 0.998 and 1. This suggests that AIC and BIC reliably identify the true model. This also suggests that both AIC and BIC perform consistently in selecting the linear model and do not favor overfitting in all conditions. Second, Table 6 shows that the model recovery rate of BIC is always larger than that of AIC across all conditions when the true model is the linear model. This indicates that BIC consistently performs better than AIC when the true model is the linear model. Third, both AIC and BIC appear to perform consistently in selecting the linear model in more than 90% of the replications under all 12 conditions when the true model is the linear model. BIC, however, tends to be more consistent than AIC, selecting the linear model in more than 99% of the replications under all 12 conditions.

Figure 9 illustrates how the model selection rates of AIC and BIC change across different conditions, when the true model is the linear model. For example, the model recovery rate of AIC is always greater than 0.915 and the model recovery rate of BIC is always greater than 0.998 under different sample size and error variance conditions. This suggests that the roles of sample size and error variance in model selection accuracy and consistency are not substantial when the true model is the linear model.

Figure 9. Change of Model Recovery Rate of AIC and BIC under Different Conditions When the True Model Is the Linear Model





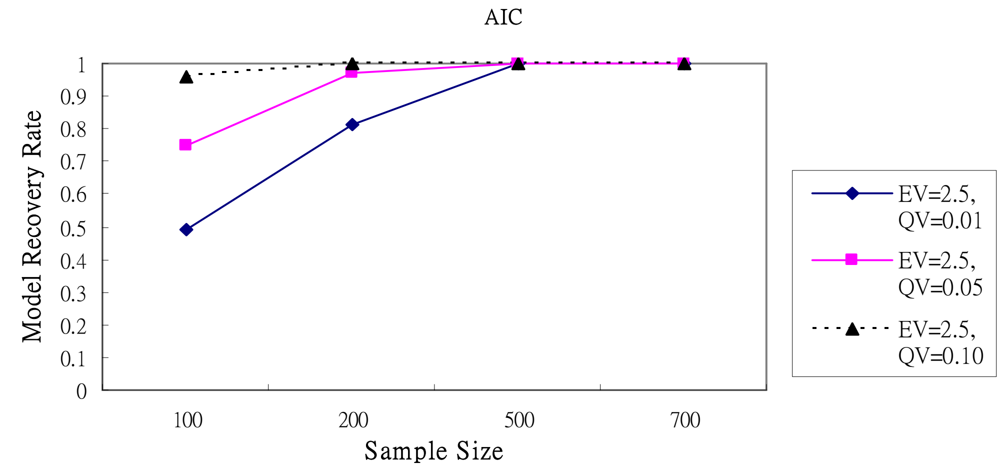
<u>Underfitting</u>

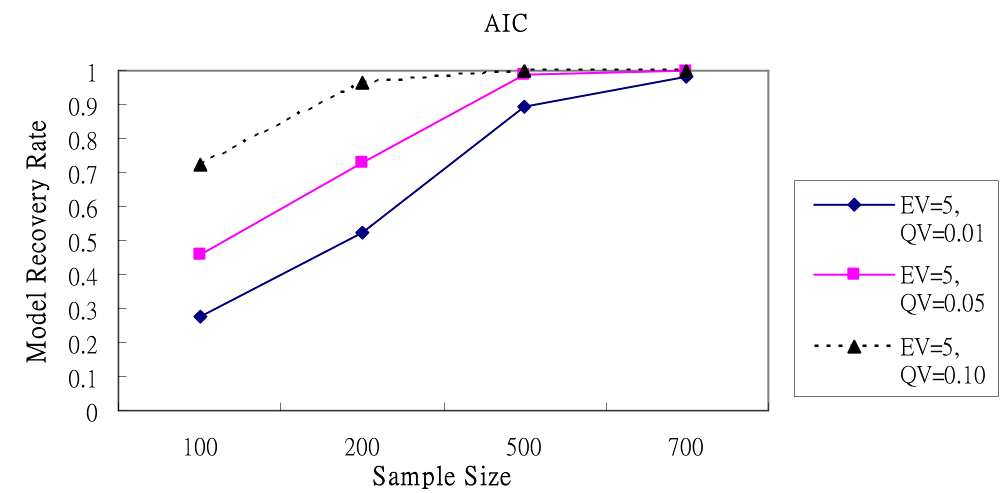
When the true model is the quadratic model (i.e., cases 13-48), the findings are as follows. First, AIC outperforms BIC in identifying the true model. As shown in Table 6, the model recovery rates of AIC (ranging from 0.182 to 0.999) are almost always much higher than those of BIC (ranging from 0.003 to 0.998) across different conditions. Second, in 28 of the 36 cases, AIC selects the quadratic model in more than 50 % of the replications, and in the remaining 8 cases (i.e., cases 13, 16, 17, 19, 20, 21, 28, and 29), it selects the linear model. The BIC, however, selects the quadratic model in more than 50 % of the replications in only 13 cases (i.e., cases 15, 24, 31, 32, 33, 35, 36, 40, 41, 42, 44, 45, and 48) and the linear model in the remaining 23 cases. This demonstrates that BIC has a preference for selection of the simpler model (i.e., underfitting), which is consistent with previous research. Third, BIC tends to be more consistent than AIC in selecting a model (true or misspecified). BIC demonstrates consistency in selecting a specific model (true or misspecified) in more than 80% of the replications in 28 out of the total 36 conditions. AIC, however, demonstrates such consistency in only 22 out of the 36 conditions.

Figures 10 and 11 illustrate how the model selection rates of AIC and BIC change across different conditions, when the quadratic model is the true model. When sample size is 100, the model recovery rates range from 0.182 to 0.958 for AIC and from 0.014 to 0.599 for BIC. When sample size becomes larger, the model recovery rates for AIC and BIC increase substantially; from 0.757 to 1.0 for AIC and from 0.159 to 0.998 for BIC. The model recovery rates of AIC and BIC also go up as the quadratic variance increases from 0.01 to 0.1. It appears that AIC and BIC are better able to identify the true model as sample size and variance of quadratic factor increase, but less able as error variance increases.

Additionally, underfitting tends to decrease as sample size and quadratic variance increase, but tends to increase when error variance increases. Underfitting is more severe for BIC than for AIC under all conditions.

Figure 10. Change of Model Recovery Rate of AIC under Different Conditions When the True Model is the Quadratic Model





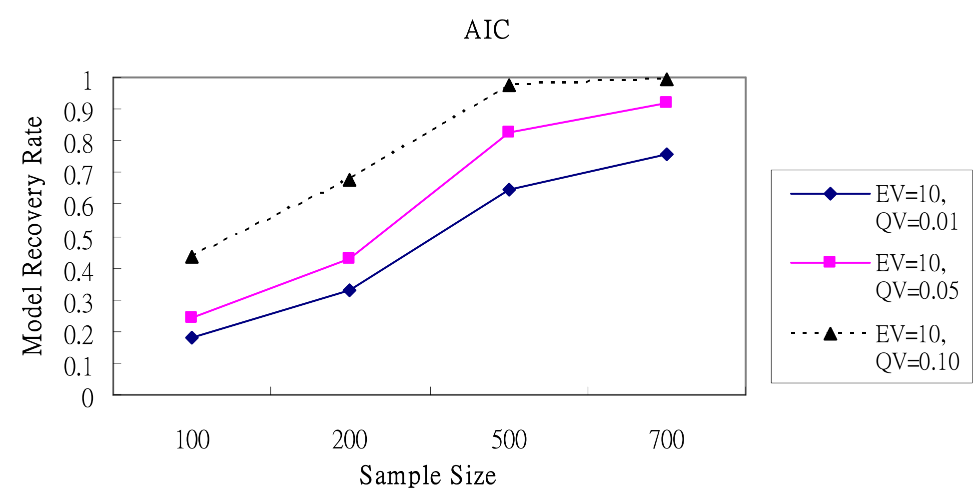
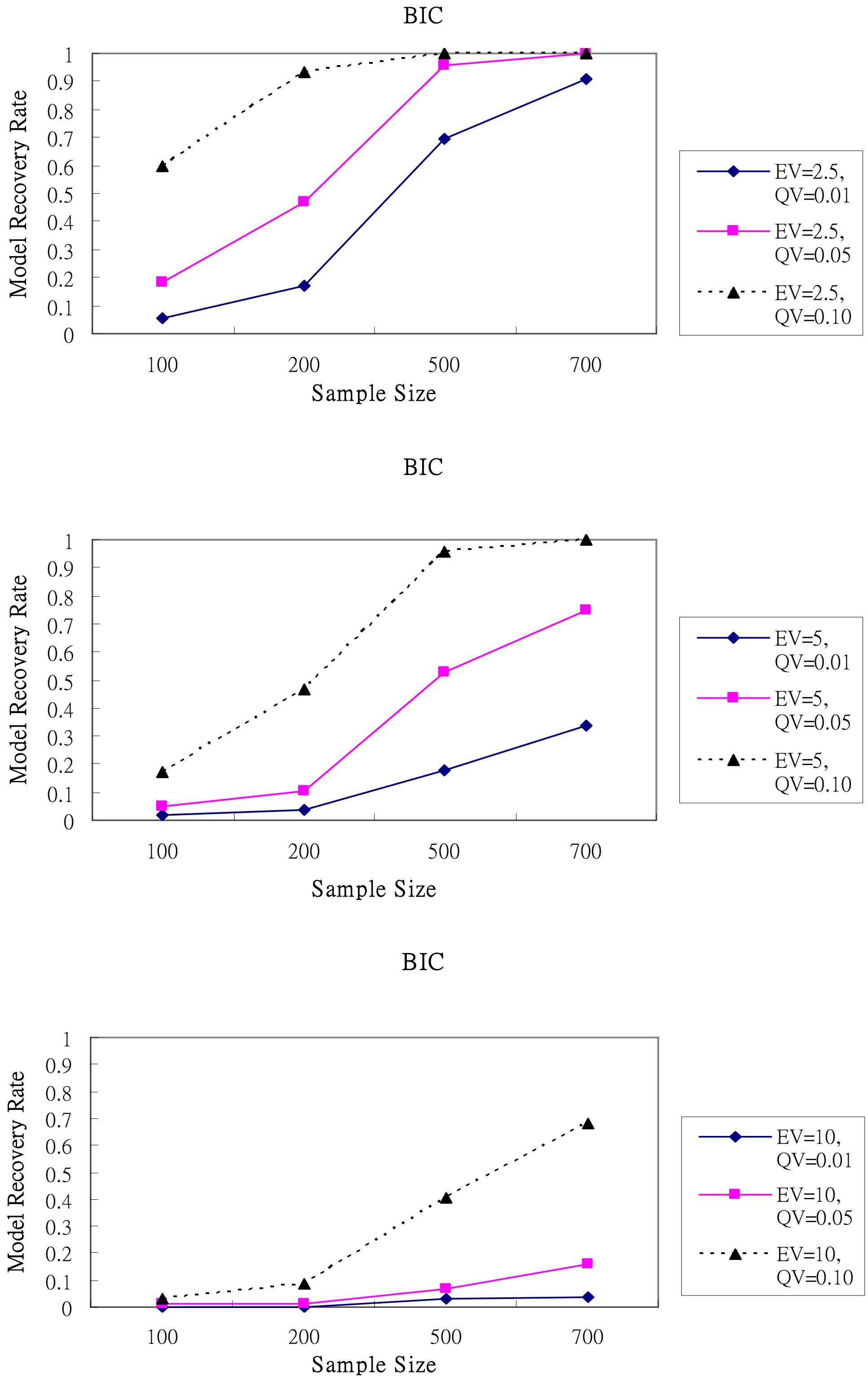


Figure 11. Change of Model Recovery Rate of BIC under Different Conditions When the True Model is the Quadratic Model



Comparing the Empirical Model Recovery Rate of AIC and the Power

Because the power for choosing a correct model was controlled in advance (as shown in Table 4 in section 3.1.1.2) in order to evaluate the impact of model selection on parameter estimates, I also checked whether the empirical model recovery rate is close to the theoretically controlled power. Table 7 provides a comparison of the empirical model recovery rate and the controlled power, when the true model is the quadratic model and the model selection index is AIC. The table shows that the empirical model recovery rate is quite close to the previously controlled power. The differences between the empirical recovery rate and the power range only from -0.047 to 0.007.

Table 7. Comparison between the Empirical Model Recovery Rate of AIC and the Power to Choose Quadratic Model by AIC When the Quadratic Model is the True Model

Sample Size	Error Variance	Empirical Recovery Rate / Power (Difference)		
		QV=0.01	QV=0.05	QV=0.10
N=100	EV= 2.5	0.490 / 0.537 (-0.047)	0.746 / 0.749 (-0.003)	0.958 / 0.959 (-0.001)
	EV= 5	0.276 / 0.322 (-0.046)	0.460 / 0.460 (0.000)	0.725 / 0.733 (-0.008)
	EV= 10	0.182 / 0.203 (-0.021)	0.240 / 0.265 (-0.025)	0.437 / 0.426 (0.011)
N=200	EV= 2.5	0.810 / 0.833 (-0.023)	0.973 / 0.966 (0.007)	0.999 / 0.999 (0.000)
	EV= 5	0.524 / 0.550 (-0.026)	0.731 / 0.750 (-0.019)	0.967 / 0.960 (0.007)
	EV= 10	0.328 / 0.326 (0.002)	0.431 / 0.448 (-0.017)	0.679 / 0.708 (-0.029)
N=500	EV= 2.5	0.999 / 0.997 (0.002)	1.000 / 0.999 (0.001)	1.000 / 0.999 (0.001)
	EV= 5	0.894 / 0.916 (-0.022)	0.991 / 0.989 (0.002)	1.000 / 0.999 (0.001)
	EV= 10	0.643 / 0.651 (-0.008)	0.829 / 0.827 (0.002)	0.975 / 0.981 (-0.006)
N=700	EV= 2.5	1.000 / 0.999 (0.001)	1.000 / 0.999 (0.001)	1.000 / 0.999 (0.001)
	EV= 5	0.982 / 0.978 (0.004)	0.999 / 0.999 (0.000)	1.000 / 0.999 (0.001)
	EV= 10	0.757 / 0.796 (-0.039)	0.920 / 0.933 (-0.013)	0.995 / 0.998 (-0.003)

4.1.2. Biases

In this section, the parameter biases and variance biases of the true-model-selection estimates, the AIC-model-selection estimates, and the BIC-model-selection estimates were examined in order to evaluate the impact of preliminary model selection on latent growth model parameter estimates. Relative bias was used as an indicator. Relative bias is the ratio of the bias to the population value (bias in this case is calculated as the parameter estimate averaged over the replications of the Monte Carlo study minus the population value). Parameter recovery is seen as adequate when the absolute relative bias rates are less than 0.1; it is seen as mediocre when the absolute relative bias rates range from 0.1 to 0.5; it is seen as poor when the absolute relative bias rates are greater than 0.5.

4.1.2.1 True-model-selection estimates

Relative Parameter Bias

Because model selection impact in this study is evaluated by using the true-model-selection estimates as the baseline, it is important to check the quality of these estimates. The quality of the true-model-selection estimates is looked at from two aspects: relative parameter bias (accuracy) and relative variance bias (variability). Relative parameter bias is calculated using the following formula,

$$\text{Relative parameter bias} = (\bar{\hat{\theta}} - \theta) / \theta,$$

where $\bar{\hat{\theta}}$ is the parameter estimate average over the replications of the Monte Carlo study, and θ is the population value. Relative variance bias is calculated using a similar equation but in this case the population variance is subtracted from the average of the squared standard errors for each parameter estimate, and this difference is divided by the population variance. In this study, since the number of replications (1,000) is large, the variance of each parameter estimate over the replications is considered to be the population variance.

Relative parameter bias and relative variance bias are examined for the following parameters: mean intercept (α_1), mean slope (α_2), mean quadratic factor (α_3), intercept variance (ϕ_{11}), slope variance (ϕ_{22}), quadratic variance (ϕ_{33}), and covariance of intercept and slope (ϕ_{21}).

Table 8 presents the relative parameter bias for each condition. For the common parameters in both the linear and quadratic models (i.e., α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}), the relative bias rates range from -0.03 to 0.01 for the linear model and from -0.02 to 0.65 for the quadratic model. This suggests that parameters for the linear model are estimated better than those for the quadratic model.

For parameters α_1, α_2, and α_3, the relative bias rate ranges from -0.028 to 0.040. This indicates that those parameters are recovered satisfactorily for the conditions simulated in this study. Recovery for ϕ_{11} appears adequate, with the relative parameter bias ranging from 0.013 to 0.088. Recovery for ϕ_{22} and ϕ_{21} is poor to adequate, with the relative parameter bias ranging from -0.026 to 0.650. Recovery for ϕ_{33}, however, is poor, ranging from -0.021 to 7.047. When the sample size and the parameter value for ϕ_{33} are larger, the relative parameter bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 100 to 700 and parameter value for ϕ_{33} increases from 0.01 to 0.1, the relative parameter bias for ϕ_{33} decreases markedly, from 7.047 to -0.004.

Relative Variance Bias

Table 9 presents the relative variance bias for each condition. In general, the relative variance bias for the linear model is smaller than that for the quadratic model. For the common parameters in both the linear and quadratic models (i.e., α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}), the relative variance bias rates range from -0.107 to 0.095 for the linear model and from -0.125 to 0.285 for the quadratic model. This suggests that parameters for the linear model are recovered better than those for the quadratic model.

For parameters α_1, α_2, and α_3, the relative bias rate ranges from -0.125 to 0.111. This indicates that those parameters are recovered satisfactorily for the conditions simulated in this study. Recovery for parameters ϕ_{11} and ϕ_{22} appears adequate, with the relative parameter bias ranging from -0.097 to 0.285. Recovery for ϕ_{33}, however, is poor, ranging from -0.111 to 10.283. When sample size and parameter value for ϕ_{33} are larger, the relative variance bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 100 to 700 and parameter value for quadratic variance increases from 0.01 to 0.1, the relative bias for quadratic variance decreases substantially, from 10.283 to 0.069. Recovery for ϕ_{21} is adequate, with the relative parameter bias ranging from -0.107 to 0.145.

Table 8. Relative Parameter Bias for True-Model-Selection Estimates

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	-0.003	-0.015	-0.012	--	-0.001	0.000	--
2	L		EV=5	--	-0.004	-0.025	-0.012	--	-0.001	0.000	--
3	L		EV=10	--	-0.009	-0.031	-0.017	--	0.001	0.000	--
4	L	N=200	EV=2.5	--	-0.003	-0.002	0.001	--	-0.001	0.000	--
5	L		EV=5	--	0.001	-0.003	0.003	--	0.001	0.000	--
6	L		EV=10	--	-0.010	-0.026	-0.012	--	0.000	0.000	--
7	L	N=500	EV=2.5	--	-0.002	-0.008	-0.004	--	0.000	0.000	--
8	L		EV=5	--	0.005	0.001	-0.003	--	-0.001	0.000	--
9	L		EV=10	--	-0.002	-0.010	-0.004	--	0.001	0.000	--
10	L	N=700	EV=2.5	--	-0.002	-0.004	-0.004	--	0.000	0.000	--
11	L		EV=5	--	-0.003	-0.007	-0.003	--	0.000	0.000	--
12	L		EV=10	--	0.003	0.006	0.000	--	0.000	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	0.025	0.168	0.120	1.696	0.000	0.000	-0.010
14	Q			QV=0.05	-0.013	0.049	0.044	0.118	0.001	0.000	-0.018
15	Q			QV=0.10	0.002	0.031	0.024	0.012	0.000	0.001	0.040
16	Q	N=100	EV=5	QV=0.01	0.033	0.256	0.272	3.456	0.001	0.000	0.018
17	Q			QV=0.05	0.031	0.187	0.158	0.424	-0.001	0.001	0.038
18	Q			QV=0.10	-0.002	0.029	0.044	0.091	0.004	-0.001	-0.028
19	Q	N=100	EV=10	QV=0.01	0.088	0.650	0.593	7.047	-0.001	0.001	0.028
20	Q			QV=0.05	0.071	0.540	0.450	1.058	-0.001	0.000	0.033
21	Q			QV=0.10	0.068	0.408	0.373	0.480	0.001	0.000	-0.017
22	Q	N=200	EV=2.5	QV=0.01	0.015	0.110	0.094	1.263	0.000	0.000	0.004
23	Q			QV=0.05	0.001	0.025	0.025	0.017	0.001	0.000	0.013
24	Q			QV=0.10	-0.003	-0.017	0.000	-0.021	-0.001	0.000	0.000
25	Q	N=200	EV=5	QV=0.01	0.021	0.185	0.169	2.309	-0.001	0.000	0.000
26	Q			QV=0.05	0.000	0.063	0.070	0.187	0.001	0.000	0.001
27	Q			QV=0.10	0.004	0.051	0.045	0.091	-0.001	0.000	0.003
28	Q	N=200	EV=10	QV=0.01	0.052	0.436	0.370	4.505	0.000	0.000	0.019
29	Q			QV=0.05	0.034	0.302	0.279	0.663	-0.001	0.000	0.014
30	Q			QV=0.10	0.057	0.405	0.285	0.280	0.002	0.000	0.025
31	Q	N=500	EV=2.5	QV=0.01	0.010	0.063	0.057	0.734	0.000	0.000	0.006
32	Q			QV=0.05	-0.006	-0.012	-0.008	-0.016	0.000	0.000	0.011
33	Q			QV=0.10	-0.003	0.000	-0.001	0.003	0.000	0.000	0.002
34	Q	N=500	EV=5	QV=0.01	0.022	0.142	0.110	1.321	0.001	0.000	-0.003
35	Q			QV=0.05	0.001	0.016	0.027	0.072	0.000	0.000	0.016
36	Q			QV=0.10	0.011	0.055	0.022	0.012	0.000	0.000	-0.010
37	Q	N=500	EV=10	QV=0.01	0.034	0.224	0.201	2.869	-0.002	0.000	0.016
38	Q			QV=0.05	0.009	0.074	0.096	0.322	0.000	0.000	-0.006
39	Q			QV=0.10	0.000	-0.006	0.015	0.023	0.000	0.000	-0.002
40	Q	N=700	EV=2.5	QV=0.01	0.007	0.053	0.044	0.514	0.000	0.000	0.003
41	Q			QV=0.05	0.002	0.007	-0.001	0.002	0.000	0.000	-0.009
42	Q			QV=0.10	-0.001	-0.016	-0.013	-0.004	0.000	0.000	-0.008
43	Q	N=700	EV=5	QV=0.01	0.011	0.093	0.082	1.165	0.000	0.000	0.005
44	Q			QV=0.05	0.003	0.008	0.009	0.045	0.000	0.000	-0.010
45	Q			QV=0.10	0.002	-0.013	-0.012	-0.008	-0.001	0.000	0.011
46	Q	N=700	EV=10	QV=0.01	0.028	0.207	0.189	2.285	0.000	0.000	-0.025
47	Q			QV=0.05	0.010	0.079	0.075	0.170	0.000	0.000	-0.003
48	Q			QV=0.10	0.005	0.027	0.035	0.044	-0.001	0.000	0.004

Table 9. Relative Variance Bias for True-Model-Selection Estimates

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	-0.079	-0.107	-0.073	--	-0.072	0.002	--
2	L		EV=5	--	-0.007	-0.067	0.062	--	0.079	-0.001	--
3	L		EV=10	--	0.002	0.006	-0.022	--	-0.018	0.008	--
4	L	N=200	EV=2.5	--	0.007	0.026	-0.012	--	-0.027	-0.012	--
5	L		EV=5	--	0.076	0.015	-0.072	--	-0.028	0.050	--
6	L		EV=10	--	-0.025	0.005	-0.056	--	-0.045	-0.025	--
7	L	N=500	EV=2.5	--	-0.059	-0.057	0.017	--	-0.085	0.016	--
8	L		EV=5	--	-0.023	0.095	0.021	--	-0.005	-0.006	--
9	L		EV=10	--	-0.002	-0.027	-0.030	--	-0.074	0.038	--
10	L	N=700	EV=2.5	--	-0.034	0.014	0.042	--	0.002	0.000	--
11	L		EV=5	--	0.038	-0.004	0.043	--	-0.010	-0.046	--
12	L		EV=10	--	0.077	0.023	-0.038	--	-0.077	-0.027	--
13	Q	N=100	EV=2.5	QV=0.01	0.109	0.081	0.107	10.283	-0.021	0.001	0.043
14	Q			QV=0.05	0.059	-0.011	0.054	2.246	-0.125	-0.103	-0.038
15	Q			QV=0.10	-0.058	-0.037	0.023	0.444	0.058	-0.034	-0.004
16	Q	N=100	EV=5	QV=0.01	0.028	0.030	0.147	1.653	-0.055	-0.066	0.022
17	Q			QV=0.05	0.021	-0.019	0.115	0.855	-0.012	-0.006	-0.040
18	Q			QV=0.10	0.031	-0.001	0.081	0.342	-0.024	-0.011	0.055
19	Q	N=100	EV=10	QV=0.01	-0.009	0.055	0.226	0.583	-0.002	0.075	0.026
20	Q			QV=0.05	0.061	0.100	0.285	0.542	-0.081	0.026	0.086
21	Q			QV=0.10	0.055	0.100	0.230	0.333	-0.043	0.027	0.059
22	Q	N=200	EV=2.5	QV=0.01	0.057	0.076	0.131	6.861	0.017	0.014	0.079
23	Q			QV=0.05	-0.054	0.033	0.085	1.125	0.004	0.014	0.009
24	Q			QV=0.10	-0.020	0.017	0.066	0.210	0.042	0.005	0.030
25	Q	N=200	EV=5	QV=0.01	-0.031	-0.046	0.113	1.447	0.037	0.007	0.016
26	Q			QV=0.05	-0.023	0.034	0.054	0.431	-0.037	0.012	-0.019
27	Q			QV=0.10	0.035	0.015	-0.031	0.150	-0.015	0.100	0.070
28	Q	N=200	EV=10	QV=0.01	0.051	0.079	0.155	0.548	-0.026	-0.053	-0.052
29	Q			QV=0.05	-0.015	0.040	0.131	0.329	-0.004	0.029	-0.003
30	Q			QV=0.10	0.045	0.033	0.100	0.162	0.000	-0.003	-0.042
31	Q	N=500	EV=2.5	QV=0.01	0.000	0.076	0.055	4.561	-0.020	-0.013	0.027
32	Q			QV=0.05	0.095	0.009	0.048	0.326	0.111	-0.032	-0.042
33	Q			QV=0.10	0.079	0.001	-0.013	0.000	0.027	-0.021	0.057
34	Q	N=500	EV=5	QV=0.01	-0.077	0.067	0.172	1.281	0.075	-0.040	-0.039
35	Q			QV=0.05	-0.007	0.018	0.100	0.308	0.092	-0.034	0.000
36	Q			QV=0.10	0.011	0.046	0.108	0.067	0.057	0.019	-0.032
37	Q	N=500	EV=10	QV=0.01	0.080	0.098	0.189	0.455	0.066	-0.028	-0.039
38	Q			QV=0.05	0.076	0.108	0.167	0.210	0.004	0.055	0.085
39	Q			QV=0.10	0.066	0.094	0.126	0.163	0.075	0.101	0.045
40	Q	N=700	EV=2.5	QV=0.01	0.060	0.145	0.175	4.267	-0.016	0.092	0.000
41	Q			QV=0.05	0.024	0.032	-0.013	0.114	-0.049	-0.063	-0.029
42	Q			QV=0.10	0.030	-0.022	-0.027	-0.073	-0.019	0.043	0.000
43	Q	N=700	EV=5	QV=0.01	-0.017	-0.027	0.044	0.908	0.015	0.009	0.039
44	Q			QV=0.05	0.049	0.042	0.022	0.158	-0.092	-0.055	0.018
45	Q			QV=0.10	-0.075	-0.040	-0.097	-0.111	-0.071	0.061	0.000
46	Q	N=700	EV=10	QV=0.01	0.015	0.036	0.137	0.469	0.055	0.011	0.020
47	Q			QV=0.05	0.076	0.056	0.084	0.194	-0.049	0.009	0.058
48	Q			QV=0.10	0.085	0.092	0.087	0.069	-0.042	0.000	0.000

In summary, the quality of true-model-selection estimates for the linear model is slightly better than that for the quadratic model. Recovery for α_1, α_2, α_3, and ϕ_{11} is adequate for all 48 conditions. Recovery for ϕ_{33}, however, is poor, especially when sample size is small and the parameter value for ϕ_{33} is small. As expected, for all parameters, relative bias goes down as the sample size increases from 100 to 700. This pattern reflects the principle that the maximum likelihood estimate is biased when sample size is finite, but converges to the parameter value as sample size goes to infinity.

4.1.2.2 AIC-model-selection estimates

<u>Relative Parameter Bias</u>

Table 10 presents the relative parameter bias for AIC-model-selection estimates. In general, the parameter estimates for the linear model are better than those for the quadratic model. This is similar to the result of the true-model-selection estimates in section 4.1.2.1. For the parameters common to both the linear and quadratic models (i.e., α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}), the relative bias rates range from -0.032 to 0.006 for the linear model, and from
-0.023 to 1.332 for the quadratic model.

Mean parameters α_1 and α_2 are recovered satisfactorily, with the relative parameter bias rate ranging only from -0.023 to 0.014. Recovery for mean parameter α_3 appears adequate when sample size is large (700), with the relative parameter bias ranging from

-0.009 to 0.099. But when sample size is small (100), recovery becomes adequate to poor, with the relative parameter bias ranging from 0.065 to 0.913.

Recovery for variance parameter ϕ_{11} also appears adequate, with the relative parameter bias ranging from -0.01 to 0.183. Recovery for parameters ϕ_{22}, ϕ_{21}, and ϕ_{33} is poor to adequate, with the relative parameter bias ranging from -0.032 to 13.108. When sample size and parameter value for ϕ_{33} are larger, the relative parameter bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 100 to 700 and parameter value for ϕ_{33} increases from 0.01 to 0.1, the relative parameter bias for ϕ_{33} decreases markedly, from 13.018 to -0.004.

Table 10. Relative Parameter Bias for AIC-Model-Selection Estimates

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	-0.003	-0.013	-0.011	--	-0.002	0.000	--
2	L		EV=5	--	-0.006	-0.032	-0.012	--	0.000	0.000	--
3	L		EV=10	--	-0.008	-0.023	-0.015	--	0.002	0.000	--
4	L	N=200	EV=2.5	--	-0.006	-0.008	0.000	--	-0.001	0.000	--
5	L		EV=5	--	0.001	-0.003	0.003	--	0.001	0.000	--
6	L		EV=10	--	-0.011	-0.032	-0.014	--	0.000	0.000	--
7	L	N=500	EV=2.5	--	-0.003	-0.006	-0.003	--	0.000	0.000	--
8	L		EV=5	--	0.006	0.004	-0.002	--	-0.001	0.000	--
9	L		EV=10	--	-0.001	-0.007	-0.003	--	0.001	0.000	--
10	L	N=700	EV=2.5	--	-0.002	-0.004	-0.003	--	0.000	0.000	--
11	L		EV=5	--	-0.003	-0.006	-0.004	--	0.000	0.000	--
12	L		EV=10	--	0.002	0.003	-0.001	--	0.000	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	0.035	0.183	0.124	2.324	-0.006	0.004	0.290
14	Q			QV=0.05	-0.010	0.056	0.017	0.239	-0.002	0.001	0.110
15	Q			QV=0.10	0.004	0.028	0.019	0.037	-0.001	0.001	0.065
16	Q	N=100	EV=5	QV=0.01	0.031	0.286	0.316	5.274	-0.012	0.010	0.625
17	Q			QV=0.05	0.024	0.103	0.097	0.882	-0.008	0.006	0.428
18	Q			QV=0.10	0.001	0.033	0.002	0.268	0.001	0.001	0.129
19	Q	N=100	EV=10	QV=0.01	0.183	1.332	1.089	13.018	-0.023	0.014	0.913
20	Q			QV=0.05	0.116	0.773	0.443	2.276	-0.016	0.011	0.706
21	Q			QV=0.10	0.071	0.351	0.227	0.966	-0.006	0.005	0.314
22	Q	N=200	EV=2.5	QV=0.01	0.013	0.098	0.086	1.382	-0.002	0.002	0.095
23	Q			QV=0.05	0.001	0.027	0.022	0.032	0.001	0.000	0.027
24	Q			QV=0.10	-0.003	-0.017	0.000	-0.020	-0.001	0.000	0.001
25	Q	N=200	EV=5	QV=0.01	0.034	0.263	0.211	3.170	-0.006	0.004	0.264
26	Q			QV=0.05	0.002	0.053	0.040	0.355	-0.002	0.002	0.150
27	Q			QV=0.10	0.006	0.050	0.038	0.114	-0.001	0.000	0.020
28	Q	N=200	EV=10	QV=0.01	0.092	0.707	0.541	7.004	-0.012	0.009	0.567
29	Q			QV=0.05	0.049	0.364	0.226	1.180	-0.008	0.006	0.393
30	Q			QV=0.10	0.063	0.398	0.220	0.505	-0.001	0.003	0.207
31	Q	N=500	EV=2.5	QV=0.01	0.010	0.063	0.057	0.736	0.000	0.000	0.007
32	Q			QV=0.05	-0.006	-0.012	-0.008	-0.016	0.000	0.000	0.011
33	Q			QV=0.10	-0.003	0.000	-0.001	0.003	0.000	0.000	0.002
34	Q	N=500	EV=5	QV=0.01	0.024	0.146	0.107	1.419	0.000	0.001	0.050
35	Q			QV=0.05	0.000	0.014	0.025	0.079	0.000	0.000	0.020
36	Q			QV=0.10	0.011	0.055	0.022	0.012	0.000	0.000	-0.010
37	Q	N=500	EV=10	QV=0.01	0.040	0.278	0.238	3.604	-0.005	0.004	0.214
38	Q			QV=0.05	0.010	0.079	0.086	0.444	-0.001	0.001	0.070
39	Q			QV=0.10	0.000	-0.012	0.003	0.032	0.000	0.000	0.011
40	Q	N=700	EV=2.5	QV=0.01	0.007	0.053	0.044	0.514	0.000	0.000	0.003
41	Q			QV=0.05	0.002	0.007	-0.001	0.002	0.000	0.000	-0.009
42	Q			QV=0.10	-0.001	-0.016	-0.013	-0.004	0.000	0.000	-0.008
43	Q	N=700	EV=5	QV=0.01	0.011	0.093	0.083	1.193	0.000	0.000	0.015
44	Q			QV=0.05	0.003	0.007	0.009	0.046	0.000	0.000	-0.010
45	Q			QV=0.10	0.002	-0.013	-0.012	-0.008	-0.001	0.000	0.011
46	Q	N=700	EV=10	QV=0.01	0.035	0.234	0.202	2.580	-0.002	0.002	0.099
47	Q			QV=0.05	0.007	0.069	0.062	0.212	-0.001	0.000	0.036
48	Q			QV=0.10	0.005	0.025	0.033	0.046	-0.001	0.000	0.006

Relative Variance Bias

Table 11 presents the relative variance bias for AIC-model-selection estimates. In general, the relative variance bias for the linear model is smaller than that for the quadratic model. This result is similar to the result of the true-model-selection estimates in section 4.1.2.1. For parameters α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}, the relative variance bias rates range from

-0.105 to 0.097 for the linear model, and from -0.125 to 0.390 for the quadratic model.

For parameters α_1, α_2, and α_3, the relative variance bias rates range from -0.125 to 0.111. This indicates that the variability of the estimates for α_1, α_2, and α_3 are estimated quite adequately. Recovery for ϕ_{11} also appears adequate, with the relative variance bias ranging from -0.078 to 0.116. Recovery for ϕ_{22} is mediocre to adequate, with the relative variance bias ranging from -0.097 to 0.390. Recovery for ϕ_{33}, however, is poor, ranging from -0.111 to 9.449. When sample size and parameter value for ϕ_{33} are larger, the relative variance bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 100 to 700 and parameter value for quadratic variance increases from 0.01 to 0.1, the relative variance bias for ϕ_{33} decreases substantially, from 9.449 to 0.069. The relative variance bias for ϕ_{21} ranges from -0.105 to 0.145, indicating the variability is estimated adequately.

Table 11. Relative Variance Bias for AIC-Model-Selection Estimates

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	-0.078	-0.105	-0.071	--	-0.071	0.003	--
2	L		EV=5	--	-0.009	-0.068	0.062	--	0.078	0.000	--
3	L		EV=10	--	0.004	0.010	-0.020	--	-0.017	0.009	--
4	L	N=200	EV=2.5	--	0.002	0.022	-0.013	--	-0.029	-0.013	--
5	L		EV=5	--	0.076	0.014	-0.073	--	-0.028	0.049	--
6	L		EV=10	--	-0.025	0.003	-0.058	--	-0.045	-0.026	--
7	L	N=500	EV=2.5	--	-0.060	-0.057	0.018	--	-0.086	0.016	--
8	L		EV=5	--	-0.022	0.097	0.022	--	-0.004	-0.006	--
9	L		EV=10	--	0.000	-0.025	-0.028	--	-0.073	0.040	--
10	L	N=700	EV=2.5	--	-0.034	0.014	0.044	--	0.002	0.000	--
11	L		EV=5	--	0.037	-0.004	0.042	--	-0.010	-0.046	--
12	L		EV=10	--	0.076	0.022	-0.039	--	-0.078	-0.027	--
13	Q	N=100	EV=2.5	QV=0.01	0.116	0.083	0.125	9.449	-0.019	-0.014	0.038
14	Q			QV=0.05	0.058	-0.022	0.040	2.038	-0.125	-0.122	-0.033
15	Q			QV=0.10	-0.059	-0.043	0.014	0.429	0.058	-0.039	-0.004
16	Q	N=100	EV=5	QV=0.01	0.018	0.029	0.195	1.650	-0.068	-0.080	0.025
17	Q			QV=0.05	0.010	-0.042	0.111	0.898	-0.018	-0.033	-0.026
18	Q			QV=0.10	0.025	-0.027	0.046	0.349	-0.023	-0.033	0.060
19	Q	N=100	EV=10	QV=0.01	0.007	0.098	0.390	0.876	0.002	0.083	0.027
20	Q			QV=0.05	0.036	0.039	0.213	0.610	-0.076	-0.013	0.085
21	Q			QV=0.10	-0.001	-0.001	0.100	0.362	-0.044	-0.015	0.059
22	Q	N=200	EV=2.5	QV=0.01	0.053	0.070	0.127	6.747	0.015	0.008	0.067
23	Q			QV=0.05	-0.053	0.033	0.086	1.098	0.004	0.012	0.009
24	Q			QV=0.10	-0.020	0.017	0.066	0.210	0.042	0.005	0.030
25	Q	N=200	EV=5	QV=0.01	-0.020	-0.031	0.149	1.456	0.039	0.003	0.011
26	Q			QV=0.05	-0.019	0.033	0.062	0.449	-0.035	-0.002	-0.019
27	Q			QV=0.10	0.036	0.014	-0.030	0.154	-0.015	0.097	0.074
28	Q	N=200	EV=10	QV=0.01	0.069	0.108	0.230	0.717	-0.025	-0.055	-0.049
29	Q			QV=0.05	-0.020	0.021	0.119	0.438	0.001	0.010	0.000
30	Q			QV=0.10	0.029	0.004	0.070	0.210	0.001	-0.021	-0.040
31	Q	N=500	EV=2.5	QV=0.01	0.000	0.076	0.055	4.561	-0.020	-0.013	0.027
32	Q			QV=0.05	0.095	0.009	0.048	0.326	0.111	-0.032	-0.042
33	Q			QV=0.10	0.079	0.001	-0.013	0.000	0.027	-0.021	0.057
34	Q	N=500	EV=5	QV=0.01	-0.074	0.071	0.180	1.281	0.077	-0.041	-0.039
35	Q			QV=0.05	-0.007	0.018	0.102	0.314	0.092	-0.035	0.000
36	Q			QV=0.10	0.011	0.046	0.108	0.067	0.057	0.019	-0.032
37	Q	N=500	EV=10	QV=0.01	0.092	0.117	0.242	0.571	0.066	-0.029	-0.039
38	Q			QV=0.05	0.081	0.115	0.187	0.255	0.005	0.052	0.085
39	Q			QV=0.10	0.065	0.093	0.125	0.166	0.075	0.100	0.045
40	Q	N=700	EV=2.5	QV=0.01	0.060	0.145	0.175	4.267	-0.016	0.092	0.000
41	Q			QV=0.05	0.024	0.032	-0.013	0.114	-0.049	-0.063	-0.029
42	Q			QV=0.10	0.030	-0.022	-0.027	-0.073	-0.019	0.043	0.000
43	Q	N=700	EV=5	QV=0.01	-0.016	-0.024	0.050	0.908	0.014	0.009	0.039
44	Q			QV=0.05	0.049	0.041	0.022	0.158	-0.092	-0.055	0.018
45	Q			QV=0.10	-0.075	-0.040	-0.097	-0.111	-0.071	0.061	0.000
46	Q	N=700	EV=10	QV=0.01	0.019	0.042	0.151	0.505	0.057	0.010	0.020
47	Q			QV=0.05	0.076	0.056	0.085	0.199	-0.049	0.008	0.058
48	Q			QV=0.10	0.085	0.092	0.086	0.069	-0.042	0.000	0.000

4.1.2.3 BIC-model-selection estimates

Relative Parameter Bias

Table 12 presents the relative parameter bias for BIC-model-selection estimates. In general, the parameter estimates for the linear model are better than those for the quadratic model. This is similar to the results of the true-model-selection estimates in section 4.1.2.1 and the AIC-model-selection estimates in section 4.1.2.2. For the parameters common to both the linear and quadratic models (i.e., α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}), the relative bias rates range from -0.031 to 0.006 for the linear model, and from -0.575 to 23.790 for the quadratic model.

Similar to the results for the true-model-selection estimates and the AIC-model-selection estimates, mean parameters α_1 and α_2 are recovered satisfactorily, with the relative parameter bias rates ranging only from -0.075 to 0.035. Recovery for mean parameter α_3, however, appears worse than that under true-model-selection or AIC-model-selection. When sample size is large (700), the relative parameter bias for α_3 ranges from -0.008 to 0.590 (vice -0.009 to 0.099 under AIC-model-selection, and -0.009 to 0.005 under true-model-selection). When sample size is small (100), the relative parameter bias recovery is even worse, with the relative parameter bias ranging from 0.460 to 2.382 (vice 0.065 to 0.913 under AIC-model-selection, and -0.028 to 0.040 under true-model-selection).

Recovery for variance parameter ϕ_{11} appears mediocre to adequate, with the relative parameter bias ranging from -0.019 to 0.419. Recovery for parameters ϕ_{21}, ϕ_{22}, and ϕ_{33} is poor to adequate, with the relative parameter bias ranging from -0.008

to 23.790. When sample size and parameter value for ϕ_{33} are larger, the relative parameter bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 100 to 700 and parameter value for ϕ_{33} increases from 0.01 to 0.1, the relative parameter bias for ϕ_{33} decreases markedly, from 23.790 to -0.004.

Table 12. Relative Parameter Bias for BIC-Model-Selection Estimates

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	-0.003	-0.015	-0.012	--	-0.001	0.000	--
2	L		EV=5	--	-0.004	-0.025	-0.011	--	-0.001	0.000	--
3	L		EV=10	--	-0.009	-0.031	-0.017	--	0.001	0.000	--
4	L	N=200	EV=2.5	--	-0.003	-0.002	0.001	--	-0.001	0.000	--
5	L		EV=5	--	0.001	-0.003	0.003	--	0.001	0.000	--
6	L		EV=10	--	-0.010	-0.026	-0.012	--	0.000	0.000	--
7	L	N=500	EV=2.5	--	-0.002	-0.008	-0.004	--	0.000	0.000	--
8	L		EV=5	--	0.005	0.001	-0.003	--	-0.001	0.000	--
9	L		EV=10	--	-0.002	-0.010	-0.004	--	0.001	0.000	--
10	L	N=700	EV=2.5	--	-0.002	-0.004	-0.004	--	0.000	0.000	--
11	L		EV=5	--	-0.003	-0.007	-0.004	--	0.000	0.000	--
12	L		EV=10	--	0.003	0.006	0.000	--	0.000	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	0.039	0.145	0.145	2.398	-0.010	0.011	0.733
14	Q			QV=0.05	-0.005	-0.048	-0.052	0.619	-0.004	0.006	0.460
15	Q			QV=0.10	0.004	-0.004	-0.025	0.186	-0.001	0.002	0.200
16	Q	N=100	EV=5	QV=0.01	-0.015	-0.575	-0.152	5.791	-0.029	0.022	1.252
17	Q			QV=0.05	0.069	0.310	0.045	1.059	-0.007	0.013	1.069
18	Q			QV=0.10	0.002	-0.006	-0.089	0.763	-0.004	0.006	0.456
19	Q	N=100	EV=10	QV=0.01	0.117	1.821	1.753	23.790	-0.075	0.035	2.382
20	Q			QV=0.05	0.116	1.350	0.590	4.463	-0.031	0.021	1.308
21	Q			QV=0.10	0.031	-0.168	-0.202	1.769	-0.021	0.012	0.700
22	Q	N=200	EV=2.5	QV=0.01	0.019	0.086	0.081	1.969	-0.009	0.006	0.390
23	Q			QV=0.05	0.007	0.042	0.003	0.259	-0.002	0.002	0.206
24	Q			QV=0.10	-0.003	-0.023	-0.006	0.011	-0.001	0.000	0.023
25	Q	N=200	EV=5	QV=0.01	0.026	0.223	0.185	2.736	-0.021	0.016	0.897
26	Q			QV=0.05	0.008	0.026	-0.026	0.980	-0.006	0.008	0.504
27	Q			QV=0.10	0.005	0.028	-0.039	0.339	-0.003	0.002	0.177
28	Q	N=200	EV=10	QV=0.01	0.419	2.348	1.159	9.470	-0.034	0.029	1.711
29	Q			QV=0.05	0.162	1.257	0.638	3.181	-0.014	0.017	1.171
30	Q			QV=0.10	0.105	0.526	0.232	1.313	-0.010	0.009	0.661
31	Q	N=500	EV=2.5	QV=0.01	0.012	0.071	0.057	0.934	-0.001	0.002	0.100
32	Q			QV=0.05	-0.005	-0.014	-0.013	-0.002	0.000	0.000	0.025
33	Q			QV=0.10	-0.003	0.000	-0.001	0.003	0.000	0.000	0.002
34	Q	N=500	EV=5	QV=0.01	0.029	0.160	0.088	1.804	-0.005	0.006	0.373
35	Q			QV=0.05	-0.003	-0.030	-0.025	0.231	-0.003	0.003	0.179
36	Q			QV=0.10	0.012	0.058	0.018	0.032	0.000	0.000	0.005
37	Q	N=500	EV=10	QV=0.01	0.077	0.403	0.380	5.863	-0.012	0.013	0.836
38	Q			QV=0.05	-0.019	-0.105	-0.074	1.068	-0.013	0.008	0.517
39	Q			QV=0.10	0.003	-0.076	-0.093	0.270	-0.003	0.002	0.184
40	Q	N=700	EV=2.5	QV=0.01	0.007	0.052	0.041	0.544	0.000	0.000	0.030
41	Q			QV=0.05	0.002	0.007	-0.002	0.003	0.000	0.000	-0.008
42	Q			QV=0.10	-0.001	-0.016	-0.013	-0.004	0.000	0.000	-0.008
43	Q	N=700	EV=5	QV=0.01	0.010	0.067	0.068	1.551	-0.005	0.004	0.236
44	Q			QV=0.05	0.001	-0.021	-0.022	0.113	-0.002	0.001	0.062
45	Q			QV=0.10	0.002	-0.014	-0.013	-0.008	-0.001	0.000	0.013
46	Q	N=700	EV=10	QV=0.01	0.011	0.170	0.165	3.145	-0.012	0.010	0.590
47	Q			QV=0.05	0.006	0.002	-0.024	0.728	-0.009	0.005	0.345
48	Q			QV=0.10	0.002	-0.020	-0.023	0.158	-0.003	0.001	0.106

Relative Variance Bias

Table 13 presents the relative variance bias for BIC-model-selection estimates. In general, the relative variance bias for the linear model is smaller than that for the quadratic model. This result is similar to the result of the true-model-selection estimates and that of the AIC-model-selection estimates. For parameters α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}, the relative variance bias rates range from -0.107 to 0.095 for the linear model, and from -0.276 to 0.880 for the quadratic model.

For mean parameters α_1, α_2, and α_3, the relative variance bias rates range from -0.165 to 0.122. This indicates that the variation of the estimates for parameters α_1, α_2, and α_3 are estimated quite adequately. This result is similar to that for the true-model-selection estimates and that for the AIC-model-selection estimates.

Recovery for ϕ_{11} appears mediocre, with the relative variance bias ranging from -0.165 to 0.359. Recovery for ϕ_{22} is poor to mediocre, with the relative variance bias ranging from -0.276 to 0.880. Recovery for ϕ_{33} is poor, ranging from -0.111 to 11.157. When sample size and parameter value for ϕ_{33} are larger, the relative variance bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 100 to 700 and parameter value for quadratic variance increases from 0.01 to 0.1, the relative variance bias for ϕ_{33} decreases substantially, from 11.157 to -0.073. The relative variance bias for ϕ_{21} ranges from -0.271 to 0.511, indicating the recovery is mediocre.

Table 13. Relative Variance Bias for BIC-Model-Selection Estimates

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	-0.079	-0.107	-0.073	--	-0.071	0.002	--
2	L		EV=5	--	-0.007	-0.067	0.062	--	0.079	0.000	--
3	L		EV=10	--	0.002	0.006	-0.022	--	-0.018	0.008	--
4	L	N=200	EV=2.5	--	0.007	0.026	-0.012	--	-0.027	-0.012	--
5	L		EV=5	--	0.076	0.015	-0.072	--	-0.028	0.050	--
6	L		EV=10	--	-0.025	0.005	-0.056	--	-0.045	-0.025	--
7	L	N=500	EV=2.5	--	-0.059	-0.057	0.017	--	-0.085	0.016	--
8	L		EV=5	--	-0.023	0.095	0.021	--	-0.005	-0.006	--
9	L		EV=10	--	-0.002	-0.027	-0.030	--	-0.074	0.038	--
10	L	N=700	EV=2.5	--	-0.034	0.014	0.042	--	0.002	0.000	--
11	L		EV=5	--	0.038	-0.004	0.043	--	-0.010	-0.046	--
12	L		EV=10	--	0.077	0.023	-0.038	--	-0.077	-0.027	--
13	Q	N=100	EV=2.5	QV=0.01	0.105	0.058	0.107	11.157	-0.031	-0.043	-0.016
14	Q			QV=0.05	0.049	-0.075	-0.027	1.326	-0.122	-0.164	-0.017
15	Q			QV=0.10	-0.070	-0.085	-0.047	0.320	0.056	-0.073	0.014
16	Q	N=100	EV=5	QV=0.01	-0.039	-0.124	-0.004	1.595	-0.063	-0.165	-0.016
17	Q			QV=0.05	0.002	-0.082	0.024	0.780	-0.017	-0.095	-0.063
18	Q			QV=0.10	-0.005	-0.099	-0.054	0.242	-0.029	-0.093	0.085
19	Q	N=100	EV=10	QV=0.01	0.036	0.213	0.580	1.009	-0.071	0.115	0.098
20	Q			QV=0.05	-0.043	-0.017	0.230	0.712	-0.112	-0.048	0.122
21	Q			QV=0.10	-0.165	-0.271	-0.276	0.095	-0.055	-0.106	0.066
22	Q	N=200	EV=2.5	QV=0.01	0.056	0.064	0.151	5.747	0.013	-0.019	0.045
23	Q			QV=0.05	-0.045	0.039	0.105	0.755	0.002	-0.019	0.009
24	Q			QV=0.10	-0.021	0.013	0.063	0.198	0.042	0.000	0.030
25	Q	N=200	EV=5	QV=0.01	-0.038	-0.064	0.062	1.330	0.024	-0.017	-0.022
26	Q			QV=0.05	-0.033	-0.020	-0.015	0.375	-0.033	-0.042	-0.019
27	Q			QV=0.10	0.019	-0.027	-0.077	0.154	-0.017	0.061	0.084
28	Q	N=200	EV=10	QV=0.01	0.359	0.511	0.880	1.791	0.023	-0.055	-0.010
29	Q			QV=0.05	-0.031	-0.033	-0.001	0.544	-0.007	-0.029	-0.008
30	Q			QV=0.10	-0.010	-0.072	-0.032	0.240	0.006	-0.059	-0.021
31	Q	N=500	EV=2.5	QV=0.01	0.005	0.084	0.072	4.317	-0.020	-0.017	0.027
32	Q			QV=0.05	0.097	0.009	0.049	0.315	0.112	-0.034	-0.042
33	Q			QV=0.10	0.079	0.001	-0.013	0.000	0.027	-0.021	0.057
34	Q	N=500	EV=5	QV=0.01	-0.074	0.059	0.153	1.258	0.080	-0.048	-0.039
35	Q			QV=0.05	-0.008	0.010	0.097	0.324	0.092	-0.049	0.000
36	Q			QV=0.10	0.012	0.046	0.108	0.067	0.058	0.016	-0.021
37	Q	N=500	EV=10	QV=0.01	0.140	0.170	0.350	0.765	0.080	-0.027	-0.039
38	Q			QV=0.05	0.043	0.059	0.150	0.383	-0.004	0.020	0.092
39	Q			QV=0.10	0.040	0.043	0.066	0.193	0.078	0.076	0.045
40	Q	N=700	EV=2.5	QV=0.01	0.059	0.141	0.171	4.267	-0.016	0.089	0.000
41	Q			QV=0.05	0.024	0.032	-0.013	0.114	-0.049	-0.063	-0.029
42	Q			QV=0.10	0.030	-0.022	-0.027	-0.073	-0.019	0.043	0.000
43	Q	N=700	EV=5	QV=0.01	-0.011	-0.016	0.075	0.921	0.012	0.000	0.039
44	Q			QV=0.05	0.049	0.039	0.022	0.171	-0.091	-0.062	0.018
45	Q			QV=0.10	-0.075	-0.040	-0.097	-0.111	-0.071	0.061	0.000
46	Q	N=700	EV=10	QV=0.01	-0.016	0.001	0.098	0.464	0.039	-0.001	0.020
47	Q			QV=0.05	0.064	0.030	0.072	0.290	-0.049	-0.014	0.058
48	Q			QV=0.10	0.078	0.078	0.075	0.089	-0.041	-0.010	0.000

As indicated above in section 4.1.2.2, the examination of AIC model selection estimates shows that recovery for parameters α_1, α_2, and ϕ_{11} is satisfactory for all conditions. Recovery for parameters α_3, ϕ_{21}, ϕ_{22}, and ϕ_{33}, however, is mediocre to poor when sample size is small, although recovery is adequate when sample size is large. As indicated in section 2.1.2.3, the examination of the BIC-model-selection estimates shows that recovery for parameters α_1 and α_2 is satisfactory for all conditions. Recovery for parameters α_3, ϕ_{11}, ϕ_{21}, ϕ_{22}, and ϕ_{33} is adequate when sample size is large, but is mediocre to poor when sample size is small. Generally, it is expected that the parameter estimates are close to the population parameter values. If the estimated parameter values are markedly different from the population parameter values, the difference might result from the preliminary model selection. Considering the true-model-selection bias to be the baseline, the deviations of AIC- and BIC-model-selection estimates from the population parameter values were further investigated in order to assess the possible impact of the preliminary model selection.

In order to assess the impact of model selection, the differences of the relative biases of the true-model-selection estimates, the AIC-model-selection estimates, and the BIC-model-selection estimates were calculated. When the absolute value of the difference is greater than 0.1, it is evidence that the preliminary model selection had impact on parameter estimates. Table 14 summarizes the results. Detailed results for each condition are presented in Appendix A, Table A1 though A6.

Table 14. Differences of Relative Bias between True-Model-Selection, AIC-Model-Selection, and BIC-Model-Selection Estimates

		ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
Difference of Relative Parameter Bias								
AIC-True	Mean	0.0049	0.0266	0.0065	0.4613	0.0024	0.0018	0.1573
	Abs>0.1	1/48	3/48	3/48	18/36	0/48	0/48	15/36
BIC-True	Mean	0.0129	0.0657	0.0086	1.2306	-0.0073	0.0056	0.4924
	Abs>0.1	2/48	11/48	14/48	27/36	0/48	0/48	26/36
BIC-AIC	Mean	0.0080	0.0391	0.0020	0.7675	-0.0049	0.0039	0.3351
	Abs>0.1	2/48	11/48	9/48	26/36	0/48	0/48	26/36
Difference of Relative Variance Bias								
AIC-True	Mean	-0.0010	0.0979	0.0265	0.1665	0.0065	-0.0012	0.1116
	Abs>0.1	3/48	14/48	9/48	27/36	2/48	3/48	13/36
BIC-True	Mean	0.0070	0.1370	0.0285	0.9894	0.0016	0.0026	0.4840
	Abs>0.1	6/48	14/48	16/48	30/36	2/48	3/48	26/36
BIC-True	Mean	0.0082	0.1403	0.0255	0.9948	0.0015	0.0079	0.4835
	Abs>0.1	6/48	15/48	15/48	30/36	3/48	4/48	26/36

AIC-Pre = difference between the AIC-model-selection estimates and the true-model-selection estimates.
BIC-Pre = difference between the BIC-model-selection estimates and the true-model-selection estimates.
BIC-AIC = difference between the BIC-model-selection estimates and the AIC-model-selection estimates.
Abs> 0.1 = the number of cases in which the absolute value of the difference is greater than 0.1.

One can see from this table that, in terms of relative parameter bias, model selection by AIC or BIC has a strong impact on parameter estimates for parameters ϕ_{33} and α_3. The mean of the differences of relative parameter bias for parameter ϕ_{33} is 0.4613 for AIC-True, 1.2306 for BIC-True, and 0.7675 for BIC-AIC. The mean of the differences of relative parameter bias for parameter α_3 is 0.1573 for AIC-True, 0.4924 for BIC-True, and 0.3351 for BIC-AIC. This indicates that in terms of accuracy, the quality of the true-model-selection estimators is better than that of the AIC-model-selection estimators and that of the BIC-model-selection estimators.

Additionally, for parameter ϕ_{33}, the absolute value of the difference is greater than 0.1 in 18 cases for AIC-True, which is 50% of the total 36 cases when the quadratic model is the true model; in 27 cases for BIC-True, which is 75% of the total 36 cases; and in 26 cases for BIC-AIC, which is 72% of the total 36 cases. For parameter α_3, the absolute value of the difference is greater than 0.1.in 15 cases for AIC-True, which is 42% of the total 36 cases when the quadratic model is the true model; in 26 cases for BIC-True, which is 72% of the total 36 cases; and in 26 cases for BIC-AIC, which is 72% of the total 36 cases.

One can also see that, in terms of relative variance bias, model selection by AIC or BIC has obvious impact on parameter estimates for parameters ϕ_{21}, ϕ_{33} and α_3. This is most obvious in the case of parameter ϕ_{33}. The mean difference of relative variance bias for parameter ϕ_{33} is especially large (0.9894) for BIC-True. Also, the absolute value of the difference is greater than 0.1 in 30 cases for AIC-True, which is 83% of the total 36 cases when the quadratic model is the true model.

4.1.3 Coverage Rate

The effect of preliminary model selection on coverage rate (i.e., the number of replications whose 95% confidence intervals contain the true population parameter) is evaluated in this section. Two coverage rates, unconditional and conditional, were calculated for parameters α_1, α_2, α_3, ϕ_{11}, ϕ_{22}, ϕ_{33}, and ϕ_{21} for a total of 48 cases. Unconditional coverage rates were calculated without model selection, and conditional coverage rates were calculated with model selection by AIC and BIC. A coverage rate was considered adequate if it was between 0.925 and 0.975. Tables 15, 16, and 17 show the descriptive statistics for the true-model-selection, the AIC-model-selection, and the BIC-model-selection coverage rates, respectively. (Detailed coverage rates for all 48 cases in each condition are presented in Appendix A, Tables A7, A8, and A9.) The unconditional and conditional coverage rates were then compared to each other to assess the impact of model selection, as shown in Table 18.

True-Model-Selection Coverage

The true-model-selection coverage rate is calculated to evaluate the ability of the analysis model to recover the population parameters. This unconditional coverage rate is used as the baseline, to which the AIC- and BIC-model-selection coverage rates are later compared. In general, the unconditional coverage rate is closer to the nominal level of 0.950 when the linear model is the true model than it is when the quadratic model is the true model. When the linear model is the true model, the coverage rates for the parameters are all between 0.925 and 0.975. When the quadratic model is the true model, however, the coverage rates are outside 0.925 and

0.975 for 15 cases for parameter ϕ_{22} and 11 cases for parameter ϕ_{33}. Table 15 shows the descriptive statistics for this true-model-selection coverage rate.

One can see from this table that coverage for parameters ϕ_{11}, α_1, α_2, and α_3 is adequate, with no coverage rates falling outside 0.925 and 0.975. In fact, the mean coverage rate for those parameters is close to the nominal level of 0.95. Coverage for parameters ϕ_{22} and ϕ_{33}, however, is not adequate, with coverage rates falling outside 0.925 and 0.975 in 15 cases for ϕ_{22} and 11 cases for ϕ_{33}.

<u>AIC-Model-Selection Coverage</u>

The AIC-model-selection coverage rate was also calculated, in order to evaluate the ability of the analysis model to recover the population parameters, conditional on selecting a correct model by AIC. As with the true-model-selection coverage rate, the AIC-model-selection coverage rate is closer to the nominal level when the linear model is the true model than when the quadratic model is the true model. When the linear model is the true model, the AIC-model-selection coverage rates are all between 0.925 and 0.975. When the quadratic model is the true model, however, there are several cases in which coverage rates fall outside 0.925 and 0.975, for all parameters except α_1. Table 16 shows the descriptive statistics for this conditional coverage rate.

One can see from Table 16 that conditional coverage for parameter α_1 remains adequate, with no cases in which coverage rates fall outside 0.925 and 0.975. Conditional coverage for parameters ϕ_{11}, α_2, and α_3, however, is mediocre. And, conditional coverage for variance parameters ϕ_{21}, ϕ_{22} and ϕ_{33}, is poor, with 15 ,19, and 22 cases respectively in which the conditional rates fall outside 0.925 and 0.975. Also, the mean conditional rates for these parameters are distant from the nominal level of 0.95.

BIC-Model-Selection Coverage

As was seen with the true-model-selection coverage rate and the AIC-model-selection coverage rate, the BIC-model-selection coverage rate is closer to the nominal level when the linear model is the true model than when the quadratic model is the true model. When the linear model is the true model, the coverage rates for the parameters are all between 0.925 and 0.975. When the quadratic model is the true model, however, there are several cases in which coverage rates fall outside 0.925 and 0.975 for each parameter.

One can see from Table 17 that the number of cases in which coverage rates fall outside 0.925 and 0.975 ranges from 10 to 27. Also, most of the mean conditional rates are distant from the nominal level of 0.95. This indicates that the BIC-model-selection coverage is not adequate. Note that a 0% coverage rate occurs for parameters α_2 and α_3, and a 100% coverage rate for parameters ϕ_{11}, ϕ_{22}, and ϕ_{21}. These are markedly abnormal and may result from the low success rate of selecting a correct model by BIC in some cases (e.g., case 28 in Table 6, which has sample size = 200, error variance = 10, quadratic variance = 0.01, and a very small success rate of 0.002). Because BIC does not perform effectively in such cases, these conditional coverage rates might be invalid.

Tables15, 16, and 17 demonstrate that there are substantial differences in the numbers of cases in which coverage rate falls outside 0.925 and 0.975 for true model selection, AIC model selection , and BIC model selection. The numbers for AIC model selection and BIC model selection are always greater than that for true model selection. This suggests that model selection has a significant impact on coverage rate.

Table 15. Descriptive Statistics for True-Model-Selection Coverage Rate

	Minimum	Maximum	Mean	Cases in which coverage rate falls outside 0.925 and 0.975	Total Cases
ϕ_{11}	.926	.965	.946	0	48
ϕ_{21}	.906	.974	.942	5	48
ϕ_{22}	.760	.957	.921	15	48
ϕ_{33}	.695	.977	.937	11	36
α_1	.932	.963	.948	0	48
α_2	.934	.967	.948	0	46
α_3	.933	.963	.950	0	36

Table 16. Descriptive Statistics for AIC-Model-Selection Coverage Rate

	Minimum	Maximum	Mean	Cases in which coverage rate falls outside 0.925 and 0.975	Total Cases
ϕ_{11}	.863	.965	.939	7	48
ϕ_{21}	.841	.974	.930	15	48
ϕ_{22}	.645	.956	.903	19	48
ϕ_{33}	.723	.976	.927	12	36
α_1	.925	.965	.947	0	48
α_2	.866	.974	.944	6	48
α_3	.883	.976	.944	9	36

Table 17. Descriptive statistics for BIC-Model-Selection Coverage Rate

	Minimum	Maximum	Mean	Cases in which coverage rate falls outside 0.925 and 0.975	Total Cases
ϕ_{11}	.786	1.000	.931	13	48
ϕ_{21}	.714	1.000	.913	22	48
ϕ_{22}	.464	1.000	.883	27	48
ϕ_{33}	.667	1.000	.904	15	36
α_1	.667	1.000	.935	10	48
α_2	.000	.968	.843	19	46
α_3	.000	.971	.771	18	36

To assess the impact of model selection even further, the differences between the true-model-selection, the AIC-model-selection, and the BIC-model-selection coverage rates were examined. Also, the partial correlations of the differences of the coverage rates with respect to sample size, error variance, and quadratic variance, respectively, were calculated. Table 18 summarizes the results.

One can see from Table 18 that model selection by AIC has an impact on the coverage rate for parameter ϕ_{22}. The absolute value of the difference for parameter ϕ_{22} is greater than 0.25 in 11 cases for AIC-True, which is 23% of the total 48 cases. The mean of the differences of coverage rates for parameter ϕ_{22} is -0.018 for AIC-True. This indicates that, on average, conditional coverage rates of AIC model selection are smaller than unconditional coverage rates. This difference correlates positively with sample size when controlling error variance and quadratic variance, with a strong correlation of 0.737 at a significance level of 0.05. This difference correlates negatively with error variance when controlling sample size and quadratic variance, with a strong correlation of -0.667 at a significance level of 0.05.

Compared to AIC model selection, BIC model selection appears to have a greater impact on coverage rates. For example, the mean of the differences of coverage rates is -0.028 for ϕ_{21}, -0.037 for ϕ_{22}, and -0.032 for ϕ_{33}. This indicates that, on average, conditional coverage rates of BIC model selection are smaller than unconditional coverage rates. These means of differences are all positively correlated with sample size. Also, the absolute value of the difference is greater than 0.25 in 18 out of the 48 cases for ϕ_{21}, in 22 out of the 48 cases for ϕ_{22}, and in 22 out of the 36 cases for ϕ_{33}.

Table 18. Differences between True-Model-Selection, AIC-Model-Selection, and BIC-Model-Selection Coverage Rates

		ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
AIC-True	Abs>0.025/ total cases	6/48	8/48	11/48	8/36	0/48	6/48	7/36
	Mean	-0.007	-0.011	-0.018	-0.009	-0.000	-0.004	-0.006
	Partial correlation with sample size	0.635**	0.748**	0.737**	0.463**	0.323*	0.557**	0.578**
	Partial correlation with error variance	-0.663**	-0.665**	-0.667**	-0.103	-0.153	-0.348**	-0.395**
	Partial correlation with quadratic variance	0.411**	0.511**	0.151	0.506**	0.039	0.145	0.188
BIC-True	Abs>0.025/ total cases	11/48	18/48	22/48	17/36	7/48	19/48	19/36
	Mean	-0.015	-0.028	-0.037	-0.032	-0.012	-0.104	-0.179
	Partial correlation with sample size	0.251	0.371**	0.294*	0.491**	0.206	0.519**	0.554**
	Partial correlation with error variance	-0.381**	-0.247	-0.330*	-0.077	-0.187	-0.652**	-0.618**
	Partial correlation with quadratic variance	-0.121	0.009	-0.237	0.269	0.270	0.646**	0.690**
BIC-AIC	Abs>0.025/ total cases	10/48	19/48	16/48	17/36	9/48	19/48	20/36
	Mean	-0.007	-0.017	-0.019	-0.017	-0.012	-0.100	-0.129
	Partial correlation with sample size	0.047	0.134	0.037	0.434**	0.167	0.471**	0.521**
	Partial correlation with error variance	-0.174	-0.049	-0.128	-0.054	-0.167	-0.642**	-0.609**
	Partial correlation with quadratic variance	-0.221	-0.121	-0.277	0.111	0.261	0.649**	0.695**

AIC-True = difference between the AIC-model-selection coverage rate and the true-model-selection coverage rate
BIC-True = difference between the BIC-model-selection coverage rate and the true-model-selection coverage rate
BIC-AIC = difference between the BIC-model-selection coverage rate and the AIC-model-selection coverage rate
Abs>0.025 = the number of cases in which the absolute value of the difference coverage rate is greater than 0.025
* indicates a significant level of 0.10; ** indicates a significant level of 0.05.

In summary, compared to the true-model-selection coverage rate, the AIC- and BIC-model-selection coverage rates have more cases in which the coverage rates substantially deviate from the nominal rate of 0.95. In addition, the means of the differences between the unconditional rate and the conditional rates are all negative. This indicates that the conditional rates tend to be underestimated. Moreover, differences between the unconditional coverage rate and the conditional coverage rates are substantial for parameters ϕ_{21}, ϕ_{22}, and ϕ_{33}. These differences are positively correlated with sample size and negatively correlated with error variance. This suggests that sample size and error variance play important roles to determine the coverage rates after the correct model is selected.

4.2. Study2: Conducting Model Selection and Parameter Estimation Using Split-Data Sets

In study 2, each original data set generated from Study 1 is randomly separated into two parts by 50% vs. 50% data splitting in order to assess data splitting method as a possible way to mitigate the effects of model uncertainty. The first part is used to select an appropriate model based on model fit evaluation. The second part is then used to estimate the parameters based on the model selected.

4.2.1 Model Selection Accuracy

Table 19 displays the model recovery rates of AIC and BIC using split-data sets under each condition. When the true model is the linear model (i.e., cases 1-12), four findings are noted. First, the model recovery rate of AIC ranges between 0.921 and 0.947 and the model recovery rate of BIC ranges between 0.997 and 1. This suggests that AIC and BIC reliably identify the true model. This also suggests that both AIC and BIC perform consistently in selecting the linear model and do not favor overfitting in all conditions. Second, Table 19 also shows that the model recovery rate of BIC is always larger than that of AIC across all conditions when the true model is the linear model. This indicates that BIC consistently performs better than AIC when the true model is the linear model. Third, both AIC and BIC appear to perform consistently in selecting the linear model in more than 90% of the replications under all 12 conditions when the true model is the linear model. BIC, however, tends to be more consistent than AIC, selecting the linear model in more than 99% of the replications under all 12 conditions. Fourth, both AIC and BIC model recovery rates correlate positively with sample size when controlling error variance and quadratic variance, with a strong partial correlation of 0.922 for AIC and 0.680 for BIC at a significance level of 0.05. They also correlate positively with quadratic variance, with a strong partial correlation of 0.798 for AIC and 0.653 for BIC. They correlate negatively, however, with error variance, with a strong partial correlation of -0.859 for AIC, and -0.706 for BIC. These results are similar to those shown in Table 6 in section 4.1.1 when the original data sets were used.

Table 19. Model Recovery Rate in Each Condition Using the Split-Data Sets

	True Model	Sample Size	Error Variance	Quadratic Variance	AIC	BIC
Case1	L	N=50	EV=2.5	--	0.938	0.997
Case2	L		EV=5	--	0.921	0.997
Case3	L		EV=10	--	0.926	0.999
Case4	L	N=100	EV=2.5	--	0.934	0.999
Case5	L		EV=5	--	0.926	1.000
Case6	L		EV=10	--	0.931	0.999
Case7	L	N=250	EV=2.5	--	0.927	1.000
Case8	L		EV=5	--	0.934	1.000
Case9	L		EV=10	--	0.933	1.000
Case10	L	N=350	EV=2.5	--	0.939	1.000
Case11	L		EV=5	--	0.947	0.999
Case12	L		EV=10	--	0.935	1.000
Case13	Q	N=50	EV=2.5	QV=0.01	0.269	0.028
Case14	Q			QV=0.05	0.446	0.087
Case15	Q			QV=0.10	0.720	0.250
Case16	Q	N=50	EV=5	QV=0.01	0.160	0.011
Case17	Q			QV=0.05	0.268	0.024
Case18	Q			QV=0.10	0.423	0.082
Case19	Q	N=50	EV=10	QV=0.01	0.138	0.004
Case20	Q			QV=0.05	0.143	0.009
Case21	Q			QV=0.10	0.241	0.022
Case22	Q	N=100	EV=2.5	QV=0.01	0.494	0.053
Case23	Q			QV=0.05	0.700	0.160
Case24	Q			QV=0.10	0.955	0.562
Case25	Q	N=100	EV=5	QV=0.01	0.330	0.024
Case26	Q			QV=0.05	0.432	0.039
Case27	Q			QV=0.10	0.711	0.184
Case28	Q	N=100	EV=10	QV=0.01	0.184	0.004
Case29	Q			QV=0.05	0.226	0.008
Case30	Q			QV=0.10	0.406	0.035
Case31	Q	N=250	EV=2.5	QV=0.01	0.908	0.261
Case32	Q			QV=0.05	0.987	0.629
Case33	Q			QV=0.10	1.000	0.980
Case34	Q	N=250	EV=5	QV=0.01	0.610	0.050
Case35	Q			QV=0.05	0.843	0.161
Case36	Q			QV=0.10	0.983	0.598
Case37	Q	N=250	EV=10	QV=0.01	0.375	0.012
Case38	Q			QV=0.05	0.514	0.024
Case39	Q			QV=0.10	0.790	0.126
Case40	Q	N=350	EV=2.5	QV=0.01	0.974	0.421
Case41	Q			QV=0.05	1.000	0.839
Case42	Q			QV=0.10	1.000	1.000
Case43	Q	N=350	EV=5	QV=0.01	0.773	0.093
Case44	Q			QV=0.05	0.941	0.286
Case45	Q			QV=0.10	0.998	0.814
Case46	Q	N=350	EV=10	QV=0.01	0.440	0.009
Case47	Q			QV=0.05	0.671	0.040
Case48	Q			QV=0.10	0.895	0.221

4.2.2 Relative Parameter Biases

After the first part of the split-data sets was used for model fitting, the second part was used to estimate the parameters based on the model selected. The relative parameter biases were calculated. Also, the differences were calculated between the true-model-selection, the AIC-model-selection using original data sets, the AIC-model-selection using split-data sets, the BIC-model-selection using original data sets, and the BIC-model-selection using split-data sets.

4.2.2.1 AIC-model-selection estimates

Relative Parameter Bias

Table 20 presents the relative parameter bias for AIC-model-selection estimates using the split-data sets. In general, the parameter estimates for the linear model are better than those for the quadratic model. For the parameters common to both the linear and quadratic models (i.e., α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}), the relative bias rates range from 0 to -0.05 for the linear model, and from -0.040 to 11.158 for the quadratic model.

Mean parameters α_1 and α_2 are recovered satisfactorily, with the relative parameter bias rate ranging only from -0.001 to 0.005. Recovery for mean parameter α_3 appears adequate when sample size is large (350), with the relative parameter bias ranging from -0.032 to 0.016. But when sample size is small (50), recovery becomes adequate to mediocre, with the relative parameter bias ranging from -0.042 to 0.177.

Recovery for variance parameter ϕ_{11} also appears adequate, with the relative parameter bias ranging from -0.022 to 0.127. Recovery for parameters ϕ_{22}, ϕ_{21}, and ϕ_{33} is poor to adequate, with the relative parameter biases ranging from -0.050 to 11.158. When sample size and parameter value for ϕ_{33} are larger, the relative parameter bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 50 to 350 and parameter value for ϕ_{33} increases from 0.01 to 0.1, the relative parameter bias for ϕ_{33} decreases markedly, from 11.158 to 0.007.

Relative Variance Bias

Table 21 presents the relative variance biases for AIC-model-selection estimates using the split-data sets. In general, the relative variance bias for both the linear model and quadratic model is large. For all parameters, the relative variance bias is greater than 0.672. This indicates that recovery for all parameters is poor when using the split-data sets.

Table 20. Relative Parameter Bias for AIC-Model-Selection Estimates Using the Split-Data Sets

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=50	EV=2.5	--	-0.012	-0.041	-0.016	--	0.000	0.000	--
2	L		EV=5	--	-0.005	-0.013	-0.001	--	-0.002	0.000	--
3	L		EV=10	--	-0.022	-0.024	-0.019	--	0.001	0.000	--
4	L	N=100	EV=2.5	--	0.001	0.006	-0.004	--	-0.002	0.000	--
5	L		EV=5	--	0.012	0.006	0.000	--	0.001	0.000	--
6	L		EV=10	--	-0.021	-0.050	-0.023	--	-0.002	0.000	--
7	L	N=250	EV=2.5	--	0.005	0.003	-0.004	--	0.000	0.000	--
8	L		EV=5	--	0.002	0.005	-0.001	--	-0.001	0.000	--
9	L		EV=10	--	0.004	0.001	-0.006	--	0.001	0.000	--
10	L	N=350	EV=2.5	--	-0.001	0.001	-0.001	--	0.000	0.000	--
11	L		EV=5	--	-0.005	-0.007	-0.006	--	0.001	0.000	--
12	L		EV=10	--	-0.002	-0.005	0.002	--	0.001	0.000	--
13	Q	N=50	EV=2.5	QV=0.01	0.045	0.249	0.197	2.147	0.003	-0.001	-0.014
14	Q			QV=0.05	-0.007	0.133	0.138	0.224	0.002	0.000	-0.028
15	Q			QV=0.10	-0.002	-0.008	0.001	-0.034	-0.001	0.001	0.038
16	Q	N=50	EV=5	QV=0.01	0.065	0.398	0.409	4.329	-0.005	0.002	-0.027
17	Q			QV=0.05	0.067	0.418	0.308	0.572	-0.004	0.002	0.177
18	Q			QV=0.10	0.021	0.252	0.203	0.210	0.005	-0.001	-0.042
19	Q	N=50	EV=10	QV=0.01	0.127	1.063	0.976	11.158	0.003	-0.001	-0.009
20	Q			QV=0.05	0.106	1.030	0.729	1.592	0.002	0.000	0.002
21	Q			QV=0.10	0.103	0.669	0.585	0.586	-0.001	0.001	-0.009
22	Q	N=100	EV=2.5	QV=0.01	0.023	0.132	0.102	1.708	0.000	0.000	0.026
23	Q			QV=0.05	0.020	0.092	0.054	0.052	-0.002	0.001	0.024
24	Q			QV=0.10	-0.006	0.000	0.013	-0.040	-0.003	0.000	0.012
25	Q	N=100	EV=5	QV=0.01	0.030	0.320	0.276	3.317	0.001	0.000	-0.044
26	Q			QV=0.05	0.010	0.245	0.251	0.540	-0.001	0.001	0.061
27	Q			QV=0.10	0.014	0.106	0.080	0.148	-0.002	0.000	0.002
28	Q	N=100	EV=10	QV=0.01	0.086	0.785	0.608	6.575	0.000	0.001	0.057
29	Q			QV=0.05	0.074	0.697	0.521	1.083	-0.004	0.001	0.040
30	Q			QV=0.10	0.068	0.567	0.450	0.528	0.000	0.002	0.125
31	Q	N=250	EV=2.5	QV=0.01	0.020	0.138	0.105	1.277	0.001	0.000	-0.001
32	Q			QV=0.05	-0.006	-0.013	0.002	0.023	-0.001	0.000	0.020
33	Q			QV=0.10	-0.013	-0.038	-0.021	-0.016	0.000	0.000	-0.004
34	Q	N=250	EV=5	QV=0.01	0.025	0.182	0.162	1.968	0.000	0.000	0.024
35	Q			QV=0.05	0.003	0.023	0.041	0.149	-0.001	0.000	0.008
36	Q			QV=0.10	0.013	0.075	0.048	0.045	0.000	0.000	-0.020
37	Q	N=250	EV=10	QV=0.01	0.031	0.304	0.347	4.137	-0.002	0.001	0.023
38	Q			QV=0.05	0.033	0.263	0.225	0.546	0.002	-0.001	-0.044
39	Q			QV=0.10	0.015	0.128	0.121	0.133	0.000	0.000	-0.002
40	Q	N=350	EV=2.5	QV=0.01	0.013	0.102	0.072	0.839	0.000	0.000	-0.004
41	Q			QV=0.05	0.000	0.005	-0.004	0.014	0.000	0.000	0.002
42	Q			QV=0.10	-0.004	-0.017	-0.019	-0.008	0.000	0.000	-0.005
43	Q	N=350	EV=5	QV=0.01	0.012	0.127	0.128	1.672	0.000	0.000	-0.010
44	Q			QV=0.05	0.010	0.038	0.048	0.132	0.000	0.000	-0.005
45	Q			QV=0.10	0.002	-0.015	-0.010	0.007	0.000	0.000	0.016
46	Q	N=350	EV=10	QV=0.01	0.042	0.279	0.279	3.386	0.000	0.000	-0.032
47	Q			QV=0.05	0.018	0.183	0.176	0.413	0.000	0.000	-0.008
48	Q			QV=0.10	0.013	0.074	0.090	0.143	0.000	0.000	-0.014

Table 21. Relative Variance Bias for AIC-Model-Selection Estimates Using the Split-Data Sets

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=50	EV=2.5	--	0.862	0.793	0.888	--	0.855	1.008	--
2	L		EV=5	--	1.026	0.908	1.198	--	1.166	1.022	--
3	L		EV=10	--	1.020	1.038	0.999	--	0.959	1.024	--
4	L	N=100	EV=2.5	--	1.053	1.072	0.985	--	0.960	0.973	--
5	L		EV=5	--	1.206	1.048	0.860	--	0.962	1.095	--
6	L		EV=10	--	0.949	0.995	0.880	--	0.903	0.939	--
7	L	N=250	EV=2.5	--	0.911	0.904	1.044	--	0.842	1.034	--
8	L		EV=5	--	0.951	1.190	1.050	--	0.986	0.988	--
9	L		EV=10	--	1.017	0.957	0.945	--	0.860	1.076	--
10	L	N=350	EV=2.5	--	0.942	1.040	1.104	--	1.007	1.006	--
11	L		EV=5	--	1.075	0.991	1.085	--	0.978	0.909	--
12	L		EV=10	--	1.151	1.046	0.936	--	0.842	0.948	--
13	Q	N=50	EV=2.5	QV=0.01	1.198	0.975	0.961	27.480	0.989	1.027	1.082
14	Q			QV=0.05	1.053	0.806	0.833	6.587	0.753	0.829	0.954
15	Q			QV=0.10	0.799	0.673	0.672	2.558	1.121	0.924	0.975
16	Q	N=50	EV=5	QV=0.01	0.980	0.867	1.038	4.266	0.927	0.918	1.090
17	Q			QV=0.05	0.979	0.802	0.963	2.639	0.994	1.004	0.930
18	Q			QV=0.10	0.941	0.783	0.837	1.451	0.956	1.011	1.143
19	Q	N=50	EV=10	QV=0.01	0.945	1.033	1.401	2.217	1.010	1.200	1.101
20	Q			QV=0.05	1.010	1.025	1.253	1.797	0.834	1.077	1.195
21	Q			QV=0.10	0.988	0.992	1.128	1.203	0.928	1.083	1.125
22	Q	N=100	EV=2.5	QV=0.01	1.102	1.060	1.111	15.671	1.046	1.016	1.157
23	Q			QV=0.05	0.888	0.950	0.971	3.625	1.032	1.026	1.009
24	Q			QV=0.10	0.900	0.888	0.897	1.664	1.077	1.015	1.052
25	Q	N=100	EV=5	QV=0.01	0.936	0.895	1.186	3.884	1.075	1.037	1.049
26	Q			QV=0.05	0.898	0.968	0.962	1.776	0.913	1.053	0.981
27	Q			QV=0.10	0.995	0.863	0.683	1.071	0.973	1.190	1.144
28	Q	N=100	EV=10	QV=0.01	1.031	1.079	1.197	2.008	0.922	0.901	0.923
29	Q			QV=0.05	0.900	0.963	1.049	1.437	0.992	1.085	1.026
30	Q			QV=0.10	1.048	1.004	1.095	1.194	1.000	1.018	0.944
31	Q	N=250	EV=2.5	QV=0.01	1.001	1.119	1.035	10.927	0.964	0.977	1.054
32	Q			QV=0.05	1.162	0.944	0.968	2.135	1.222	0.939	0.917
33	Q			QV=0.10	1.126	0.950	0.897	1.040	1.044	0.956	1.094
34	Q	N=250	EV=5	QV=0.01	0.833	1.092	1.233	3.517	1.153	0.936	0.935
35	Q			QV=0.05	0.947	0.947	1.040	1.519	1.184	0.927	0.988
36	Q			QV=0.10	0.992	1.020	1.087	1.032	1.113	1.038	0.947
37	Q	N=250	EV=10	QV=0.01	1.132	1.152	1.319	1.877	1.118	0.964	0.934
38	Q			QV=0.05	1.115	1.128	1.151	1.210	1.019	1.129	1.183
39	Q			QV=0.10	1.076	1.079	1.043	1.072	1.159	1.224	1.109
40	Q	N=350	EV=2.5	QV=0.01	1.111	1.242	1.244	10.600	0.970	1.185	1.000
41	Q			QV=0.05	1.027	1.008	0.883	1.586	0.899	0.873	0.912
42	Q			QV=0.10	1.053	0.939	0.915	0.875	0.961	1.083	1.000
43	Q	N=350	EV=5	QV=0.01	0.954	0.916	1.015	2.803	1.026	1.028	1.078
44	Q			QV=0.05	1.054	0.987	0.873	1.191	0.816	0.891	1.053
45	Q			QV=0.10	0.821	0.851	0.700	0.700	0.859	1.122	1.015
46	Q	N=350	EV=10	QV=0.01	1.005	1.018	1.158	1.776	1.114	1.027	1.049
47	Q			QV=0.05	1.089	1.011	0.992	1.167	0.897	1.023	1.115
48	Q			QV=0.10	1.089	1.046	0.958	0.933	0.917	1.001	1.009

4.2.2.2 BIC-model-selection estimates

Relative Parameter Bias

Table 22 presents the relative parameter biases for BIC-model-selection estimates using the split-data sets. In general, the parameter estimates for the linear model are better than those for the quadratic model. This is similar to the results of AIC-model-selection estimates. For the parameters common to both the linear and quadratic models (i.e., α_1, α_2, ϕ_{11}, ϕ_{22}, and ϕ_{21}), the relative bias rates range from -0.002 to 0.008 for the linear model, and from -1.512 to 8.005 for the quadratic model.

Similar to the results of AIC-model-selection estimates, mean parameters α_1 and α_2 are recovered satisfactorily, with the relative parameter bias rates ranging only from -0.025 to 0.059. Recovery for mean parameter α_3, however, appears worse than that under AIC model selection. When sample size is large (350), the relative parameter bias for α_3 ranges from -0.06 to 0.003. When sample size is small (100), the relative parameter bias recovery is even worse, with the relative parameter bias ranging from -1.512 to 0.121. Recovery for variance parameter ϕ_{11} appears mediocre to adequate, with the relative parameter bias ranging from -0.041 to 0.436. Recovery for parameters ϕ_{21}, ϕ_{22}, and ϕ_{33} is poor to adequate, with the relative parameter biases ranging from -0.024 to 8.005. When sample size and parameter value for ϕ_{33} are larger, the relative parameter bias for the estimate of ϕ_{33} is smaller. For example, when sample size increases from 50 to 350 and parameter value for ϕ_{33} increases from 0.01 to 0.1, the relative parameter bias for ϕ_{33} decreases markedly, from 4.953 to 0.

Relative Variance Bias

Table 23 presents the relative variance bias for BIC-model-selection estimates using the split-data sets. In general, the relative variance bias for both the linear model and quadratic model is large. For all parameters, the relative variance bias is greater than 0.507. This indicates that recovery for all parameters is poor when using the split-data sets.

Table 22. Relative Parameter Bias for BIC-Model-Selection Estimates Using the Split-Data Sets

Case	**True Model**	**Sample Size**	**Error Variance**	**Quadratic Variance**	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=50	EV=2.5	--	-0.009	-0.037	-0.017	--	0.000	0.000	--
2	L		EV=5	--	-0.003	-0.017	-0.004	--	-0.002	0.000	--
3	L		EV=10	--	-0.026	-0.030	-0.019	--	0.001	0.000	--
4	L	N=100	EV=2.5	--	-0.001	0.001	-0.002	--	-0.001	0.000	--
5	L		EV=5	--	0.008	0.002	0.000	--	0.001	0.000	--
6	L		EV=10	--	-0.021	-0.054	-0.022	--	-0.001	0.000	--
7	L	N=250	EV=2.5	--	0.004	0.003	-0.001	--	0.000	0.000	--
8	L		EV=5	--	0.002	0.002	-0.002	--	-0.001	0.000	--
9	L		EV=10	--	0.002	-0.001	-0.004	--	0.001	0.000	--
10	L	N=350	EV=2.5	--	0.000	0.000	-0.001	--	-0.001	0.000	--
11	L		EV=5	--	-0.006	-0.006	-0.005	--	0.001	0.000	--
12	L		EV=10	--	-0.001	-0.004	0.002	--	0.001	0.000	--
13	Q	N=50	EV=2.5	QV=0.01	0.033	0.001	-0.002	1.256	0.003	0.002	0.124
14	Q			QV=0.05	-0.041	0.037	0.091	0.113	-0.007	0.002	-0.031
15	Q			QV=0.10	-0.018	-0.005	-0.005	-0.033	0.000	0.001	-0.009
16	Q	N=50	EV=5	QV=0.01	0.093	0.039	0.229	2.911	-0.025	0.009	0.269
17	Q			QV=0.05	0.058	0.829	0.657	0.954	0.004	0.003	0.264
18	Q			QV=0.10	0.051	0.285	0.218	0.242	0.011	-0.003	-0.036
19	Q	N=50	EV=10	QV=0.01	0.436	2.172	1.045	4.953	0.059	-0.026	-1.512
20	Q			QV=0.05	0.296	2.525	1.622	2.755	0.016	0.004	0.121
21	Q			QV=0.100	0.128	0.579	0.612	0.605	-0.023	0.000	-0.246
22	Q	N=100	EV=2.5	QV=0.01	0.030	0.017	0.028	1.808	0.007	-0.003	-0.040
23	Q			QV=0.05	0.012	0.079	0.075	0.021	-0.002	0.001	0.002
24	Q			QV=0.10	-0.005	-0.017	0.002	-0.053	-0.002	0.000	0.015
25	Q	N=100	EV=5	QV=0.01	0.008	0.213	0.259	3.236	-0.007	0.001	0.068
26	Q			QV=0.05	0.032	0.405	0.364	0.613	-0.005	0.001	0.010
27	Q			QV=0.10	0.019	0.208	0.126	0.094	-0.002	0.001	0.047
28	Q	N=100	EV=10	QV=0.01	0.132	1.021	1.249	8.005	0.008	-0.013	-0.945
29	Q			QV=0.05	0.181	1.416	1.057	2.267	0.012	0.001	0.086
30	Q			QV=0.10	0.160	0.944	0.373	0.515	0.008	-0.001	-0.033
31	Q	N=250	EV=2.5	QV=0.01	0.028	0.191	0.129	1.357	0.000	0.001	0.021
32	Q			QV=0.05	-0.005	-0.011	0.002	0.044	0.000	0.000	0.027
33	Q			QV=0.10	-0.013	-0.043	-0.024	-0.018	0.000	0.000	-0.003
34	Q	N=250	EV=5	QV=0.01	-0.020	-0.014	0.093	1.278	0.005	0.000	-0.049
35	Q			QV=0.05	-0.009	0.004	0.059	0.153	-0.001	0.000	0.023
36	Q			QV=0.10	0.010	0.063	0.052	0.056	0.000	0.000	-0.017
37	Q	N=250	EV=10	QV=0.01	0.131	0.752	0.444	6.397	-0.002	0.001	0.058
38	Q			QV=0.05	-0.035	-0.030	0.164	0.637	0.003	-0.004	-0.211
39	Q			QV=0.10	0.010	0.086	0.122	0.160	0.000	-0.001	-0.018
40	Q	N=350	EV=2.5	QV=0.01	0.019	0.100	0.066	0.797	-0.001	0.000	-0.006
41	Q			QV=0.05	0.004	0.006	-0.007	0.009	-0.001	0.000	0.001
42	Q			QV=0.10	-0.004	-0.017	-0.019	-0.008	0.000	0.000	-0.005
43	Q	N=350	EV=5	QV=0.01	0.011	0.116	0.124	1.652	-0.002	0.000	-0.023
44	Q			QV=0.05	0.024	0.082	0.062	0.165	0.000	0.000	-0.002
45	Q			QV=0.10	0.004	-0.007	-0.012	0.000	0.000	0.000	0.003
46	Q	N=350	EV=10	QV=0.01	0.074	0.452	0.482	2.706	-0.006	0.002	-0.006
47	Q			QV=0.05	-0.014	-0.093	0.014	0.313	0.000	-0.002	-0.060
48	Q			QV=0.10	0.003	-0.012	0.054	0.122	-0.002	0.000	0.003

Table 23. Relative Variance Bias for BIC-Model-Selection Estimates Using the Split-Data Sets

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=50	EV=2.5	--	0.874	0.797	0.882	--	0.861	1.006	--
2	L		EV=5	--	1.029	0.904	1.187	--	1.167	1.018	--
3	L		EV=10	--	1.007	1.029	0.995	--	0.952	1.022	--
4	L	N=100	EV=2.5	--	1.045	1.070	0.992	--	0.956	0.977	--
5	L		EV=5	--	1.194	1.042	0.861	--	0.956	1.095	--
6	L		EV=10	--	0.951	0.997	0.884	--	0.904	0.941	--
7	L	N=250	EV=2.5	--	0.907	0.905	1.052	--	0.840	1.038	--
8	L		EV=5	--	0.952	1.189	1.047	--	0.986	0.986	--
9	L		EV=10	--	1.012	0.955	0.948	--	0.858	1.078	--
10	L	N=350	EV=2.5	--	0.945	1.041	1.103	--	1.009	1.006	--
11	L		EV=5	--	1.072	0.990	1.086	--	0.977	0.909	--
12	L		EV=10	--	1.152	1.046	0.933	--	0.842	0.948	--
13	Q	N=50	EV=2.5	QV=0.01	1.199	0.779	0.507	28.764	1.019	0.956	1.136
14	Q			QV=0.05	0.957	0.701	0.746	6.538	0.733	0.851	0.962
15	Q			QV=0.10	0.769	0.645	0.626	2.553	1.106	0.935	0.979
16	Q	N=50	EV=5	QV=0.01	1.088	0.884	1.174	3.289	1.011	0.860	0.997
17	Q			QV=0.05	0.826	0.732	0.927	2.814	0.946	1.125	0.963
18	Q			QV=0.10	1.028	0.825	0.802	1.496	0.985	0.998	1.143
19	Q	N=50	EV=10	QV=0.01	1.078	0.908	0.585	1.006	1.126	0.982	0.779
20	Q			QV=0.05	1.044	1.243	1.979	2.144	0.812	1.033	0.999
21	Q			QV=0.10	0.916	0.851	0.849	0.966	1.000	1.126	1.136
22	Q	N=100	EV=2.5	QV=0.01	1.093	0.945	0.935	16.684	1.059	0.937	1.124
23	Q			QV=0.05	0.880	0.961	0.994	3.630	1.026	1.060	1.009
24	Q			QV=0.10	0.900	0.888	0.905	1.672	1.082	1.018	1.052
25	Q	N=100	EV=5	QV=0.01	0.911	0.854	1.135	3.981	1.045	1.000	0.951
26	Q			QV=0.05	0.852	0.924	0.932	1.689	0.885	1.039	0.981
27	Q			QV=0.10	0.989	0.867	0.684	1.040	0.963	1.199	1.112
28	Q	N=100	EV=10	QV=0.01	1.154	1.467	2.098	3.592	0.759	0.895	0.959
29	Q			QV=0.05	1.001	1.230	1.537	2.111	0.922	1.103	1.160
30	Q			QV=0.10	1.159	1.091	1.177	1.430	1.059	0.964	0.946
31	Q	N=250	EV=2.5	QV=0.01	1.016	1.131	1.027	11.415	0.973	0.991	1.054
32	Q			QV=0.05	1.169	0.959	0.990	2.067	1.222	0.939	0.917
33	Q			QV=0.10	1.125	0.949	0.897	1.040	1.044	0.955	1.094
34	Q	N=250	EV=5	QV=0.01	0.729	0.970	1.100	3.494	1.113	0.938	0.935
35	Q			QV=0.05	0.927	0.931	1.021	1.470	1.163	0.924	0.976
36	Q			QV=0.10	0.988	1.025	1.098	1.039	1.105	1.036	0.947
37	Q	N=250	EV=10	QV=0.01	1.316	1.366	1.627	2.422	1.161	0.931	0.947
38	Q			QV=0.05	0.991	0.950	0.854	0.848	0.989	1.139	1.190
39	Q			QV=0.10	0.999	0.987	0.924	0.952	1.134	1.199	1.096
40	Q	N=350	EV=2.5	QV=0.01	1.129	1.246	1.232	10.867	0.982	1.188	1.000
41	Q			QV=0.05	1.037	1.012	0.883	1.600	0.906	0.873	0.912
42	Q			QV=0.10	1.053	0.939	0.915	0.875	0.961	1.083	1.000
43	Q	N=350	EV=5	QV=0.01	0.933	0.902	1.009	2.829	1.012	1.015	1.078
44	Q			QV=0.05	1.084	1.010	0.891	1.204	0.826	0.887	1.053
45	Q			QV=0.10	0.824	0.853	0.700	0.700	0.863	1.125	1.015
46	Q	N=350	EV=10	QV=0.01	0.991	0.937	0.843	0.984	1.175	1.119	1.108
47	Q			QV=0.05	1.006	0.902	0.905	1.238	0.870	0.976	1.087
48	Q			QV=0.10	1.073	1.027	0.945	0.947	0.912	0.995	1.009

To further assess whether data splitting mitigate the problems arising from model selection, I examined the differences between the true-model-selection, the AIC-model-selection using original data sets, the AIC-model-selection using split-data sets, the BIC-model-selection using original data sets, and the BIC-model-selection using split-data sets. Table 24 summarizes the results.

One can see from this table that, in terms of relative parameter bias, data splitting has an impact on parameter estimates for parameters ϕ_{21} and ϕ_{22}. The number of cases in which the absolute value of the relative parameter bias difference is greater than 0.1 is larger in SAIC-True than in AIC-True for parameters ϕ_{21}, ϕ_{22}, and ϕ_{33}, but is smaller for parameter α_3. Therefore, in terms of relative parameter bias, no general conclusion could be made regarding whether data splitting mitigates the problem arising from model selection by AIC or by BIC.

Note that, in terms of relative variance bias, data splitting has a substantial impact on parameter estimates for all parameters. In fact, the numbers of cases in which the absolute value of the difference relative parameter bias is greater than 0.1 is markedly larger in SAIC-True than in AIC-True for all parameters. This suggests that data-splitting may greatly increase the variability in estimates without the reward of eliminating parameter bias. The comparison of SBIC-True and BIC-True leads to the same conclusion.

Table 24. Differences of Relative Parameter Bias between True-Model-Selection, AIC-Model-Selection, and BIC-Model-Selection Estimates

		ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
Difference of Relative Parameter Bias								
AIC-True	Mean	0.0049	0.0266	0.0065	0.4613	-0.0024	0.0018	0.1573
	Abs>0.1/ total cases	1/48	3/48	3/48	18/36	0/48	0/48	15/36
BIC-True	Mean	0.0129	0.0657	0.0086	1.2306	-0.0073	0.0056	0.4924
	Abs>0.1/ total cases	2/48	11/48	14/48	27/36	0/48	0/48	26/36
SAIC-True	Mean	0.0089	0.0857	0.0695	0.3402	-0.0002	0.0001	0.0036
	Abs>0.1/ total cases	0/48	14/48	15/48	23/36	0/48	0/48	1/36
SBIC-True	Mean	0.0243	0.1558	0.1135	0.2677	0.0009	-0.0005	-0.0472
	Abs>0.1/ total cases	4/48	18/48	12/48	22/36	0/48	0/48	7/36
Difference of Relative Variance Bias								
AIC-True	Mean	-0.0010	0.0979	0.0265	0.1665	0.0065	-0.0012	0.1116
	Abs>0.1/ total cases	3/48	14/48	9/48	27/36	2/48	3/48	13/36
BIC-True	Mean	0.0070	0.1370	0.0285	0.9894	0.0016	0.0026	0.4840
	Abs>0.1/ total cases	6/48	14/48	16/48	30/36	2/48	3/48	26/36
SAIC-True	Mean	0.9887	0.9526	0.9458	1.8036	0.9942	1.0128	0.7668
	Abs>0.1/ total cases	48/48	48/48	48/48	36/36	48/48	48/48	36/36
SBIC-True	Mean	0.9893	0.9483	0.9396	1.8641	0.9934	1.0057	0.7570
	Abs>0.1/ total cases	48/48	48/48	48/48	36/36	48/48	48/48	36/36

AIC-True = difference between the AIC-model-selection coverage rate using original data sets and the true-model-selection coverage rate
BIC-True = difference between the BIC-model-selection coverage rate using original data sets and the true-model-selection coverage rate
SAIC-True = difference between the AIC-model-selection coverage rate using split- data sets and the true-model-selection coverage rate
SBIC-True = difference between the BIC-model-selection coverage rate using split- data sets and the true-model-selection coverage rate
Abs > 0.1 = the number of cases in which the absolute value of the difference is greater than 0.1.

4.2.3 Coverage Rate

Table 25 shows the descriptive statistics for AIC-model-selection coverage rate using the 50% split-data sets. One can see from Table 25 that conditional coverage for parameter α_1 remains adequate, with only one case in which coverage rates fall outside 0.925 and 0.975. Conditional coverage for parameters ϕ_{11}, α_2, and α_3, however, is mediocre. Conditional coverage for variance parameters ϕ_{21}, ϕ_{22} and ϕ_{33}, is poor, with 15 ,25, and 8 cases respectively in which the conditional rates fall outside 0.925 and 0.975.

Comparing Table 16 in section 4.1.3 and Table 25, one can see that there are marked differences in the numbers of cases in which coverage rate falls outside 0.925 and 0.975 for the AIC-model-selection coverage rates using original data sets and using 50% split-data sets. In general, the number is bigger when using original data sets than when using split-data sets for all parameters except ϕ_{22}. This suggests that data splitting might mitigate the impact of model selection on coverage rate when AIC is used for model selection.

Table 26 shows the descriptive statistics for the BIC-model-selection coverage rate using the 50% split-data sets. One can see from Table 26 that the number of cases in which coverage rates fall outside 0.925 and 0.975 ranges from 12 to 28. This indicates that the BIC-model-selection coverage is not adequate.

Comparing Table 17 in section 4.1.3 and Table 26, one can see that there are differences in the numbers of cases in which coverage rate falls outside 0.925 and 0.975 for the BIC-model-selection coverage rates using original data sets and using 50% split-data sets. In general, the number is smaller when using original data sets than when using split-data sets for all parameters except α_2 and α_3. This suggests that data splitting might worsen the impact of model selection on coverage rate when BIC is used for model selection.

To further assess whether data splitting mitigates the problem arising from model selection, I examined the differences between the true-model-selection, the AIC-model-selection using original data sets, the AIC-model-selection using split-data sets, the BIC-model-selection using original data sets, and the BIC-model-selection using split-data sets. Table 27 summarizes the results.

Table 25. Descriptive Statistics for AIC-Model-Selection Coverage Rate Using 50% Split-Data Sets

	Minimum	Maximum	Mean	Cases in which coverage rate falls outside 0.925 and 0.975	Total Cases
ϕ_{11}	.900	.964	.94085	5	48
ϕ_{21}	.863	.961	.93065	15	48
ϕ_{22}	.693	.953	.89988	25	48
ϕ_{33}	.790	.986	.95081	8	36
α_1	.924	.972	.94848	1	48
α_2	.919	.971	.94608	3	46
α_3	.919	.971	.94764	2	36

Table 26. Descriptive Statistics for BIC-Model-Selection Coverage Rate Using 50% Split-Data Sets

	Minimum	Maximum	Mean	Cases in which coverage rate falls outside 0.925 and 0.975	Total Cases
ϕ_{11}	.727	1.000	.93767	16	48
ϕ_{21}	.667	1.000	.92235	24	48
ϕ_{22}	.636	1.000	.90138	28	48
ϕ_{33}	.746	1.000	.95317	18	36
α_1	.750	1.000	.94877	12	48
α_2	.750	1.000	.93827	17	46
α_3	.750	1.000	.94047	13	36

Comparing AIC-True and SAIC-True, one can see that the number of cases in which the absolute value of the coverage rate difference is greater than 0.025 is smaller in SAIC-True than in AIC-True for parameters ϕ_{11}, ϕ_{21}, α_2, and α_3, but is for parameters ϕ_{22}, and α_1. Therefore, no general conclusion could be made regarding whether data splitting mitigates the problem arising from model selection by AIC.

Also, comparing BIC-True and SBIC-True, one can see that the number of cases in which the absolute value of the coverage rate difference is greater than 0.025 is smaller in SBIC-True than in BIC-True for parameters ϕ_{33}, α_2, and α_3, but is larger for parameters ϕ_{11}, ϕ_{21}, and α_1. Therefore, no general conclusion could be made regarding whether data splitting mitigates the problem arising from model selection by BIC.

Table 27. Differences of Model Selection Coverage Rates

		ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
AIC-True	Abs>0.025/ total cases	6/48	8/48	11/48	8/36	0/48	6/48	7/36
	Mean	-0.007	-0.011	-0.018	-0.009	-0.000	-0.004	-0.006
SAIC-True	Abs>0.025/ total cases	3/48	7/48	15/48	8/36	1/48	2/48	2/36
	Mean	-0.005	-0.011	-0.021	0.013	0.000	-0.002	-0.003
SAIC-AIC	Abs>0.025/ total cases	4/48	5/48	14/48	14/36	2/48	6/48	11/36
	Mean	0.001	0.000	-0.003	0.022	0.000	0.001	0.002
BIC-True	Abs>0.025/ total cases	11/48	18/48	22/48	17/36	7/48	19/48	19/36
	Mean	-0.015	-0.028	-0.037	-0.032	-0.012	-0.104	-0.179
SBIC-True	Abs>0.025/ total cases	17/48	19/48	22/48	14/36	11/48	17/48	14/36
	Mean	-0.009	-0.019	-0.020	0.016	0.000	-0.010	-0.010
SBIC-BIC	Abs>0.025/ total cases	14/48	20/48	19/48	23/36	12/48	20/48	19/36
	Mean	0.006	0.009	0.017	0.048	0.013	0.094	0.168

AIC-True = difference between the AIC-model-selection coverage rate using original data sets and the true-model-selection coverage rate
BIC-True = difference between the BIC-model-selection coverage rate using original data sets and the true-model-selection coverage rate
SAIC-True = difference between the AIC-model-selection coverage rate using split- data sets and the true-model-selection coverage rate
SBIC-True = difference between the BIC-model-selection coverage rate using split- data sets and the true-model-selection coverage rate
SAIC-AIC= difference between the AIC-model-selection coverage rate using split- data sets and the AIC-model-selection coverage rate using original data sets
ABIC-BIC= difference between the BIC-model-selection coverage rate using split- data sets and the BIC-model-selection coverage rate using original data sets
Abs>0.025 = the number of cases in which the absolute value of the difference coverage rate is greater than 0.025

Chapter 5 Discussion and Conclusion

In behavioral and social sciences, model selection and parameter estimation are treated as two separate steps of data analysis. The second step, parameter estimation, is generally conducted on the assumption that the model selected in step one is a correct model, and thus it is performed using the same data set that was used in step one. This two-step process ignores the effects of model uncertainty on parameter estimation and statistical inference, and thus may ultimately lead to misleading or invalid inferences. The problems arising from the use of this two-step process have been well investigated in the context of regression. In the case of latent growth modeling (LGM), however, there have been no such published studies. This present study was thus designed to investigate the possible problems arising from the use of this two-step process in LGM. The goals of this study were:

(1) To examine the subsequent impact of preliminary model selection using information criteria on LGM parameter estimates;

(2) To assess the data splitting method as a possible way to mitigate the effects of model uncertainty.

To achieve these goals, I conducted two Monte Carlo simulation studies. Study 1 conducted both model selection and parameter estimation using the same data set, to investigate the possible impact of preliminary model selection in terms of model selection accuracy, relative parameter biases, and coverage rate. Study 2 conducted model selection and parameter estimation using different split-data sets, in order to assess the data splitting method as a possible way to mitigate the effects of model uncertainty.

5.1 Summary of Major Findings and Discussion

Model Selection Accuracy

The performances of AIC and BIC were evaluated from the aspects of model selection accuracy and model selection consistency by examining model recovery rate under different conditions (e.g., different sample sizes, underfitting vs overfitting). The following observations were made:

First, when the linear model is the true model, both AIC and BIC accurately and consistently identify the true model, although BIC appears to be somewhat more consistent than AIC. Second, when the quadratic model is the true model, AIC appears to be more accurate than BIC, which tends, in these cases, to select the simpler, linear model (i.e., underfitting). BIC, however,,appears to be more consistent than AIC in selecting a model (true or misspecified). This agrees with previous research (Hurvich & Tsai, 1990; Kang & Cohen, 2007; Zhang, 2008). Third, both AIC and BIC model selection are more accurate with larger sample sizes, and smaller quadratic variances.

The above findings provide evidence that model uncertainty (i.e., that the selected model might be wrong) does exist. Use of the two-step process in LGM ignores the effects of this uncertainty on parameter estimation and statistical inference and may therefore ultimately result in misleading and invalid inferences.

Relative Parameter Biases

The impact of model selection on post-model-selection point estimators was evaluated from the aspects of accuracy and variability by examining relative parameter bias and relative variance bias, respectively. The following observations were made:

First, the recovery for parameters is better when the linear model is the true model than when the quadratic model is the true model. Recovery for parameters α_1 and α_2 is satisfactory for all conditions. Recovery for parameters α_3, ϕ_{11}, ϕ_{21}, ϕ_{22}, and ϕ_{33} is adequate when sample size is large, but is mediocre to poor when sample size is small. Second, comparison of the relative biases of the true-model-selection estimates, the AIC model selection estimates, and the BIC model selection estimates shows that model selection has an impact on parameter point estimates. In fact, both the AIC and BIC model selections substantially compromise the accuracy of post-model-selection estimators and lead to additional bias in the estimates for parameters α_3 and ϕ_{33}. Also, both the AIC and BIC model selections overestimate the sampling variability of the estimates for parameters α_3, ϕ_{21}, and ϕ_{33}.

The above findings suggest that the use of the two-step process in LGM ignores the effects of model uncertainty on parameter estimation. Therefore, inference based on AIC or BIC model selection leads to additional bias in, and overestimates the sampling variability of, the parameter estimates. Even when the magnitude of the mean of the quadratic slope factor was intentionally set small in this study to make the data generated by the quadratic growth model not dramatically different from those generated by the linear growth model, simulation results still showed that the post-model-selection parameter estimator had larger relative parameter biases and larger relative variance biases. This provides strong evidence that model selection has a non-ignorable impact on LGM parameter estimates. Bias becomes substantially smaller when sample size becomes larger. Conditions with larger sample sizes are deemed to have greater power to choose the right model; thus the problem of model selection creating additional bias in the parameter estimates is mitigated when sample size is large.

Coverage Rate

The impact of model selection on post-model-selection interval estimators was evaluated by examining unconditional and conditional coverage rates. The following observations were made:

First, both the unconditional and the conditional coverage rates were closer to the nominal level of 0.950 when the linear model was the true model than when the quadratic model was the true model. Second, compared to the true-model-selection coverage rate, the AIC and BIC model selection coverage rates had more cases in which the coverage rates deviated substantially from the nominal rate of 0.95. Also, conditional coverage rates of the AIC and BIC model selections were, on average, smaller than unconditional coverage rates. This difference correlated positively with sample size, but negatively with error variance. These results indicate that the conditional rates tend to be underestimated and sample size and error variance play important roles in determining the conditional coverage rates. Third, extreme BIC conditional coverage rates (0% for parameters α_2 and α_3, and 100% for parameters ϕ_{11}, ϕ_{22}, and ϕ_{21}) occur in some cases, e.g., when sample size = 100 or 200, error variance = 10, quadratic variance = 0.01, and when the model selection success rate is very small. Because BIC does not perform effectively in such cases, these conditional coverage rates might be invalid.

Assessment of Data Splitting

Hurvich and Tsai (1992) stated that data splitting provides a possible substitute for a true replicate sample in model validation and thus suggested a possible remedy based on data splitting to solve the problem resulting from the use of the same data set for both structural identification and inference. Therefore, in the current study the data splitting technique was assessed as a possible way to mitigate the effects of model uncertainty.

Unfortunately, however, according to the simulation results, the post-model-selection problems due to model uncertainty still existed when the data splitting method was applied. It seems that fitting a model to just part of the data results in a loss of efficiency. Data splitting may greatly increase the variability in estimates without the reward of eliminating bias.

In summary, the above findings provide evidence that model uncertainty exists and that ignoring the effects of model uncertainty compromises the quality of both the post-model-selection point estimators and interval estimators. This result is consistent with previous research in the context of regression (Breiman, 1988; Hurvich & Tsai, 1990; Pötscher, 1991; Rencher & Pun, 1980). In the simulation portion of this current study, even when the data generated by the quadratic growth model are very similar to the data generated by the linear growth model, the post-model-selection parameter estimator still has larger relative parameter biases, larger relative variance biases, and smaller coverage rates for a 95% confidence interval than those of the true-model-selection estimator. This provides strong evidence that model selection has a non-ignorable impact on LGM parameter point and interval estimates. This impact is not mitigated even after the data splitting method was applied.

Reasons for the problems arising from model selection might be: (1) because the estimation procedure depends upon the outcome of model selection, the properties of the post-model-selection estimators and related statistics (e.g., the estimates of mean squared error) might be different from those when the model is known *a priori*. Consequently, bias may exist; (2) because the use of a model selection procedure affects the asymptotic distribution of parameter estimators and related statistics, the validity of the subsequent inference procedures may be severely affected (Miller, 1984; Zhang, 1992).

5.2 Limitations and Suggestions for Future Research

There are limitations inherent in any research, and this study is no exception. As can be seen from the results of this study, inconsistencies and inaccuracies were shown in AIC and BIC model selection in at least some of the simulated conditions. AIC appeared to function better under certain conditions in the simulation study. Deciding whether a linear or a quadratic growth model is more appropriate for a particular data set, however, is difficult at best, because the true model is not known for real-world data. Therefore, it is important to note that this current study was intended to examine model selection from a statistical perspective. In practice, when selecting a model, it is also important to consider nonstatistical perspectives.

Ideally, the results discussed in this study should have been based on 1,000 replications with proper parameter estimates. Unfortunately, this was not the case. Negative estimates of variances (i.e., Heywood Cases) occurred in a great number of replications for all 48 cases, when fitting the data to the quadratic model. This problem was encountered more frequently (on average, in 50% of the replications) with the smaller sample sizes and the smaller quadratic growth factor variances. This problem, however, was substantially less severe when fitting the data to the linear model. In fact, it occurred only in the cases with sample size of 100. Table 28 shows the number of cases without negative estimates of variance when fitting the data to the quadratic model. Table 29 shows the number of cases without negative estimates of variance when fitting the data to the linear model. Those numbers are the same even when conducting the parameter estimates with different computer software, EQS or Mplus.

There are three possible reasons for these Heywood cases. First, nonconvergence (Kolenikov & Bollen, 2008); in our study, however, all the replications converged in less than 20 iterations. Second, structurally misspecified models (Dillon, Kumar & Mulani, 1987; Bollen, 1989); this would not be the cause of the problem in the current study, however, because the true model is set in advance. Third, sampling fluctuations (Anderson & Gerbing, 1984); this seems to be the most likely cause of the Heywood cases in our study. I gave the quadratic growth factor variance a population value of 0.01, which is small. With small sample size, random draws may end up with a negative variance simply due to chance.

Table 28. The Number of Cases Without Negative Estimates of Variance When Fitting the Data to the Quadratic Model.

Case	True Model	Sample Size	Error Variance	Quadratic Variance	Total	S1	S2	S3	S4
1	L	N=100	EV=2.5	--	400	183	311	250	279
2	L		EV=5	--	415	206	389	307	307
3	L		EV=10	--	459	237	425	356	348
4	L	N=200	EV=2.5	--	490	266	471	395	391
5	L		EV=5	--	493	315	469	419	419
6	L		EV=10	--	480	358	455	442	455
7	L	N=500	EV=2.5	--	497	388	488	482	485
8	L		EV=5	--	507	488	489	495	503
9	L		EV=10	--	492	447	481	471	476
10	L	N=700	EV=2.5	--	484	442	498	471	474
11	L		EV=5	--	485	485	482	489	488
12	L		EV=10	--	485	476	493	479	493
13	Q	N=100	EV=2.5	QV=0.01	428	160	368	335	273
14	Q			QV=0.05	603	222	495	380	393
15	Q			QV=0.1	672	268	606	483	434
16	Q	N=100	EV=5	QV=0.01	459	211	394	341	331
17	Q			QV=0.05	562	221	477	378	377
18	Q			QV=0.1	599	245	520	387	425
19	Q	N=100	EV=10	QV=0.01	468	235	416	343	349
20	Q			QV=0.05	504	229	472	370	372
21	Q			QV=0.1	537	213	470	381	368
22	Q	N=200	EV=2.5	QV=0.01	561	321	499	470	445
23	Q			QV=0.05	762	387	710	586	591
24	Q			QV=0.1	900	436	832	713	704
25	Q	N=200	EV=5	QV=0.01	520	336	501	441	422
26	Q			QV=0.05	663	357	623	531	557
27	Q			QV=0.1	801	417	730	646	609
28	Q	N=200	EV=10	QV=0.01	502	339	484	456	460
29	Q			QV=0.05	599	365	577	523	501
30	Q			QV=0.1	676	391	646	564	562
31	Q	N=500	EV=2.5	QV=0.01	636	488	611	578	572
32	Q			QV=0.05	934	669	911	838	823
33	Q			QV=0.1	989	781	966	911	904
34	Q	N=500	EV=5	QV=0.01	577	485	573	543	534
35	Q			QV=0.05	836	618	780	749	737
36	Q			QV=0.1	958	685	924	862	876
37	Q	N=500	EV=10	QV=0.01	548	498	532	515	532
38	Q			QV=0.05	697	546	660	621	611
39	Q			QV=0.1	819	563	775	700	717
40	Q	N=700	EV=2.5	QV=0.01	653	509	631	616	597
41	Q			QV=0.05	968	731	939	886	888
42	Q			QV=0.1	997	851	988	965	960
43	Q	N=700	EV=5	QV=0.01	598	523	569	562	566
44	Q			QV=0.05	881	672	836	792	769
45	Q			QV=0.1	968	770	947	900	888
46	Q	N=700	EV=10	QV=0.01	555	499	539	547	511
47	Q			QV=0.05	741	564	694	662	659
48	Q			QV=0.1	873	671	833	764	764

Table 29. The Number of Cases Without Negative Estimates of Variance When Fitting the Data to the Linear Model.

Case	True Model	Sample Size	Error Variance	Quadratic Variance	Total	S1	S2	S3	S4
1	L	N=100	EV=2.5	--	999	883	993	968	977
2	L		EV=5	--	1000	926	1000	992	992
3	L		EV=10	--	1000	953	1000	994	997
4	L	N=200	EV=2.5	--	1000	972	1000	1000	1000
5	L		EV=5	--	1000	992	1000	999	1000
6	L		EV=10	--	1000	998	1000	1000	1000
7	L	N=500	EV=2.5	--	1000	1000	1000	1000	1000
8	L		EV=5	--	1000	1000	1000	1000	1000
9	L		EV=10	--	1000	1000	1000	1000	1000
10	L	N=700	EV=2.5	--	1000	1000	1000	1000	1000
11	L		EV=5	--	1000	1000	1000	1000	1000
12	L		EV=10	--	1000	1000	1000	1000	1000
13	Q	N=100	EV=2.5	QV=0.01	1000	882	999	988	987
14	Q			QV=0.05	1000	921	1000	992	988
15	Q			QV=0.1	999	874	993	970	970
16	Q	N=100	EV=5	QV=0.01	999	925	998	996	988
17	Q			QV=0.05	1000	938	999	996	992
18	Q			QV=0.1	1000	943	1000	994	996
19	Q	N=100	EV=10	QV=0.01	1000	937	1000	996	996
20	Q			QV=0.05	1000	949	1000	996	997
21	Q			QV=0.1	1000	966	1000	998	998
22	Q	N=200	EV=2.5	QV=0.01	1000	987	1000	1000	1000
23	Q			QV=0.05	1000	991	1000	999	1000
24	Q			QV=0.1	1000	967	1000	999	997
25	Q	N=200	EV=5	QV=0.01	1000	995	1000	1000	1000
26	Q			QV=0.05	1000	997	1000	1000	1000
27	Q			QV=0.1	1000	992	1000	1000	1000
28	Q	N=200	EV=10	QV=0.01	1000	995	1000	1000	1000
29	Q			QV=0.05	1000	997	1000	1000	1000
30	Q			QV=0.1	1000	999	1000	1000	1000
31	Q	N=500	EV=2.5	QV=0.01	1000	1000	1000	1000	1000
32	Q			QV=0.05	1000	1000	1000	1000	1000
33	Q			QV=0.1	1000	1000	1000	1000	1000
34	Q	N=500	EV=5	QV=0.01	1000	1000	1000	1000	1000
35	Q			QV=0.05	1000	1000	1000	1000	1000
36	Q			QV=0.1	1000	1000	1000	1000	1000
37	Q	N=500	EV=10	QV=0.01	1000	1000	1000	1000	1000
38	Q			QV=0.05	1000	1000	1000	1000	1000
39	Q			QV=0.1	1000	1000	1000	1000	1000
40	Q	N=700	EV=2.5	QV=0.01	1000	1000	1000	1000	1000
41	Q			QV=0.05	1000	1000	1000	1000	1000
42	Q			QV=0.1	1000	1000	1000	1000	1000
43	Q	N=700	EV=5	QV=0.01	1000	1000	1000	1000	1000
44	Q			QV=0.05	1000	1000	1000	1000	1000
45	Q			QV=0.1	1000	1000	1000	1000	1000
46	Q	N=700	EV=10	QV=0.01	1000	1000	1000	1000	1000
47	Q			QV=0.05	1000	1000	1000	1000	1000
48	Q			QV=0.1	1000	1000	1000	1000	1000

According to Kolenikov and Bollen (2008), one possible way to avoid Heywood cases is to restrict the range of estimates to be $[0, +\infty]$. Therefore, I constrained the negative estimates of variance whenever they occurred to the lower boundary in order to keep the estimates in the interior of parameter space. The constrained estimates for the total 1,000 replications were used for two reasons. First, to retain the randomness of sampling. Second, to avoid larger biases in the parameter estimates. The results show that using only the replications without any Heywood cases end up with larger biases in the parameter estimates. However, if the estimates are at the boundary (e.g., the quadratic growth factor variance is equal to zero), the estimates and statistical tests might behave in unusual ways (Andrews, 2001). As shown in Table 7 in section 4.1.1, the empirical model recovery rate is close to the theoretical power level. Therefore, model recovery seems not to be influenced by the Heywood cases in this study.

It would be interesting to further examine the possible impact of Heywood cases in model selection. Thus, future research studies may want to compare the constrained and unconstrained estimates to see whether the power in detecting structural misspecification is affected. In addition, our results showed that the problems arising from the two-step process are not substantially mitigated by the data splitting method. Future research might assess whether bootstrapping or Bayesian model averaging is a better alternative to mitigate the problems of model selection in LGM. In our study, only the linear and quadratic latent growth models were investigated. It would be worthwhile to evaluate the possible impact on more complicated models in SEM (e.g., mixture latent growth model).

Appendices

Table A1. Difference between the AIC-Model-Selection and the True-Model-Selection Relative Parameter Bias (AIC-True)

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.000	0.002	0.001	--	-0.001	0.000	--
2	L		EV=5	--	-0.002	-0.007	0.000	--	0.001	0.000	--
3	L		EV=10	--	0.001	0.008	0.002	--	0.001	0.000	--
4	L	N=200	EV=2.5	--	-0.003	-0.006	-0.001	--	0.000	0.000	--
5	L		EV=5	--	0.000	0.000	0.000	--	0.000	0.000	--
6	L		EV=10	--	-0.001	-0.006	-0.002	--	0.000	0.000	--
7	L	N=500	EV=2.5	--	-0.001	0.002	0.001	--	0.000	0.000	--
8	L		EV=5	--	0.001	0.003	0.001	--	0.000	0.000	--
9	L		EV=10	--	0.001	0.003	0.001	--	0.000	0.000	--
10	L	N=700	EV=2.5	--	0.000	0.000	0.001	--	0.000	0.000	--
11	L		EV=5	--	0.000	0.001	-0.001	--	0.000	0.000	--
12	L		EV=10	--	-0.001	-0.003	-0.001	--	0.000	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	0.010	0.015	0.004	0.628	-0.006	0.004	0.300
14	Q			QV=0.05	0.003	0.007	-0.027	0.121	-0.003	0.001	0.128
15	Q			QV=0.10	0.002	-0.003	-0.005	0.025	-0.001	0.000	0.025
16	Q	N=100	EV=5	QV=0.01	-0.002	0.030	0.044	1.818	-0.013	0.010	0.607
17	Q			QV=0.05	-0.007	-0.084	-0.061	0.458	-0.007	0.005	0.390
18	Q			QV=0.10	0.003	0.004	-0.042	0.177	-0.003	0.002	0.157
19	Q	N=100	EV=10	QV=0.01	0.095	0.682	0.496	5.971	-0.022	0.013	0.885
20	Q			QV=0.05	0.045	0.233	-0.007	1.218	-0.015	0.011	0.673
21	Q			QV=0.10	0.003	-0.057	-0.146	0.486	-0.007	0.005	0.331
22	Q	N=200	EV=2.5	QV=0.01	-0.002	-0.012	-0.008	0.119	-0.002	0.002	0.091
23	Q			QV=0.05	0.000	0.002	-0.003	0.015	0.000	0.000	0.014
24	Q			QV=0.10	0.000	0.000	0.000	0.001	0.000	0.000	0.001
25	Q	N=200	EV=5	QV=0.01	0.013	0.078	0.042	0.861	-0.005	0.004	0.264
26	Q			QV=0.05	0.002	-0.010	-0.030	0.168	-0.003	0.002	0.149
27	Q			QV=0.10	0.002	-0.001	-0.007	0.023	0.000	0.000	0.017
28	Q	N=200	EV=10	QV=0.01	0.040	0.271	0.171	2.499	-0.012	0.009	0.548
29	Q			QV=0.05	0.015	0.062	-0.053	0.517	-0.007	0.006	0.379
30	Q			QV=0.10	0.006	-0.007	-0.065	0.225	-0.003	0.003	0.182
31	Q	N=500	EV=2.5	QV=0.01	0.000	0.000	0.000	0.002	0.000	0.000	0.001
32	Q			QV=0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000
33	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
34	Q	N=500	EV=5	QV=0.01	0.002	0.004	-0.003	0.098	-0.001	0.001	0.053
35	Q			QV=0.05	-0.001	-0.002	-0.002	0.007	0.000	0.000	0.004
36	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
37	Q	N=500	EV=10	QV=0.01	0.006	0.054	0.037	0.735	-0.003	0.004	0.198
38	Q			QV=0.05	0.001	0.005	-0.010	0.122	-0.001	0.001	0.076
39	Q			QV=0.10	0.000	-0.006	-0.012	0.009	0.000	0.000	0.013
40	Q	N=700	EV=2.5	QV=0.01	0.000	0.000	0.000	0.000	0.000	0.000	0.000
41	Q			QV=0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000
42	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
43	Q	N=700	EV=5	QV=0.01	0.000	0.000	0.001	0.028	0.000	0.000	0.010
44	Q			QV=0.05	0.000	-0.001	0.000	0.001	0.000	0.000	0.000
45	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
46	Q	N=700	EV=10	QV=0.01	0.007	0.027	0.013	0.295	-0.002	0.002	0.124
47	Q			QV=0.05	-0.003	-0.010	-0.013	0.042	-0.001	0.000	0.039
48	Q			QV=0.10	0.000	-0.002	-0.002	0.002	0.000	0.000	0.002

Table A2. Difference between the BIC-Model-Selection and the True-Model-Selection Relative Parameter Bias (BIC-True)

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.000	0.000	0.000	--	0.000	0.000	--
2	L		EV=5	--	0.000	0.000	0.001	--	0.000	0.000	--
3	L		EV=10	--	0.000	0.000	0.000	--	0.000	0.000	--
4	L	N=200	EV=2.5	--	0.000	0.000	0.000	--	0.000	0.000	--
5	L		EV=5	--	0.000	0.000	0.000	--	0.000	0.000	--
6	L		EV=10	--	0.000	0.000	0.000	--	0.000	0.000	--
7	L	N=500	EV=2.5	--	0.000	0.000	0.000	--	0.000	0.000	--
8	L		EV=5	--	0.000	0.000	0.000	--	0.000	0.000	--
9	L		EV=10	--	0.000	0.000	0.000	--	0.000	0.000	--
10	L	N=700	EV=2.5	--	0.000	0.000	0.000	--	0.000	0.000	--
11	L		EV=5	--	0.000	0.000	-0.001	--	0.000	0.000	--
12	L		EV=10	--	0.000	0.000	0.000	--	0.000	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	0.014	-0.023	0.025	0.702	-0.010	0.011	0.743
14	Q			QV=0.05	0.008	-0.097	-0.096	0.501	-0.005	0.006	0.478
15	Q			QV=0.10	0.002	-0.035	-0.049	0.174	-0.001	0.001	0.160
16	Q	N=100	EV=5	QV=0.01	-0.048	-0.831	-0.424	2.335	-0.030	0.022	1.234
17	Q			QV=0.05	0.038	0.123	-0.113	0.635	-0.006	0.012	1.031
18	Q			QV=0.10	0.004	-0.035	-0.133	0.672	-0.008	0.007	0.484
19	Q	N=100	EV=10	QV=0.01	0.029	1.171	1.160	16.743	-0.074	0.034	2.354
20	Q			QV=0.05	0.045	0.810	0.140	3.405	-0.030	0.021	1.275
21	Q			QV=0.10	-0.037	-0.576	-0.575	1.289	-0.022	0.012	0.717
22	Q	N=200	EV=2.5	QV=0.01	0.004	-0.024	-0.013	0.706	-0.009	0.006	0.386
23	Q			QV=0.05	0.006	0.017	-0.022	0.242	-0.003	0.002	0.193
24	Q			QV=0.10	0.000	-0.006	-0.006	0.032	0.000	0.000	0.023
25	Q	N=200	EV=5	QV=0.01	0.005	0.038	0.016	0.427	-0.020	0.016	0.897
26	Q			QV=0.05	0.008	-0.037	-0.096	0.793	-0.007	0.008	0.503
27	Q			QV=0.10	0.001	-0.023	-0.084	0.248	-0.002	0.002	0.174
28	Q	N=200	EV=10	QV=0.01	0.367	1.912	0.789	4.965	-0.034	0.029	1.692
29	Q			QV=0.05	0.128	0.955	0.359	2.518	-0.013	0.017	1.157
30	Q			QV=0.10	0.048	0.121	-0.053	1.033	-0.012	0.009	0.636
31	Q	N=500	EV=2.5	QV=0.01	0.002	0.008	0.000	0.200	-0.001	0.002	0.094
32	Q			QV=0.05	0.001	-0.002	-0.005	0.014	0.000	0.000	0.014
33	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
34	Q	N=500	EV=5	QV=0.01	0.007	0.018	-0.022	0.483	-0.006	0.006	0.376
35	Q			QV=0.05	-0.004	-0.046	-0.052	0.159	-0.003	0.003	0.163
36	Q			QV=0.10	0.001	0.003	-0.004	0.020	0.000	0.000	0.015
37	Q	N=500	EV=10	QV=0.01	0.043	0.179	0.179	2.994	-0.010	0.013	0.820
38	Q			QV=0.05	-0.028	-0.179	-0.170	0.746	-0.013	0.008	0.523
39	Q			QV=0.10	0.003	-0.070	-0.108	0.247	-0.003	0.002	0.186
40	Q	N=700	EV=2.5	QV=0.01	0.000	-0.001	-0.003	0.030	0.000	0.000	0.027
41	Q			QV=0.05	0.000	0.000	-0.001	0.001	0.000	0.000	0.001
42	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
43	Q	N=700	EV=5	QV=0.01	-0.001	-0.026	-0.014	0.386	-0.005	0.004	0.231
44	Q			QV=0.05	-0.002	-0.029	-0.031	0.068	-0.002	0.001	0.072
45	Q			QV=0.10	0.000	-0.001	-0.001	0.000	0.000	0.000	0.002
46	Q	N=700	EV=10	QV=0.01	-0.017	-0.037	-0.024	0.860	-0.012	0.010	0.615
47	Q			QV=0.05	-0.004	-0.077	-0.099	0.558	-0.009	0.005	0.348
48	Q			QV=0.10	-0.003	-0.047	-0.058	0.114	-0.002	0.001	0.102

Table A3. Difference between the BIC-Model-Selection and the AIC-Model-Selection Relative Parameter Bias (BIC-AIC)

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.000	-0.002	-0.001	--	0.001	0.000	--
2	L		EV=5	--	0.002	0.007	0.001	--	-0.001	0.000	--
3	L		EV=10	--	-0.001	-0.008	-0.002	--	-0.001	0.000	--
4	L	N=200	EV=2.5	--	0.003	0.006	0.001	--	0.000	0.000	--
5	L		EV=5	--	0.000	0.000	0.000	--	0.000	0.000	--
6	L		EV=10	--	0.001	0.006	0.002	--	0.000	0.000	--
7	L	N=500	EV=2.5	--	0.001	-0.002	-0.001	--	0.000	0.000	--
8	L		EV=5	--	-0.001	-0.003	-0.001	--	0.000	0.000	--
9	L		EV=10	--	-0.001	-0.003	-0.001	--	0.000	0.000	--
10	L	N=700	EV=2.5	--	0.000	0.000	-0.001	--	0.000	0.000	--
11	L		EV=5	--	0.000	-0.001	0.000	--	0.000	0.000	--
12	L		EV=10	--	0.001	0.003	0.001	--	0.000	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	0.004	-0.038	0.021	0.074	-0.004	0.007	0.443
14	Q			QV=0.05	0.005	-0.104	-0.069	0.380	-0.002	0.005	0.350
15	Q			QV=0.10	0.000	-0.032	-0.044	0.149	0.000	0.001	0.135
16	Q	N=100	EV=5	QV=0.01	-0.046	-0.861	-0.468	0.517	-0.017	0.012	0.627
17	Q			QV=0.05	0.045	0.207	-0.052	0.177	0.001	0.007	0.641
18	Q			QV=0.10	0.001	-0.039	-0.091	0.495	-0.005	0.005	0.327
19	Q	N=100	EV=10	QV=0.01	-0.066	0.489	0.664	10.772	-0.052	0.021	1.469
20	Q			QV=0.05	0.000	0.577	0.147	2.187	-0.015	0.010	0.602
21	Q			QV=0.10	-0.040	-0.519	-0.429	0.803	-0.015	0.007	0.386
22	Q	N=200	EV=2.5	QV=0.01	0.006	-0.012	-0.005	0.587	-0.007	0.004	0.295
23	Q			QV=0.05	0.006	0.015	-0.019	0.227	-0.003	0.002	0.179
24	Q			QV=0.10	0.000	-0.006	-0.006	0.031	0.000	0.000	0.022
25	Q	N=200	EV=5	QV=0.01	-0.008	-0.040	-0.026	-0.434	-0.015	0.012	0.633
26	Q			QV=0.05	0.006	-0.027	-0.066	0.625	-0.004	0.006	0.354
27	Q			QV=0.10	-0.001	-0.022	-0.077	0.225	-0.002	0.002	0.157
28	Q	N=200	EV=10	QV=0.01	0.327	1.641	0.618	2.466	-0.022	0.020	1.144
29	Q			QV=0.05	0.113	0.893	0.412	2.001	-0.006	0.011	0.778
30	Q			QV=0.10	0.042	0.128	0.012	0.808	-0.009	0.006	0.454
31	Q	N=500	EV=2.5	QV=0.01	0.002	0.008	0.000	0.198	-0.001	0.002	0.093
32	Q			QV=0.05	0.001	-0.002	-0.005	0.014	0.000	0.000	0.014
33	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
34	Q	N=500	EV=5	QV=0.01	0.005	0.014	-0.019	0.385	-0.005	0.005	0.323
35	Q			QV=0.05	-0.003	-0.044	-0.050	0.152	-0.003	0.003	0.159
36	Q			QV=0.10	0.001	0.003	-0.004	0.020	0.000	0.000	0.015
37	Q	N=500	EV=10	QV=0.01	0.037	0.125	0.142	2.259	-0.007	0.009	0.622
38	Q			QV=0.05	-0.029	-0.184	-0.160	0.624	-0.012	0.007	0.447
39	Q			QV=0.10	0.003	-0.064	-0.096	0.238	-0.003	0.002	0.173
40	Q	N=700	EV=2.5	QV=0.01	0.000	-0.001	-0.003	0.030	0.000	0.000	0.027
41	Q			QV=0.05	0.000	0.000	-0.001	0.001	0.000	0.000	0.001
42	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
43	Q	N=700	EV=5	QV=0.01	-0.001	-0.026	-0.015	0.358	-0.005	0.004	0.221
44	Q			QV=0.05	-0.002	-0.028	-0.031	0.067	-0.002	0.001	0.072
45	Q			QV=0.10	0.000	-0.001	-0.001	0.000	0.000	0.000	0.002
46	Q	N=700	EV=10	QV=0.01	-0.024	-0.064	-0.037	0.565	-0.010	0.008	0.491
47	Q			QV=0.05	-0.001	-0.067	-0.086	0.516	-0.008	0.005	0.309
48	Q			QV=0.10	-0.003	-0.045	-0.056	0.112	-0.002	0.001	0.100

Table A4. Difference between the AIC-Model-Selection and the True-Model-Selection Relative Variance Bias (AIC-True)

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.001	0.002	0.002	--	0.001	0.001	--
2	L		EV=5	--	-0.002	-0.001	0.000	--	-0.001	0.001	--
3	L		EV=10	--	0.002	0.004	0.002	--	0.001	0.001	--
4	L	N=200	EV=2.5	--	-0.005	-0.004	-0.001	--	-0.002	-0.001	--
5	L		EV=5	--	0.000	-0.001	-0.001	--	0.000	-0.001	--
6	L		EV=10	--	0.000	-0.002	-0.002	--	0.000	-0.001	--
7	L	N=500	EV=2.5	--	-0.001	0.000	0.001	--	-0.001	0.000	--
8	L		EV=5	--	0.001	0.002	0.001	--	0.001	0.000	--
9	L		EV=10	--	0.002	0.002	0.002	--	0.001	0.002	--
10	L	N=700	EV=2.5	--	0.000	0.000	0.002	--	0.000	0.000	--
11	L		EV=5	--	-0.001	0.000	-0.001	--	0.000	0.000	--
12	L		EV=10	--	-0.001	-0.001	-0.001	--	-0.001	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	0.007	0.002	0.018	-0.835	0.002	-0.015	-0.006
14	Q			QV=0.05	-0.001	-0.011	-0.014	-0.208	0.000	-0.019	0.005
15	Q			QV=0.10	-0.001	-0.006	-0.009	-0.015	0.000	-0.005	0.000
16	Q	N=100	EV=5	QV=0.01	-0.010	-0.001	0.048	-0.003	-0.013	-0.014	0.003
17	Q			QV=0.05	-0.011	-0.023	-0.004	0.043	-0.006	-0.027	0.014
18	Q			QV=0.10	-0.006	-0.027	-0.035	0.007	0.001	-0.022	0.005
19	Q	N=100	EV=10	QV=0.01	0.016	0.043	0.164	0.293	0.004	0.008	0.001
20	Q			QV=0.05	-0.025	-0.061	-0.072	0.068	0.005	-0.039	-0.001
21	Q			QV=0.10	-0.056	-0.101	-0.130	0.029	-0.001	-0.042	0.000
22	Q	N=200	EV=2.5	QV=0.01	-0.005	-0.006	-0.004	-0.114	-0.002	-0.006	-0.012
23	Q			QV=0.05	0.001	0.001	0.001	-0.027	0.000	-0.002	0.000
24	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
25	Q	N=200	EV=5	QV=0.01	0.011	0.015	0.036	0.009	0.002	-0.004	-0.005
26	Q			QV=0.05	0.004	-0.001	0.008	0.018	0.002	-0.015	0.000
27	Q			QV=0.10	0.001	-0.001	0.001	0.005	0.000	-0.003	0.004
28	Q	N=200	EV=10	QV=0.01	0.018	0.029	0.075	0.169	0.001	-0.002	0.003
29	Q			QV=0.05	-0.005	-0.019	-0.012	0.109	0.005	-0.019	0.003
30	Q			QV=0.10	-0.016	-0.030	-0.030	0.048	0.001	-0.018	0.002
31	Q	N=500	EV=2.5	QV=0.01	0.000	0.000	0.000	0.000	0.000	-0.001	0.000
32	Q			QV=0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000
33	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
34	Q	N=500	EV=5	QV=0.01	0.003	0.004	0.008	0.000	0.002	-0.001	0.000
35	Q			QV=0.05	0.000	0.000	0.002	0.006	0.000	-0.001	0.000
36	Q			QV=0.10	0.000	-0.001	0.000	0.000	0.000	0.000	0.000
37	Q	N=500	EV=10	QV=0.01	0.012	0.019	0.053	0.116	0.001	-0.001	0.001
38	Q			QV=0.05	0.005	0.007	0.020	0.045	0.001	-0.003	0.001
39	Q			QV=0.10	-0.001	-0.002	-0.001	0.003	0.000	-0.001	0.000
40	Q	N=700	EV=2.5	QV=0.01	0.000	0.000	0.000	0.000	0.000	0.000	0.000
41	Q			QV=0.05	0.000	0.000	0.000	0.000	0.001	0.000	0.000
42	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
43	Q	N=700	EV=5	QV=0.01	0.001	0.003	0.006	0.000	-0.001	0.000	0.000
44	Q			QV=0.05	0.000	-0.001	0.000	0.000	0.000	0.000	0.001
45	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
46	Q	N=700	EV=10	QV=0.01	0.005	0.006	0.014	0.036	0.002	-0.001	0.000
47	Q			QV=0.05	0.000	0.000	0.001	0.005	0.000	-0.001	0.000
48	Q			QV=0.10	0.000	0.000	-0.001	0.000	0.000	0.000	0.000

Table A5. Difference between the BIC-Model-Selection and the True-Model-Selection Relative Variance Bias (BIC-True)

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.000	0.000	0.000		0.001	0.000	
2	L		EV=5	--	0.000	0.000	0.000		0.000	0.001	
3	L		EV=10	--	0.000	0.000	0.000		0.000	0.000	
4	L	N=200	EV=2.5	--	0.001	0.000	0.000		0.000	0.000	
5	L		EV=5	--	0.000	0.000	0.000		0.000	0.000	
6	L		EV=10	--	0.000	0.001	0.000		0.000	0.000	
7	L	N=500	EV=2.5	--	0.000	0.000	0.000		0.000	0.000	
8	L		EV=5	--	0.000	0.000	0.000		0.000	0.000	
9	L		EV=10	--	0.000	0.000	-0.001		0.000	0.000	
10	L	N=700	EV=2.5	--	0.000	0.000	0.000		0.000	0.000	
11	L		EV=5	--	0.000	0.000	0.000		0.000	0.000	
12	L		EV=10	--	0.000	0.000	0.000		0.000	0.000	
13	Q	N=100	EV=2.5	QV=0.01	-0.004	-0.024	0.000	0.874	-0.010	-0.044	-0.060
14	Q			QV=0.05	-0.010	-0.064	-0.081	-0.920	0.003	-0.061	0.021
15	Q			QV=0.10	-0.012	-0.048	-0.070	-0.124	-0.002	-0.039	0.018
16	Q	N=100	EV=5	QV=0.01	-0.067	-0.154	-0.151	-0.058	-0.008	-0.099	-0.038
17	Q			QV=0.05	-0.019	-0.063	-0.091	-0.075	-0.005	-0.089	-0.023
18	Q			QV=0.10	-0.036	-0.099	-0.135	-0.100	-0.005	-0.082	0.030
19	Q	N=100	EV=10	QV=0.01	0.045	0.158	0.354	0.426	-0.069	0.040	0.072
20	Q			QV=0.05	-0.104	-0.117	-0.055	0.170	-0.031	-0.074	0.036
21	Q			QV=0.10	-0.220	-0.371	-0.506	-0.238	-0.012	-0.133	0.007
22	Q	N=200	EV=2.5	QV=0.01	-0.002	-0.012	0.020	-1.114	-0.004	-0.033	-0.034
23	Q			QV=0.05	0.009	0.007	0.020	-0.370	-0.002	-0.033	0.000
24	Q			QV=0.10	-0.001	-0.004	-0.003	-0.012	0.000	-0.005	0.000
25	Q	N=200	EV=5	QV=0.01	-0.007	-0.018	-0.051	-0.117	-0.013	-0.024	-0.038
26	Q			QV=0.05	-0.010	-0.054	-0.070	-0.056	0.004	-0.055	0.000
27	Q			QV=0.10	-0.016	-0.042	-0.046	0.005	-0.002	-0.039	0.014
28	Q	N=200	EV=10	QV=0.01	0.308	0.432	0.725	1.243	0.049	-0.002	0.042
29	Q			QV=0.05	-0.016	-0.073	-0.132	0.215	-0.003	-0.058	-0.005
30	Q			QV=0.10	-0.055	-0.106	-0.132	0.078	0.006	-0.056	0.021
31	Q	N=500	EV=2.5	QV=0.01	0.005	0.008	0.017	-0.244	0.000	-0.005	0.000
32	Q			QV=0.05	0.002	0.000	0.001	-0.011	0.001	-0.002	0.000
33	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
34	Q	N=500	EV=5	QV=0.01	0.003	-0.008	-0.019	-0.023	0.005	-0.008	0.000
35	Q			QV=0.05	-0.001	-0.008	-0.003	0.016	0.000	-0.015	0.000
36	Q			QV=0.10	0.001	-0.001	0.000	0.000	0.001	-0.003	0.011
37	Q	N=500	EV=10	QV=0.01	0.060	0.072	0.161	0.310	0.015	0.001	0.001
38	Q			QV=0.05	-0.033	-0.049	-0.017	0.173	-0.008	-0.035	0.008
39	Q			QV=0.10	-0.026	-0.052	-0.060	0.030	0.003	-0.025	0.000
40	Q	N=700	EV=2.5	QV=0.01	-0.001	-0.004	-0.004	0.000	0.000	-0.003	0.000
41	Q			QV=0.05	0.000	0.000	0.000	0.000	0.001	0.000	0.000
42	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
43	Q	N=700	EV=5	QV=0.01	0.006	0.011	0.031	0.013	-0.003	-0.009	0.000
44	Q			QV=0.05	0.000	-0.003	0.000	0.013	0.001	-0.007	0.001
45	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
46	Q	N=700	EV=10	QV=0.01	-0.031	-0.035	-0.039	-0.005	-0.017	-0.012	0.000
47	Q			QV=0.05	-0.012	-0.026	-0.012	0.096	0.000	-0.023	0.000
48	Q			QV=0.10	-0.007	-0.014	-0.012	0.020	0.001	-0.010	0.000

Table A6. Difference between the BIC-Model-Selection and the AIC-Model-Selection Relative Variance Bias (BIC-AIC)

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	-0.001	-0.002	-0.002	--	0.000	-0.001	--
2	L		EV=5	--	0.002	0.001	0.000	--	0.001	0.000	--
3	L		EV=10	--	-0.002	-0.004	-0.002	--	-0.001	-0.001	--
4	L	N=200	EV=2.5	--	0.005	0.004	0.001	--	0.002	0.001	--
5	L		EV=5	--	0.000	0.001	0.001	--	0.000	0.001	--
6	L		EV=10	--	0.000	0.002	0.002	--	0.000	0.001	--
7	L	N=500	EV=2.5	--	0.001	0.000	-0.001	--	0.001	0.000	--
8	L		EV=5	--	-0.001	-0.002	-0.001	--	-0.001	0.000	--
9	L		EV=10	--	-0.002	-0.002	-0.002	--	-0.001	-0.002	--
10	L	N=700	EV=2.5	--	0.000	0.000	-0.002	--	0.000	0.000	--
11	L		EV=5	--	0.001	0.000	0.001	--	0.000	0.000	--
12	L		EV=10	--	0.001	0.001	0.001	--	0.001	0.000	--
13	Q	N=100	EV=2.5	QV=0.01	-0.011	-0.025	-0.018	1.708	-0.012	-0.029	-0.054
14	Q			QV=0.05	-0.009	-0.053	-0.067	-0.712	0.003	-0.042	0.016
15	Q			QV=0.10	-0.011	-0.042	-0.061	-0.109	-0.002	-0.034	0.018
16	Q	N=100	EV=5	QV=0.01	-0.057	-0.153	-0.199	-0.055	0.005	-0.085	-0.041
17	Q			QV=0.05	-0.008	-0.040	-0.087	-0.118	0.001	-0.062	-0.037
18	Q			QV=0.10	-0.030	-0.072	-0.100	-0.107	-0.006	-0.060	0.025
19	Q	N=100	EV=10	QV=0.01	0.029	0.115	0.190	0.133	-0.073	0.032	0.071
20	Q			QV=0.05	-0.079	-0.056	0.017	0.102	-0.036	-0.035	0.037
21	Q			QV=0.10	-0.164	-0.270	-0.376	-0.267	-0.011	-0.091	0.007
22	Q	N=200	EV=2.5	QV=0.01	0.003	-0.006	0.024	-1.000	-0.002	-0.027	-0.022
23	Q			QV=0.05	0.008	0.006	0.019	-0.343	-0.002	-0.031	0.000
24	Q			QV=0.10	-0.001	-0.004	-0.003	-0.012	0.000	-0.005	0.000
25	Q	N=200	EV=5	QV=0.01	-0.018	-0.033	-0.087	-0.126	-0.015	-0.020	-0.033
26	Q			QV=0.05	-0.014	-0.053	-0.077	-0.074	0.002	-0.040	0.000
27	Q			QV=0.10	-0.017	-0.041	-0.047	0.000	-0.002	-0.036	0.010
28	Q	N=200	EV=10	QV=0.01	0.290	0.403	0.650	1.074	0.048	0.000	0.039
29	Q			QV=0.05	-0.011	-0.054	-0.120	0.106	-0.008	-0.039	-0.008
30	Q			QV=0.10	-0.039	-0.076	-0.102	0.030	0.005	-0.038	0.019
31	Q	N=500	EV=2.5	QV=0.01	0.005	0.008	0.017	-0.244	0.000	-0.004	0.000
32	Q			QV=0.05	0.002	0.000	0.001	-0.011	0.001	-0.002	0.000
33	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
34	Q	N=500	EV=5	QV=0.01	0.000	-0.012	-0.027	-0.023	0.003	-0.007	0.000
35	Q			QV=0.05	-0.001	-0.008	-0.005	0.010	0.000	-0.014	0.000
36	Q			QV=0.10	0.001	0.000	0.000	0.000	0.001	-0.003	0.011
37	Q	N=500	EV=10	QV=0.01	0.048	0.053	0.108	0.194	0.014	0.002	0.000
38	Q			QV=0.05	-0.038	-0.056	-0.037	0.128	-0.009	-0.032	0.007
39	Q			QV=0.10	-0.025	-0.050	-0.059	0.027	0.003	-0.024	0.000
40	Q	N=700	EV=2.5	QV=0.01	-0.001	-0.004	-0.004	0.000	0.000	-0.003	0.000
41	Q			QV=0.05	0.000	0.000	0.000	0.000	0.000	0.000	0.000
42	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
43	Q	N=700	EV=5	QV=0.01	0.005	0.008	0.025	0.013	-0.002	-0.009	0.000
44	Q			QV=0.05	0.000	-0.002	0.000	0.013	0.001	-0.007	0.000
45	Q			QV=0.10	0.000	0.000	0.000	0.000	0.000	0.000	0.000
46	Q	N=700	EV=10	QV=0.01	-0.035	-0.041	-0.053	-0.041	-0.018	-0.011	0.000
47	Q			QV=0.05	-0.012	-0.026	-0.013	0.091	0.000	-0.022	0.000
48	Q			QV=0.10	-0.007	-0.014	-0.011	0.020	0.001	-0.010	0.000

Table A7. True-Model-Selection Coverage Rate

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.930	0.929	0.929		0.944	0.949	
2	L		EV=5	--	0.948	0.931	0.945		0.956	0.956	
3	L		EV=10	--	0.942	0.942	0.929		0.950	0.947	
4	L	N=200	EV=2.5	--	0.947	0.947	0.945		0.942	0.948	
5	L		EV=5	--	0.960	0.952	0.944		0.947	0.950	
6	L		EV=10	--	0.941	0.950	0.937		0.946	0.940	
7	L	N=500	EV=2.5	--	0.934	0.946	0.948		0.937	0.951	
8	L		EV=5	--	0.951	0.956	0.955		0.948	0.952	
9	L		EV=10	--	0.947	0.950	0.950		0.938	0.952	
10	L	N=700	EV=2.5	--	0.945	0.950	0.948		0.958	0.947	
11	L		EV=5	--	0.945	0.955	0.955		0.939	0.952	
12	L		EV=10	--	0.954	0.954	0.939		0.944	0.953	
13	Q	N=100	EV=2.5	QV=0.01	0.958	0.947	0.938	0.954	0.953	0.949	0.953
14	Q			QV=0.05	0.931	0.925	0.910	0.958	0.937	0.938	0.933
15	Q			QV=0.10	0.926	0.906	0.893	0.948	0.958	0.938	0.950
16	Q	N=100	EV=5	QV=0.01	0.947	0.932	0.937	0.963	0.949	0.934	0.956
17	Q			QV=0.05	0.947	0.925	0.880	0.962	0.949	0.944	0.937
18	Q			QV=0.10	0.941	0.910	0.836	0.964	0.944	0.952	0.954
19	Q	N=100	EV=10	QV=0.01	0.933	0.924	0.860	0.967	0.942	0.966	0.961
20	Q			QV=0.05	0.941	0.930	0.820	0.976	0.932	0.947	0.949
21	Q			QV=0.10	0.943	0.926	0.760	0.855	0.937	0.947	0.946
22	Q	N=200	EV=2.5	QV=0.01	0.957	0.948	0.942	0.963	0.947	0.951	0.963
23	Q			QV=0.05	0.942	0.955	0.934	0.968	0.949	0.947	0.952
24	Q			QV=0.10	0.940	0.949	0.946	0.965	0.952	0.948	0.950
25	Q	N=200	EV=5	QV=0.01	0.941	0.940	0.947	0.969	0.950	0.947	0.935
26	Q			QV=0.05	0.931	0.938	0.916	0.964	0.951	0.947	0.952
27	Q			QV=0.10	0.951	0.917	0.881	0.963	0.953	0.963	0.954
28	Q	N=200	EV=10	QV=0.01	0.955	0.939	0.913	0.977	0.946	0.944	0.940
29	Q			QV=0.05	0.945	0.936	0.886	0.977	0.943	0.956	0.947
30	Q			QV=0.10	0.951	0.935	0.834	0.789	0.948	0.942	0.940
31	Q	N=500	EV=2.5	QV=0.01	0.952	0.959	0.938	0.973	0.955	0.939	0.955
32	Q			QV=0.05	0.965	0.946	0.939	0.972	0.963	0.944	0.945
33	Q			QV=0.10	0.956	0.945	0.953	0.954	0.962	0.947	0.955
34	Q	N=500	EV=5	QV=0.01	0.941	0.954	0.948	0.977	0.961	0.937	0.948
35	Q			QV=0.05	0.949	0.948	0.937	0.976	0.961	0.943	0.955
36	Q			QV=0.10	0.945	0.947	0.942	0.937	0.952	0.954	0.954
37	Q	N=500	EV=10	QV=0.01	0.951	0.951	0.957	0.974	0.953	0.952	0.953
38	Q			QV=0.05	0.953	0.951	0.937	0.695	0.950	0.958	0.960
39	Q			QV=0.10	0.943	0.924	0.893	0.858	0.960	0.967	0.956
40	Q	N=700	EV=2.5	QV=0.01	0.953	0.974	0.947	0.975	0.952	0.954	0.958
41	Q			QV=0.05	0.953	0.953	0.945	0.973	0.945	0.940	0.950
42	Q			QV=0.10	0.958	0.952	0.948	0.944	0.948	0.960	0.953
43	Q	N=700	EV=5	QV=0.01	0.949	0.939	0.941	0.968	0.953	0.957	0.956
44	Q			QV=0.05	0.955	0.955	0.931	0.974	0.939	0.936	0.946
45	Q			QV=0.10	0.944	0.945	0.924	0.932	0.937	0.959	0.951
46	Q	N=700	EV=10	QV=0.01	0.944	0.950	0.945	0.969	0.953	0.948	0.954
47	Q			QV=0.05	0.957	0.949	0.925	0.729	0.941	0.940	0.952
48	Q			QV=0.10	0.951	0.937	0.921	0.874	0.943	0.950	0.955

Table A8. AIC-Model-Selection Coverage Rate

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.930	0.927	0.930	--	0.945	0.952	--
2	L		EV=5	--	0.945	0.930	0.944	--	0.957	0.959	--
3	L		EV=10	--	0.942	0.941	0.925	--	0.951	0.950	--
4	L	N=200	EV=2.5	--	0.945	0.947	0.944	--	0.941	0.947	--
5	L		EV=5	--	0.958	0.951	0.943	--	0.949	0.950	--
6	L		EV=10	--	0.938	0.953	0.939	--	0.944	0.939	--
7	L	N=500	EV=2.5	--	0.935	0.948	0.945	--	0.940	0.953	--
8	L		EV=5	--	0.952	0.954	0.956	--	0.951	0.955	--
9	L		EV=10	--	0.945	0.949	0.949	--	0.935	0.951	--
10	L	N=700	EV=2.5	--	0.943	0.949	0.951	--	0.959	0.949	--
11	L		EV=5	--	0.946	0.957	0.955	--	0.938	0.951	--
12	L		EV=10	--	0.958	0.956	0.940	--	0.943	0.956	--
13	Q	N=100	EV=2.5	QV=0.01	0.955	0.922	0.912	0.914	0.941	0.941	0.945
14	Q			QV=0.05	0.929	0.914	0.890	0.945	0.936	0.937	0.934
15	Q			QV=0.10	0.927	0.904	0.888	0.946	0.956	0.939	0.951
16	Q	N=100	EV=5	QV=0.01	0.920	0.877	0.880	0.917	0.964	0.866	0.899
17	Q			QV=0.05	0.930	0.887	0.807	0.924	0.950	0.922	0.909
18	Q			QV=0.10	0.928	0.890	0.788	0.950	0.949	0.957	0.957
19	Q	N=100	EV=10	QV=0.01	0.863	0.841	0.775	0.874	0.934	0.901	0.901
20	Q			QV=0.05	0.892	0.863	0.688	0.917	0.925	0.883	0.883
21	Q			QV=0.10	0.913	0.886	0.645	0.876	0.929	0.911	0.915
22	Q	N=200	EV=2.5	QV=0.01	0.953	0.941	0.935	0.958	0.946	0.967	0.972
23	Q			QV=0.05	0.941	0.956	0.933	0.967	0.949	0.948	0.957
24	Q			QV=0.10	0.940	0.949	0.946	0.965	0.952	0.948	0.950
25	Q	N=200	EV=5	QV=0.01	0.926	0.910	0.918	0.947	0.947	0.945	0.924
26	Q			QV=0.05	0.922	0.926	0.895	0.951	0.953	0.945	0.962
27	Q			QV=0.10	0.951	0.916	0.877	0.962	0.951	0.962	0.953
28	Q	N=200	EV=10	QV=0.01	0.918	0.884	0.848	0.936	0.930	0.909	0.896
29	Q			QV=0.05	0.912	0.893	0.812	0.956	0.928	0.937	0.921
30	Q			QV=0.10	0.943	0.918	0.784	0.841	0.951	0.937	0.940
31	Q	N=500	EV=2.5	QV=0.01	0.952	0.959	0.938	0.973	0.955	0.939	0.956
32	Q			QV=0.05	0.965	0.946	0.939	0.972	0.963	0.944	0.945
33	Q			QV=0.10	0.956	0.945	0.953	0.954	0.962	0.947	0.955
34	Q	N=500	EV=5	QV=0.01	0.940	0.952	0.943	0.974	0.965	0.953	0.971
35	Q			QV=0.05	0.949	0.948	0.936	0.976	0.961	0.945	0.959
36	Q			QV=0.10	0.945	0.947	0.942	0.937	0.952	0.954	0.954
37	Q	N=500	EV=10	QV=0.01	0.944	0.930	0.936	0.963	0.955	0.960	0.952
38	Q			QV=0.05	0.946	0.943	0.925	0.723	0.955	0.965	0.965
39	Q			QV=0.10	0.944	0.922	0.891	0.861	0.960	0.969	0.959
40	Q	N=700	EV=2.5	QV=0.01	0.953	0.974	0.947	0.975	0.952	0.954	0.958
41	Q			QV=0.05	0.953	0.953	0.945	0.973	0.945	0.940	0.950
42	Q			QV=0.10	0.958	0.952	0.948	0.944	0.948	0.960	0.953
43	Q	N=700	EV=5	QV=0.01	0.948	0.938	0.940	0.967	0.953	0.964	0.970
44	Q			QV=0.05	0.955	0.955	0.931	0.974	0.939	0.937	0.947
45	Q			QV=0.10	0.944	0.945	0.924	0.932	0.937	0.959	0.951
46	Q	N=700	EV=10	QV=0.01	0.931	0.941	0.933	0.959	0.962	0.974	0.976
47	Q			QV=0.05	0.953	0.945	0.918	0.730	0.945	0.947	0.961
48	Q			QV=0.10	0.951	0.937	0.921	0.874	0.943	0.950	0.957

Table A9. BIC-Model-Selection Coverage Rate

Case	True Model	Sample Size	Error Variance	Quadratic Variance	ϕ_{11}	ϕ_{21}	ϕ_{22}	ϕ_{33}	α_1	α_2	α_3
1	L	N=100	EV=2.5	--	0.931	0.929	0.929	--	0.944	0.950	--
2	L		EV=5	--	0.948	0.931	0.945	--	0.956	0.956	--
3	L		EV=10	--	0.942	0.942	0.929	--	0.950	0.947	--
4	L	N=200	EV=2.5	--	0.947	0.947	0.945	--	0.942	0.948	--
5	L		EV=5	--	0.960	0.952	0.944	--	0.947	0.950	--
6	L		EV=10	--	0.941	0.950	0.937	--	0.946	0.940	--
7	L	N=500	EV=2.5	--	0.934	0.946	0.948	--	0.937	0.951	--
8	L		EV=5	--	0.951	0.956	0.955	--	0.948	0.952	--
9	L		EV=10	--	0.947	0.950	0.950	--	0.938	0.952	--
10	L	N=700	EV=2.5	--	0.945	0.950	0.948	--	0.958	0.947	--
11	L		EV=5	--	0.945	0.955	0.955	--	0.939	0.952	--
12	L		EV=10	--	0.954	0.954	0.939	--	0.944	0.953	--
13	Q	N=100	EV=2.5	QV=0.01	0.947	0.895	0.860	0.930	0.947	0.772	0.596
14	Q			QV=0.05	0.917	0.878	0.839	0.856	0.917	0.900	0.817
15	Q			QV=0.10	0.923	0.885	0.876	0.918	0.960	0.932	0.945
16	Q	N=100	EV=5	QV=0.01	0.889	0.778	0.889	0.944	0.944	0.500	0.333
17	Q			QV=0.05	0.936	0.830	0.809	0.830	0.936	0.809	0.489
18	Q			QV=0.10	0.920	0.902	0.753	0.862	0.931	0.891	0.925
19	Q	N=100	EV=10	QV=0.01	1.000	1.000	1.000	0.667	0.667	0.333	0.333
20	Q			QV=0.05	0.786	0.786	0.714	0.786	0.929	0.714	0.643
21	Q			QV=0.10	0.857	0.821	0.464	0.857	0.893	0.786	0.857
22	Q	N=200	EV=2.5	QV=0.01	0.935	0.871	0.900	0.929	0.918	0.906	0.876
23	Q			QV=0.05	0.953	0.936	0.919	0.940	0.942	0.932	0.934
24	Q			QV=0.10	0.938	0.946	0.943	0.963	0.954	0.951	0.951
25	Q	N=200	EV=5	QV=0.01	0.944	0.889	0.917	0.944	0.833	0.472	0.278
26	Q			QV=0.05	0.886	0.867	0.810	0.829	0.895	0.800	0.838
27	Q			QV=0.10	0.940	0.889	0.840	0.927	0.949	0.944	0.936
28	Q	N=200	EV=10	QV=0.01	1.000	1.000	1.000	1.000	1.000	0.000	0.000
29	Q			QV=0.05	0.786	0.714	0.500	0.929	1.000	0.643	0.571
30	Q			QV=0.10	0.875	0.818	0.625	0.818	0.943	0.807	0.795
31	Q	N=500	EV=2.5	QV=0.01	0.951	0.945	0.917	0.963	0.950	0.951	0.966
32	Q			QV=0.05	0.966	0.945	0.940	0.971	0.964	0.948	0.953
33	Q			QV=0.10	0.956	0.945	0.953	0.954	0.962	0.947	0.955
34	Q	N=500	EV=5	QV=0.01	0.928	0.917	0.900	0.944	0.972	0.844	0.883
35	Q			QV=0.05	0.934	0.928	0.913	0.964	0.962	0.928	0.960
36	Q			QV=0.10	0.946	0.946	0.941	0.938	0.950	0.956	0.961
37	Q	N=500	EV=10	QV=0.01	0.933	0.933	0.933	0.867	0.933	0.567	0.233
38	Q			QV=0.05	0.914	0.900	0.900	0.800	0.914	0.829	0.729
39	Q			QV=0.10	0.935	0.896	0.821	0.886	0.973	0.965	0.938
40	Q	N=700	EV=2.5	QV=0.01	0.954	0.975	0.942	0.972	0.955	0.968	0.971
41	Q			QV=0.05	0.953	0.953	0.945	0.973	0.945	0.941	0.951
42	Q			QV=0.10	0.958	0.952	0.948	0.944	0.948	0.960	0.953
43	Q	N=700	EV=5	QV=0.01	0.932	0.897	0.906	0.947	0.941	0.935	0.932
44	Q			QV=0.05	0.949	0.948	0.923	0.967	0.937	0.940	0.956
45	Q			QV=0.10	0.944	0.945	0.924	0.932	0.937	0.959	0.952
46	Q	N=700	EV=10	QV=0.01	0.917	0.944	0.917	0.944	0.889	0.556	0.556
47	Q			QV=0.05	0.925	0.881	0.824	0.780	0.925	0.868	0.855
48	Q			QV=0.10	0.944	0.917	0.901	0.889	0.943	0.955	0.956

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *AC-19*, 716-723.

Akaike, H. (1978). A Bayesian analysis of the minimum ACI procedure. *Annals of the Institute of Statistical Mathematics, Series A*, *30*, 9-14.

Anderson, J. C. & Gerbing, D. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49*, 155-173.

Andrew, D. W. K. (2001). Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica 69*, 683-734.

Anscombe, F. J. (1967). Topics in the investigation of linear relations fitted by the method of least squares (with discussion). *Journal of the Royal Statistical Society*: *Series B, 29*, 1-52.

Bancroft, T. A. (1944). On biases in estimation due to the use of preliminary tests of significance. *Annals of Mathematical Statistics, 15*, 190-204.

Bancroft, T. A., & Han, C. P. (1977). Inference based on conditional specification: a note and a bibliography. *International Statistical Review*, *45*, 117-127.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246.

Bentler, P. M. & Wu, E. J. C. (1995). EQS program manual.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

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

Breiman, L. (1988). Submodel selection and evaluation in regression – The conditional case and little bootstrap, *Technical Report 169*, University of California, Berkeley, Dept. of Statistics.

Breiman, L. (1992). The little bootstrap and other methods for dimensionality selection in regression: X-fixed prediction error. *Journal of the American Statistical Association*, *87*, 738-354.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101*, 147-158.

Buckland, S. T., Burnham, K. P., & Augustin, N.H. (1997). Model selection: An integral part of inference. *Biometrics, 53*, 603-618.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2^{nd} ed.), New York: Springer-Verlag.

Burnham, K. P., & Anderson, D. R. (2004). Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*, 261-304.

Chatfield, C. (1995). Model uncertainty, data mining and statistical inference (with discussion). *Journal of the Royal Statistical Society, Series A, 158*, 419-466.

Chiou, P., & Han, C. P. (1995). Conditional interval estimation of the exponential location parameter following rejection of a pre-test. *Communications in Statistics-Theory and Methods*, *24*, 1481–1492.

Coffman, D. L., & Millsap, R. E. (2006). Evaluating latent growth models using individual fit statistics. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*, 1-27.

Collins, L. M., & Sayer, A. (2001). *New methods for the analysis of change*. Washington, DC: American Psychological Association.

Cronbach, L. J., & Furby, L. (1970). How we should measure 'change': Or should we? *Psychological Bulletin, 74*, 68-80.

Cudeck, R., & Browne, M. W. (1983). Cross-validation of covariance structures. *Multivariate Behavioral Research*, *18*, 147-167.

Cudeck, R., & Henly, S. J. (1991) Model selection in covariance structures analysis and the "problem" of sample size: A clarification. *Psychological Bulletin*, *109*, 512-519.

Danilov, D., & Magnus, J.R. (2004). On the harm that ignoring pretesting can cause. *Journal of Econometrics, 122*, 27-46.

Dillon, W. R., Kumar, A. & Mulani, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes and solutions to Heywood cases. *Psychological Bulletin*, 101, 126-135.

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

Duncan, S. C., & Duncan, T. E. (1994). Modeling incomplete longitudinal substance use data using latent variable growth curve methodology. *Multivariate Behavioral Research, 29*, 313-338.

Duncan, S. C., & Duncan, T. E. (1996). A multivariate latent growth curve analysis of adolescent substance use. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*, 323-347.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika, 56*, 495-515.

Faraway, J. J. (1992). On the cost of data analysis. *Journal of Computational and Graphical Statistics, 1*, 213-229.

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. New York: Chichester: Wiley-Interscience.

Freedman, D., Navidi, W., & Peters, S. (1988). On the impact of variable selection in fitting regression equations. *In Lecture notes in Economics and Mathematical systems* (pp.1-16). New York: Springer-Verlag.

Giles, J. A., & Giles, D. E. A. (1993). Pre-test estimation and testing in econometrics: Recent developments. *Journal of Economic Surveys, 7*, 145-197.

Giles, J. A., & Srivastava, V. (1993). The exact distribution of a least-squares regression coefficient after a preliminary t-test. *Statistics and Probability Letters*, *16*, 59-64.

Goldstein, H. (2003). *Multilevel statistical models* (3rd ed.). London: Arnold.

Gottman, J. M. (1995). *The analysis of change*. Mahwah, N J: Lawrence Erlbaum Associates.

Hancock, G. R. (2006). Power analysis in covariance structure modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (pp. 69-115). Greenwood, CT: Information Age.

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

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.

Hjort, N. L., & Claeskens, G (2003). Frequentist model average estimators. *Journal of the American Statistical Association, 98*, 879–899.

Hjorth, U. (1982). Model selection and forward validation. *Scandinavian Journal of Statistics, 9*, 95-105.

Hjorth, U. (1990). Model selection needs resampling methods. *Technical Report LiTH-MAT-R-1990-12*. Linköping University, Linköping.

Hjorth, U. (1994). *Computer intensive statistical methods: validation, model selection and bootstrap*. London: Chapman and Hall.

Hodges, J. S. (1987). Uncertainty, policy analysis and statistics. *Statistical Science*, *2*, 259-275.

Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian model averaging: A tutorial (with discussion). *Statistical Science*, *19*, 382-417.

Hoover, K.D., & Perez, S. J. (1999). Data mining reconsidered; encompassing and the general-to-specific approach to specification search. *The Econometrics Journal, 2*, 1-25.

Hoyle, R. H. (Ed.). (1995). *Structural equation modeling: concepts, issues, and applications.* Thousand Oaks, CA: Sage.

Hox, J. J. (2000). Multilevel analyses of grouped and longitudinal data. In T. D. Little, K. U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and nultilevel data* (pp. 15-32). Mahwah, NJ: Erlbaum.

Huntsberger, D. V. (1955). A generalization of a preliminary testing procedure for pooling data. *Annals of Mathematical Statistics, 26*, 734-743.

Hurvich, C. M., & Tsai, C. L. (1990). The impact of model selection on inference in linear regression. *The American Statistician, 44*, 214-217.

Ichikawa, M., & Konishi, S. (1995). Application of the bootstrap methods in factor analysis. Psychometrika, 60, 77-93.

Kabaila, P. (1995). The effect of model selection on confidence regions and prediction regions. *Econometric Theory, 11*, 537-549.

Kabaila, P. (1998). Valid confidence intervals in regression after variable selection. *Econometric Theory, 14*, 463-482.

Kabaila, P. (2006). On the coverage probability of confidence intervals in regression after variable selection. *Australian and New Zealand Journal of Statistics, 47*, 549-562.

Kabaila, P., & Leeb, H. (2006). On the large-sample minimal coverage probability of confidence intervals after model selection. *Journal of the American Statistical Association, 101*, 619-629, Kline, R. B. (2004). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Leeb, H. (2005). The distribution of a linear predictor after model selection: conditional finite-sample distributions and asymptotic approximations. *Journal of Statistical Planning and Inference, 134*, 64-89.

Leeb, H., & Pötcher, B. M. (2003). The finite-sample distribution of post-model-selection estimators, and uniform versus non-uniform approximations. *Econometric Theory, 19*, 100-142.

Leeb, H., & Pötcher, B. M. (2005). Model selection and inference: Facts and fiction, *Econometric Theory, 21*, 21-59.

Lovell, M. C. (1983). Data mining. *The Review of Economics and Statistics, 65*, 1-12.

MacCallum, R. C., Kim, C., Malarkey, W. B., & Kiecolt-Glaser, J. K. (1997). Studying multivariate change using multilevel models and latent curve models. *Multivariate Behavioral Research*, *32*, 215-253.

Madigan, D., & Raftery, A. E. (1994). Model selection and accounting for model uncertainty in graphical models using Occam's window. *Journal of the American Statistical Association*, *89*, 1535-1546.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107-122.

Mehta, P. D., & West, S. G. (2000). Putting the individual back in individual growth curves. *Psychological Methods*, *5*, 23-43.

Miller, A. J. (1984). Selection of subsets of regression variables (with discussion). *Journal of the Royal Statistical Society, Series. A, 147*, 389-425.

Miller, A. J. (1990) *Subset selection in regression*. London: Chapman and Hall.

Mosteller, F. (1948). On pooling data. *Journal of the American Statistical Association, 43*, 231-242.

Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual difference in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods*, *2*, 371-402.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 599-620.

Muthén, L. K., & Muthén, B. O. (1998-2007). Mplus user's guide (5th ed.). Los Angeles: Author.

Nevitt, J., & Hancock, G. R. (2001). Performance of bootstrapping approaches to model test statistics and parameter standard error estimation in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 8,* 353-377.

Picard, R. R., & Cook, R. D. (1984). Cross-validation of regression models. *Journal of American Statistical Association*, *79*, 575-583.

Patterson, G. R. (1993). Orderly change in a stable world: The antisocial trait as a chimera. *Journal of Consulting and Clinical Psychology*, *61*, 911-919.

Pötcher, B. M. (1991). Effects of model selection on inference. *Econometric Theory, 7*, 163-185.

Pötcher, B. M. (1995). Comment on 'The effect of model selection on confidence regions and prediction regions' by P. Kabaila. *Econometric Theory,* 11, 550–559.

Pötcher, B. M., & Novak, A.J. (1998). The distribution of estimators after model selection: Large and small sample results. *Journal of Statistical Computation and Simulation, 60*, 19-56.

Raftery, A. E. (1995) Rejoinder: Model selection is unavoidable in social research. *Sociological Methodology*, *25*, 185-195.

Raftery, A. E. (1996). Approximate Bayes factors and accounting for model uncertainty in generalized linear models. *Biometrika*, *83*, 251-266.

Raftery, A. E., Madigan, D. M. & Hoeting, J. (1993). Model selection and accounting for model uncertainty in linear regression models. *Technical Report 262*, Department of Statistics, University of Washington, Seattle.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2^{nd} ed.). Thousand Oaks: Sage.

Preacher, K. J. (2006). Quantifying parsimony in structural equation modeling. *Multivariate Behavioral Research*, *4*, 227-259.

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling.* Thousand Oaks, CA: Sage.

Regal, R. R., & Hook, E. B. (1991) The effects of model selection on confidence intervals for the size of a closed population. *Statistics in Medicine*, *10*, 717-721.

Rencher, A. C., & Pun, F. C. (1980). Inflation of R^2 in best subset regression. *Technometrics*, *22*, 49-53.

Rogosa, D., & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. *Psychology*, *61*, 911-919.

Saris, W. E., & Satorra, A. (1993). Power evaluations in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 181-204). Newbury Park, CA: Sage.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399-419). Thousand Oaks, CA: Sage.

Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*, 507-514.

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

Shao, J. (1993). Linear model selection via cross validation. *Journal of the American Statistical Association, 88*, 486-494.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, *24*, 323-355.

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

Sobel, M. E., & Bohrnstedt, G. W. (1985). Use of null models in evaluating fit of covariance structure models. *Sociological Methodology*, *15*, 152-178.

Taylor, J. M. G, Siqueira, A. L., & Weiss, R. T. (1996). The cost of adding parameters to a Model. *Journal of the Royal Statistical Society. Series B (Methodological), 58*, 593-607.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1-10.

Tukey, J. W. (1980). Methodological comments focused on opportunities. In L. V. Jones (Ed.). *The collected works of John W. Tukey*. Murray Hill, NJ: Bell Telephone Laboratories.

Weiss, R. E. (2005). *Modeling longitudinal data.* New York: Springer.

Whittaker, T. A., & Stapleton, L. M. (2006). The performance of cross-validation indices used to select among competing covariance structure models under multivariate nonnormality conditions. *Multivariate Behavioral Research*, *41*, 295-335.

Willett, J. B. (1988). Questions and answers in the measurement of change. *Review of Research in Education, 15*, 345-422.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116*, 363-381.

Wu, W., West, S. G., & Taylor, A. B. (2009). Evaluating model fit for growth curve models: Integration of fit indices from SEM and MLM frameworks. *Psychological Methods*, *14*, 183-201.

Zhang, P. (1992), Influence after variable selection in linear regression models, *Biometrika*, *79*, 741-746.

Zhang, P. (1992). On the distributional properties of model selection criteria. *Journal of the American Statistical Association, 87*, 732-737.