ABSTRACT

Title of Dissertation: Effects of Model Selection on the Coverage Probability of Confidence Intervals in Binary-Response Logistic Regression

Dongquan Zhang, Doctor of Philosophy, 2008

Dissertation Directed by: Professor C. Mitchell Dayton
Department of Measurement, Statistics and Evaluation

While model selection is viewed as a fundamental task in data analysis, it imposes considerable effects on the subsequent inference. In applied statistics, it is common to carry out a data-driven approach in model selection and draw inference conditional on the selected model, as if it is given *a priori*. Parameter estimates following this procedure, however, generally do not reflect uncertainty about the model structure. As far as confidence intervals are concerned, it is often misleading to report estimates based upon conventional $1-\alpha$ without considering possible post-model-selection impact. This paper addresses the coverage probability of confidence intervals of logit coefficients in binary-response logistic regression. We conduct simulation studies to examine the performance of automatic model selectors AIC and BIC, and their subsequent effects on actual coverage probability of interval estimates. Important considerations (e.g. model structure, covariate correlation, etc.) that may have key influence are investigated. This study contributes in terms of understanding quantitatively how the post-model-selection confidence intervals perform in terms of coverage in binary-response logistic regression models.

A major conclusion was that while it is usually below the nominal level, there is no simple predictable pattern with regard to how and how far the actual coverage probability of confidence intervals may fall. The coverage probability varies given the effects of multiple factors:

(1) While the model structure always plays a role of paramount importance, the covariate correlation significantly affects the interval's coverage, with the tendency that a higher correlation indicates a lower coverage probability.

(2) No evidence shows that AIC inevitably outperforms BIC in terms of achieving higher coverage probability, or *vice versa*. The model selector's performance is dependent upon the uncertain model structure and/or the unknown parameter vector θ.

(3) While the effect of sample size is intriguing, a larger sample size does not necessarily achieve asymptotically more accurate inference on interval estimates.

(4) Although the binary threshold of the logistic model may affect the coverage probability, such effect is less important. It is more likely to become substantial with an unrestricted model when extreme values along the dimensions of other factors (e.g. small sample size, high covariate correlation) are observed.

EFFECTS OF MODEL SELECTION ON THE COVERAGE PROBABILITY OF CONFIDENCE INTERVALS IN BINARY-RESPONSE LOGISTIC REGRESSION

By

Dongquan Zhang

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
2008

Advisory Committee:

Professor C. Mitchell Dayton, Chair/Advisor
Associate Professor Robert G. Croninger
Professor Gregory R. Hancock
Professor Robert Mislevy
Professor George Macready

Table of Contents

List of Tables

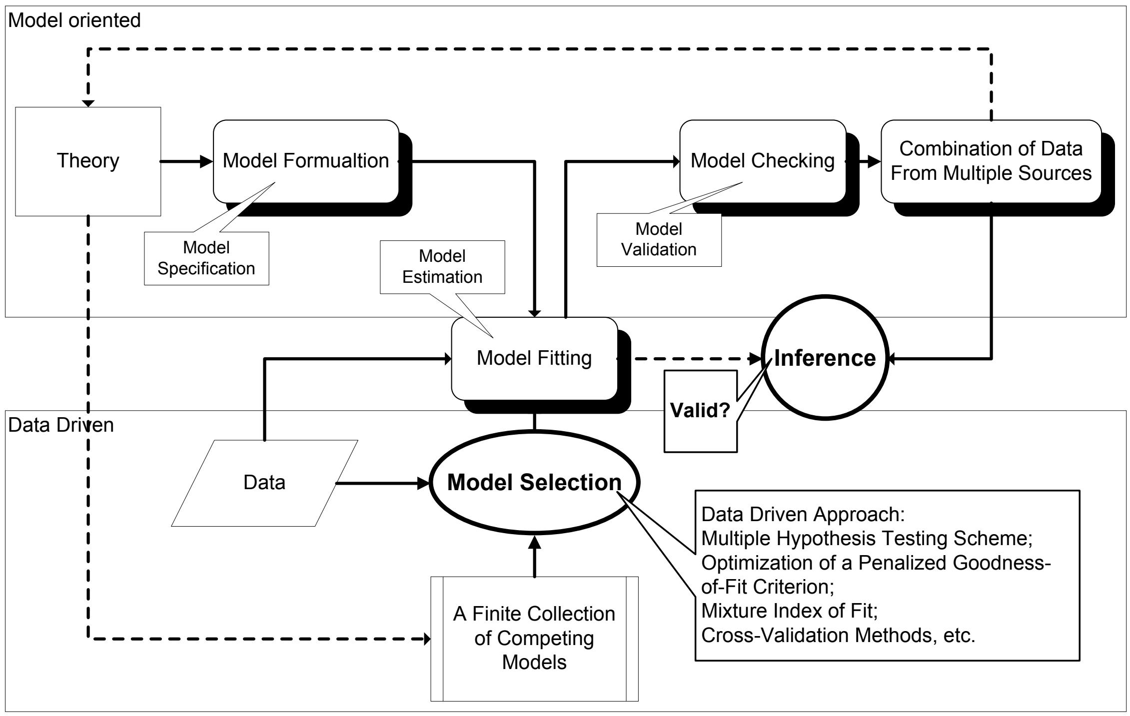
List of Figures

Chapter I: Introduction

Statistical Inference: A Process View

In theoretical statistics, it is typical that a proposed probability model is given as a starting point, following which standard parametric methods are applied to examine statistical properties such as an estimator's asymptotic distribution. In applied statistics, however, it is common to carry out a data-driven approach to select a statistical model. While the sampling data are collected from a target population, a group of competing models, which may or may not be nested in structure, are proposed. With certain stopping rules or benchmarks, standard procedures, for instance preliminary hypothesis tests, optimization of penalized goodness-of-fit criteria, mixture index of fit, or any cross-validation methods, are performed to select a "best" model. Conditional on the selected model and following some optimal estimation process such as MLE, inferences with regard to the properties of underlying parameters are drawn. Following Chatfield's (1995) summary on the "whole model building process" and Leeb and Pötcher's (2005) description about model selection, Figure 1 illustrates the usual practices associated with model selection and statistical inference. While the upper level is based upon the theoretical statistics, the lower level shows the approach in the statistical applications.

As far as the applied statistics is concerned, while model selection is viewed as a fundamental task in data analysis, far less attention has been addressed to how the presence of the selection itself impacts parameter estimates. The statistical properties of quantities of interest, for example post-model-selection estimators, or confidence intervals that are constructed with the standard parametric methods, are often reported as "unbiased" without considering that the same data set has been used for both model selection and inferences.



The figure depicts ideas following Chatfield (1995) and Leeb and Pötcher (2005).

Figure 1: Statistical inference: a process view

One problem associated with the process is that researchers ignore, to a large extent, structure uncertainty in the selected model. When researchers proceed as if the selected model is given *a priori*, it raises concerns about the validity of the inferences. In such cases, however, the effects of model selection are subtle and insidious. Although some research studies have been observed in this area, investigations with regard to such concerns are still less common and less well understood.

As Hodges (1987) and Draper *et al.* (1987) summarize, for any statistical problem there are three major sources of uncertainty (Chatfield, 1995):

(a) structure of the model;

(b) estimates of the model parameters assuming (a) is known;

(c) unexplained random variation in observed variables assuming (a) and (b) are known.

While recognition and research concerned with (b) and (c) have been pervasive for decades, comparatively little attention was paid to (a). As Chatfield (1995) claims, "This is very strange given that errors arising from (a) are likely to be far worse than those arising from other sources" (p. 421). The root of the problem is the failure to distinguish between a selected model and a "true" model known *a priori*, which, from a broader but more scrupulous point of view, may still be incorrect or "at best an approximation" (p. 421) of the real world situation. Due to the complicated nature of the problem, in the following discussions we do not touch the most rigorous portion but assume, at most of the time, that the model given *a priori* is true. Although it is arguable whether such an assumption is justifiable, it does not affect our analysis when we try to simulate the typical model building practices.

Sen (1979) explains the issue of model uncertainty from a perspective with regard to parameter space: For a parametric model with the parameter (vector) $\theta \in \Omega$ (where Ω defines a parameter space), a usual practice is to obtain the (unrestricted) maximum likelihood estimator (MLE) $\tilde{\theta}$ of θ by maximizing (over $\theta \in \Omega$) the likelihood function of the sample observations, with the assumption that the underlying distributions are of specified forms. When ω, a proper subspace of Ω, can be identified from certain practical considerations, a restricted MLE $\hat{\theta}$ of θ may serve as (asymptotically) a better estimator than $\tilde{\theta}$ given $\theta \in \omega$. It is very difficult, however, to establish with 100% certainty whether $\theta \in \omega$. If the answer is negative, " $\hat{\theta}$ may not only lose its optimality but also may be a biased (or even an inconsistent) estimator" (p. 1019). This is often the case when a selected model is applied for parameter estimates without sufficient *a priori* evidence.

As Breiman (1988) suggests, models selected by various data-driven methods can produce strongly biased estimates of mean squared prediction error. Miller (1990) calls such bias "selection bias" in the estimates of regression coefficients. It has been well established, particularly with studies in recent years (e.g. Leeb, 2002; Danilov & Magnus, 2004; Kabalia, 2006; among others), that the model selection stage can "severely affect the validity" of the following procedures (Hurvich & Tsai, 1990) and that the usual practice is "logically unsound and practically misleading" (Zhang, 1992). However, in many cases the standard statistical theory is applied prior to the verification of underlying assumptions, as if there is no uncertainty about the model structure, or the uncertain structure can, without further justification, be assumed as having "*a prespecified known form*" (Chatfield, 1995) following model selection. In statistical applications, very few studies take a position as Hocking (1976) states in selecting the best subsets of regressor variables: "The properties described here are dependent on the assumption that the subset of variables under consideration has been selected without reference to the data. Since this is contrary to normal practice, the results should be used with caution." (p. 5)

Breiman (1992) comments the wide ignorance of model selection effects as "a quiet scandal in the statistical community" suggesting the fact that even the effects are dramatic, post-model-selection estimates serve as the essential objective of data analysis, and the validity of reported inferences is often, intentionally or unintentionally, overlooked.

Confidence Intervals

A confidence interval defines, given a specified probability, a space in which a measurement or trial falls. As far as the interval estimates are concerned, a typical approach is to construct, with nominal coverage probability of $1-\alpha$, the interval around a quantity of interest, θ, based upon the selected model. For a proposed coverage probability of 95%, for example, usually it takes the form

$$\theta \in \hat{\theta}_{\hat{S}} \pm 1.96\hat{\tau}_{\hat{S}} / \sqrt{n} \tag{1}$$

where $\hat{S}$ represents the chosen model, *n* is the sample size, and $\hat{\tau}_{\hat{S}} / \sqrt{n}$ is an estimator of the *SE* for $\hat{\theta}_{\hat{S}}$, without considering model uncertainty and possibly inflated variance in the model selection step.

Kabaila (2006) and Kabaila and Leeb (2006) refer to this construction as the "naive confidence interval," as most of the time this procedure leads to inaccurate and misleading inferences (Kabaila & Leeb, 2006). Research studies in the relevant areas have suggested that the naive confidence interval has an actual coverage probability that is less than the nominal $1-\alpha$, and that the usual statistical practices adopted by people often result in too optimistic conclusions. Given the fact that the naive confidence interval is widely applied in the statistical framework, research concerned with the quantification of its post-model-selection coverage probability is still far from adequate. Major reasons for this lack of study may include, but not limited to, confounded effects of multiple factors within an incorporated model selection and inference procedure, barriers in identifying and distinguishing various types of biases and demanding requirements for intensive numerical computations, especially when the number of available models for selection becomes large.

Binary Logistic Regression

Among the analyses on the coverage probability of naive confidence intervals, most investigate the model selection effects with linear regression models. Studies that address the problem (including relevant discussions on point estimators) with nonlinear models (e.g. logistic regression models), however, are comparatively rare. A couple of examples in this category include Pötcher (1991) and Pötcher and Novak (1998), where asymptotic distribution of post-model-selection estimators in the context of general nonlinear parametric models is derived, following which their properties are studied with small samples.

Following the notation of Hosmer and Lemeshow(1989) and McCullagh and Nelder (1989), a binary logit model can be written as

$$\ln[\Psi] = \sum_{k=0}^{K} \beta_k' X_{k,n}, \tag{2}$$

where Ψ is the odds ratio $\frac{P(Y=1)/[1-P(Y=1)]}{P(Y=0)/[1-P(Y=0)]} = \frac{\pi}{1-\pi}$, and the outcome parameters are the natural logarithm of the odds ratio of the binary variable *Y,* which is a random *n*-vector of responses. The *X* is a $n \times k$ matrix with *k* linearly independent columns, or observed scores on *K* independent variables, and the β is an unknown *k*-vector of linear regression coefficients. In an unstandardized environment, typically $X_{0,n} = 1_n$ are specified so that β_0 indicates an intercept term. Given that the *i*th subject ($i = 1,...,n$) has Bernoulli Response Y_i and *p*-dimensional covariate vector x_i, and that the subjects are independent, the logistic regression model for subject *i* is

$$E(Y_i \mid x_i) = \Pr(Y_i = 1 \mid x_i) = \frac{\exp(\beta_i' x)}{1 + \exp(\beta_i' x)} \tag{3}$$

Compared with multiple linear regression models, logistic regression does not assume a linear relationship between the dependent variable and the independent variable, but assumes that a linear combination of the predictor variables determines the probability of outcome response through the link of the logit function. While the covariates in logistic regression can take any form, the dependent variable is not assumed to be normally distributed and homoscedastic for each level of the independent variable, as would be the case in linear regression. Moreover, logistic regression makes no assumption about normally distributed residuals but depends upon large sample theory on distribution properties.

In a post-model-selection scenario, while the relationship between the logit coefficients and the outcome variable turns out to be log-linear, the inferences are usually drawn based upon the sampling distribution of the ML estimation $\ln(\hat{\Psi})$.Suppose that the quantity of interest, θ, is given as $\theta = a'\beta$, where $a\ (\neq 0)$ is a known *k*-vector and β is the logit coefficient. To construct the naive confidence interval for θ with a pre-specified value $1-\alpha$ (e.g. .95), it is usual to first calculate the end points, which are denoted as $\hat{\beta} \pm z_{1-\alpha/2} \times SE(\hat{\beta})$, of the confidence interval for the coefficient β, and then exponentiate the results (Hosmer and Lemeshow, 1989, p. 44).

Purpose and Organization

The goal of this study is to distinguish effects of data-driven model selection, specifically on the interval estimates when a binary logistic model is applied in areas such as education. The statistical applications, which entail procedures such as pretests and drawing inference with the selected model, always raise concerns with regard to model uncertainty. Nevertheless, in applied statistics it is almost infeasible, due to practical restrictions, to eliminate the pretest procedures. We do not intend to involve discussions on some relevant important problems such as how to modify the model selection stage (e.g. the pros and cons for model averaging), or how to provide alternative solutions (e.g. split data) to obtain valid confidence intervals, but rather to take an approach from a realistic perspective, focusing on the current broadly applied procedures to understand the overoptimistic results, and to gauge the gap between the usually obtained conclusions and unknown truth.

This paper is organized as follows: Chapter 2 reviews major research in the relevant fields, and summarizes main accomplishments that have been achieved by previous studies. Chapter 3 addresses the research questions and elaborates methods of this study. Following Chapter 4, which reports the simulation results, Chapter 5 details data analysis, discussions and conclusions.

Chapter II: Literature Review

Preliminary Test Estimator

Giles and Giles (1993), Chatfield (1995), and Magnus (1999) summarize the extensive discussions on the structure uncertainty that is related to model selection and pretests. Specifically, Chatfield (1995) provides a broad view of statistical inference and reviews various ways of assessing and overcoming the effects of model selection, which include simulation, resampling methods and data splitting. Extensive peer discussions on the relevant topics such as model formulation, data mining and the subsequent consequences on statistical inferences follow Chatfield's comments. Buckland *et al.* (1997) suggest that the statistical inference should take extra model selection uncertainty into account, and recommend different strategies, for example weighting contending models, to achieve such goals as modified but more accurate confidence intervals. Hoover and Perez (1999) compare various model selection strategies and procedures that accommodate the effects of pretest estimators, and recommend general-to-specific procedure. Most of the extensive discussions about model uncertainty have originated from the study of pretest estimators.

In estimation theory, it is usual to estimate θ, the parameter (vector) of interest, through some optimal procedures such as maximum likelihood estimation. When it is undetermined, however, whether the unrestricted MLE $\tilde{\theta}$ or the restricted MLE $\hat{\theta}$ is an optimal estimator, researchers propose to perform a preliminary likelihood ratio test for

$$H_0 : \theta \in \varpi \tag{4}$$

where $\omega \subset \Omega$ is a restricted parameter space. Depending upon whether H_0 is retained or rejected, the preliminary test estimator (PTE) θ^* is then identified to be $\hat{\theta}$ or $\tilde{\theta}$.

Early Studies

In past decades, the properties of PTE θ^* have been studied along with the considerations of model uncertainty. Bancroft (1944), whose work is viewed as one of the earliest studies on uncertainty in formulating the statistical model, examines the biases when the statistical estimation is based upon pretests. The second example in Bancroft's paper, which deals with a simple linear regression function, $y = \beta_1 x_1 + \beta_2 x_2 + e$ (where y is the dependent variable, x_1 and x_2 are independent variables, β_1 and β_2 are regression coefficients, respectively, and *e* stands for the error term), is concerned with the estimator properties of the regression coefficient β_1 when it is uncertain whether the other independent variable, x_2, should be included in the regression. The consideration involves examination of the biased post-model-selection parameter estimate. Bancroft indicates that variations of β_2 and ρ, the correlation between x_1 and x_2, have direct effects on the bias, or

$$Bias = \rho\beta_2[1 - \sum_{i=0}^{\infty} \frac{a^i e^{-a}}{i!} I_{x_0}(\frac{n-3}{2}, \frac{3}{2} + i)], \tag{5}$$

where $x_0 = \dfrac{1}{\dfrac{\lambda}{n-3} + 1}$, $a = \dfrac{(1-\rho^2)(n-1)\beta_2^2}{2}$, I_{x_0} is an expression of an infinite series of incomplete integrals of the *F* distribution, and λ is the desired percentage point of the *F* distribution, which determines if the term containing x_2 should be included in the model, for 1 and (n-3) degrees of freedom corresponding to some assigned significance level. The derivation suggests some significant findings, which include that zero correlation between x_1 and x_2 leads to zero bias and that the bias in estimating β_1 is independent of the magnitude of β_1.

Mosteller (1948) extends the study in regression by pooling data conditional on whether β_2 equals zero, and calculating the mean square error of the pretest estimator. The procedure is later generalized by Huntsberger (1955), who proposes that the pretest estimator can be written as a weighted average of the available estimators, following that the model is considered as either restricted ($\beta_2 = 0$) or unrestricted ($\beta_2 \neq 0$) with pooling data. Meanwhile, Huntsberger also considers the situation that the parameters of interest, θ_1 and θ_2, are equal, where the a pooled estimator g($\hat{\theta}_1, \hat{\theta}_2$) provides better estimation for θ_1 than the estimator of $\hat{\theta}_1$. When it is not certain about whether θ_1 and θ_2 are equal, a statistical test *T* is conducted on the equality and examined with a weight function

$$W(T) = \phi(T)\hat{\theta}_1 + [1 - \phi(T)]g(\hat{\theta}_1, \hat{\theta}_2) \tag{6}$$

where $\phi(T)$ is defined as a function of *T* with

$$\begin{aligned} &\phi(T) = 0, T \subset A_\alpha \text{ and} \\ &\phi(T) = 1, T \subset R_\alpha \end{aligned} \tag{7}$$

where A_α and R_α stand for acceptance and rejection regions of the hypothesis H_0 with *Type I Error* α . Huntsberger suggests that the weight function offers some advantage over the estimator of $\hat{\theta}_1$ with either pooling or not pooling data.

Sclove *et al.* (1972) further extend the regression study to a wider multivariate analysis, and investigate the non-optimal properties of the preliminary-test estimators. Rencher and Pun (1980) study the inflation of R^2 in the best-subsets selection method, and examine the validity of standard errors of the selected model. Miller (1984) shows that regression estimators are biased and standard hypothesis tests may be invalid when the coefficients of subsets of regressor variables are estimated from the data used for model selection, and specifically identifies three types of biases, namely those due to omission, competition and application of a stopping rule.

Lovell (1983), in his discussion on the consequence of data mining, compares the significance levels of random data given the number of nominally significant explanatory variables in a *t*-test and proposes a rule of thumb, which states that a regression coefficient appearing to be significant at the level $\hat{\alpha}$ should be taken as significant only at the level $\alpha = 1-(1-\hat{\alpha})^{c/k}$, when the best k out of c candidate exploratory variables have been selected in a model. Mittelhammer (1984) is concerned with the risk functions of the pretest estimators when the model is misspecified. Other studies in this period include Leamer (1978), Hjorth and Holmqvist (1981), Hodges (1987), Dijkstra (1988), Miller (1990), and the work of Hjorth (1982, 1987, 1989, 1990, 1994). Bancroft and Han (1977) review the practices of unconditional and conditional specification in applied statistics, and compile a bibliography of early literatures. Judge and Bock (1978; 1983) summarize the findings in the research on pretest estimators, most of which address model selection in the field of econometrics.

Asymptotic Properties

Sen (1979) derives the asymptotic bias and dispersion matrix of the pretest estimators based on the maximum likelihood estimation and compares the estimator expressions when the model is either restricted or unrestricted. Defining MLE $\tilde{\theta}$ and $\hat{\theta}$ as the unrestricted and restricted estimators, and θ^* as the preliminary test estimator (PTE) of the quantity of interest, $\theta \in \Omega$, for a sequence of K_n independent samples (i =1,....k) , a comparison among the asymptotic biases of the three estimators $\tilde{\beta}(\gamma), \hat{\beta}(\gamma)$ and $\beta^*(\gamma)$, which take the form $n^{1/2}(\theta' - \theta)$ (where $\gamma = 0$ following null hypothesis of (4), and θ' stands for $\tilde{\theta}, \hat{\theta}$ and θ^*, respectively), indicates that under the null hypothesis, $\{\theta_n^*\}$ is asymptotically at least as good as $\{\tilde{\theta}_n\}$. When θ lies near the boundary of ω, θ_n^* has a smaller bias than $\hat{\theta}_n$. When $\|\gamma\| \to \infty$ (which suggests that the null hypothesis is definitely rejected), $\theta_n^* \to 0$ and $\hat{\theta}_n \to \infty$. Therefore, asymptotically, θ^* always offers some advantageous properties over the unrestricted and restricted estimators.

Following Sen's work, Pötcher (1991) and Giles and Srivastava (1993) investigate the asymptotic properties of pretest estimators. Pötcher (1991) studies consistent model selectors with time series models. He derives the asymptotic distribution of a post-model-selection estimator, both unconditional and conditional on selecting a correct model. For the (zero) restricted estimator $\hat{\theta}$ (which can be generalized to restrictions other than the zero restriction), Pötcher suggests that the asymptotic distribution of bias $n^{1/2}(\hat{\theta} - \theta)$ is normal with mean 0 and covariance matrix D_{uv}^0 (u and v define size of the matrix with $u \le v$), where

$$D_{uv}^0 = \begin{bmatrix} (A_{p_0+u-1}^0)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \tag{8}$$

The superscript and subscript "0"s stand for those matrix or statistics corresponding to the "true" model. $A(\theta) = \lim_{n\to\infty} n^{-1} EL_{n,PP}(\theta)$ is a matrix equivalent to the limit of second partial derivatives of an objective function $L_n(\theta)$, from which the estimator is obtained as an optimizer. p defines the parameters in the model M_p with $0 \le p \le P$, and P is for the full model. Therefore, the model selector, along with the unknown parameter θ, has important effect on the bias. Pötcher states that although asymptotically consistent selectors such as BIC typically lead to superefficient estimators near the true parameter value, there are unpleasant properties with regard to the restricted estimators. The finite sample distribution does not uniformly converge to the asymptotic distribution, which is undesirable as most statistical applications involve finite samples.

Pötcher and Novak (1998) evaluate the accuracy of the approximation provided by the asymptotic distribution in small samples, and compare the results of conservative and consistent model selectors. They consider the bias effects due to underestimation of θ and find discrepancies between the large sample and small sample conditional distributions when the minimal true model order is selected. Leeb and Pötcher (2001) discuss whether the post-model-selection estimators can be uniformly consistent. Moreover, Leeb and Pötcher (2003) obtain the uniform approximations to the finite-sample distributions, unconditional and conditional on selecting a correct model. However, as Leeb and Pötcher notice, the uniform approximation and asymptotic distribution cannot be used directly for purposes of inference, as both of them depend upon the unknown parameter value of θ.

Model Averaging

The philosophy for model averaging methods is that estimators after model selection are actually formed as mixtures of estimators from different potential models that may be selected with certain probability. Therefore, compared to the traditional method of sticking to the estimators of a single model, it is advantageous to smooth estimators across models (Hjort & Claeskens, 2003). Buckland *et al.* (1997) propose a strategy to use weighting method in model selection, which is later adopted in Burnham and Anderson (2002). Raftery *et al.* (1993) and Draper (1995) note that, even there is no true relationship for randomly generate data, traditional methods of model selection can lead to models that appear to have strong predictive power. Both of the research studies demonstrate that a Bayesian Model Averaging (BMA) framework, with which the extra estimator variability stemming from not knowing the correct model *a priori* is taken into account, can resolve this difficulty.

The BMA approach is concerned with delivering posterior distribution of interest parameters, given that prior probabilities for a list of potential models and those for the parameters of each model are set up. In recent years, Bayesian model averaging methods have gained considerable attentions (e.g. Schervish & Tsay, 1988; Le *et al*., 1993; Kass & Raftery, 1995; Madigan & York, 1995; Draper, 1995; Chatfield, 1995; among others), among which Hoeting *et al.* (1999) discuss the machinery of Markov Chain Monte Carlo (MCMC) in their tutorial, where within the Bayesian framework multiple chains move among different potential models. Hjort and Claeskens (2003) indicate major problems associated with the BMA approach such as difficulties in setting up correct priors and unclear consequences stemming from the mix of conflicting prior opinions on interest parameters, and propose "Compromise Estimators" denoted with the form

$$\hat{\theta} = \sum_{p} c(P \mid D_n)\hat{\theta}_p \tag{9}$$

where $\hat{\theta}_p$ is the estimator of model M_p, which is selected by certain model selection criterion, and $c(P \mid D_n)$ is some form of the weight function with D_n being a function of discrepancy between the full model and the null model.

Mean Squared Error

Following Roehrig's (1984) work, Magnus and Durbin (1999) assess the relative performance of post model selection estimators in linear regression models with the mean squared error (MSE). The model under study is a linear regression model $y = X\beta + Z\gamma + e$, where X (which is a $n \times k$ matrix) contains explanatory variables that are included in the model on theoretical or other grounds, Z (which is a $n \times m$ matrix) contains additional explanatory variables of which the researchers are less certain, β is a vector of nonrandom parameters, γ is a vector of nuisance parameters, and e is the random vector of unobservable residuals or disturbances that are *i.i.d* $N(0, \sigma^2)$. When $\gamma = 0$, the model is restricted.

By defining the matrices

$$M = I_n - X(X'X)^{-1}X' \text{ and } Q = (X'X)^{-1}X'Z(Z'MZ)^{-1/2} \tag{10}$$

and the parameter vector $\theta = (Z'MZ)^{1/2}\gamma$, the Least Square Estimators of β and γ are represented as $b_u = b_r - Q\hat{\theta}$ and $\hat{\gamma} = (Z'MZ)^{-1}Z'My$. The subscripts '$u$' and '$r$' stand for ‘unrestricted’ and ‘restricted’, respectively. The pretest estimator, b, is proposed as a Weighted-Average Least-Squares (WALS) estimator $b = \lambda b_u + (1-\lambda)b_r$ from those of the unrestricted and restricted models. Let $\tilde{\theta} = \lambda(\hat{\theta}, s^2)\hat{\theta}$, the Equivalence Theorem shows that MSE of the pretest estimator is

$$MSE(b) = \sigma^2 (X'X)^{-1} + QMSE(\tilde{\theta})Q' \qquad (11)$$

where $\lambda = \lambda(\hat{\theta}, s^2)$ is a scalar function of $\hat{\theta}$ defining weight.

Magnus and Durbin’s (1999) findings demonstrate that the MSE of b, the WALS estimator, is dependent upon the MSE of $\tilde{\theta}$, as long as a λ-function can be defined such that $\lambda\hat{\theta}$ is an optimal estimator of θ. One unresolved issue, however, is that there are 2^m models to consider. With the increase of m, the number of columns in the matrix of Z, the number of partially restricted models increases dramatically, which makes the settlement of the WALS estimator a more challenging task.

Danilov and Magnus (2004) derive the first and second moments of the WALS estimator and show that the error related to the pretest estimator (and more generally, the WALS estimator) vary substantially with different model selection procedures. Following the notation above, they call the columns of X “Focus” regressors and those of Z “auxiliary” regressors. Danilov and Magnus redefine the estimator of β to be $b = \sum_i \lambda_i b_{(i)}$, where i denotes the ith partially restricted model. With an idempotent matrix W, which is determined by X, Z and a selection matrix S_i, the first and second moments are derived as

$$\begin{aligned} E(b) &= \beta - \sigma QE(W\hat{\eta} - \eta), \\ Var(b) &= \sigma^2((X'X)^{-1} + Q\,\mathrm{var}(W\hat{\eta})Q' \end{aligned} \qquad (12)$$

and the Mean Squared Error of the estimator can be generalized as

$$MSE(b) = \sigma^2((X'X)^{-1} + QMSE(W\hat{\eta}))Q') \tag{13}$$

Where $\hat{\eta}$ is the optimal estimator of η, the scaled and normalized parameter vector that is defined as $\eta = (Z'MZ)^{1/2}\gamma / \sigma$. The generalization establishes a relationship between the MSE and the number of "auxiliary" regressors and allows obtaining bounds for the error.

Coverage Probability

While studies suggest that the naïve confidence intervals have an actual coverage probability smaller than the nominal level, results concerned with its quantification after model selection are limited. Important literature in this category consists of large-sample limit analyses of confidence intervals in Sen (1979), Saleh and Sen (1983), Pötcher (1991), Zhang (1992), Kabaila (1995) and Pötcher (1995), Hjort and Claeskens (2003), and Kabaila and Leeb (2006). Other studies with regard to coverage probability include Kabaila (1998; 2006), who investigate the finite-sample minimal coverage probability of confidence intervals with different number of competing models; Arabatzis, Gregoire, and Reynolds (1989), Chiou and Han (1995), and Han (1998), who work on conditional coverage probability when the preliminary test is rejected (i.e. $\theta \notin \varpi$ in (4)); Hurvich and Tsai (1990), Regal and Hook (1991), and Hook and Regal (1997), who perform simulation studies on actual coverage probability. Meanwhile, Kabaila (2006) develops a Monte Carlo simulator to improve the efficiency in simulating the coverage probability, and applies the methodology in model selection procedures with up to 20 candidate models.

Large Samples

Large samples results have been discussed along with the studies on asymptotic properties on pointwise estimation. In the relevant context, Sen (1979), Saleh and Sen (1983), Zhang (1992), Kabaila (1995), Pötcher (1991) and Pötcher (1995) conduct large-sample limit analyses of confidence intervals after model selection.

Zhang's Theorem

Zhang (1992) studies the fixed-parameter large-sample limit coverage probability. According to Zhang's Theorem 3, for any $0 < \alpha < 1$:

$$\lim_{n\to\infty} \text{pr}\{\beta \in S_{1-\alpha}(\hat{k}_\lambda)\} < 1-\alpha \tag{14}$$

where λ is a parameter in defining the generalized final prediction error (FPE) criterion (Shibata, 1984), $S_{1-\alpha}$ stands for the confidence region of β at level $1-\alpha$, and the model with $\hat{k}_\lambda$ parameters is selected. With Theorem 3 Zhang states that, for model selected by FPE, the large-sample limit coverage probability of a certain naive confidence set is strictly below the nominal level, as "…the actual variance-covariance matrix of $\hat{\beta}(\hat{k}_\lambda)$ is greater than the variance-covariance matrix of $\hat{\beta}(k)$ ..." and "if one regards $\hat{k}_\lambda$ as fixed and derives nominal confidence interval…the size of the confidence set is bound to be smaller because a smaller variance term has been used…" (p. 744). As far as the Zhang's statement is concerned, Kabaila and Leeb (2006) argue that in the general context, the Theorem only holds in the cases where the overall model is not the most parsimonious correct model for the underlying parameter, or the parameter belongs to one of the lower-dimensional models.

Kabaila's Study on Asymptotic Properties

Based on Pötcher's (1991) results, Kabaila (1995) derives the asymptotic properties of the confidence regions after model selection. By considering a time series model $\{Y_t\}$ whose distribution is determined by (θ_0, ψ_0) (where θ_0 is a parameter vector of interest and ψ_0 is a nuisance parameter vector, (θ_0, ψ_0) is supposed to belong to some set A allowable values), Kabaila suggests there are two conditions. Condition 1 can be shown as: for each $(\theta_0, \psi_0) \in A$,

$$\underset{n\to\infty}{Lim} P_{\theta_0,\psi_0}\{\theta_0 \in r_n(Y_1,....,Y_n)\} = 1-\alpha \tag{15}$$

where $P_{\theta,\psi}\{B\}$ denotes the probability of the event B when the parameter values are (θ,ψ) , and $\{r_i\}$ is a sequence of set-valued functions that should be satisfied by the asymptotic distribution of the estimator θ_0. Kabaila suggests that it is an illusion that Condition 1 implies that $\{r_n(Y_1,...,Y_n)\}$ is a confidence region for θ_0 with minimum coverage probability asymptotically approaching $1-\alpha$, reason being that no matter how large the sample size is, a $\delta > 0$ may exist so that there is a $(\theta_0, \psi_0) \in A$ for which $\underset{n\to\infty}{Lim} P_{\theta_0,\psi_0}\{\theta_0 \in r_n(Y_1,....,Y_n)\} < 1-\alpha-\delta$. Kabaila further indicates that $\{r_i\}$ should instead be required to satisfy Condition 2:

$$\underset{n\to\infty}{Lim} \inf_{(\theta_0,\psi_0)\in A} P_{\theta_0,\psi_0}\{\theta_0 \in r_n(Y_1,....,Y_n)\} = 1-\alpha \tag{16}$$

with an *infimum* term suggesting that the greatest lower bound is imposed. Kabaila illustrates examples to show that the chance Condition 2 is satisfied is very low when $\{r_n(Y_1,...,Y_n)\}$ is determined by a model selection procedure.

Hjort and Claeskens's Alternative Confidence Interval

For large samples, Hjort and Claeskens (2003) propose an alternative confidence interval after model selection based on model averaging approach. The intervals are denoted as

$$\begin{aligned} low_n &= \hat{\mu} - \hat{\omega}^t \{D_n - \hat{\delta}(D_n)\} / \sqrt{n} - u\hat{k} / \sqrt{n} \\ up_n &= \hat{\mu} - \hat{\omega}^t \{D_n - \hat{\delta}(D_n)\} / \sqrt{n} + u\hat{k} / \sqrt{n} \end{aligned} \tag{17}$$

where $\hat{\mu}$ is a compromise estimator, and D_n is a function of discrepancy between the full model and null model as explained in (9). $\hat{\delta}(D)$ is viewed as an estimator of δ based upon D, $\hat{\omega}$ is an estimator of ω, function of partial derivatives evaluated at the null model (θ_0, γ_0), $\hat{k}$ is a consistent estimator of k, standard deviation of the full model, and u is a normal quantile. The confidence interval is claimed as asymptotically precise as the intended $1-\alpha$ level. However, Kabaila and Leeb (2006) have a comment on this confidence interval, indicating it "is essentially the same as the standard confidence interval based on overall model…in the sense that corresponding endpoints of the two intervals differ from each other by terms of order $o_p(1/\sqrt{n})$ …" (p. 628) More discussions on this alternative confidence interval can be referred to Remark 4.2 of Kabaila and Leeb (2006).

Kabaila and Leeb's Upper Bound

Leeb and Pötcher (2004; 2005) identify the problems associated with an incorrect view in favor of consistent model selectors, which argues that the selection procedures utilize the standard asymptotic distributions and thus do not have significant impact on the inference (including the coverage of the interval estimates). With the studies on multiple linear regression models, Kabaila and Leeb (2006) analyze the large-sample limit minimal coverage probability of the confidence intervals. Kabaila and Leeb's (2006) major contribution is the derivation of an upper bound for the large-sample limit minimal coverage probability when the model is chosen by conservative model selection procedure, which is denote as (where $c = \sqrt{2}$ for AIC):

$$\Psi_{\alpha}(\rho_p, c) = \inf_{x \in R} \Psi_{\alpha}(x, \rho_p, c) \tag{18}$$

with $\Psi_{\alpha}(x, \rho_p, c)$ defined by

$$\Delta\left(\frac{\rho_p x}{\sqrt{1-\rho_p^2}}, t_{\alpha/2}\right)\Delta(x,c) + \int_{-\alpha/2}^{\alpha/2}\left(1-\Delta\left(\frac{x+\rho_p z}{\sqrt{1-\rho_p^2}}, \frac{c}{\sqrt{1-\rho_p^2}}\right)\right)\phi(z)dz \tag{19}$$

where ρ_p is the correlation coefficient between the first (which is always in the selected model) and $p-th$ component of the least-squares estimator for the parameter vector of interest, c is a finite constant, $x \in R,$ and R is the user-specified subset of $\{0,1\}^p$ with which a model from the set of competing models $M_r (r \in R)$ is chosen by the data driven model selector, $\Delta(a,b)$ denotes $\Phi(a+b) - \Phi(a-b)$ with $\Phi(\cdot)$ being the standard Gaussian cdf, $\phi(\cdot)$ is the corresponding Lebesgue-density, and $t_{\alpha/2}$ is the $1-\alpha/2$ quantile of $\Phi(\cdot)$.

Kabaila and Leeb propose that the upper bound is (1) the first analytical result for the large-sample limit minimal coverage probability when a "conservative" (or "over-consistent") model selection procedure (e.g. AIC) is used; (2) easy to compute numerically, even with large amount of potential models in selection; (3) depends only on model selection procedure and regressor matrix but not other unknown parameters; (4) statistically meaningful even for the simulated relatively small samples. For the conservative model selection procedure, they mathematically prove that the upper bound for the large sample limit minimal coverage probability depends only on ρ_p, the correlation between the independent variables.

Finite Samples

Leeb and Pötcher (2005, p. 122) assert,

"…Regardless of whether a consistent or a conservative model selection procedure is used, the finite-sample distributions of a post-model-selection estimator are typically not uniformly close to the respective (pointwise) asymptotic distributions…The finite-sample distributions of post-model-selection estimators are typically complicated and depend on unknown parameters. Estimation of these finite-sample distributions is 'impossible' (even in large samples)…"

Therefore, due to the non-uniformity to the large sample distributions, the application of the finite sample distributions (for both point and interval estimates) with the asymptotic results is limited. It was not until recent years that some studies appear in the literature. From a theoretical point of view, Pötcher (1995) and Kabaila (1995) have discussions on the construction of valid confidence intervals based on post-model-selection estimators or more general shrinkage estimators. Leeb and Pötcher (2003) obtain limited approximations to the finite-sample distributions with respect to the point estimate parameters. Other examples include Arabatzis, Gregoire, and Reynolds (1989), Chiou and Han (1995), and Han (1998), where the researchers discuss the conditional coverage probability following the rejection of a preliminary test. For coverage rate of confidence regions, Hurvich & Tsai (1990) conduct a Monte Carlo study on linear regression models with selection procedures of AIC and BIC. They consider five nested competing models in linear regression with quite small sample sizes (up to 50). Their findings suggest that the coverage probabilities are almost all significantly smaller than the nominal $1-\alpha$, with AIC performs better than BIC. Based on the results of the simulation study, they recommend splitting the data, using a small portion for the model selection stage and the rest large portion for parameter estimates. Such an approach is similar to the exploratory and confirmatory analysis with different set of data.

Kabaila's Modified Confidence Intervals

Kabaila (1998) proposes modified confidence intervals with the minimum coverage probability $1-\alpha$ for simple variable selection procedures involving two competing models. Given the unrestricted model $Y = X\theta + \varepsilon$, where X is a known $n \times p$ matrix and θ is an unknown vector $[\theta_{1,\dots,}\theta_p]^T$ with p parameters, and the restricted model $Y = C\psi + \varepsilon$, where C is a known $n \times (p-1)$ matrix and ψ is an unknown vector $[\theta_{1,\dots,}\theta_{p-1}]^T$ with the first $(p-1)$ parameters of the unrestricted model, the naïve confidence intervals of the two models (for θ_1)can be denoted as:

$$[\hat{\Theta}_1 - c_{n-p} S\sqrt{v_x(1)}, \hat{\Theta}_1 + c_{n-p} S\sqrt{v_x(1)}] \tag{20}$$

$$[\hat{\Psi}_1 - c_{n-p+1} \tilde{S}\sqrt{v_c(1)}, \hat{\Psi}_1 + c_{n-p+1} \tilde{S}\sqrt{v_c(1)}] \tag{21}$$

where $\hat{\Theta}_1$ and $\hat{\Psi}_1$ are the Least Square estimators of θ_1, S is the estimator of σ, the standard deviation for θ_1, v_a is the $(1,1)^{\text{th}}$ element of the $(A^T A)^{-1}$ with c_{n-p} satisfying $P(|T| \le c_{n-p}) = 1-\alpha$, T_{n-p} being the *t*-distribution with $(n-p)$ degrees of freedom. The idea to propose modified confidence intervals with minimum coverage probability $(=1-\alpha)$ is to first introduce a positive constant q against which the quantity

$|\hat{\Theta}_p| / S\sqrt{v_x(p)}$ is evaluated. If the quantity is equal to or larger than q, the inferences will be drawn based on the unrestricted model, ignoring the fact that the-model has been selected by a data-driven procedure, and the confidence interval of (20) applies. If the quantity is smaller than q, the inferences will be drawn based on the restricted model, and the modified confidence interval is

$$[\hat{\Psi}_1 - c\tilde{S}\sqrt{v_c(1)}, \hat{\Psi}_1 + c\tilde{S}\sqrt{v_c(1)}] \tag{22}$$

where c is usually chosen as smaller than c_{n-p+1} subject to the modified interval with minimum coverage probability $1-\alpha$. It is shown that this interval is preferable given that θ_p / σ is small, which, while limited, builds a foundation for the study on typically more complicated situations that involve selection among more models.

<u>Kabaila's Monte Carlo Simulation Estimator</u>

Kabaila (2006) further proposes that in certain conditions, techniques such as large sample results, upper bounds and Monte Carlo estimates can be used to accommodate parameter estimates. He derives a Monte Carlo simulation estimator of the coverage probability of naïve confidence intervals conditional on variance reduction. By applying the "leaps and bounds" algorithm (Furnival & Wilson, 1974), the Monte Carlo methodology involves implementation of limited search through the parameter space for a relatively small coverage probability, and comparisons between the results and the nominal $1-\alpha$. It is claimed that the new Monte Carlo method increases the efficiency due to the conditioning on variance reduction (Kabaila, 2006; Kabaila and Leeb, 2006). In the application with real life data, Kabaila's results suggest that for the regression parameters, the coverage probability of the naïve confidence interval (.95) is only between .70 and .80, with the results from AIC higher than those from BIC.

Generalized Logistic Regression

For logistic regression models, one relevant domain of studies focus on power-transformation, which is concerned with introducing an extra shape parameter, λ, or multiple parameters, to improve the model fit to data. In the power-transformation family, an early literature goes back to Box and Cox (1964). Peers (1965) conducts a study on confidence points with regard to different models with focus on the Bayesian probability points. More studies on power-transformation appear in Bickel and Doksum (1981), Carroll and Ruppert (1981), Aranda-Ordaz (1981), Guerrero and Johnson (1982), Box and Cox (1982), Hinkley and Runger (1984) and Taylor (1986). For instance, Aranda-Ordaz (1981) proposes two families (asymmetric and symmetric) of functions involving a linear combination of covariates. The probability of an event (success=1) can be denoted as:

$$\Pr(Y=1\mid x)=\begin{cases}1-(1+\lambda e^{\beta' X})^{-1/\lambda}, & \lambda e^{\beta' X}>-1\\ 1, & \text{otherwise}\end{cases}\tag{23}$$

for the asymmetric family and

$$\Pr(Y=1\mid x)=\begin{cases}0, & \frac{1}{2}\lambda(\beta' X)<=-1\\ \dfrac{\left(1+\frac{1}{2}\lambda(\beta' X)\right)^{1/\lambda}}{\left(1+\frac{1}{2}\lambda(\beta' X)\right)^{1/\lambda}+\left(1-\frac{1}{2}\lambda(\beta' X)\right)^{1/\lambda}}, & \frac{1}{2}\lambda(\beta' X)<1\\ 1, & \frac{1}{2}\lambda(\beta' X)>=1\end{cases}\tag{24}$$

for the symmetric family. Analyses performed conditional on model selection (based on λ) generally reduce bias of the parameter estimates due to the improvement in model data fit.

However, other than those from the vector of parameters of interest, there are variations in estimating the additional shape parameters. Consequently the parameter variance is inflated. Researchers call the inflation of variance "cost" (e.g. Taylor, 1988) or trade-off to achieve smaller bias, as the inferences are drawn conditionally on the selected model without considering the fact that the "true" model may be unknown. Taylor (1988) studies the inflation in variance caused by including extra parameter in the binary-response regression model, and assesses how much the variability is being affected. Further related studies include Taylor *et al.* (1996), which argues that inflation in variance associated with adding parameters is "directly proportional to the number of parameters", and Siqueira and Taylor (1999), which addresses relevant problems on treatment effects of the Box-Cox Transformation.

Given the large amount of applications of log-linear models, the research with regard to the actual coverage probability of confidence intervals has appeared to be infrequent in the literature.

Chapter III: Methodology

The purpose of this study is to investigate the actual coverage probability of confidence intervals after model selection with finite samples, unconditional and conditional on selecting the correct or incorrect logistic model. The effect of model structure uncertainty is one of the important factors we want to address. Relevant studies on large sample (e.g. Leeb, 2005; Kabaila 2006; Kabaila & Leeb, 2006) for linear regression models suggest that the upper bound of the coverage probability depends "crucially" (Kabaila & Leeb, 2006) on covariate correlation. We are interested in whether the relationship applies to logistic models. Moreover, a comparison between the conservative (AIC) and consistent (BIC) selection procedures raises concern in at least two aspects: (1) to what extent the model selection procedures differ in selecting the correct model; (2) what is the difference by using these procedures on the coverage probability with regard to the logit coefficient. As a feature of the binary logistic regression, whether the threshold in dichotomizing the outcome response influences the interval coverage is also investigated.

Research Design

The study assumes that the true model is one of the candidate models and that the covariates are continuous and normally distributed. With this assumption, we generate data from the known true models and focus on β_1, the first component in the vector of the logit coefficients. This approach is similar to Hurvich and Tsai's (1990) study on confidence intervals with linear regression models and Leeb's (2004) work in simulating the effects of model selection on the distribution of a linear predictor.

Manipulated Factors

Five factors are incorporated into the design: data-generating model, correlation between covariates, model selection procedure, sample size, and an offset term in defining the binary threshold.

Data-Generating Model

The logistic model in the study is typical and widely applied in the educational framework. The model is based on McArdle and Hamagami's (1994) research on predicting college graduation of freshman student-athletes with their academic performance. In Mcardle and Hamagami's study, two commonly used covariates, ACT/SAT scores and core GPA scores in high school, were chosen for the basic prediction model. The standardized model can be denoted as:

$$Logit(GRADRATE) = \beta_1 ACTSAT + \beta_2 GPACORE \tag{25}$$

The significance test indicates that the covariate *ACTSAT* (x_1) always accounts for a considerable amount of variance with regard to the probability of success in students' graduation rates, and it is included in all the candidate models. Our model specifications mimic the real life situation in that it is not certain whether the covariate *GPACORE* (x_2) should be included in the model. We take $M_8, M_2,$ and M_9 in Table 5 of McArdle and Hamagami (1994) as the basis for our study. For simplicity, we rename the models as $M_1, M_2,$ and M_3. The model of M_1 (which can be viewed as the unrestricted model) contains x_1 and x_2, both predicting to some extent the *Logit (GRADRATE)*. In this study, we specify $\beta_1 = 2$ and $\beta_2 = 1$ as the underlying parameters for the unrestricted model.

Special formulations are addressed in other two models. In M_2, only x_1 predicts the outcome variable ($\beta_1 = 2$ and $\beta_2 = 0$). This is typical when a restricted model is considered. In M_3, the two β coefficients are restricted to be of equal size ($\beta_1 = \beta_2 = 2$). While most studies attend to the typical restriction with decreased order, M_3 is unique as although both covariates are taken into account, it in fact changes the number of parameters and degrees of freedom, which directly affect how the model is selected. To accommodate the study for M_3, we define an academic variable *ACADE* as the sum of the two covariates *ACTSAT* and *GPACORE*. The model of M_3 can then be specified as:

$$\text{Prob}(GRADRATE = 1) = \frac{\exp(\beta_1 ACADE)}{1 + \exp(\beta_1 ACADE)}. \tag{26}$$

In summary, the specifications of the data-generating model are:

$$M_1 : LOGIT(GRADURATE) = 2 * ACTSAT + 1 * GPACORE$$
$$M_2 : LOGIT(GRADURATE) = 2 * ACTSAT + 0 * GPACORE$$
$$M_3 : LOGIT(GRADURATE) = 2 * ACTSAT + 2 * GPACORE = 2 * ACADE$$

In addition, to investigate the effects of the binary threshold an offset term is imposed on the models, which is specified in the subsection of Offset.

<u>Covariate Correlation</u>

As Kabaila and Leeb (2006) suggest, in linear regression models covariate correlation plays a key role in determining the upper bound of the coverage probability. Generally speaking, the higher the correlation, the lower the coverage probability. Five levels of covariate correlation, 0.1, 0.3, 0.5, 0.7 and 0.9 are specified to cover a wide range of covariate relationship in the context of logistic regression.

For the restricted models, this manipulation is of particular purpose. In M_3, the covariates are restricted to predict the outcome at the same level, no matter what the covariate correlation is. The manipulation can generate some interesting results (e.g. when the covariate correlation is low). In M_2, although the variable GPACORE does not predict the outcome, it correlates with ACTSAT and, in fact, serves as a suppressor variable. A suppressor may increase the total variance explained by the model when it only has negligible correlation with the dependent variable. Horst (1966) indicates:

> "A suppressor variable may be defined as those predictor variables which do not measure variance in the criterion measures, but which do measure some of the variance in the predictor measures which is not found in the criterion measure. They measure invalid variance in the predictor measures and serve to suppress this invalid variance." (p. 363)

Conger (1974) provides another definition of suppressor variables as "...a variable which increases the predictive validity of another variable (or set of variables) by its inclusion in a regression equation." (p. 36-37) As the manipulation covers a wide range, we consider various conditions when the suppressor is involved in the model.

Model Selection Procedure

While model building methods such as stepwise variable selection are available for logistic regression, this study analyzes the effects of automatic model selection procedures. Being the most popular automatic selectors, AIC and BIC are chosen for the study. Both model selectors penalize the log-likelihood by the number of covariates in the model. As suggested by the pervious studies (e.g. Leeb and Pötcher, 2003; Kabaila, 2005), although BIC involves the sample size in its penalty term and serves as a consistent estimator, with finite samples the distribution of post-model-selection estimators does not uniformly converge to the asymptotic distribution. Therefore, to understand how the coverage deviates with the finite framework is one objective of the design.

For the binary logistic model, the number of levels of the outcome variable is fixed to be 2, and the analysis is focused on the results of:

(1) AIC [Akaike information criterion; Akaike (1973)] given by:

$$AIC = -2\ln(L) + 2(p+1) \tag{27}$$

where *L* is the maximum likelihood for the model, and p is the number of covariates estimated in the selected model.

(2) BIC (or SC) [Bayesian information criterion; Akaike (1978) and Schwarz (1978)] given by

$$BIC = -2\ln(L) + \times\ln(n)(p+1) \tag{28}$$

where n is the sample size, and *L* and *p* are defined as those in AIC.

With different penalty terms, automatic model selection procedures can be roughly classified as either "conservative" (e.g. AIC), or "consistent" (e.g. BIC) (Leeb, 2005), where "consistent" indicates "the probability of choosing the most parsimonious correct model converges to one as sample size increases" and "conservative" suggests "not consistent but such that the probability of choosing only correct models converges to one" (Kabaila & Leeb, 2006, p. 620). By comparing AIC and BIC in terms of their impact on the coverage probability, we investigate the effects of conservative and consistent model selection procedures. As suggested by the generalized final prediction error (FPE) criterion that minimizes the residual sum of squares (Shibata, 1984), the essential difference between these two selection procedures is the parameter λ (Zhang, 1992). Specifically, while the parameter is 2 for AIC, it equals $\ln(n)$ for BIC, with which sample size is explicitly considered in the penalty.

Sample Size

Sample sizes of 50, 100, 200, 350 and 500 are selected based upon the following considerations: (1) for practical purposes, sample sizes of 50 and 100 are typical in many empirical logistic regression studies; (2) sample sizes of 200, 350 and 500 are considered to investigate, with adequate but not too much distance, robustness of coverage rate when sample increases; (3) Although the model selection procedures that involve prediction with categorical data are often engaged with large sample size in the sector of social science (e.g. survey, large scale standardized testing, etc.), a pre-study suggests that there is no obviously significant difference on the coverage probability when the sample size goes up to 1000 or 2000. Meanwhile, Kabaila and Leeb (2006) suggest that the upper bound of the coverage probability with the conservative selection procedure is not asymptotically dependent upon the sample size. The range from 50 to 500 is adequate to approximate the effects of sample size with regard to the finite sample conditions.

Offset

In applied binary-response logistic regression, it is most common to define the binary threshold with arbitrarily specified probability of 0.5. We consider possible variations in the dichotomization process so that the threshold can deviate from 0.5, which are observed from time to time in the statistical applications. The approach we take is to impose a fixed offset term on the standardized model, which specifies the *a priori* defined component of binary threshold. The offset causes fitting of the model as

$$Logit(GRADRATE = 1) = \beta' X + W \tag{29}$$

where *W* is the offset term without estimated coefficient. As we apply standardized model in the study ($\beta_0 = 0$), *W*, while distinct from the β parameters, plays a role as an intercept. The *Logit* with specified probabilities is selected as the offset. When all the components in β are zero, *W* serves as a constant (in this study *W* is a $n \times 1$ vector) that enables the threshold to be equal to the specified π, the binary threshold. As Table 1 illustrates, when $\pi = 0.5$, the logit is 0. As the offset is symmetric for the rest probabilities, we select three offset values to examine the effect of binary threshold: 0, 0.84729786 and 2.197224577, which correspond to the probabilities of 0.5, 0.7 and 0.9, respectively.

Table1: LOGIT with different probabilities

π	$1-\pi$	$\frac{\pi}{1-\pi}$	$\ln\left(\frac{\pi}{1-\pi}\right)$
0.1	0.9	0.111111	-2.197224577
0.3	0.7	0.428571	-0.84729786
0.5	0.5	1	0
0.7	0.3	2.333333	0.84729786
0.9	0.1	9	2.197224577

The simulation study is a balanced experimental design with the five factors fully crossed. Therefore, there are $3 \times 5 \times 2 \times 5 \times 3 = 450$ design conditions. Within each condition, 1000 random replications are implemented to simulate whether the underlying parameter of interest is covered by the naïve confidence interval that is constructed after model selection and parameter estimates. For each replication, two alpha levels (.01 and .05) are considered in constructing naïve confidence intervals with regard to the logit coefficient of interest.

Data Generation

Generation of the Covariates

SAS 9.1 (SAS Institute, Inc. 2005) was used for data generation and simulation. For the covariates (*ACTSAT* and *GPACORE*) in the study, the matrix decomposition procedure (Kaiser and Dickman, 1962) is applied to generate data from a bivariate normal distribution. The procedure imposes a specified correlation matrix on a set of otherwise uncorrelated random normal variables, as if the data were sampled from a population with specified population correlations as represented by the imposed correlations matrix (Fan *et al*., 2001). With the random generator of RANNOR, we first specify a correlation matrix, which can be varied in accordance with the design conditions, and generate two random normal uncorrelated variables with the specified sample size *N*. The means and standard deviations (SDs) of the variables are defined as 0 and 1, approximating McArdle and Hamagami's (1994) national *Z* scores. The variables are then transformed, with the pre-specified correlation pattern matrix, into correlated covariates.

Generation of the Outcome variable

With the pre-specified sample size *N*, the response of prediction is then generated from the assumed true model with the logistic functions (while β_1 is equal to 2, β_2 is equal to 1 for M_1, 0 for M_2 and 2 for M_3):

$$PROB(GRADRATE = 1) = \frac{\exp(\beta_1 ACTSAT + \beta_2 GPACORE + W)}{1 + \exp(\beta_1 ACTSAT + \beta_2 GPACORE + W)} \quad (30)$$

The variable *PROB* is ranged from 0 to 1. By applying the generator RANUNI, a random uniform variable *Y* is then generated. By comparing the value of *PROB* and *Y* for each observation, the binary variable *GRADRATE* is obtained, whose correlations with the covariates are driven by the model from which the data is generated.

Model Selection, Parameter Estimate and Interval Construction

The three models were used to fit the same data generated in the iteration, after which the model selection procedure was conducted. Conditional on the variation of the specified factors, the “best” model was selected with minimal AIC and BIC, which may or may not correspond to the true model. For the model selected, the PROC LOGISTIC statement, which applies Fisher’s scoring method in ML algorithm, was executed to estimate the parameters. The 95% and 99% naïve confidence intervals were then obtained and compared with the underlying parameter of the true model. Macros were specified for iterations in data generation, model selection and parameter estimates. The PROC LOGISTIC results of confidence intervals were contained in SAS BASE, and exported when the last iteration was converged. Table 2 summarizes the manipulated factors in data generation.

Table 2: Factor manipulation and data generation

FACTORS	NO. OF CONDITIONS	CONDITION SPECIFICATION	WITHIN MACRO?
Data-generating Model	3	See *	N
Covariate Correlation	5	.1, .3, .5, .7, .9	N
Selection Criteria	2	AIC, BIC	Y
Sample size	5	50, 100, 200, 350, 500	Y
Offset	3	0, 0.8473, 2.1972	N
Total	**450**		
Alpha Level	2	.05, .01	Y

***TRUE MODEL: LOGIT(GRADRATE)=Beta1*ACTSAT+Beta2*GPACORE+W**

M_1: LOGIT(GRADRATE)=2*ACTSAT+1*GPACORE+W

M_2: LOGIT(GRADRATE)=2*ACTSAT+0*GPACORE+W

M_3: LOGIT(GRADRATE)=2*ACTSAT+2*GPACORE=2*ACADE+W

Note: (1) The parameter of interest is BETA1 (ACTSAT);
(2) The correlations between GRADRATE and the covariates are determined by the true model;
(3) 1000 iterations are conducted for each design condition.

Data Analysis

There are basically two major purposes on analyzing the data: (1) with varied model structure, how the automatic selectors perform in choosing the correct (and incorrect) model; (2) what is the actual coverage probability of confidence intervals after model selection, and how it is affected by various factors including (1).

While the selectors' performance is investigated with the success rate on selecting correct and incorrect models in the iterative process of model selections, we focus on the coverage probability of the confidence intervals around $\hat{\beta}_{1\_population}$, the logit coefficient for *ACTSAT* (X_1), which is present in all three models. The coverage of the confidence interval is defined as

$$\exp(\hat{\beta}_{1\_population}) \in \exp[\hat{\beta}_{1\_sample} \pm z_{1-\alpha/2} \times SE(\hat{\beta}_{1\_sample})] \tag{31}$$

where $\hat{\beta}_{1\_population} = 2$, as specified in the true models.

For the interval coverage, we investigate both the unconditional and the conditional coverage probability. The unconditional coverage is defined as the coverage of confidence intervals after model selection, without considering if a true model is selected. It corresponds to the overall coverage that is reported in most post-model-selection interval estimates, which do not take model uncertainty into account. The conditional coverage is conditional on selecting a correct or incorrect model. From a sampling point of view, the unconditional coverage probability can be viewed as a weighted average of the conditional coverage probability. To determine the coverage probability, the proportion of the naïve confidence intervals that cover the "true" parameter is counted for each of the design conditions.

Unconditional Coverage Probability

First, we obtain the actual coverage probability, which is counted as the proportion of the overall iterations in which the confidence interval covers $\hat{\beta}_{1\_population}$. We are concerned with the deviation rate of the overall coverage probability from the conventional expectations, or how far the actual coverage probability falls below the nominal level, given various conditions crossed and combined.

With the null hypothesis: $H_{0:}$ $Actual\ Coverage\ probability = 1-\alpha$, the deviation is analyzed with z-score, which is calculated based on normal approximation to the Binomial Distribution:

$$Z = \frac{C - n(1-\alpha)}{\sqrt{n(1-\alpha)\alpha}} \tag{32}$$

where C is the count that θ_1 is covered by the confidence interval, n is the number of replicates, and $1-\alpha$ is the nominal coverage probability (95% and 99%, respectively).

Conditional Coverage Probability

With the simulation results, the conditional coverage probability (CCP) is computed as the proportion that the confidence interval covers the parameter when the correct model (*C*) or an incorrect model (*I*) is selected as the best model, or:

$$\begin{aligned} CCP_C &= \frac{(T_{vc} \mid T_{sc})}{T_{sc}} \times 100\%, \text{ and} \\ CCP_I &= \frac{(T_{vi} \mid T_{si})}{T_{si}} \times 100\% \end{aligned} \tag{33}$$

where T_s stands for count on selecting the correct or incorrect model, T_v stands for count on the parameter is covered, with condition either the correct or the incorrect model is selected. In this study, the coverage probability conditional on selecting each of the two incorrect models is analyzed.

Generally speaking, when the true model is selected, the conditional coverage probability is reasonably close to $1-\alpha$ (Leeb and Pötcher, 2003; Hurvich and Tsai, 1990). However, when the model is misspecified, the conditional coverage probability is prone to fall far below the nominal level, and thus imposes major adverse effects on the overall coverage probability. Therefore, the conditional coverage probability differentiates, to certain extent, the impact of “successful” and “unsuccessful” model selections.

Chapter IV: Results

The simulation results are grouped into three major categories: the model selector performance, the unconditional coverage probability, and the conditional coverage probability.

For the model selector performance, we count the frequency that the selection procedures chose the correct and incorrect model, and compare the difference between AIC and BIC. As indicated in Chapter III, the unconditional coverage probability is standardized with the z-score, with which we gauge the actual coverage rate without considering whether a correct or incorrect model is selected. The results are directly associated with the validity of those usually reported on naïve confidence intervals. Based on selecting the correct or incorrect model, we investigate further the effects on the conditional coverage probability, with the correct model being either unrestricted or restricted.

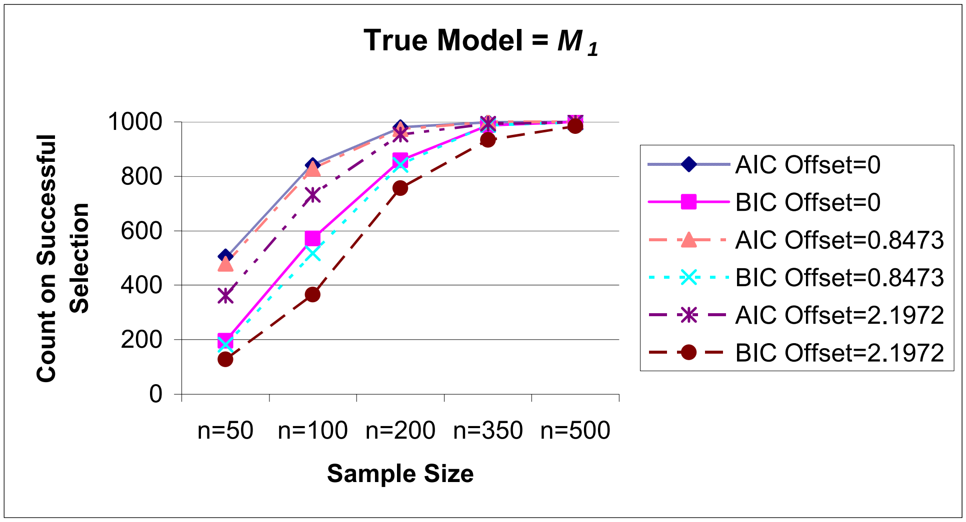
Model Selector Performance

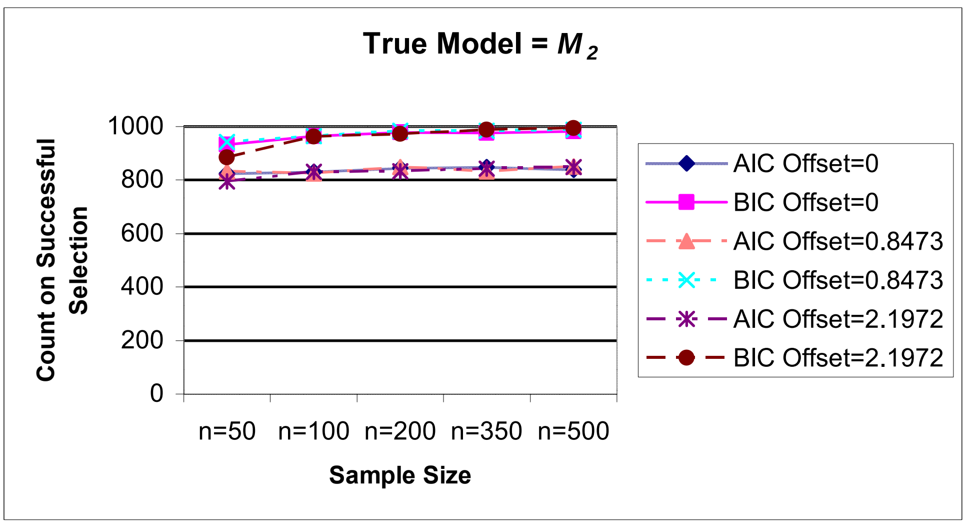
We first examine how the model selectors of AIC and BIC perform in selecting the correct model. As covariate correlation is one of the most important factors that may affect the post-model-selection coverage probability, we compare AIC and BIC through the main dimension of covariate correlation. When the true model is M_1 (the unrestricted model), the difference between AIC and BIC is most evident, and AIC performs constantly better than BIC. When the true model is M_2 and M_3 (restricted), BIC enjoys a higher selection success rate. While the bold values in Table 12 to Table 56 illustrate the numerical results of count on selecting the correct model under various conditions, Figure 2 to Figure 6 convey aggregated graphic information to facilitate the comparison.

Corr=0.1

Figure 2 illustrates the model selectors' performance in selecting the correct model when the covariate correlation is 0.1. When the true model is M_1, sample size plays a key role in determining the selection success rate for both AIC and BIC. With the sample size of $N = 50$ and zero offset, AIC selects the correct model for 50.6% of the time, and BIC chooses the correct model for 19.6% of the time. The results suggest a considerable difference between the two selectors out of the 1000 iterations. When the sample size goes up, the selection success rate for both AIC and BIC increases, and the peer difference gets smaller. The biggest improvement occurs between N=50 and N=200, where at N=200 the selection success rate reaches 98% for AIC and 85.8% for BIC. Among the false selections, the model selectors favor M_3. Both selectors achieve 100% success rate when the sample size increases to 500.

When the true model is one of the restricted models (M_2 or M_3), the role of sample size is not so substantial. For instance, when N=50, the selection success rate is 82.4% for AIC and 93.3% for BIC (M_2), and 81.3% for AIC and 93.1% for BIC (M_3), which are pretty high compared to the corresponding cases when M_1 is the correct model. In either of the restricted models, both selectors do not achieve 100% success rate when N=500, with that of AIC being around 84% and BIC being around 98%. Among the false selections, both selectors favor M_1, no matter whether M_2 or M_3 is true.





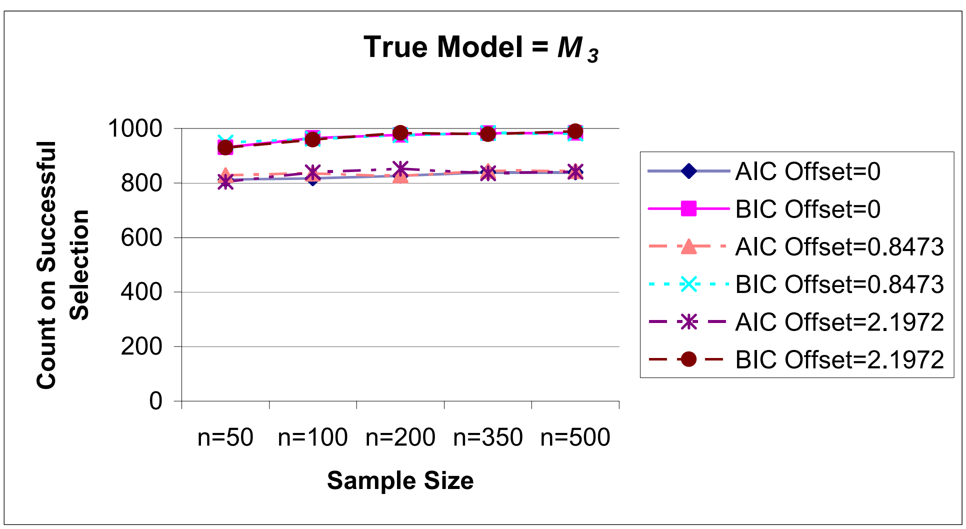


Figure 2: Model selector performance when covariate correlation is 0.1

The effect of the offset is most noteworthy when the correct model is M_1, where the more skewed the dichotomization, the lower the success rate. The difference is not substantial when one of the restricted models is true.

Corr=0.3

Figure3 illustrates the model selector performance in selecting the correct model when the covariate correlation is 0.3. When the true model is M_1, The role of sample size is similar to that when the covariate correlation is 0.1. For instance, with $\mathrm{N} = 50$ and zero offset, the success rate for selecting the correct model is 34.2% for AIC and 11.1% for BIC. When the sample size gets larger, the selection success rate increases substantially, and the difference between AIC and BIC is smaller. When $\mathrm{N} = 500$, the success rate for both of the selectors is close to 100%, with BIC a little bit lower.

When the true model is one of the restricted models (M_2 or M_3), BIC performs consistently better, and the role of sample size is minor. Taking M_2 as an example, the range of selection success rate is between 79% and 85.1% for AIC and between 90.4% and 99.2% for BIC (for the case of zero-offset). BIC gains relatively higher selection success rate.

Similar to the corresponding cases when $\mathrm{Corr} = 0.1$, the effect of the response dichotomization is observed only when the true model is M_1, where a larger offset term indicates a lower success rate. When the sample size is small, it is noticeable that the selection success goes down compared to the results of 0.1. The gap is gradually closed as the sample size approaches 500.

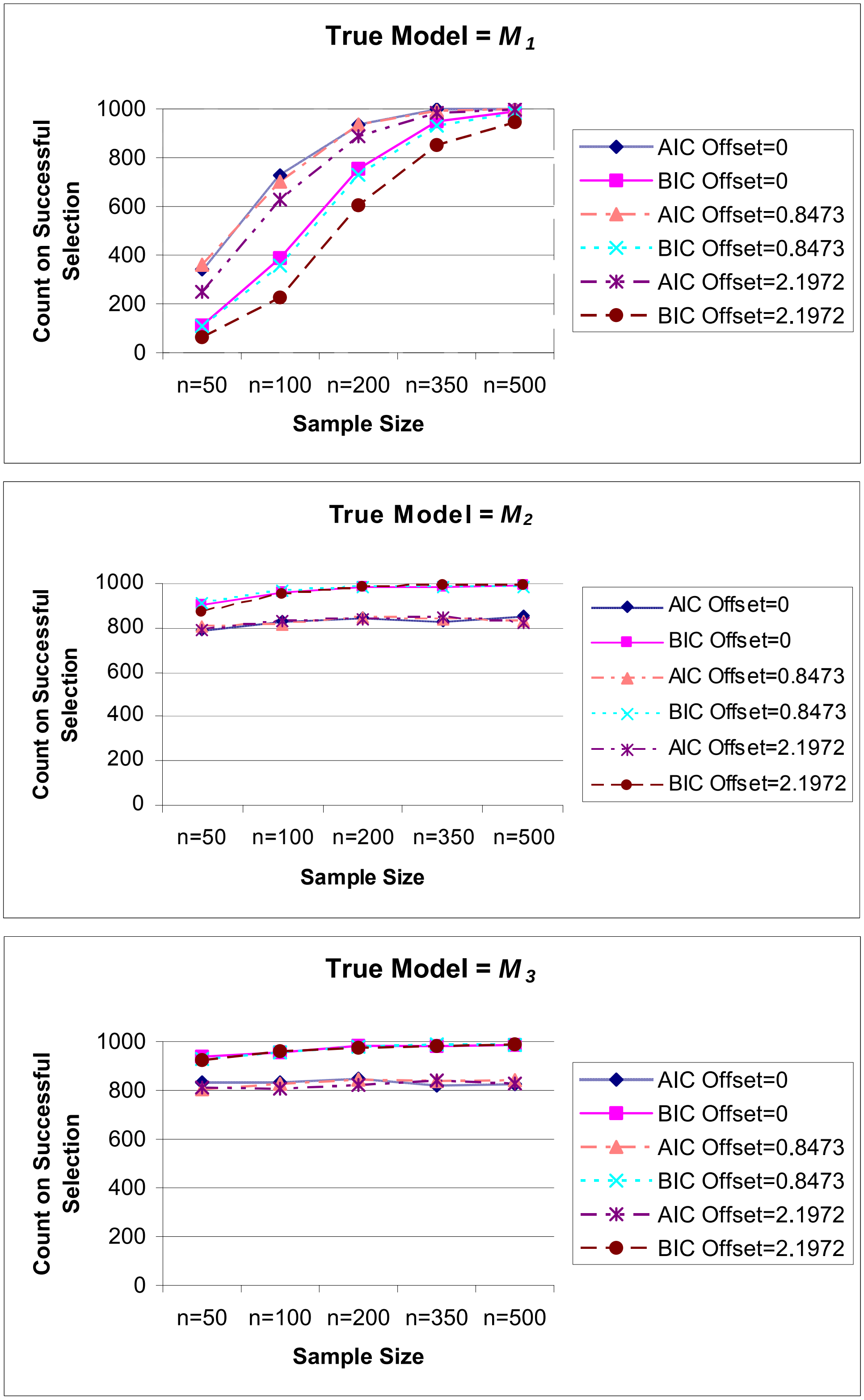


Figure 3: Model selector performance when covariate correlation is 0.3

Corr=0.5

Figure 4 illustrates the model selector performance in selecting the correct model when the covariate correlation is 0.5. As indicated by Figure 4, the success rate in model selection is lower when the correlation gets higher. Compared to that when the correlation is 0.3, the success rate heads down further, with the results of BIC in small sample sizes most impressive. When the true model is M_1, with $\mathrm{N} = 50$ and zero offset, the rate is 24.6% for AIC and 3.8% for BIC. When $\mathrm{N} = 500$, the success rate for AIC is close to 100%, and that for BIC is only 91.6%. The results suggest that AIC outperforms BIC in any sample size.

When the true model is one of the restricted models (M_2 or M_3), again BIC performs consistently better, and the role of sample size is comparatively small. When the true model is M_2 (with zero-offset), the range of success rate is between 82% and 84.8% for AIC and between 88.7% and 99.3% for BIC, for which BIC gains, though small, relatively more benefits with a larger sample size.

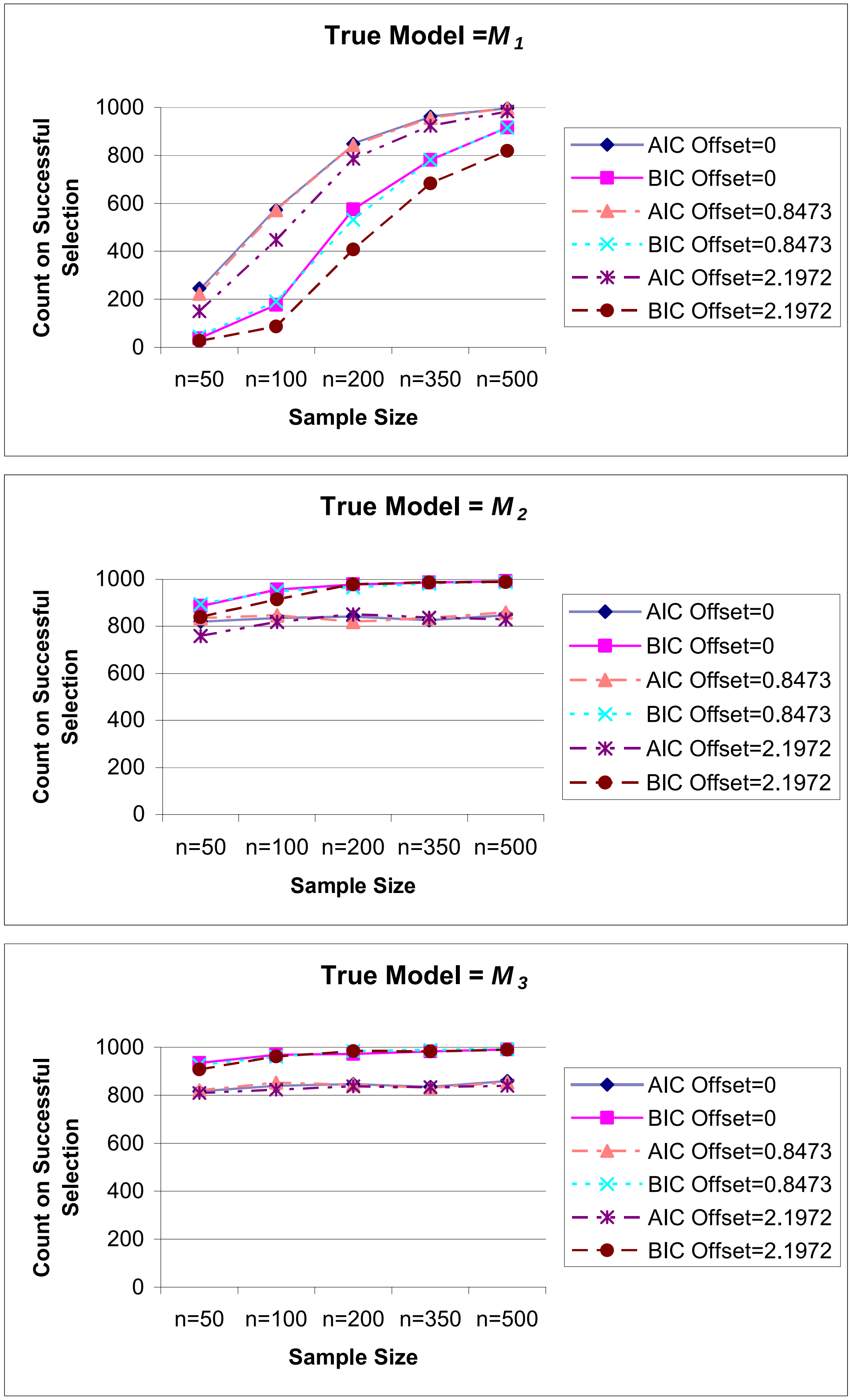


Figure 4: Model selector performance when covariate correlation is 0.5

Corr=0.7

Figure 5 illustrates the model selector performance in selecting the correct model when the covariate correlation is 0.7. One most significant change for the zero offset condition occurs when the true model is M_1 with $\mathrm{N} = 50$, for which the selection success rate falls to 8.3% for AIC and 1.2% for BIC, respectively, which are surprisingly low. For AIC the selection success rate increases when the sample size gets larger, achieving 94% at $\mathrm{N} = 500$. The selection success for BIC, however, is severely affected by the high covariate correlation with the maximum rate of 69.5%. When the true model is one of the restricted models (M_2 or M_3), again BIC performs consistently better, the role of sample size is minor, and the success rate is smaller compared to those with lower correlation.

Among the incorrect selections, both AIC and BIC select M_3 more frequently than M_2 when the true model is M_1. When M_2 is true, with small sample size the selectors favor M_3, and with large sample size the selectors favor M_1.

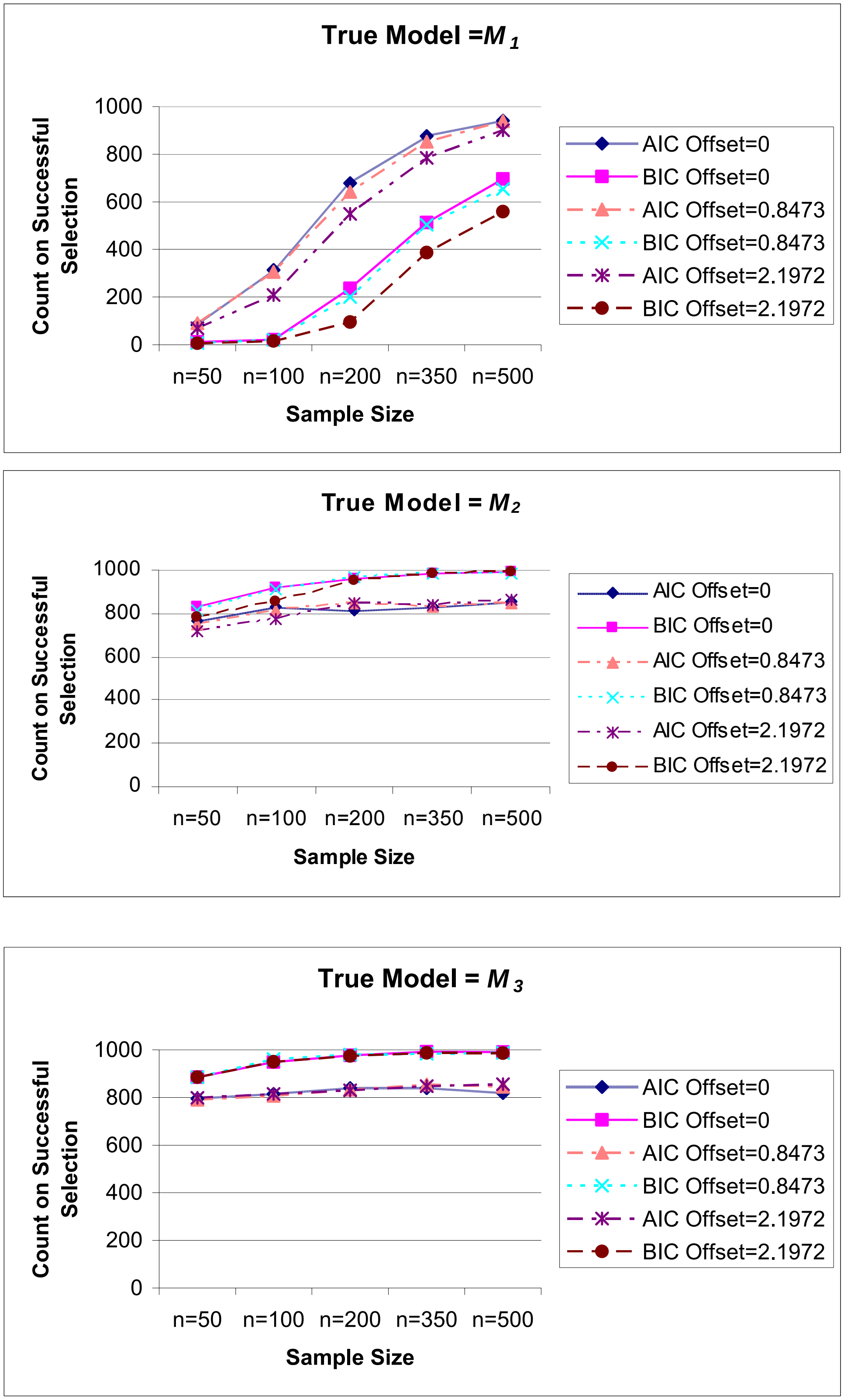
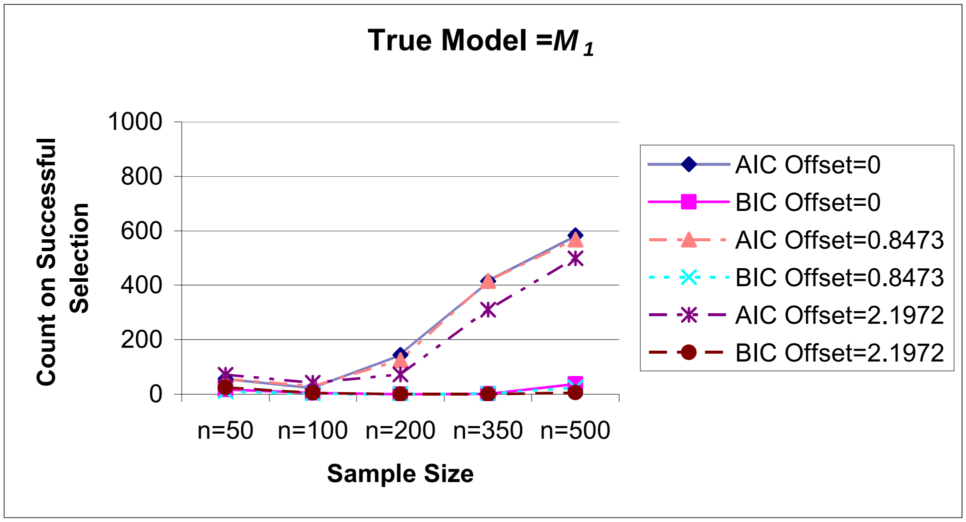


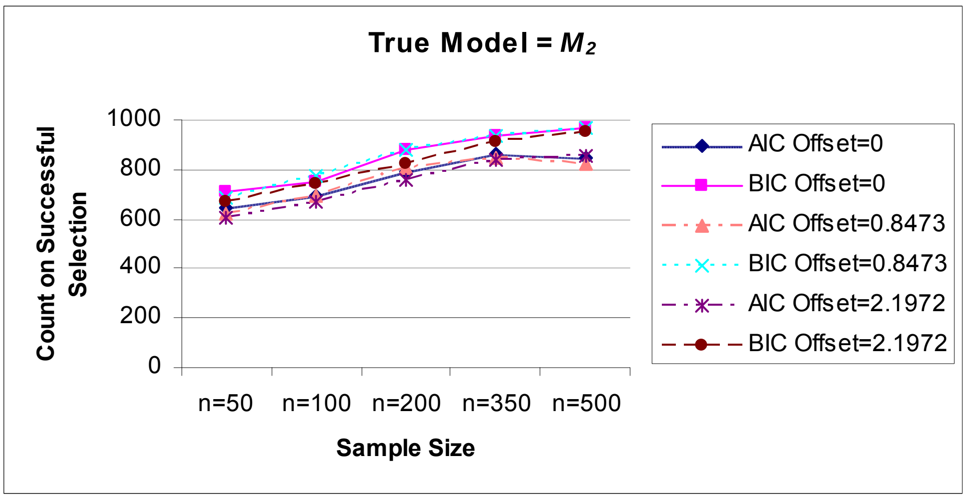
Figure 5: Model selector performance when covariate correlation is 0.7

Corr=0.9

Figure 6 illustrates the model selector performance in selecting the correct model when the covariate correlation is 0.9, the extreme value along this dimension. The selection success in model selection gets even lower. When the true model is M_1, for AIC the rate falls below 20% for any sample size that is smaller or equal to 200. The results for BIC, astonishingly, are close to 0%, no matter what sample size is applied. Therefore, to increase sample size seems not to improve the selection success rate at all, as most of the time BIC favors the incorrect model. The maximum values for AIC and BIC are 58.3% and 3.8%, respectively. In contrast, when the true model is one of the restricted models (M_2 or M_3), again BIC performs consistently better and achieves very high success rate (96.5% and 98.9%) when $\mathrm{N} = 500$ and $\mathrm{Offset} = 0$.

Among the incorrect selections, both AIC and BIC select M_3 more frequently than M_2 when the true model is M_1. When true model is M_2, AIC favors M_3 with small sample size and M_1 with large sample size, and BIC always selects M_3 more frequently. When M_3 is true, AIC favors M_1 (except for $\mathrm{N} = 50$), and BIC selects M_2 more frequently when sample size is small.





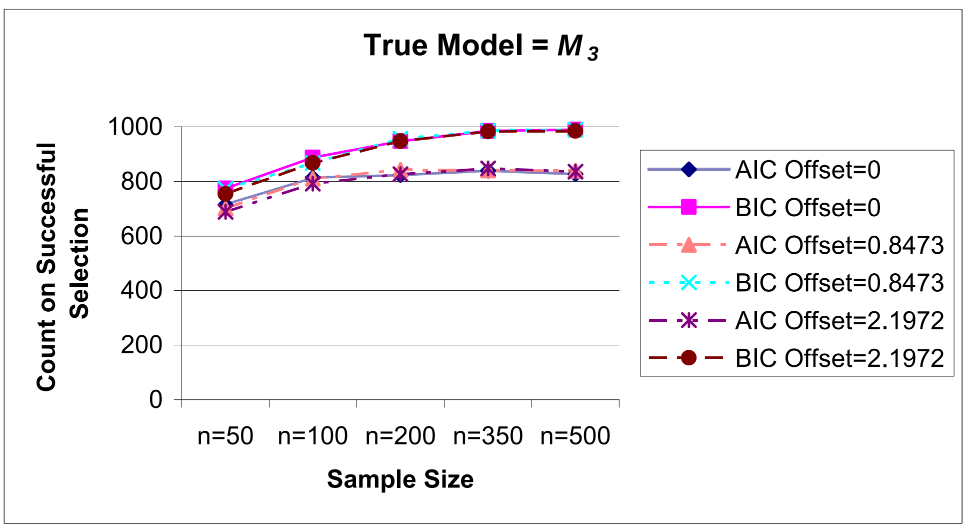


Figure 6: Model selector performance when covariate correlation is 0.9

Summary

Other than the model selector itself, we notice that three out of four manipulated factors have important influence on the success or failure in model selection. First, the variation of model structure itself is noteworthy: while AIC performs consistently better with the unrestricted model, BIC outperforms AIC with the restricted models. Second, high covariate correlation imposes consistently negative effects on the selection success rate, the degree of which is related to the true model structure. Third, while a larger sample size is generally helpful to improve the selection success, such improvement is subject to the constraints of covariate correlation and model structure. While the way to dichotomize the response probability has limited effect, such effect is observed when the true model is unrestricted, where a larger offset term suggests a lower selection success rate.

Unconditional Coverage Probability

Generally speaking, the actual unconditional coverage probabilities are below the nominal level, indicating that the model selection procedure imposes remarkable impact on the coverage probability of the confidence intervals.

Correct Model = M_1 ***(Unrestricted)***

When the correct model is the unrestricted model, the unconditional coverage probability is severely compromised. Table 3 to Table 5 (in Appendix A) illustrate the coverage probabilities of the naïve confidence intervals when the nominal levels are .95 and .99, and the corresponding z-scores when M_1 is true. Among the three tables, Table

3 depicts the results for $\text{offset} = 0$, Table 4 shows the results for $\text{offset} = 0.8473$, and Table 5 demonstrates the results for $\text{offset} = 2.1972$. Meanwhile, Figure 7 and Figure 8 present the coverage probability and the z-score when the symmetric dichotomization is considered, which correspond to the results in Table 3.

1- α **=0.95**

AIC

It is apparent that the coverage probability and the covariate correlation are negatively correlated. Other than a couple of exceptions (e.g. when $\text{Corr} = .3$ and $\text{N} = 200$), the higher the covariate correlation, the lower the coverage probability. The change of the z-score indicates that the coverage probability varies slightly when the correlation is between 0.1 and 0.3, decreases at a moderate rate when the correlation increases from 0.3 to 0.7, and drops sharply thereafter. The extreme z-score (-69.791) is observed in Table 4, where the coverage probability is as low as 0.469 with a sample size of 200 and covariate correlation of 0.9.

When the sample size is small (up to 200), all the z-scores are beyond negative three SDs (except for the case in Table 4, where the sample size is 200 and the correlation is 0.1), and most are beyond negative five SDs, suggesting that the null hypothesis is rejected at any conventional level, and that the actual coverage probability is significantly different from the nominal rate. When the sample size is equal to 350 and 500, the coverage probability is close to 0.95 with low covariate correlations (0.1 and 0.3), and gets significantly lower when the covariate correlation is larger than 0.5.

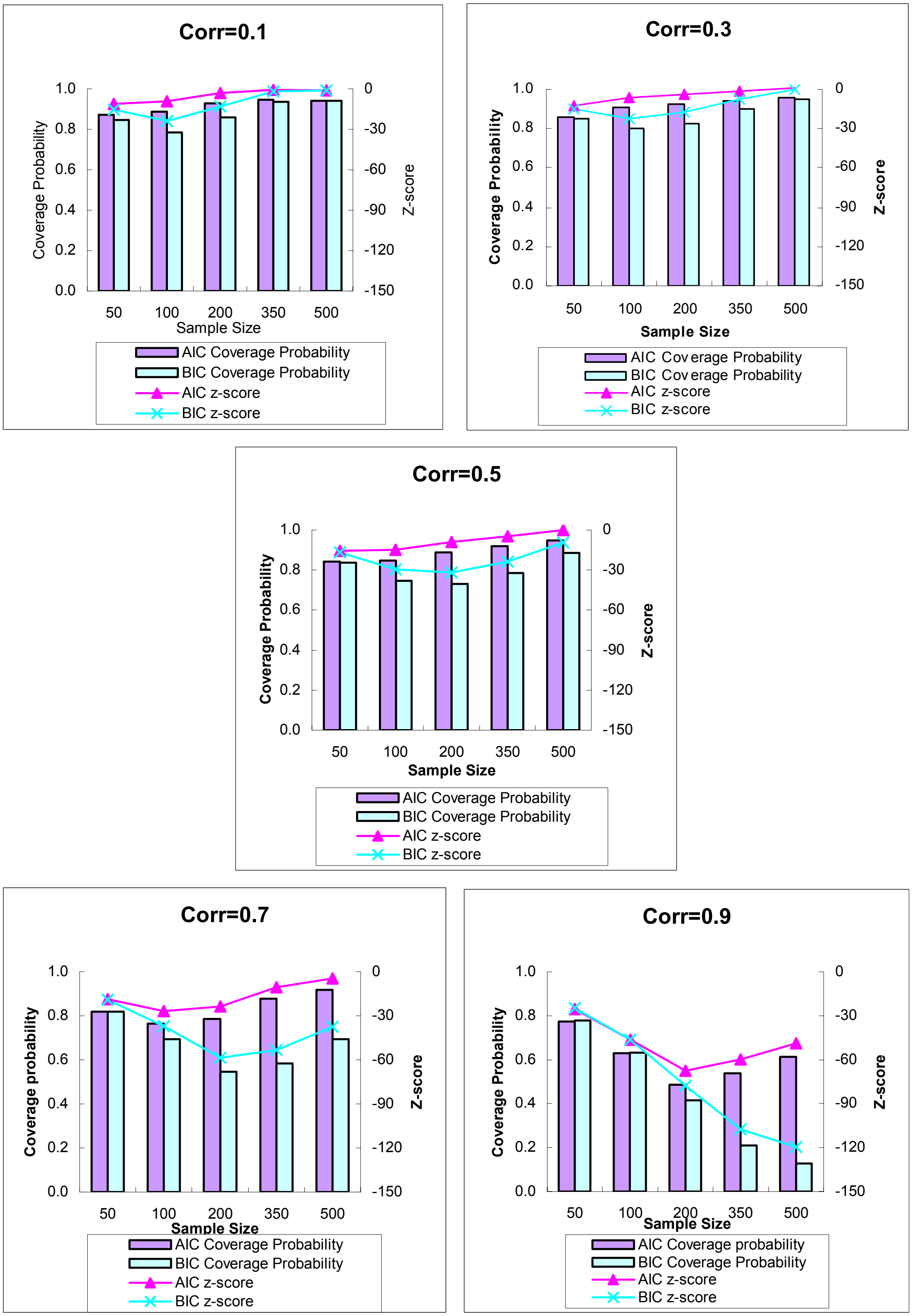


Figure 7: 95% Coverage Probability when the correct model is M_1 (Offset=0)

In the tables if we draw a diagonal from the upper left to the lower right, it is clear that all those values above the diagonal are bold (significantly lower than 0.95). It is noticeable, though, that the increase of sample size alone is not necessarily helpful to retain a high coverage probability. As a matter of fact, when the correlation is extremely high (0.9), a larger sample size brings about a lower coverage probability. When the correlation is 0.7, with the increase of the sample size the coverage probability first goes down (with N=100 and 200), and then goes up (when N=350 and 500). Similar situation occurs at $\text{Corr} = 0.3/0.5$ in Table 5 (the coverage probability touches the bottom when the sample size is 100), where a large offset term is imposed and the model applies a highly skewed cut value to dichotomize the response probability.

BIC

While a higher covariate correlation unanimously suggests a lower coverage probability, generally speaking the performance of BIC tends to be worse than AIC, as indicated by more cells with bold z-score, and larger absolute values in the corresponding cells. With a large sample size and high covariate correlation, the discrepancy of the coverage probability between AIC and BIC is particularly evident. The extreme z-score (-122.315) is observed in Table 4, where the coverage probability is as astonishingly low as 0.107 with a sample size of 500 and covariate correlation of 0.9.

It seems that BIC, the consistent model selector, performs better with larger sample size only if the covariate correlation is low. Taking the case of $N{=}500$ in Figure 7 as an example, when the correlation is either 0.1 or 0.3, the coverage probability is very close to the nominal level of 0.95 (in contrast most values in the cells of small sample size are significantly below the nominal level). However, when the covariate correlation goes up, the larger the sample size, the worse the coverage probability. The results suggest that the coverage variability increases with the increase of sample size.

The effect of the dichotomization threshold is more detectable when we examine the results of BIC. Specifically, while in Table 3 we still observe a few cells in which values are close to the nominal level (e.g. Corr=0.1 and N=500), in Table 5 all the z-scores are beyond -3.5.

1-α =0.99

<u>AIC</u>

Similar to that with the nominal level of 0.95, the higher the covariate correlation, the lower the coverage probability. In the mean time, the trend of z-score change remains the same, where the coverage probability varies slightly when the correlation is between 0.1 and 0.3, decreases at a moderate rate when the correlation increases from 0.3 to 0.7, and drops sharply thereafter. The extreme z-score (-110.602) is observed in Table 5, where the coverage probability is as low as 0.642 with a sample size of 350 and covariate correlation of 0.9.

All the z-scores are beyond negative three SDs when the sample size is small (up to 200), and most are beyond negative five SDs. When the sample size is equal to 350, the coverage probability is close to 0.99 with low covariate correlations (0.1 and 0.3), and gets significantly lower when the covariate correlation is larger than 0.5, as those suggested when the nominal coverage probability is 0.95. When the sample size is 500, the coverage probability is close to .99 when covariate correlation goes up to 0.5, and then drops significantly. The pattern that a larger sample size alone does not improve the coverage probability is similar to the case when $1-\alpha=0.95$, and a larger sample size indicates a lower coverage probability when the correlation is 0.9. When the correlation is between 0.5 and 0.7, the coverage probability goes down and then up when the sample size constantly increases, indicating that a smaller sample size of $\mathrm{N}=50$ even results in better coverage than 100 or 200. Such cases are more frequently observed when a none-zero offset is imposed.

BIC

By examining the z-score, the 99% naïve confidence intervals perform even worse. In all the three tables only three cells (out of 75 cells) report values close to the nominal level. The extreme z-value (-234.234) is observed in Table 4, where the coverage probability is as low as 0.253 with a sample size of 500 and covariate correlation of 0.9.

It seems that the asymptotic benefits (brought by a larger sample size) only present provided that at least two conditions are satisfied: (1) the covariate correlation is small enough; (2) certain sample size "limit" is bypassed given (1). For instance, when $\mathrm{N}=50$, the coverage probability is never the worst. If the covariate correlation is 0.1, the "limit" is around 100 (200 in Table 5); if the correlation is 0.3, the "limit" is 200 (100 in Table 3); if the correlation is 0.9, the "limit" is never reached within 500. Moreover, if to bypass the "limit" sooner is viewed desirable to achieve asymptotical consistency, the variations among different tables suggest that some cost has to be paid for the skewed dichotomization: the more skewed, the smaller covariate correlation is required. Such a trend also presents when $1-\alpha=0.95$.

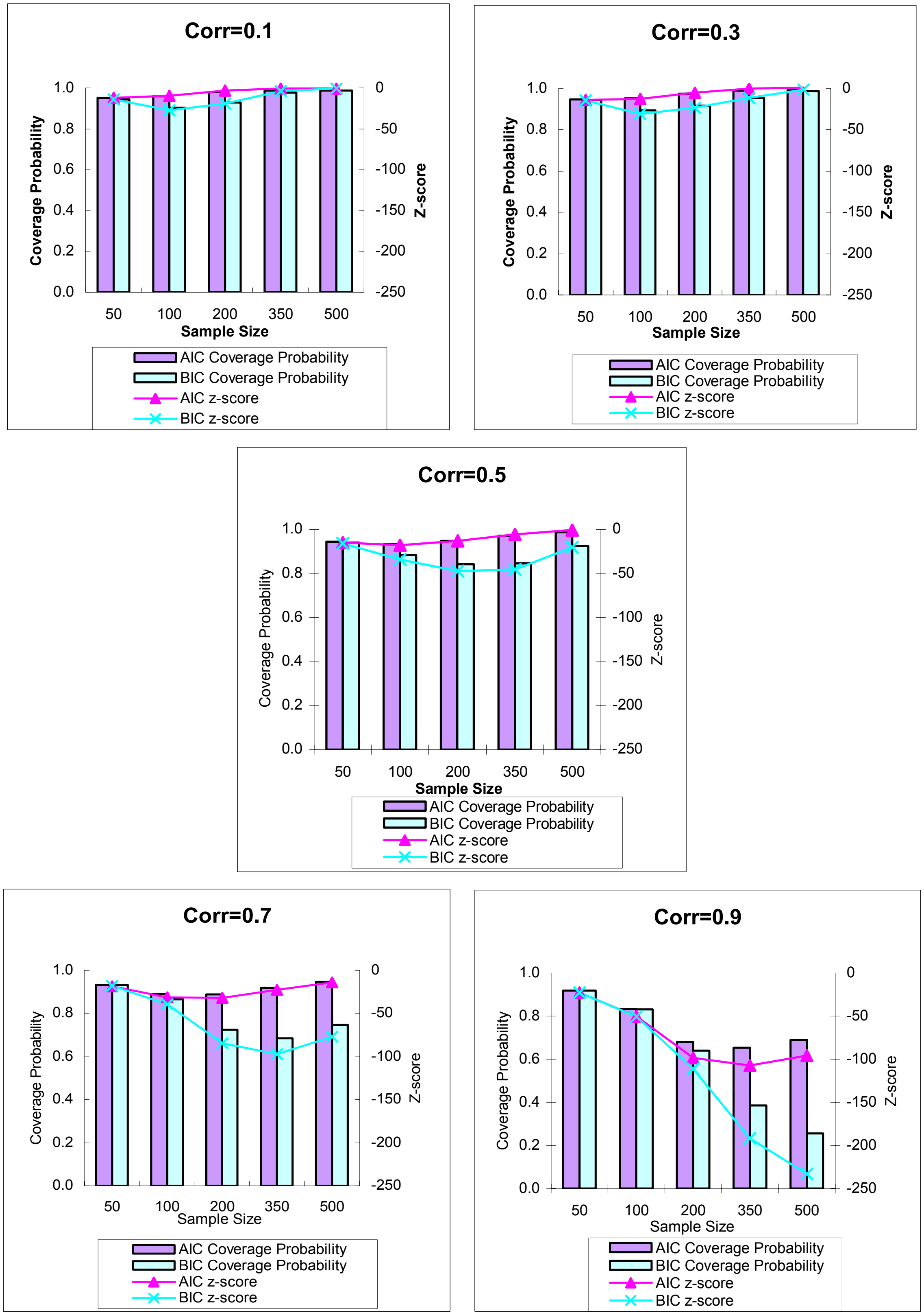


Figure 8: 99% Coverage Probability when the correct model is M_1 (Offset=0)

Correct Model = M_2 ***(Zero-restricted)***

When the correct model is the restricted model of M_2, the confidence intervals perform better in terms of unconditional coverage probability than those when the correct model is M_1. Compared to that with the unrestricted true model, the variation of the coverage probability due to the variation of the covariate correlation is enlarged. Table 6 to Table 8 illustrate the coverage probabilities of the naïve confidence intervals when the nominal levels are .95 and .99, and the corresponding z-scores when M_2 is true. Among the three tables, Table 6 depicts the results for $\text{offset}=0$, Table 7 shows the results for $\text{offset}=0.8473$, and Table 8 demonstrates the results for $\text{offset}=2.1972$. Meanwhile, Figure 9 and Figure 10 present the coverage probability and the z-score when the symmetric dichotomization is considered, which correspond to the results in Table 6.

1- α **=0.95**

<u>AIC</u>

While a lower covariate correlation suggests a higher coverage probability, it is observable that the confidence intervals perform pretty well when the covariate correlation is 0.1 and 0.3. As a matter of fact, when the response probability is symmetrically dichotomized (Figure 9), the coverage probability is close to the nominal level even with a correlation of 0.5 (with one exception when $N=50$). Notice that when the unrestricted model is true, the corresponding results in Figure 7 are well below 0.9.

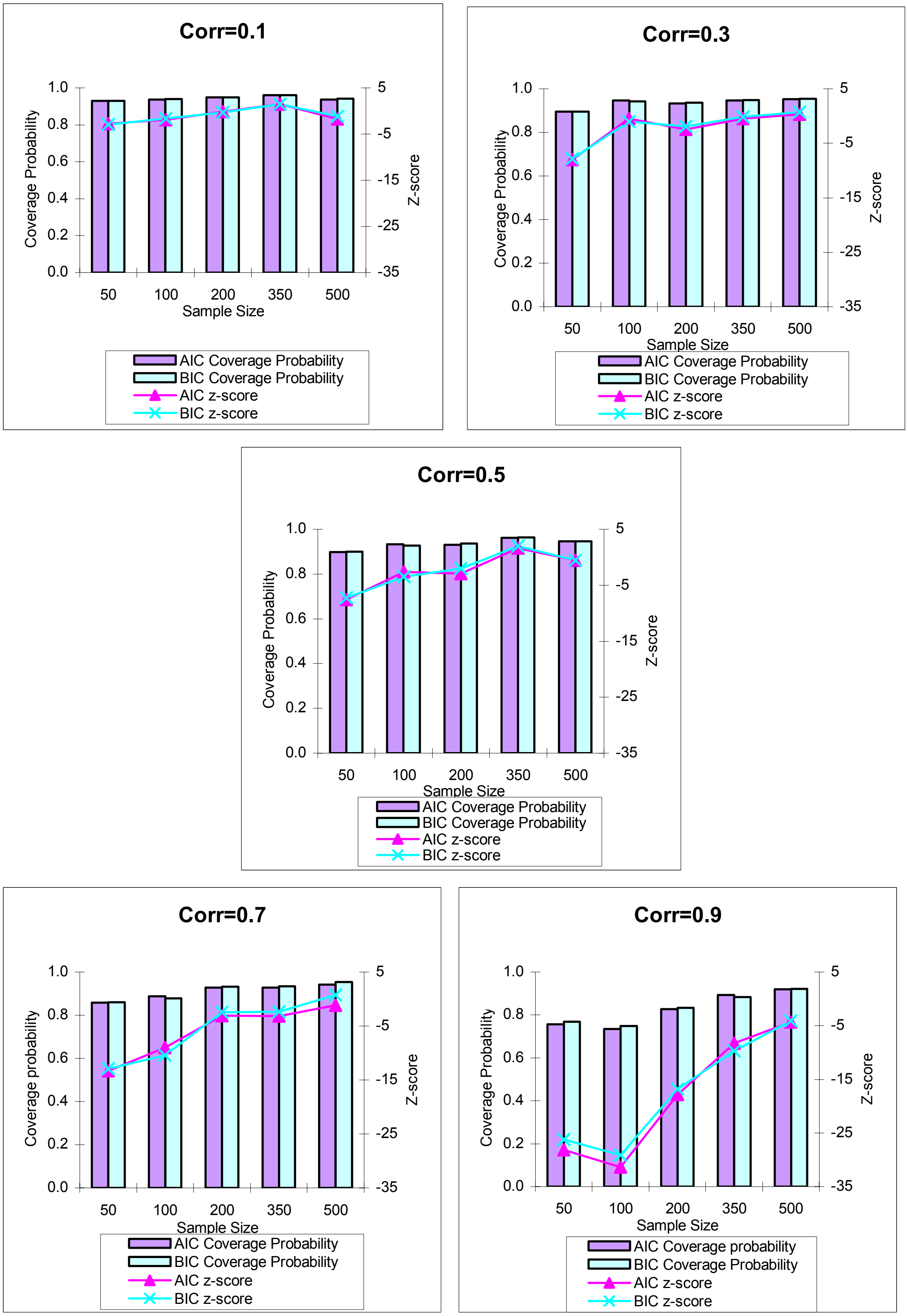


Figure 9: 95% Coverage Probability when the correct model is M_2 (Offset=0)

The change of the z-score indicates that the coverage probability varies slightly when the correlation is between 0.1 and 0.3, decreases at a moderate rate when the correlation increases from 0.3 to 0.7, and drops sharply thereafter. However, when the sample size is 500, such trend is to some extent mitigated. The extreme z-score (-31.341) is observed in Table 6, where the coverage probability is as low as 0.734 with a sample size of 100 and covariate correlation of 0.9.

The z-scores beyond negative three SDs are a little bit less than half of the total observations, with Table 6 (offset=0) being the least and Table 8 (Offset=2.1972) being the most. Therefore, with the true (zero) restricted model it seems that moderate sample size and covariate correlation are adequate to secure the validity of the coverage probability.

BIC

The performance of BIC in terms of coverage probability is significantly different from the scenario with the unrestricted true model. Among the three tables there is no such situation that all the coverage probability is significantly below the nominal level, even when the offset term of 2.1972 is imposed. The bold values indicating beyond negative 3.5 SDs of the nominal level are less than half of the total observations. As a matter of fact, the unconditional coverage probabilities are not so different from those after AIC is applied for the model selection.

The extreme z-score (-29.164) is observed in Table 6, where the coverage probability is as low as 0.749 with a sample size of 100 and covariate correlation of 0.9 (in contrast to the extreme value of 0.107 with a sample size of 500 in Table 4). The extreme coverage probability is even a little bit larger than that from AIC. The aforementioned sample size "limit" does not even exist, or only exists when the covariate correlation is 0.9, where the limit is reached at around 100. In general, the coverage probability demonstrates asymptotic properties matching the consistent model selector.

1-α =0.99

AIC

Similar to the results when the true model is M_1, the absolute values of z-score are inflated compared with those under the nominal level of 0.95. However, the inflation is relatively modest, which can be examined from the absolute values. The extreme z-score (-57.347) is observed in Table 7, where the coverage probability is as low as 0.819 with a sample size of 100 and covariate correlation of 0.9.

BIC

The extreme z-score (-57.208) is observed in Table 6, where the coverage probability is as low as 0.810 with a sample size of 100 and covariate correlation of 0.9. The extreme value is very close to that resulting from AIC, indicating that the effects of model selection procedure are not significantly different. However, Compared with the corresponding extreme value in Table 4 (0.253 with $z = -234.234$), tremendous difference is observed. Therefore, in the context of BIC, the effect of the true model is remarkable.

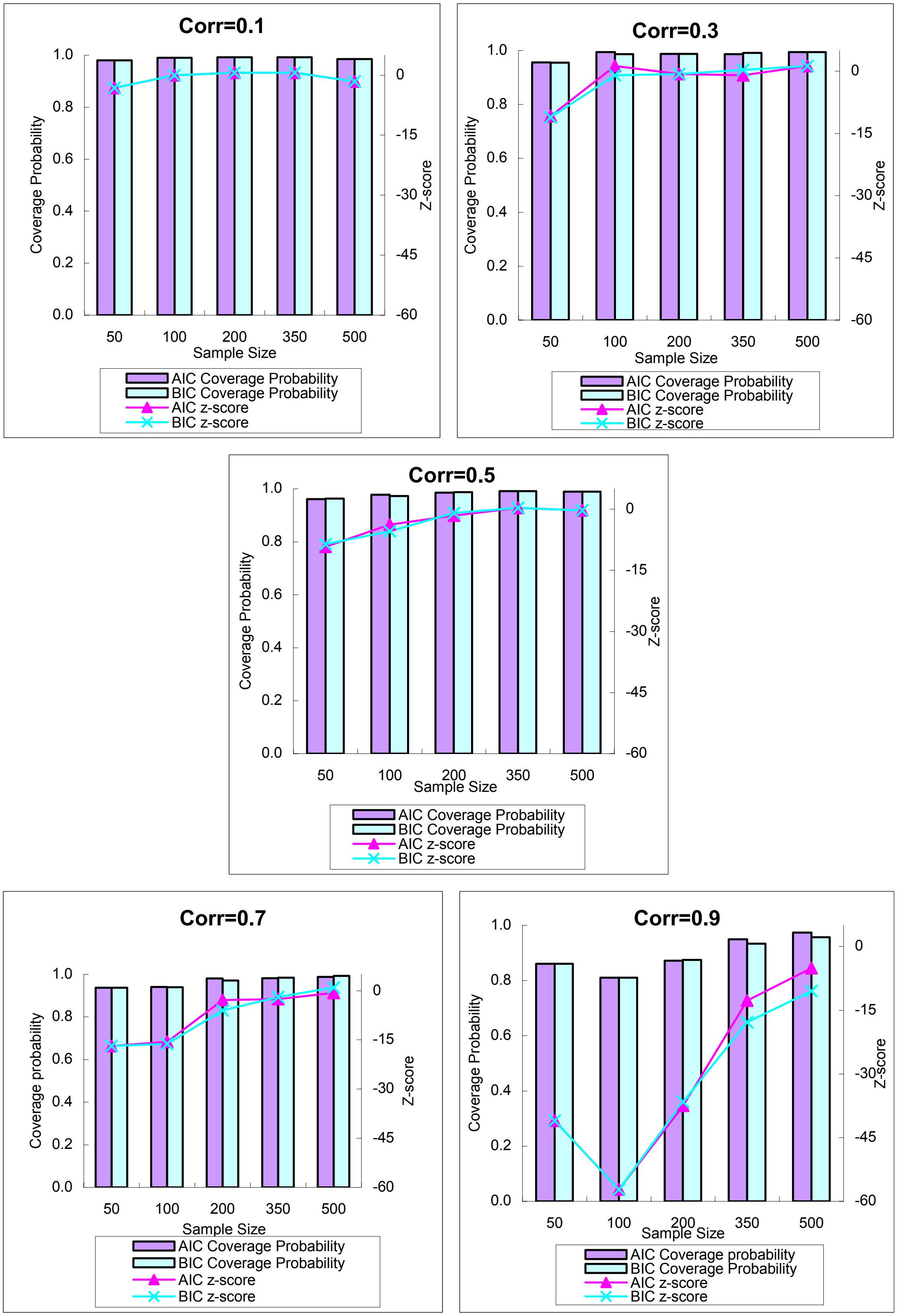


Figure 10: 99% Coverage Probability when the correct model is M_2 (Offset=0)

Correct Model = M_3 ***(Equality-restricted)***

When the correct model is M_3, on which an equal restriction is imposed, some unusual observations are presented, which is different from the pattern shown in the aforementioned models. Compared to that with the true model of M_1, the variation of the coverage probability due to the variation of the covariate correlation is enlarged. Moreover, it seems that the confidence intervals perform better in terms of unconditional coverage probability when BIC is applied in model selection. Table 9 to Table 11 illustrate the coverage probabilities of the naïve confidence intervals when the nominal levels are .95 and .99, and the corresponding z-scores when M_3 is true. Among the three tables, Table 9 depicts the results for $\text{offset}=0$, Table 10 shows the results for $\text{offset}=0.8473$, and Table 11 demonstrates the results for $\text{offset}=2.1972$. Meanwhile, Figure 11 and Figure 12 present the coverage probability and the z-score when the symmetric dichotomization is considered, which correspond to the results in Table 9.

1-α=0.95

<u>AIC</u>

Generally speaking, the coverage probabilities for the 95% confidence intervals are between those for the true models of M_1 and M_2. Although there is a general trend that a lower covariate correlation suggests a higher coverage probability, we find more exceptions, especially when an offset term is imposed. For instance, from Table 10 and Table 11 it is difficult to conclude whether the coverage probability is higher when the correlation is 0.3 or when the correlation is 0.5, as the results for different sample size vary. The extreme z-score (-24.811) is observed in Table 11, where the coverage probability is as low as 0.779 with a sample size of 50 and covariate correlation of 0.9. With a few exceptions, it seems that to increase sample size is helpful to improve the coverage probability, which presents a somewhat different picture from the aforementioned situations.

BIC

Again covariate correlation has important impact on the coverage probability, and the performance of BIC in terms of coverage probability is significantly better than that with the unrestricted true model. The bold values indicating beyond negative 3.5 SDs of the nominal level account for two fifth of the total observations. With a couple of exceptions, sample size plays important role in improving the unconditional coverage probability, which is similar to the situation with the true model of M_2.

1-α =0.99

For AIC, The extreme z-score (-29.240) is observed in Table 10, where the coverage probability is as low as 0.898 with a sample size of 50 and covariate correlation of 0.9.Taking the negative 3.5 SDs as a cut point, the 99% confidence intervals perform better than the 95% confidence intervals, as the cells with bold values are less when 1-α =0.99. This is particularly evident when the offset term is zero. Compared with the results from Table 3 to Table 5, the coverage probability is much higher, demonstrating that the uncertainty in determining the true model may have key effect on the performance of confidence intervals. For BIC, the results are similar to those with the 95% nominal coverage.

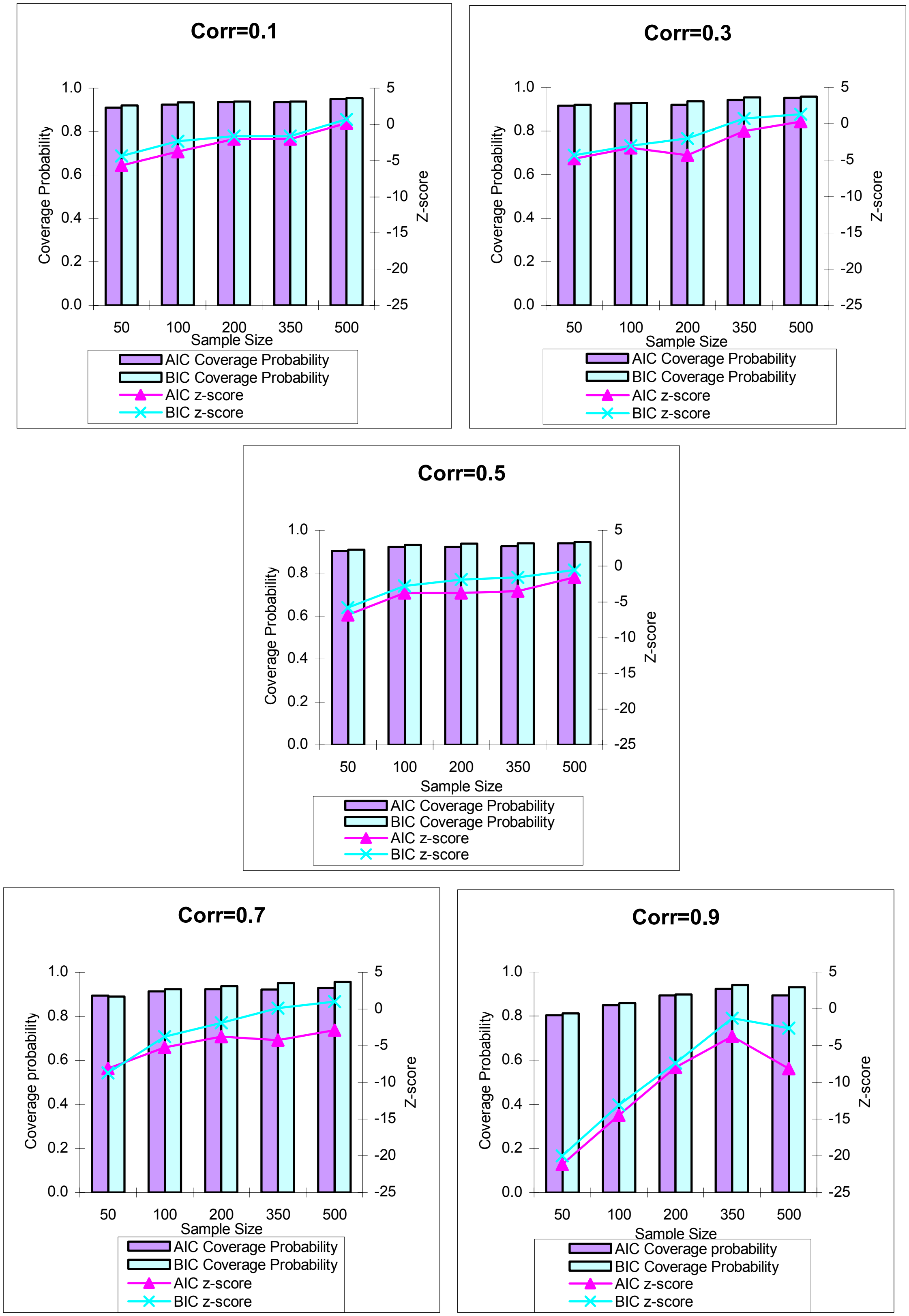


Figure 11: 95% Coverage Probability when the correct model is M_3 (Offset=0)

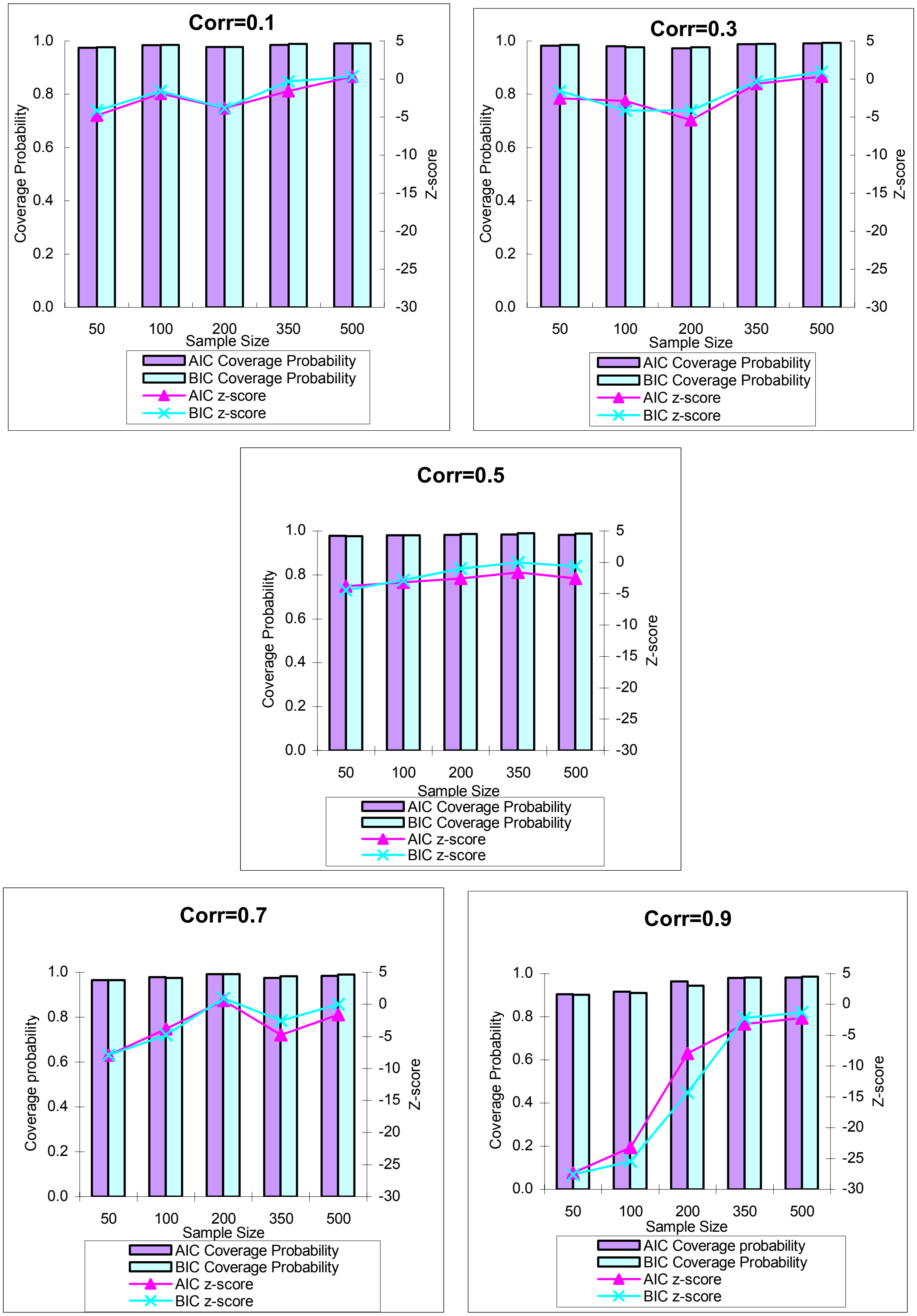


Figure 12: 99% Coverage Probability when the correct model is M_3 (Offset=0)

Conditional Coverage Probability

There are two types of conditional probabilities: conditional on selecting the correct model and on selecting an incorrect model, which, as suggested by previous studies (e.g. Leeb and Pötcher, 2003), may behave with different statistical properties. Although the simulation results suggest some small variations due to the dichotomization threshold, compared to other factors, these effects are minor. One example is that when M_1 is the correct model and the sample size is 50, for both AIC and BIC lower coverage probabilities conditional on selecting the correct model are observed with asymmetric threshold.

We report all the simulation results concerned with the conditional coverage probability in the Appendix, and detail the comparisons with the conditions that do not consider asymmetric response dichotomization (e.g. the results shown in the figures). Table 12 to Table 26 illustrate the numerical results for count on coverage and conditional coverage probability when the offset term is zero. Specifically, Table 12 to Table 16 show the results when the correct model is M_1, Table 17 to Table 21 present the results when the correct model is M_2, and Table 22 to Table 26 demonstrate the results when the correct model is M_3. In the same way, Table 27 to Table 56 show the corresponding simulation outcome when a non-zero offset term is involved.

Conditional on Selecting the Correct Model

The coverage probability conditional on selecting the correct model is high when compared to the corresponding probability conditional on selecting an incorrect model. In many cases the probability is close to the nominal level. However, variations are observed with specific conditions. To facilitate the comparison, Figure 13 to Figure 15 provide graphic information on the conditional coverage probability when Offset=0.

Correct Model = M_1 ***(Unrestricted)***

Covariate Correlation

The covariate correlation imposes major impact on the conditional coverage probability. When the correlation is low (0.1 and 0.3), the probability is most likely close to the nominal level, as long as the sample size is adequately large ($\mathrm{N} \geq 100$). When the correlation is 0.5, coverage probability falls moderately subject to the specific sample size. When the correlation goes higher, irregularities are observed. With $\mathrm{Corr}=0.7$, a zero coverage probability ($1-\alpha=0.95$) after the selection of BIC is observed with $\mathrm{N}=50$. Notice that the corresponding success rate on selecting M_1 is low (1.2%). With $\mathrm{Corr}=0.9$, zero coverage probability occurs at the sample size of 100 ($1-\alpha=0.95$) with the corresponding success rate on selecting M_1 being 0.5%. Therefore, the irregularities are due to extremely low success rate in selecting the correct model. Similarly, the 100% coverage probability after selection by BIC in the case of $\mathrm{N}=350$ does not have much significance , as it is based on only 0.1% of the selection success rate. In such cases, however, we can conclude that the coverage probability conditional on selecting the correct model is highly invalid, as the model selectors do not perform effectively in selecting the correct model.

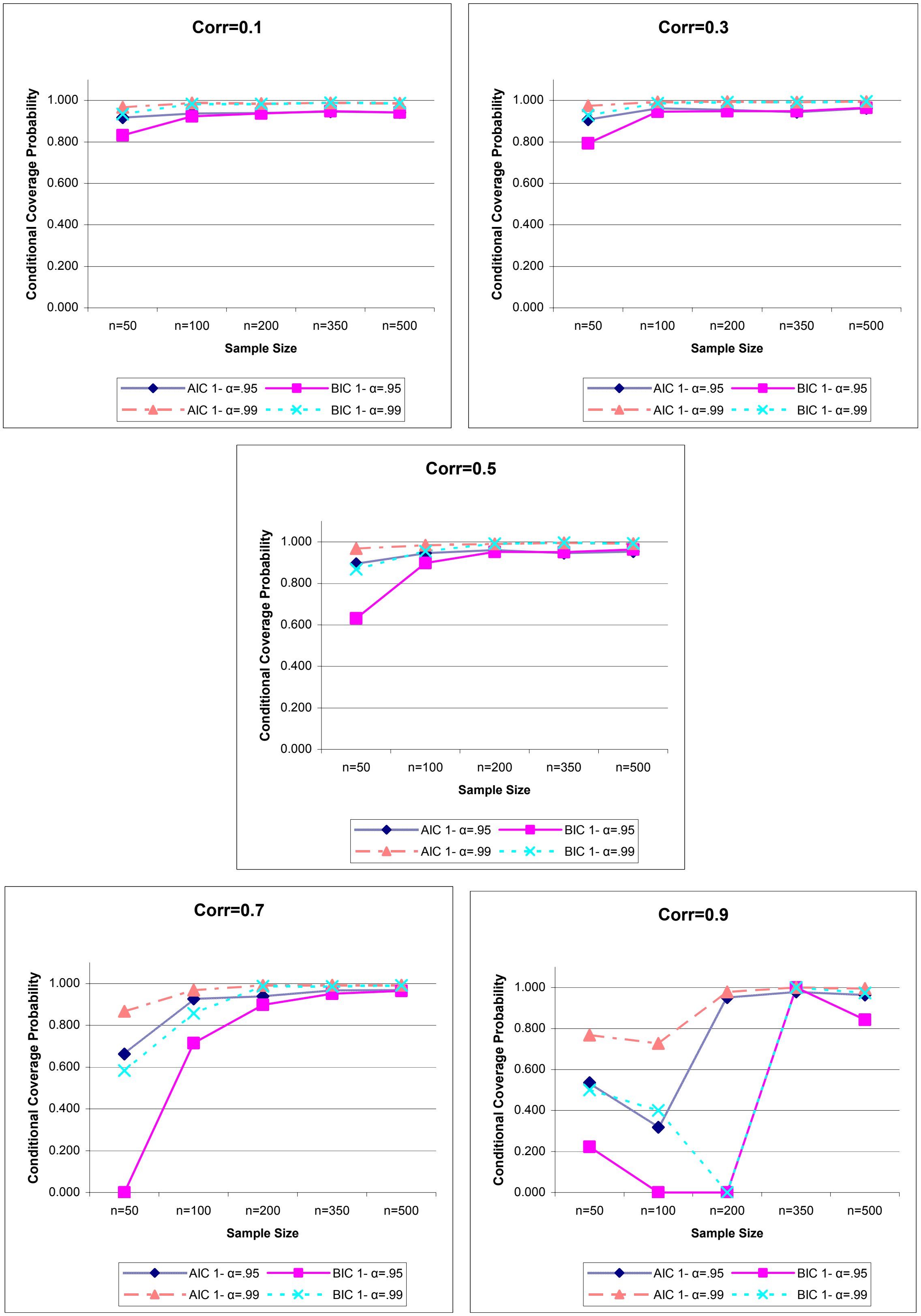


Figure 13: Coverage Probability conditional on selecting the correct model M_1 (Offset=0)

Sample Size

Although the sample size alone does not determine the coverage probability after the correct model is selected, it plays an important role to improve the conditional coverage probability, as shown in Figure 13. Even when the covariate correlation is low, a sample size of 100 or above is required to approximate the nominal coverage probability conditional on selecting the correct model. The prerequisite is more demanding when the correlation is high.

Model Selectors

AIC performs consistently better than BIC in terms of the conditional coverage probability. In the meantime, the results from AIC are much more stable than those from BIC when a higher covariate correlation strongly distorts the picture of coverage. The coverage probability conditional on selecting the correct model based on AIC is most likely close to the nominal level, as long as the sample size is reasonably adequate. The outcome resulted from BIC, however, is subject to more stringent constraints in sample size and covariate correlation in order to obtain commensurate coverage probability as AIC.

Correct Model = M_2 and Correct Model = M_3 (Restricted)

The results for the correct models of M_2 and M_3 are very similar, where high coverage close to the nominal $1-\alpha$ is observed. The fluctuations at various levels of the manipulated factors are relatively small, demonstrating a totally different picture compared to that when the true model is unrestricted. For both of the model selectors, the coverage probability of the 95% confidence interval is roughly between 93% and 96%, no matter what the covariate correlation is. The fluctuation of the 99% confidence interval is even ignorable considering possible sampling error, suggesting that the covariate correlation does not affect the coverage probability conditional on selecting the restricted true model.

AIC and BIC achieve similar results. Moreover, the effect of sample size is also small. Although a sample size of 100 or larger is somewhat more desirable, it seems that the sample size of 50 does not hurt much. The results suggest that provided that the true model is restricted, the lower-than-expected unconditional coverage probability does not stem from the condition when a true model is selected. Therefore, the major source of lower coverage is related to the model selectors' performance in selecting the incorrect model.

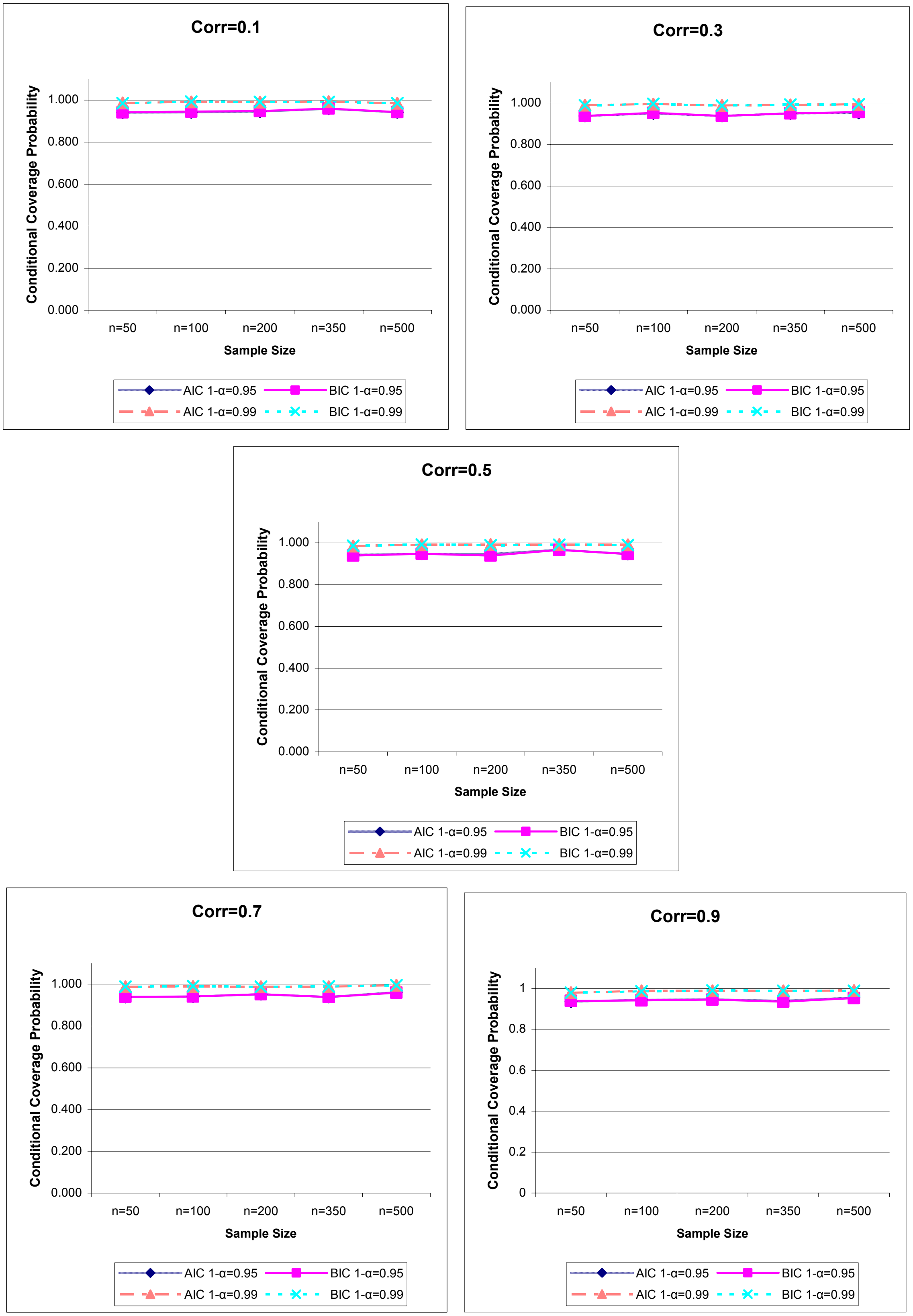


Figure 14: Coverage Probability conditional on selecting the correct model M_2 (Offset=0)

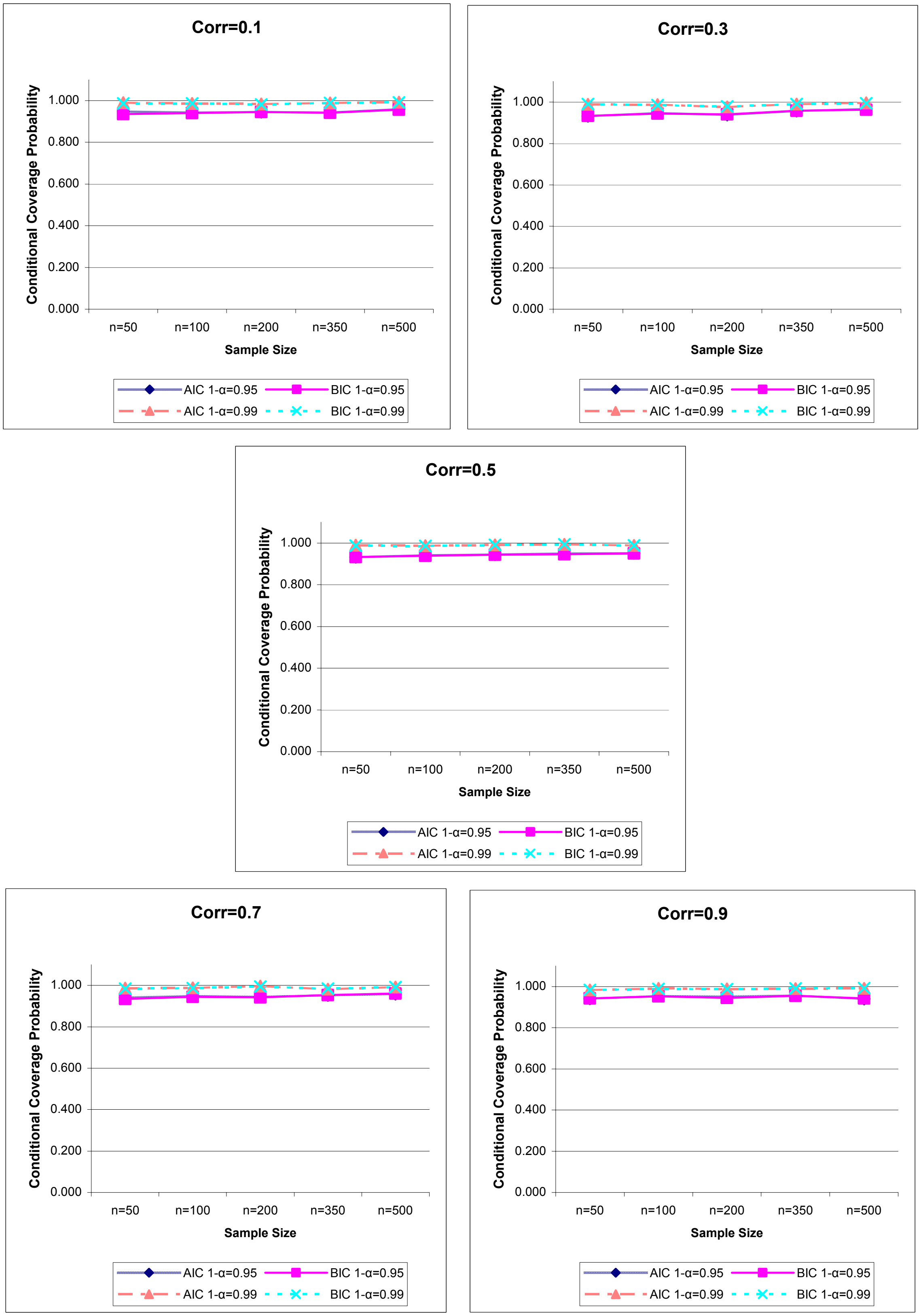


Figure 15: Coverage Probability conditional on selecting the correct model M_3 (Offset=0)

Conditional on Selecting an Incorrect Model

While there is a smaller chance the incorrect models are selected, the volatility of coverage probability conditional on selecting an incorrect model is apt to be larger than that conditional on selecting the correct model. In certain conditions, it is highly possible that the chance to select an incorrect model is close to 0 (e.g. when the sample size is large), which eliminates the chance to measure the conditional coverage probability. An overview suggests that the coverage probability is generally lower compared to that conditional on selecting the correct model.

Correct Model = M_1 *(Unrestricted)*

As indicated in the previous section, when an incorrect model is selected, the chance to select M_3 is higher. Table 12 to Table 16 contain the coverage probability results when the correct model is M_1. While the difference between AIC and BIC is minor, the effects of covariate correlation and sample size are noteworthy.

<u>**Covariate Correlation**</u>

We first investigate the cases when M_3 is chosen by the model selectors. When the correlation is low, for AIC only at small sample sizes ($N < 350$) the probability is observed, which is much lower than the nominal level. Similar observations can be found for BIC, where the count to select M_3 increases to some extent. When the correlation moves up, the chance to select M_3 increases, but the coverage probability remains low.

The highest coverage probability for the 95% confidence interval is 0.842 for AIC and 0.848 for BIC. These extreme values are observed when the covariate correlation is high.

When M_2 is chosen by the model selectors the trend is the other way: the higher the correlation, the lower the coverage probability, which is similar to the coverage probability conditional on selecting the correct model. However, as the chance to be chosen is smaller, the effect on the unconditional coverage probability is less than that when M_3 is falsely selected.

Sample Size

Sample size imposes some constraints on evaluating the coverage probability, as when sample size goes up, it is highly possible that an incorrect model is not selected. However, generally the larger the sample size, the lower the coverage probability conditional on selecting an incorrect model. Therefore, to increase sample size may not necessarily impose positive effects on the unconditional coverage probability. Such asymptotic properties conditional on selecting the incorrect model are especially undesirable when the covariate correlation is high, where the model selectors are more likely to be driven to chose the incorrect models.

Correct Model = M_2 ***and Correct Model =*** M_3 ***(Restricted)***

The coverage probability is usually much lower than that conditional on selecting the correct model. This suggests that the overall lower-than-expected coverage is mainly due to this portion of deviation. When an incorrect model is selected, the chance to select M_1 is higher, with exception in high correlation and small sample size. This is reasonable considering that M_1 offers more flexibility than the rest two models in estimating the parameter of β_2. Table 17 to Table 21 contain the results for conditional coverage probability when the correct model is M_2, and Table 22 to Table 26 contain the results when the correct model is M_3.

Covariate Correlation

When M_1 is falsely chosen, the conditional coverage probability is negatively correlated with the covariate correlation. The relationship applies in most conditions, even when the chance to falsely select M_1 is small. When another model is selected, the effects of covariate correlation display in two aspects: (1) the lower the correlation, the less chance that model is selected; (2) given that the model is chosen at certain chance (e.g. $\geq 5\%$), the negative relationship between the coverage probability and the covariate correlation applies. The condition for (2) is necessary since if the chance is too small, hardly the result can be reliable. In the meantime, the smaller the selection chance, the smaller the weight that portion of possible selections can impose on the unconditional coverage probability. In general, the outcome suggests that a lower covariate correlation is useful to improve the coverage probability, no matter the coverage is conditional on the correct or incorrect model selection.

Sample Size

When M_1 is falsely chosen, the effect of sample size is mixed with that of the covariate correlation, which makes the results difficult to interpret. In general, the coverage probability is less influenced by the sample size if AIC is applied in the model selection. For BIC, it is noticeable that a moderate sample size (e.g. 350) represents the best coverage probability most of time, with a couple of exceptions when the true model is M_3 and the covariate correlation is extreme (e.g. 0.9 or 0.1), where the chance to select M_1 is small.

When another model (either M_3 or M_2, depending on the true model) is falsely chosen, the coverage probability usually falls far below the nominal level. Moreover, the larger the sample size, the less likely that model is selected, and the lower the coverage probability conditional on that model is selected, which, although still influences the unconditional coverage probability, may only have small effects.

Model Selectors

Some differences are observed between AIC and BIC. Most of the time AIC represents a higher coverage probability if M_1 is falsely selected. The most significant difference is that the outcome from BIC seems to be more vulnerable to the high correlations (e.g. Table 26). A closer examination indicates, however, these cases occur when the chance of false selection is small, and therefore does not influence much the unconditional coverage probability.

Chapter V: Summary and Discussion

Coverage Probability Summary

While the coverage probability falls below the conventional $1-\alpha$, the distance of the gap fluctuates along the dimensions of the manipulated factors, a combination of which make the validation of the actual interval estimates after model selection intricate. The most influential factors, which can never be ignored when analyzing the post-model-selection inference, are the model structure and covariate correlation.

When the suppressor variable in M_2 or the restriction in M_3 is involved in the model formulation, the outcome of the coverage probability significantly departs from the corresponding result of M_1, which indicates that the variation of β_2 imposes direct impact on the inference.

The negative relationship between the correlation and the coverage probability accounts for a large portion of the coverage probability variance. The simulation results demonstrate that the covariate correlation plays a role of great importance in validating the coverage probability. Such results are consistent with the previous studies with regard to linear models (e.g. Kaibala & Leeb, 2006).

Unconditional Coverage Probability

As shown by the comparisons among the three true models, the uncertainty in the model structure can lead to highly diverse results, with an underlying unrestricted true model attaining lower coverage. In case of the symmetric dichotomization, z-scores in Table 3, Table 6 and Table 9 specify such disparity among M_1, M_2 and M_3, respectively.

The coverage probability with the true M_2 appears higher (compared to that with M_1) even when the suppressor variable shares very high correlation with the variable in the model. When the restriction involves equality (M_3), the situation is more complicated with some results difficult to interpret.

The unconditional coverage probability is closer to the nominal $1-\alpha$ when the correlation between the covariates is lower. The coverage probability decreases, at a moderate rate from the nominal level when the correlation increases up to around 0.5, and drops at an accelerated rate thereafter.

The coverage variation following sample size needs detailing with other factors: when the true model is M_1, with small sample size AIC and BIC prefer selecting the incorrect models, and the conditional coverage probability is low. When the sample size increases to certain level, the rate to select the true model increases, and the coverage probability goes up. The effects are confounded with the variation of covariate correlation: with lower correlation, the coverage probability is closer to nominal level. Meanwhile, exceptions occur when the covariate correlation is extremely high, where the increase of sample size does not improve the coverage probability. When the true model is M_2 or M_3, even with small sample size AIC and BIC enjoy high success rate in selecting the correct model, which enhances the coverage. In such cases, the effect of sample size on the coverage probability is mitigated.

The finite sample analysis suggests that the unconditional coverage probability does not unanimously follow the usual asymptotic properties of many statistical substances, which generally achieve better or more accurate outcome with a larger sample size. The performance of sample size is not consistent given various underlying true models and covariate correlations, a fact that sometimes makes the effort to alleviate model selection effects through enlarging the sample size ineffective.

Although there have been many studies on the performance of AIC and BIC, we are particularly interested in their difference in affecting the interval coverage. There is no universal evidence which model selector is superior, as the comparative advantage differs to a large extent across the factor of model structure. When the true model is M_1, AIC performs consistently better than BIC in terms of the coverage probability. When the true model is M_2 or M_3, the difference from the AIC and BIC is relatively small. However, the coverage probability gained through BIC seems to be more volatile.

As for the dichotomization offset, the effects are mostly secondary, especially when the true model is a restricted model. An exceedingly skewed dichotomization may facilitate more extraordinary results when extreme values are observed along the dimensions of covariate correlation and sample size. However, if a reasonably moderate dichotomization threshold is applied (e.g. 0.7), the effect to the coverage probability is not significant.

Conditional Coverage Probability

The weighted average of the conditional coverage probabilities establishes a basis to assess the validity of the overall coverage. For the confidence intervals conditional on selecting the correct model, most of the time the underlying parameter is prone to be covered, with which the probability is closer to nominal coverage rate, compared to that conditional on selecting an incorrect model. While we observe effects of the covariate correlation and sample size when the true model is M_1, with the restricted models such effects are to a large extent lessoned. Considering that only a couple of covariates are introduced in this study, a true model involving a large number of parameters seems to be more vulnerable to the effects of covariate correlation, sample size and model selection procedures.

As far as the coverage conditional on selecting an incorrect model is concerned, it is apparent that this portion contributes negatively to the validity of confidence intervals. However, in many cases the selection on incorrect model is relatively rare, which lowers the reliability of the measure. A coverage probability of 0 or 1 does not have much practical significance if the chance of model selection is only 0.1%. As a result the unpredictability of the coverage probability in various conditions is large.

The actual effects of the conditional coverage probability are highly dependent upon the extent to which the incorrect model is likely to be chosen. Given the confounded effects of different factors, it is difficult to summarize and predict to what extent the influence is enforced, and the detailed discussion within the specific environment in Chapter IV is more meaningful.

Discussion

The goal of this research is to investigate, in the context of binary logistic regression, how model structure uncertainty imposes important impact on the statistical inferences. More specifically, the study measures the contraction of the actual interval estimates compared to the nominal coverage of post-model-selection confidence intervals. Three important interrelated questions of this study are: (1) how the unawareness of the model structure uncertainty puts in force significant consequences; (2) whether the application of different model selectors influences the intervals' validity; (3) whether the feature of logistic regression, specifically the log-linear function through which the probability of outcome response is dichotomized, plays a role in the coverage probability.

To study the first question, we propose three true interconnected models with pre-specified parameter configurations, based upon which substantial variations of the coverage probability are observed. To work on the second question, we investigate two major automatic model selectors' performance in choosing the correct model, and measure the subsequent effects on the coverage probability of confidence intervals. To answer the third question, we design varied offsets that are commensurate with applying certain dichotomization thresholds, and examine the results.

Model Structure

Model structure is one of the most important considerations in shaping the actual coverage probability. We investigate the problem through several layers. The first layer, which approximates the truth, is that the model is given *a priori*. Under our assumption, we take such a model as true.

In the real world applications, however, the model structure is unknown to the researchers, and the assumption for the first layer mostly likely cannot hold. Under such circumstance, the second layer applies, where the model selection procedure determines the chance, or the extent to which the researchers are certain about the true model structure. Given our assumption, such chance is represented by the selection success rate. Needless to say there are losses in model selection, as indicated by the success rate, which usually does not achieve 100%. For the finite samples, the simpler models gain stable success rate at comparatively high level, and the unrestricted model is more vulnerable to the effects of multiple factors, no matter AIC or BIC is applied in the model selection. At this stage we review the overall coverage probability and can see that the unrestricted model suffers more.

The issue is somewhat similar to the analysis of Taylor (1986; 1988) and Taylor *et al.* (1996) on the generalized linear models, where extra parameter(s) are involved to improve model data fit. One difference in our study, however, is that we know the model structure. i.e. we are certain about λ, the shape parameter(s) in (23) and (24), by generating data from the true model. In such circumstance we still observe that a simpler (true) model structure enjoys some advantages concerned with the confidence interval coverage, even in many cases the coverage is below the nominal level.

While discussions with such phenomenon are not frequent, there are at least two aspects with regard to the effects. First, if the model is true, with more complicated model structure, the variance due to sampling error is more likely to be inflated. Such inflation presents an illusion on the naïve confidence intervals, which actually achieve even narrower coverage. The second aspect, which is more relevant to this study, is concerned with the model selectors' performance: to enhance selection success is helpful to gain higher coverage. The model selectors impose penalty to more complicated model (even the model is not overparameterized), especially when sample size is not adequate. Therefore, an unrestricted true model loses more in terms of confidence coverage due to comparatively less chance to be selected.

The third layer is more specific: it details the effects conveyed by selecting the correct and incorrect models, a combination of which contributes to the variation of the unconditional (overall) coverage probability. While the coverage probability conditional on selecting the correct model is supposed to be close to the nominal $1-\alpha$, the effects of other factors are observed most with the unrestricted model, for which the coverage drops with small sample size and high covariate correlation. It is noticeable, however, that such effects are still closely related to the model selection: the near to ground coverage probability is inevitably accompanied by the low selection success rate.

The effects conveyed by selecting the incorrect models are generally unfavorable to the overall confidence coverage as it is negatively influenced. While the conditional coverage probability is highly volatile, its negative contribution may or may not be important, depending fundamentally upon the chance the incorrect model is selected. For the unrestricted model, as the model selection success rate is comparatively low (which means the chance to select the incorrect model is high), the coverage probability is further compromised. The coverage disparity between the true M_2 and M_3 is not substantial, and is mainly driven by the difference of the model selectors, which are discussed below.

Model Selectors

Among the infrequently observed comparisons between AIC and BIC on the interval coverage, Hurvich and Tsai's (1990) study on linear models concludes that AIC is superior to BIC in terms of unconditional coverage probability (overall coverage probability in Hurvich & Tsai, 1990), where they investigate the coverage with five candidate models, among which two are manipulated as true models($p_0 = 3$ and $p_0 = 4$, respectively, where p_0 stands for the order of the true model). They assert such difference as "intriguing" as "BIC is a consistent model order selector and AIC is not" and “BIC is uniformly better than AIC at selecting the correct model for the simulation given (in their study).” (p. 215) However, a closer look at Hurvich and Tsai’s (1990) results indicate that when $p_0 = 3$ (for the restricted model), AIC and BIC are not significantly different (with 90%, 95% and 99% confidence intervals, the overall coverage probability is .806, .900 and .960 for AIC, and .820, .902 and .958 for BIC, which suggest little difference).

Different from Hurvich and Tsai’s work, this study on logistic model details the model specification with various covariate correlations, and applies larger sample sizes that cover a wide range of the finite samples’ effects. Considering that Hurvich and Tsai work on linear models and use small sample size (20 for the restricted model, 30 and 50 for the unrestricted model) without considering the correlation of the independent variables, we do not compare our outcome with theirs directly. However, we conclude that the superiority of AIC only applies with certain model structure, with high-order model favoring AIC. When the true model is a restricted model, the effect of different model selection procedures is comparative small.

To better understand the difference, we look into the model selectors’ asymptotic properties. In finite sample analysis, “…conservative model selection procedures are more powerful than consistent model selection procedures in the sense that they are less likely to erroneously select an incorrect model for large sample sizes”, and “…this advantage of the conservative procedure is paid for by a larger probability of selecting an overparameterized model” (Leeb & Pötcher, 2005, p. 36). Therefore, when the true model is unrestricted (M_1), AIC’s advantage is obvious as it gains in the first aspect and does not lose in the second aspect. When the true model is restricted, however, such tradeoff applies and neutralizes the difference between AIC and BIC. When the true model is M_3, BIC even outperforms, especially with a large sample size. In such a case the advantage of consistent model order selector seems to appear.

Dichotomization

For the binary logistic models, the effect of dichotomization threshold is model specific. For a simple true model, different strategies in setting the threshold seem not harmful. When the true model is a more complicated model, which involves more parameters, the skewed dichotomization may compromise the model selection success rate, which in turn imposes detrimental effects on the coverage probability. Such effects may be enlarged when extreme values are observed in other factors (e.g. covariate correlation, sample size, etc). In real applications, as the true model is unknown, an extremely skewed dichotomization (e.g. 0.8/0.2) is not recommended. In most cases a balanced threshold (between 0.3 and 0.7) seems more desirable.

Covariate Correlation and Sample Size

We have confirmed the importance of covariate correlation at any conventional coverage level. It is recommended that a correlation larger than 0.3 can generally be taken to trigger a cautious investigation of the coverage probability. If the covariates involved in any model share considerable portion in predicting the outcome variable, the coverage probability is inevitably compromised. Such a rule applies to models in which a suppressor variable does not predict.

The role of sample size is intriguing. While a larger sample size is helpful to improve the model selector performance, it does not necessarily gain advantage on the coverage probability. “Regardless of sample size… the sampling properties of post-model-selection estimators are typically significantly different from the nominal distributions that arise if a fixed model is supposed” (Leeb & Pötcher, 2005). Such a statement seems also suitable to the coverage of confidence intervals. Coverage does not always improve with a larger sample size. It can decline. Therefore, any attempt to enhance the coverage with large sample size has to consider carefully multiple factors that may influence the ultimate outcome.

Conclusions

While it is usually below the nominal level, there is no simple predictable pattern with regard to how and how far the actual coverage probability of confidence intervals may fall. The coverage probability of confidence intervals varies with multiple factors. While the model structure always plays a role of paramount importance, the covariate correlation significantly affects the interval’s coverage, with the tendency that a higher correlation indicates a lower coverage probability. Meanwhile, no evidence shows that AIC inevitably outperforms BIC in terms of achieving higher coverage probability, or vice versa. The model selector’s performance is dependent upon the uncertain model structure and/or the unknown parameter vector θ , against which AIC may outperform in choosing the correct model, and in turn, this influences the coverage. The effect of sample size is in intriguing, and a larger sample size does not necessarily achieve asymptotically more accurate inference on interval estimates. Although the binary threshold of the logistic model may affect the coverage probability, such effect is less important. It is most likely to become substantial with an unrestricted model when extreme values along the dimensions (e.g. small sample size, high covariate correlation) of other factors are observed.

Concluding Remarks

Logistic regression models are widely applied in educational studies. While the statistical model formulation, selection and prediction are viewed as generic means to understand and solve problems, the unawareness of model uncertainty after selection is not uncommon. An example is the model involved in this study, with which the original literature predicts academic success with such variables as students' school performance. In the original research, detailed analyses are carried out with hypothesis tests on the significance of potential variables, model comparison, subsequent inferences and minimization of estimate errors. However, very little attention has been paid to relevant important problems stemmed from the model structure itself. i.e. what is the implication of the model structure uncertainty to the significance of conclusions drawn based upon the data-driven model selection, how do model selection procedures affect the validity of the parameter estimates, how much do the actual confidence intervals under-perform below the expected coverage, etc.

In other fields that apply similar statistical framework (e.g. econometrics), major studies on model uncertainty after model selection have taken linear regression models into consideration. However, the usual assumptions for least squares estimators in linear regression do not apply, in a general sense, to logistic regression. Meanwhile, the linear regression models do not consider categorical dependent variables, which are widely observed in social science.

This study intends to make a progress in the above two aspects. We investigate the post-model-selection impact with binary logistic models that are commonly applied in educational studies, and try to understand quantitatively how the actual coverage of confidence intervals departs from the one without considering model uncertainty. By examining the behavior of the model selectors in choosing the "best" model under various circumstances, the investigation gains a basis to understand the relationship between the selectors' performance and the coverage probability. By distinguishing the deviations that contribute to the distortion of the unconditional coverage probability conditional on the success or failure of model selection, the study provides a comprehensive view that is helpful to identify and gauge the specific coverage departure behind the general lower-than-expected coverage probability.

The simulation study obtains results that are commensurate with the real life applications, the discrepancy between which and the expectations may raise serious concern about the confidence intervals' actual coverage. Given the usual suggestions that the coverage rate of naïve confidence intervals is less than $1-\alpha$, this study contributes in terms of gauging the degree the coverage probability falls, and how it falls given such conditions as varied model structure and covariate correlations.

The study has its limitations: due to the complicated nature of the problem, we only consider a small group of three typical models, which involves two types of commonly applied model restrictions. The generalization of the empirical research may have to entail applicable methods to accommodate a larger number of candidate models, which should serve as one of the goals for further study. More complicated cases, which entail the selection among non-nested models, need to be investigated. Meanwhile, the unattended issues such as the validity of post-model-selection point estimators in the context of logistic models also call for intensive research. Moreover, solutions with regard to accommodating such problems as model uncertainty are still far from satisfactory, and consistent efforts have to be made in the future.

No matter how the statistical model is formulated, as suggested by Buckland *et al*. (1997), when there is not yet a perfect solution to minimize the effects of statistical uncertainty, one principle is that the statistical inference should incorporate such effects. However, we also notice the reality "Estimation of (these) finite-sample distributions is 'impossible' (even in large samples). No resampling scheme whatsoever can help to alleviate this situation." (Leeb & Pötcher, 2005, p. 23) As a gap between the inference and the undetermined truth due to the fact of model structure uncertainty has been recognized (even the precise distance of the gap is unknown), the statistical inferences should be used in a more conservative manner.

Appendix A: Results of Simulation

Table 3: Unconditional Coverage Probability when the correct model is M_1 with Offset $= 0$.

AIC		**$1-\alpha=.95$**					**$1-\alpha=.99$**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.872	0.861	0.843	0.820	0.774	0.951	0.947	0.944	0.932	0.918
	Z	**-11.317**	**-12.913**	**-15.525**	**-18.862**	**-25.537**	**-12.395**	**-13.666**	**-14.620**	**-18.434**	**-22.883**
100	*UCP*	0.887	0.910	0.848	0.765	0.630	0.960	0.951	0.934	0.891	0.832
	Z	**-9.141**	**-5.804**	**-14.800**	**-26.843**	**-46.430**	**-9.535**	**-12.395**	**-17.798**	**-31.464**	**-50.216**
200	*UCP*	0.929	0.924	0.888	0.787	0.485	0.980	0.975	0.949	0.890	0.680
	Z	***-3.047***	**-3.772**	**-8.996**	**-23.651**	**-67.469**	***-3.178***	**-4.767**	**-13.031**	**-31.782**	**-98.524**
350	*UCP*	0.945	0.942	0.919	0.878	0.538	0.987	0.990	0.973	0.919	0.653
	Z	-0.725	-1.161	**-4.498**	**-10.447**	**-59.779**	-0.953	0.000	**-5.403**	**-22.565**	**-107.106**
500	*UCP*	0.942	0.961	0.949	0.918	0.613	0.987	0.993	0.988	0.947	0.689
	Z	-1.161	1.596	-0.145	**-4.643**	**-48.897**	-0.953	0.953	-0.636	**-13.666**	**-95.664**

BIC		**$1-\alpha=.95$**					**$1-\alpha=.99$**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.845	0.851	0.836	0.819	0.779	0.945	0.945	0.941	0.933	0.918
	Z	**-15.235**	**-14.364**	**-16.541**	**-19.007**	**-24.811**	**-14.302**	**-14.302**	**-15.573**	**-18.116**	**-22.883**
100	*UCP*	0.785	0.798	0.746	0.695	0.631	0.904	0.893	0.883	0.866	0.831
	Z	**-23.941**	**-22.054**	**-29.599**	**-36.999**	**-46.285**	**-27.333**	**-30.829**	**-34.007**	**-39.410**	**-50.534**
200	*UCP*	0.859	0.830	0.731	0.545	0.414	0.929	0.917	0.842	0.724	0.640
	Z	**-13.204**	**-17.411**	**-31.776**	**-58.764**	**-77.771**	**-19.387**	**-23.201**	**-47.037**	**-84.540**	**-111.237**
350	*UCP*	0.936	0.902	0.786	0.583	0.210	0.978	0.955	0.847	0.685	0.386
	Z	-2.031	**-6.965**	**-23.796**	**-53.250**	**-107.370**	**-3.814**	**-11.124**	**-45.448**	**-96.935**	**-191.964**
500	*UCP*	0.942	0.954	0.885	0.693	0.127	0.987	0.987	0.926	0.747	0.256
	Z	-1.161	0.580	**-9.431**	**-37.289**	**-119.413**	-0.953	-0.953	**-20.341**	**-77.230**	**-233.281**

Note:

(1) *N*=Sample Size. *CORR* = Covariate Correlation. *UCP* =Unconditional Coverage Probability (Overall Coverage Probability). *Z* = z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.

(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.

(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 4: Unconditional Coverage Probability when the correct model is M_1 with Offset=0.8473 .

AIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.863	0.854	0.846	0.820	0.762	0.973	0.945	0.947	0.945	0.909
	Z	**-12.623**	**-13.929**	**-15.090**	**-18.862**	**-27.278**	**-5.403**	**-14.302**	**-13.666**	**-14.302**	**-25.743**
100	*UCP*	0.900	0.862	0.842	0.768	0.646	0.962	0.952	0.943	0.901	0.830
	Z	**-7.255**	**-12.768**	**-15.670**	**-26.407**	**-44.109**	**-8.899**	**-12.077**	**-14.938**	**-28.286**	**-50.851**
200	*UCP*	0.938	0.924	0.875	0.778	0.469	0.977	0.974	0.937	0.881	0.665
	Z	-1.741	**-3.772**	**-10.882**	**-24.956**	**-69.791**	**-4.132**	**-5.085**	**-16.845**	**-34.642**	**-103.292**
350	*UCP*	0.939	0.945	0.918	0.844	0.531	0.988	0.985	0.965	0.914	0.648
	Z	-1.596	-0.725	**-4.643**	**-15.380**	**-60.795**	-0.636	-1.589	**-7.946**	**-24.154**	**-108.695**
500	*UCP*	0.945	0.953	0.945	0.897	0.596	0.987	0.988	0.987	0.944	0.673
	Z	-0.725	0.435	-0.725	**-7.690**	**-51.364**	-0.953	-0.636	-0.953	**-14.620**	**-100.749**

BIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.844	0.842	0.839	0.819	0.772	0.965	0.945	0.945	0.943	0.906
	Z	**-15.380**	**-15.670**	**-16.106**	**-19.007**	**-25.827**	**-7.946**	**-14.302**	**-14.302**	**-14.938**	**-26.697**
100	*UCP*	0.815	0.768	0.768	0.703	0.651	0.916	0.910	0.913	0.875	0.828
	Z	**-19.588**	**-26.407**	**-26.407**	**-35.839**	**-43.383**	**-23.519**	**-25.426**	**-24.472**	**-36.549**	**-51.487**
200	*UCP*	0.873	0.807	0.703	0.565	0.400	0.930	0.893	0.828	0.766	0.616
	Z	**-11.172**	**-20.749**	**-35.839**	**-55.862**	**-79.802**	**-19.069**	**-30.829**	**-51.487**	**-71.192**	**-118.865**
350	*UCP*	0.933	0.898	0.804	0.570	0.219	0.979	0.950	0.865	0.694	0.402
	Z	-2.467	**-7.545**	**-21.184**	**-55.136**	**-106.065**	***-3.496***	**-12.713**	**-39.728**	**-94.075**	**-186.879**
500	*UCP*	0.944	0.942	0.879	0.647	0.107	0.986	0.977	0.930	0.714	0.253
	Z	-0.871	-1.161	**-10.302**	**-43.964**	**-122.315**	-1.271	**-4.132**	**-19.069**	**-87.719**	**-234.234**

Note:

(1) *N*=Sample Size. *CORR*= Covariate Correlation. *UCP*=Unconditional Coverage Probability (Overall Coverage Probability). *Z*= z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.

(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.

(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 5: Unconditional Coverage Probability when the correct model is M_1 with Offset=2.1972.

AIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.863	0.861	0.859	0.853	0.765	0.953	0.959	0.949	0.945	0.904
	Z	**-12.623**	**-12.913**	**-13.204**	**-14.074**	**-26.843**	**-11.759**	**-9.852**	**-13.031**	**-14.302**	**-27.333**
100	*UCP*	0.879	0.837	0.836	0.777	0.668	0.961	0.938	0.922	0.907	0.851
	Z	**-10.302**	**-16.396**	**-16.541**	**-25.101**	**-40.917**	**-9.217**	**-16.527**	**-21.612**	**-26.379**	**-44.177**
200	*UCP*	0.919	0.909	0.878	0.776	0.488	0.973	0.965	0.944	0.880	0.716
	Z	**-4.498**	**-5.949**	**-10.447**	**-25.247**	**-67.034**	**-5.403**	**-7.946**	**-14.620**	**-34.960**	**-87.083**
350	*UCP*	0.945	0.940	0.914	0.827	0.487	0.991	0.982	0.952	0.893	0.642
	Z	-0.725	-1.451	**-5.223**	**-17.847**	**-67.179**	0.318	-2.543	**-12.077**	**-30.829**	**-110.602**
500	*UCP*	0.930	0.942	0.937	0.881	0.562	0.978	0.987	0.986	0.926	0.666
	Z	-2.902	-1.161	-1.886	**-10.012**	**-56.297**	**-3.814**	-0.953	-1.271	**-20.341**	**-102.974**

BIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.857	0.863	0.851	0.854	0.776	0.953	0.962	0.949	0.943	0.905
	Z	**-13.494**	**-12.623**	**-14.364**	**-13.929**	**-25.247**	**-11.759**	**-8.899**	**-13.031**	**-14.938**	**-27.015**
100	*UCP*	0.807	0.752	0.789	0.752	0.674	0.932	0.902	0.903	0.888	0.855
	Z	**-20.749**	**-28.729**	**-23.360**	**-28.729**	**-40.046**	**-18.434**	**-27.968**	**-27.650**	**-32.418**	**-42.906**
200	*UCP*	0.827	0.774	0.704	0.561	0.468	0.909	0.888	0.839	0.769	0.695
	Z	**-17.847**	**-25.537**	**-35.693**	**-56.442**	**-69.936**	**-25.743**	**-32.418**	**-47.991**	**-70.238**	**-93.757**
350	*UCP*	0.909	0.848	0.744	0.555	0.269	0.956	0.909	0.830	0.699	0.477
	Z	**-5.949**	**-14.800**	**-29.890**	**-57.313**	**-98.810**	**-10.806**	**-25.743**	**-50.851**	**-92.486**	**-163.042**
500	*UCP*	0.923	0.904	0.809	0.591	0.137	0.969	0.951	0.870	0.693	0.288
	Z	**-3.918**	**-6.674**	**-20.458**	**-52.089**	**-117.962**	**-6.674**	**-12.395**	**-38.139**	**-94.393**	**-223.110**

Note:

(1) *N*=Sample Size. *CORR*= Covariate Correlation. *UCP*=Unconditional Coverage Probability (Overall Coverage Probability). *Z*= z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.
(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.
(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 6: Unconditional Coverage Probability when the correct model is M_2 with Offset=0.

	AIC	1-α=.95					1-α=.99				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.931	0.895	0.898	0.858	0.756	0.980	0.956	0.961	0.937	0.861
	Z	-2.757	**-7.980**	**-7.545**	**-13.349**	**-28.148**	***-3.178***	**-10.806**	**-9.217**	**-16.845**	**-40.999**
100	*UCP*	0.937	0.947	0.932	0.888	0.734	0.990	0.994	0.978	0.941	0.810
	Z	-1.886	-0.435	-2.612	**-8.996**	**-31.341**	0.000	1.271	**-3.814**	**-15.573**	**-57.208**
200	*UCP*	0.949	0.933	0.930	0.929	0.827	0.992	0.988	0.985	0.981	0.872
	Z	-0.145	-2.467	-2.902	***-3.047***	**-17.847**	0.636	-0.636	-1.589	-2.860	**-37.503**
350	*UCP*	0.960	0.947	0.961	0.928	0.893	0.992	0.987	0.991	0.982	0.950
	Z	1.451	-0.435	1.596	***-3.192***	**-8.270**	0.636	-0.953	0.318	-2.543	**-12.713**
500	*UCP*	0.938	0.952	0.946	0.942	0.920	0.985	0.994	0.989	0.988	0.974
	Z	-1.741	0.290	-0.580	-1.161	**-4.353**	-1.589	1.271	-0.318	-0.636	**-5.085**

	BIC	1-α=.95					1-α=.99				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.930	0.896	0.900	0.861	0.769	0.980	0.955	0.963	0.937	0.861
	Z	-2.902	**-7.835**	**-7.255**	**-12.913**	**-26.262**	***-3.178***	**-11.124**	**-8.581**	**-16.845**	**-40.999**
100	*UCP*	0.939	0.943	0.926	0.878	0.749	0.990	0.987	0.973	0.939	0.810
	Z	-1.596	-1.016	***-3.482***	**-10.447**	**-29.164**	0.000	-0.953	**-5.403**	**-16.209**	**-57.208**
200	*UCP*	0.948	0.937	0.936	0.933	0.834	0.992	0.988	0.987	0.971	0.875
	Z	-0.290	-1.886	-2.031	-2.467	**-16.831**	0.636	-0.636	-0.953	**-6.039**	**-36.549**
350	*UCP*	0.960	0.949	0.964	0.934	0.883	0.992	0.991	0.991	0.984	0.934
	Z	1.451	-0.145	2.031	-2.322	**-9.721**	0.636	0.318	0.318	-1.907	**-17.798**
500	*UCP*	0.942	0.955	0.946	0.955	0.922	0.985	0.994	0.989	0.993	0.957
	Z	-1.161	0.725	-0.580	0.725	**-4.063**	-1.589	1.271	-0.318	0.953	**-10.488**

Note:

(1) *N*=Sample Size. *CORR*= Covariate Correlation. *UCP*=Unconditional Coverage Probability (Overall Coverage Probability). *Z*= z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.
(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.

(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 7: Unconditional Coverage Probability when the correct model is M_2 with Offset=0.8473 .

N	AIC CORR	1-α=.95 0.1	0.3	0.5	0.7	0.9	1-α=.99 0.1	0.3	0.5	0.7	0.9
50	*UCP*	0.937	0.923	0.896	0.848	0.771	0.985	0.970	0.957	0.926	0.883
	Z	-1.886	**-3.918**	**-7.835**	**-14.800**	**-25.972**	-1.589	**-6.356**	**-10.488**	**-20.341**	**-34.007**
100	*UCP*	0.937	0.944	0.920	0.885	0.753	0.985	0.984	0.975	0.939	0.819
	Z	-1.886	-0.871	**-4.353**	**-9.431**	**-28.584**	-1.589	-1.907	**-4.767**	**-16.209**	**-54.347**
200	*UCP*	0.945	0.951	0.929	0.932	0.827	0.987	0.989	0.988	0.980	0.870
	Z	-0.725	0.145	***-3.047***	-2.612	**-17.847**	-0.953	-0.318	-0.636	***-3.178***	**-38.139**
350	*UCP*	0.956	0.937	0.930	0.941	0.897	0.990	0.986	0.991	0.992	0.951
	Z	0.871	-1.886	-2.902	-1.306	**-7.690**	0.000	-1.271	0.318	0.636	**-12.395**
500	*UCP*	0.946	0.949	0.946	0.944	0.910	0.984	0.984	0.985	0.991	0.972
	Z	-0.580	-0.145	-0.580	-0.871	**-5.804**	-1.907	-1.907	-1.589	0.318	**-5.721**

N	BIC CORR	1-α=.95 0.1	0.3	0.5	0.7	0.9	1-α=.99 0.1	0.3	0.5	0.7	0.9
50	*UCP*	0.938	0.923	0.898	0.853	0.781	0.986	0.970	0.957	0.926	0.883
	Z	-1.741	**-3.918**	**-7.545**	**-14.074**	**-24.521**	-1.271	**-6.356**	**-10.488**	**-20.341**	**-34.007**
100	*UCP*	0.938	0.937	0.913	0.877	0.768	0.985	0.979	0.969	0.934	0.823
	Z	-1.741	-1.886	**-5.369**	**-10.592**	**-26.407**	-1.589	***-3.496***	**-6.674**	**-17.798**	**-53.076**
200	*UCP*	0.946	0.952	0.937	0.924	0.835	0.987	0.989	0.988	0.967	0.870
	Z	-0.580	0.290	-1.886	**-3.772**	**-16.686**	-0.953	-0.318	-0.636	**-7.310**	**-38.139**
350	*UCP*	0.957	0.937	0.937	0.953	0.894	0.990	0.989	0.993	0.996	0.937
	Z	1.016	-1.886	-1.886	0.435	**-8.125**	0.000	-0.318	0.953	1.907	**-16.845**
500	*UCP*	0.947	0.948	0.947	0.952	0.915	0.984	0.987	0.987	0.991	0.959
	Z	-0.435	-0.290	-0.435	0.290	**-5.078**	-1.907	-0.953	-0.953	0.318	**-9.852**

Note:

(1) *N*=Sample Size. *CORR*= Covariate Correlation. UCP=Unconditional Coverage Probability (Overall Coverage Probability). Z = *z*-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.

(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.

(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 8: Unconditional Coverage Probability when the correct model is M_2 with Offset=2.1972.

AIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.912	0.900	0.859	0.825	0.788	0.972	0.958	0.951	0.918	0.886
	Z	**-5.514**	**-7.255**	**-13.204**	**-18.137**	**-23.505**	**-5.721**	**-10.170**	**-12.395**	**-22.883**	**-33.053**
100	*UCP*	0.931	0.927	0.910	0.845	0.735	0.985	0.983	0.960	0.907	0.829
	Z	-2.757	***-3.337***	**-5.804**	**-15.235**	**-31.195**	-1.589	-2.225	**-9.535**	**-26.379**	**-51.169**
200	*UCP*	0.946	0.941	0.931	0.931	0.777	0.990	0.984	0.983	0.975	0.833
	Z	-0.580	-1.306	-2.757	-2.757	**-25.101**	0.000	-1.907	-2.225	**-4.767**	**-49.898**
350	*UCP*	0.942	0.953	0.947	0.943	0.845	0.991	0.989	0.991	0.991	0.904
	Z	-1.161	0.435	-0.435	-1.016	**-15.235**	0.318	-0.318	0.318	0.318	**-27.333**
500	*UCP*	0.945	0.959	0.936	0.932	0.920	0.989	0.992	0.987	0.986	0.964
	Z	-0.725	1.306	-2.031	-2.612	**-4.353**	-0.318	0.636	-0.953	-1.271	**-8.263**

BIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.915	0.903	0.864	0.837	0.800	0.973	0.960	0.953	0.919	0.885
	Z	**-5.078**	**-6.819**	**-12.478**	**-16.396**	**-21.764**	**-5.403**	**-9.535**	**-11.759**	**-22.565**	**-33.371**
100	*UCP*	0.929	0.926	0.899	0.848	0.755	0.984	0.977	0.952	0.907	0.830
	Z	***-3.047***	***-3.482***	**-7.400**	**-14.800**	**-28.294**	-1.907	**-4.132**	**-12.077**	**-26.379**	**-50.851**
200	*UCP*	0.947	0.941	0.933	0.920	0.787	0.990	0.983	0.980	0.955	0.834
	Z	-0.435	-1.306	-2.467	**-4.353**	**-23.651**	0.000	-2.225	***-3.178***	**-11.124**	**-49.580**
350	*UCP*	0.944	0.953	0.953	0.952	0.852	0.991	0.986	0.994	0.989	0.905
	Z	-0.871	0.435	0.435	0.290	**-14.219**	0.318	-1.271	1.271	-0.318	**-27.015**
500	*UCP*	0.945	0.961	0.941	0.943	0.908	0.988	0.993	0.989	0.989	0.937
	Z	-0.725	1.596	-1.306	-1.016	**-6.094**	-0.636	0.953	-0.318	-0.318	**-16.845**

Note:

(1) *N*=Sample Size. *CORR*= Covariate Correlation. *UCP*=Unconditional Coverage Probability (Overall Coverage Probability). *Z*= z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.
(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.
(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 9: Unconditional Coverage Probability when the correct model is M_3 with Offset=0 .

AIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.911	0.917	0.903	0.894	0.804	0.975	0.982	0.978	0.965	0.904
	Z	**-5.659**	**-4.788**	**-6.819**	**-8.125**	**-21.184**	**-4.767**	-2.543	**-3.814**	**-7.946**	**-27.333**
100	*UCP*	0.924	0.927	0.924	0.914	0.850	0.984	0.981	0.980	0.978	0.917
	Z	**-3.772**	***-3.337***	**-3.772**	**-5.223**	**-14.510**	-1.907	-2.860	***-3.178***	**-3.814**	**-23.201**
200	*UCP*	0.936	0.920	0.924	0.924	0.895	0.978	0.973	0.982	0.992	0.965
	Z	-2.031	**-4.353**	**-3.772**	**-3.772**	**-7.980**	**-3.814**	**-5.403**	-2.543	0.636	**-7.946**
350	*UCP*	0.936	0.943	0.926	0.921	0.924	0.985	0.988	0.985	0.975	0.980
	Z	-2.031	-1.016	***-3.482***	**-4.208**	**-3.772**	-1.589	-0.636	-1.589	**-4.767**	***-3.178***
500	*UCP*	0.951	0.952	0.939	0.930	0.894	0.991	0.991	0.982	0.985	0.983
	Z	0.145	0.290	-1.596	-2.902	**-8.125**	0.318	0.318	-2.543	-1.589	-2.225

BIC		**1-α=.95**					**1-α=.99**				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.920	0.920	0.910	0.890	0.812	0.977	0.985	0.976	0.965	0.903
	Z	**-4.353**	**-4.353**	**-5.804**	**-8.706**	**-20.023**	**-4.132**	-1.589	**-4.449**	**-7.946**	**-27.650**
100	*UCP*	0.934	0.929	0.931	0.924	0.860	0.985	0.977	0.981	0.975	0.910
	Z	-2.322	***-3.047***	-2.757	**-3.772**	**-13.059**	-1.589	**-4.132**	-2.860	**-4.767**	**-25.426**
200	*UCP*	0.939	0.936	0.937	0.937	0.899	0.978	0.977	0.987	0.993	0.945
	Z	-1.596	-2.031	-1.886	-1.886	**-7.400**	**-3.814**	**-4.132**	-0.953	0.953	**-14.302**
350	*UCP*	0.939	0.955	0.939	0.951	0.941	0.989	0.989	0.990	0.982	0.983
	Z	-1.596	0.725	-1.596	0.145	-1.306	-0.318	-0.318	0.000	-2.543	-2.225
500	*UCP*	0.955	0.959	0.946	0.957	0.932	0.991	0.993	0.988	0.990	0.986
	Z	0.725	1.306	-0.580	1.016	-2.612	0.318	0.953	-0.636	0.000	-1.271

Note:

(1) *N*=Sample Size. *CORR*= Covariate Correlation. *UCP*=Unconditional Coverage Probability (Overall Coverage Probability). *Z*= z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.

(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.

(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 10: Unconditional Coverage Probability when the correct model is M_3 with Offset=0.8473 .

AIC		1-α=.95					1-α=.99				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.914	0.917	0.904	0.877	0.790	0.975	0.981	0.977	0.958	0.898
	Z	**-5.223**	**-4.788**	**-6.674**	**-10.592**	**-23.215**	**-4.767**	-2.860	**-4.132**	**-10.170**	**-29.240**
100	*UCP*	0.916	0.918	0.920	0.933	0.836	0.981	0.975	0.980	0.983	0.919
	Z	**-4.933**	**-4.643**	**-4.353**	-2.467	**-16.541**	-2.860	**-4.767**	***-3.178***	-2.225	**-22.565**
200	*UCP*	0.936	0.919	0.943	0.915	0.908	0.988	0.979	0.989	0.977	0.969
	Z	-2.031	**-4.498**	-1.016	**-5.078**	**-6.094**	-0.636	***-3.496***	-0.318	**-4.132**	**-6.674**
350	*UCP*	0.956	0.930	0.924	0.914	0.918	0.989	0.983	0.984	0.977	0.986
	Z	0.871	-2.902	**-3.772**	**-5.223**	**-4.643**	-0.318	-2.225	-1.907	**-4.132**	-1.271
500	*UCP*	0.933	0.918	0.930	0.927	0.916	0.984	0.979	0.983	0.979	0.981
	Z	-2.467	**-4.643**	-2.902	***-3.337***	**-4.933**	-1.907	***-3.496***	-2.225	***-3.496***	-2.860

BIC		1-α=.95					1-α=.99				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.926	0.930	0.911	0.882	0.794	0.977	0.982	0.978	0.959	0.899
	Z	***-3.482***	-2.902	**-5.659**	**-9.866**	**-22.635**	**-4.132**	-2.543	**-3.814**	**-9.852**	**-28.922**
100	*UCP*	0.925	0.928	0.929	0.943	0.836	0.982	0.976	0.984	0.983	0.914
	Z	**-3.627**	***-3.192***	***-3.047***	-1.016	**-16.541**	-2.543	**-4.449**	-1.907	-2.225	**-24.154**
200	*UCP*	0.949	0.932	0.951	0.938	0.906	0.989	0.982	0.989	0.983	0.955
	Z	-0.145	-2.612	0.145	-1.741	**-6.384**	-0.318	-2.543	-0.318	-2.225	**-11.124**
350	*UCP*	0.958	0.942	0.950	0.931	0.944	0.993	0.986	0.988	0.981	0.985
	Z	1.161	-1.161	0.000	-2.757	-0.871	0.953	-1.271	-0.636	-2.860	-1.589
500	*UCP*	0.943	0.929	0.937	0.934	0.941	0.986	0.983	0.985	0.983	0.984
	Z	-1.016	***-3.047***	-1.886	-2.322	-1.306	-1.271	-2.225	-1.589	-2.225	-1.907

Note:

(1) *N*=Sample Size. *CORR* = Covariate Correlation. *UCP*=Unconditional Coverage Probability (Overall Coverage Probability). *Z*= z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.

(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.

(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 11: Unconditional Coverage Probability when the correct model is M_3 with Offset=2.1972.

AIC		1-α=.95					1-α=.99				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.908	0.911	0.900	0.893	0.779	0.967	0.971	0.964	0.959	0.899
	Z	**-6.094**	**-5.659**	**-7.255**	**-8.270**	**-24.811**	**-7.310**	**-6.039**	**-8.263**	**-9.852**	**-28.922**
100	*UCP*	0.935	0.917	0.922	0.900	0.843	0.980	0.981	0.985	0.974	0.914
	Z	-2.176	**-4.788**	**-4.063**	**-7.255**	**-15.525**	***-3.178***	-2.860	-1.589	**-5.085**	**-24.154**
200	*UCP*	0.944	0.926	0.938	0.924	0.900	0.990	0.977	0.983	0.979	0.968
	Z	-0.871	***-3.482***	-1.741	**-3.772**	**-7.255**	0.000	**-4.132**	-2.225	***-3.496***	**-6.992**
350	*UCP*	0.940	0.923	0.937	0.929	0.911	0.983	0.988	0.988	0.987	0.986
	Z	-1.451	**-3.918**	-1.886	***-3.047***	**-5.659**	-2.225	-0.636	-0.636	-0.953	-1.271
500	*UCP*	0.938	0.932	0.932	0.934	0.896	0.986	0.981	0.985	0.990	0.978
	Z	-1.741	-2.612	-2.612	-2.322	**-7.835**	-1.271	-2.860	-1.589	0.000	**-3.814**

BIC		1-α=.95					1-α=.99				
N	***CORR***	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**	**0.1**	**0.3**	**0.5**	**0.7**	**0.9**
50	*UCP*	0.915	0.919	0.907	0.897	0.780	0.971	0.973	0.969	0.957	0.901
	Z	**-5.078**	**-4.498**	**-6.239**	**-7.690**	**-24.666**	**-6.039**	**-5.403**	**-6.674**	**-10.488**	**-28.286**
100	*UCP*	0.940	0.936	0.937	0.912	0.850	0.980	0.983	0.985	0.969	0.912
	Z	-1.451	-2.031	-1.886	**-5.514**	**-14.510**	***-3.178***	-2.225	-1.589	**-6.674**	**-24.790**
200	*UCP*	0.953	0.937	0.946	0.933	0.899	0.991	0.982	0.987	0.979	0.947
	Z	0.435	-1.886	-0.580	-2.467	**-7.400**	0.318	-2.543	-0.953	***-3.496***	**-13.666**
350	*UCP*	0.956	0.939	0.950	0.949	0.929	0.986	0.988	0.991	0.990	0.981
	Z	0.871	-1.596	0.000	-0.145	***-3.047***	-1.271	-0.636	0.318	0.000	-2.860
500	*UCP*	0.947	0.946	0.948	0.951	0.927	0.988	0.987	0.990	0.991	0.983
	Z	-0.435	-0.580	-0.290	0.145	***-3.337***	-0.636	-0.953	0.000	0.318	-2.225

Note:
(1) *N*=Sample Size. *CORR* = Covariate Correlation. *UCP*=Unconditional Coverage Probability (Overall Coverage Probability). *Z* = z-score based upon normal approximation to binomial distribution with the null hypothesis that the Actual Coverage Probability is equal to the Nominal Coverage Probability.

(2) The Unconditional Coverage Probabilities are obtained across 1000 iterations at the nominal level of .95 and .99.

(3) **Bold** values (other than the Column Titles) indicate *Z* scores smaller than or equal to -3.5; ***bold italic*** values indicate z-scores between -3.0 and -3.5.

Table 12: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.1, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***
N=50	T_s	121	**506**	373	121	**506**	373	239	**196**	565	239	**196**	565
	T_v	110	464	298	117	489	345	225	163	457	234	183	528
	CCP	0.909	***0.917***	0.799	0.967	***0.966***	0.925	0.941	***0.832***	0.809	0.979	***0.934***	0.935
N=100	T_s	18	**841**	141	18	**841**	141	76	**571**	353	76	**571**	353
	T_v	18	788	81	18	831	111	74	527	184	76	561	267
	CCP	1.000	***0.937***	0.574	1.000	***0.988***	0.787	0.974	***0.923***	0.521	1.000	***0.982***	0.756
N=200	T_s	0	**980**	20	0	**980**	20	6	**858**	136	6	**858**	136
	T_v	0	920	9	0	965	15	6	804	49	6	843	80
	CCP	.	***0.939***	0.450	.	***0.985***	0.750	1.000	***0.937***	0.360	1.000	***0.983***	0.588
N=350	T_s	0	**999**	1	0	**999**	1	0	**987**	13	0	**987**	13
	T_v	0	945	0	0	987	0	0	935	1	0	976	2
	CCP	.	***0.946***	0.000	.	***0.988***	0.000	.	***0.947***	0.077	.	***0.989***	0.154
N=500	T_s	0	**1000**	0	0	**1000**	0	0	**1000**	0	0	**1000**	0
	T_v	0	942	0	0	987	0	0	942	0	0	987	0
	CCP	.	***0.942***	.	.	***0.987***	.	.	***0.942***	.	.	***0.987***	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 13: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.3, Offset=0)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	164	**342**	494	164	**342**	494	250	**111**	639	250	**111**	639
	T_v	151	310	400	161	333	453	232	88	531	247	103	595
	CCP	0.921	***0.906***	0.810	0.982	***0.974***	0.917	0.928	***0.793***	0.831	0.988	***0.928***	0.931
N=100	T_s	31	**727**	242	31	**727**	242	121	**388**	491	121	**388**	491
	T_v	28	699	183	28	722	201	111	367	320	118	383	392
	CCP	0.903	***0.961***	0.756	0.903	***0.993***	0.831	0.917	***0.946***	0.652	0.975	***0.987***	0.798
N=200	T_s	0	**935**	65	0	**935**	65	10	**754**	236	10	**754**	236
	T_v	0	892	32	0	929	46	9	715	106	10	748	159
	CCP	.	***0.954***	0.492	.	***0.994***	0.708	0.900	***0.948***	0.449	1.000	***0.992***	0.674
N=350	T_s	0	**999**	1	0	**999**	1	0	**948**	52	0	**948**	52
	T_v	0	942	0	0	990	0	0	898	4	0	940	15
	CCP	.	***0.943***	0.000	.	***0.991***	0.000	.	***0.947***	0.077	.	***0.992***	0.288
N=500	T_s	0	**999**	1	0	**999**	1	0	**989**	11	0	**989**	11
	T_v	0	961	0	0	993	0	0	953	1	0	984	3
	CCP	.	***0.962***	0.000	.	***0.994***	0.000	.	***0.964***	0.091	.	***0.995***	0.273

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 14: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.5, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	***T_s***	200	**246**	554	200	**246**	554	280	**38**	682	280	**38**	682
	T_v	177	220	446	192	238	514	246	24	566	268	33	640
	CCP	0.885	***0.894***	0.805	0.960	***0.967***	0.928	0.879	***0.632***	0.830	0.957	***0.868***	0.938
N=100	***T_s***	51	**573**	376	51	**573**	376	185	**177**	638	185	**177**	638
	T_v	42	542	264	49	564	321	154	159	433	176	169	538
	CCP	0.824	***0.946***	0.702	0.961	***0.984***	0.854	0.832	***0.898***	0.679	0.951	***0.955***	0.843
N=200	***T_s***	0	**848**	152	0	**848**	152	20	**575**	405	20	**575**	405
	T_v	0	814	74	0	840	109	17	547	167	17	570	255
	CCP	.	***0.960***	0.487	.	***0.991***	0.717	0.850	***0.951***	0.412	0.850	***0.991***	0.630
N=350	***T_s***	0	**961**	39	0	**961**	39	0	**781**	219	0	**781**	219
	T_v	0	908	11	0	956	17	0	742	44	0	778	69
	CCP	.	***0.945***	0.282	.	***0.995***	0.436	.	***0.950***	0.201	.	***0.996***	0.315
N=500	***T_s***	0	**996**	4	0	**996**	4	0	**916**	84	0	**916**	84
	T_v	0	949	0	0	987	1	0	882	3	0	909	17
	CCP	.	***0.953***	0.000	.	***0.991***	0.250	.	***0.963***	0.036	.	***0.992***	0.202

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 15: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.7, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	***T_s***	308	**83**	609	308	**83**	609	335	**12**	653	335	**12**	653
	T_v	252	55	513	289	72	571	265	0	554	312	7	614
	CCP	0.818	***0.663***	0.842	0.938	***0.867***	0.938	0.791	***0.000***	0.848	0.931	***0.583***	0.940
N=100	***T_s***	149	**314**	537	149	**314**	537	263	**21**	716	263	**21**	716
	T_v	101	291	373	126	304	461	166	15	514	223	18	625
	CCP	0.678	***0.927***	0.695	0.846	***0.968***	0.858	0.631	***0.714***	0.718	0.848	***0.857***	0.873
N=200	***T_s***	14	**678**	308	14	**678**	308	112	**236**	652	112	**236**	652
	T_v	7	637	143	11	672	207	62	212	271	88	233	403
	CCP	0.500	***0.940***	0.464	0.786	***0.991***	0.672	0.554	***0.898***	0.416	0.786	***0.987***	0.618
N=350	***T_s***	0	**876**	124	0	**876**	124	15	**513**	472	15	**513**	472
	T_v	0	848	30	0	869	50	2	488	93	9	506	170
	CCP	.	***0.968***	0.242	.	***0.992***	0.403	0.133	***0.951***	0.197	0.600	***0.986***	0.360
N=500	***T_s***	0	**940**	60	0	**940**	60	0	**695**	305	0	**695**	305
	T_v	0	911	7	0	933	14	0	671	22	0	689	58
	CCP	.	***0.969***	0.117	.	***0.993***	0.233	.	***0.965***	0.072	.	***0.991***	0.190

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 16: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.9, Offset=0)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	367	**56**	577	367	**56**	577	376	**18**	606	376	**18**	606
	T_v	258	30	486	327	43	548	263	4	512	334	9	575
	CCP	0.703	***0.536***	0.842	0.891	***0.768***	0.950	0.699	***0.222***	0.845	0.888	***0.500***	0.949
N=100	T_s	350	**22**	628	350	**22**	628	354	**5**	641	354	**5**	641
	T_v	179	7	444	269	16	547	179	0	452	269	2	560
	CCP	0.511	***0.318***	0.707	0.769	***0.727***	0.871	0.506	***0.000***	0.705	0.760	***0.400***	0.874
N=200	T_s	238	**144**	618	238	**144**	618	290	**0**	710	290	**0**	710
	T_v	62	137	286	133	141	406	65	0	349	151	0	489
	CCP	0.261	***0.951***	0.463	0.559	***0.979***	0.657	0.224	***.***	0.492	0.521	***.***	0.689
N=350	T_s	86	**415**	499	86	**415**	499	228	**1**	771	228	**1**	771
	T_v	7	406	125	15	415	223	17	1	192	34	1	351
	CCP	0.081	***0.978***	0.251	0.174	***1.000***	0.447	0.075	***1.000***	0.249	0.149	***1.000***	0.455
N=500	T_s	26	**583**	391	26	**583**	391	192	**38**	770	192	**38**	770
	T_v	1	562	50	1	580	108	1	32	94	7	37	212
	CCP	0.038	***0.964***	0.128	0.038	***0.995***	0.276	0.005	***0.842***	0.122	0.036	***0.974***	0.275

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 17: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.1, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**824**	164	12	**824**	164	12	**933**	46	21	**933**	46	21
	T_v	775	150	6	813	160	7	878	41	11	919	46	15
	CCP	***0.941***	0.915	0.500	***0.987***	0.976	0.583	***0.941***	0.891	0.524	***0.985***	1.000	0.714
N=100	T_s	**831**	169	0	**831**	169	0	**964**	35	1	**964**	35	1
	T_v	782	155	0	824	166	0	911	28	0	957	33	0
	CCP	***0.941***	0.917	.	***0.992***	0.982	.	***0.945***	0.800	0.000	***0.993***	0.943	0.000
N=200	T_s	**844**	156	0	**844**	156	0	**978**	22	0	**978**	22	0
	T_v	798	151	0	836	156	0	927	21	0	970	22	0
	CCP	***0.945***	0.968	.	***0.991***	1.000	.	***0.948***	0.955	.	***0.992***	1.000	.
N=350	T_s	**849**	151	0	**849**	151	0	**977**	23	0	**977**	23	0
	T_v	814	146	0	843	149	0	938	22	0	969	23	0
	CCP	***0.959***	0.967	.	***0.993***	0.987	.	***0.960***	0.957	.	***0.992***	1.000	.
N=500	T_s	**839**	161	0	**839**	161	0	**982**	18	0	**982**	18	0
	T_v	790	148	0	827	158	0	925	17	0	967	18	0
	CCP	***0.942***	0.919	.	***0.986***	0.981	.	***0.942***	0.944	.	***0.985***	1.000	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 18: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_2$, Corr=0.3, Offset=0)

		AIC						BIC					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**790**	174	36	**790**	174	36	**904**	47	49	**904**	47	49
	T_v	741	150	4	782	165	9	848	35	13	894	41	20
	CCP	***0.938***	0.862	0.111	***0.990***	0.948	0.250	***0.938***	0.745	0.265	***0.989***	0.872	0.408
N=100	T_s	**832**	168	0	**832**	168	0	**959**	32	9	**959**	32	9
	T_v	791	156	0	828	166	0	914	29	0	955	31	1
	CCP	***0.951***	0.929	.	***0.995***	0.988	.	***0.953***	0.906	0.000	***0.996***	0.969	0.111
N=200	T_s	**845**	155	0	**845**	155	0	**980**	20	0	**980**	20	0
	T_v	792	141	0	835	153	0	919	18	0	968	20	0
	CCP	***0.937***	0.910	.	***0.988***	0.987	.	***0.938***	0.900	.	***0.988***	1.000	.
N=350	T_s	**833**	167	0	**833**	167	0	**982**	18	0	**982**	18	0
	T_v	791	156	0	825	162	0	933	16	0	974	17	0
	CCP	***0.950***	0.934	.	***0.990***	0.970	.	***0.950***	0.889	.	***0.992***	0.944	.
N=500	T_s	**851**	149	0	**851**	149	0	**992**	8	0	**992**	8	0
	T_v	811	141	0	845	149	0	949	6	0	986	8	0
	CCP	***0.953***	0.946	.	***0.993***	1.000	.	***0.957***	0.750	.	***0.994***	1.000	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 19: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_2$, Corr=0.5, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	T_s	**820**	116	64	**820**	116	64	**887**	36	77	**887**	36	77
	T_v	773	97	28	807	110	44	833	27	40	874	33	56
	CCP	***0.943***	0.836	0.438	***0.984***	0.948	0.688	***0.939***	0.750	0.519	***0.985***	0.917	0.727
N=100	T_s	**836**	155	9	**836**	155	9	**957**	21	22	**957**	21	22
	T_v	792	139	1	828	149	1	907	16	3	949	19	5
	CCP	***0.947***	0.897	0.111	***0.990***	0.961	0.111	***0.948***	0.762	0.136	***0.992***	0.905	0.227
N=200	T_s	**842**	158	0	**842**	158	0	**978**	20	2	**978**	20	2
	T_v	797	133	0	835	150	0	918	18	0	967	20	0
	CCP	***0.947***	0.842	.	***0.992***	0.949	.	***0.939***	0.900	0.000	***0.989***	1.000	0.000
N=350	T_s	**827**	173	0	**827**	173	0	**987**	13	0	**987**	13	0
	T_v	800	161	0	820	171	0	953	11	0	978	13	0
	CCP	***0.967***	0.931	.	***0.992***	0.988	.	***0.966***	0.846	.	***0.991***	1.000	.
N=500	T_s	**848**	152	0	**848**	152	0	**993**	7	0	**993**	7	0
	T_v	803	143	0	840	149	0	940	6	0	982	7	0
	CCP	***0.947***	0.941	.	***0.991***	0.980	.	***0.947***	0.857	.	***0.989***	1.000	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 20: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.7, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**770**	97	133	**770**	97	133	**832**	32	136	**832**	32	136
	T_v	723	81	54	761	92	84	782	22	57	822	28	87
	CCP	***0.939***	0.835	0.406	***0.988***	0.948	0.632	***0.940***	0.688	0.419	***0.988***	0.875	0.640
N=100	T_s	**829**	123	48	**829**	123	48	**916**	21	63	**916**	21	63
	T_v	780	106	2	820	116	5	862	11	5	907	18	14
	CCP	***0.941***	0.862	0.042	***0.989***	0.943	0.104	***0.941***	0.524	0.079	***0.990***	0.857	0.222
N=200	T_s	**815**	184	1	**815**	184	1	**962**	21	17	**962**	21	17
	T_v	776	153	0	805	176	0	916	17	0	951	20	0
	CCP	***0.952***	0.832	0.000	***0.988***	0.957	0.000	***0.952***	0.810	0.000	***0.989***	0.952	0.000
N=350	T_s	**831**	169	0	**831**	169	0	**981**	17	2	**981**	17	2
	T_v	780	148	0	821	161	0	921	13	0	970	14	0
	CCP	***0.939***	0.876	.	***0.988***	0.953	.	***0.939***	0.765	0.000	***0.989***	0.824	0.000
N=500	T_s	**856**	144	0	**856**	144	0	**989**	11	0	**989**	11	0
	T_v	824	118	0	852	136	0	949	6	0	984	9	0
	CCP	***0.963***	0.819	.	***0.995***	0.944	.	***0.960***	0.545	.	***0.995***	0.818	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 21: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.9, Offset=0)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**647**	87	266	**647**	87	266	**707**	23	270	**707**	23	270
	T_v	605	60	91	634	80	147	664	11	94	693	18	150
	CCP	***0.935***	0.690	0.342	***0.980***	0.920	0.553	***0.939***	0.478	0.348	***0.980***	0.783	0.556
N=100	T_s	**691**	85	224	**691**	85	224	**752**	22	226	**752**	22	226
	T_v	653	48	33	683	75	52	708	8	33	741	17	52
	CCP	***0.945***	0.565	0.147	***0.988***	0.882	0.232	***0.941***	0.364	0.146	***0.985***	0.773	0.230
N=200	T_s	**792**	93	115	**792**	93	115	**876**	7	117	**876**	7	117
	T_v	750	77	0	783	88	1	829	5	0	867	7	1
	CCP	***0.947***	0.828	0.000	***0.989***	0.946	0.009	***0.946***	0.714	0.000	***0.990***	1.000	0.009
N=350	T_s	**860**	105	35	**860**	105	35	**939**	8	53	**939**	8	53
	T_v	808	85	0	850	100	0	879	4	0	928	6	0
	CCP	***0.940***	0.810	0.000	***0.988***	0.952	0.000	***0.936***	0.500	0.000	***0.988***	0.750	0.000
N=500	T_s	**847**	147	6	**847**	147	6	**965**	8	27	**965**	8	27
	T_v	810	110	0	838	136	0	920	2	0	954	3	0
	CCP	***0.956***	0.748	0.000	***0.989***	0.925	0.000	***0.953***	0.250	0.000	***0.989***	0.375	0.000

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 22: Count on Model Selection and Conditional Coverage Probability (the correct model = M_3, Corr=0.1, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	1	186	**813**	1	186	**813**	5	64	**931**	5	64	**931**
	T_v	1	140	770	1	169	805	4	45	871	5	54	918
	CCP	1.000	0.753	***0.947***	1.000	0.909	***0.990***	0.800	0.703	***0.936***	1.000	0.844	***0.986***
N=100	T_s	0	183	**817**	0	183	**817**	0	35	**965**	0	35	**965**
	T_v	0	154	770	0	178	806	0	27	907	0	33	952
	CCP	.	0.842	***0.942***	.	0.973	***0.987***	.	0.771	***0.940***	.	0.943	***0.987***
N=200	T_s	0	173	**827**	0	173	**827**	0	23	**977**	0	23	**977**
	T_v	0	154	782	0	165	813	0	16	923	0	20	958
	CCP	.	0.890	***0.946***	.	0.954	***0.983***	.	0.696	***0.945***	.	0.870	***0.981***
N=350	T_s	0	161	**839**	0	161	**839**	0	17	**983**	0	17	**983**
	T_v	0	145	791	0	156	829	0	14	925	0	17	972
	CCP	.	0.901	***0.943***	.	0.969	***0.988***	.	0.824	***0.941***	.	1.000	***0.989***
N=500	T_s	0	160	**840**	0	160	**840**	0	16	**984**	0	16	**984**
	T_v	0	146	805	0	158	833	0	14	941	0	16	975
	CCP	.	0.913	***0.958***	.	0.988	***0.992***	.	0.875	***0.956***	.	1.000	***0.991***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 23: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.3, Offset=0)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	4	164	**832**	4	164	**832**	8	54	**938**	8	54	**938**
	T_v	1	141	775	4	155	823	5	40	875	8	48	929
	CCP	0.250	0.860	***0.931***	1.000	0.945	***0.989***	0.625	0.741	***0.933***	1.000	0.889	***0.990***
N=100	T_s	0	167	**833**	0	167	**833**	2	42	**956**	2	42	**956**
	T_v	0	138	789	0	158	823	1	26	902	2	33	942
	CCP	.	0.826	***0.947***	.	0.946	***0.988***	0.500	0.619	***0.944***	1.000	0.786	***0.985***
N=200	T_s	0	152	**848**	0	152	**848**	0	18	**982**	0	18	**982**
	T_v	0	125	795	0	145	828	0	13	923	0	16	961
	CCP	.	0.822	***0.938***	.	0.954	***0.976***	.	0.722	***0.940***	.	0.889	***0.979***
N=350	T_s	0	180	**820**	0	180	**820**	0	19	**981**	0	19	**981**
	T_v	0	159	784	0	176	812	0	15	940	0	17	972
	CCP	.	0.883	***0.956***	.	0.978	***0.990***	.	0.789	***0.958***	.	0.895	***0.991***
N=500	T_s	0	175	**825**	0	175	**825**	0	14	**986**	0	14	**986**
	T_v	0	155	797	0	169	822	0	10	949	0	13	980
	CCP	.	0.886	***0.966***	.	0.966	***0.996***	.	0.714	***0.962***	.	0.929	***0.994***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 24: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.5, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	***T_s***	10	174	**816**	10	174	**816**	18	47	**935**	18	47	**935**
	T_v	7	135	761	8	162	808	11	27	872	16	37	923
	CCP	0.700	0.776	***0.933***	0.800	0.931	***0.990***	0.611	0.574	***0.933***	0.889	0.787	***0.987***
N=100	***T_s***	0	160	**840**	0	160	**840**	2	30	**968**	2	30	**968**
	T_v	0	133	791	0	151	829	2	20	909	2	25	954
	CCP	.	0.831	***0.942***	.	0.944	***0.987***	1.000	0.667	***0.939***	1.000	0.833	***0.986***
N=200	***T_s***	0	153	**847**	0	153	**847**	0	28	**972**	0	28	**972**
	T_v	0	123	801	0	143	839	0	20	917	0	24	963
	CCP	.	0.804	***0.946***	.	0.935	***0.991***	.	0.714	***0.943***	.	0.857	***0.991***
N=350	***T_s***	0	165	**835**	0	165	**835**	0	17	**983**	0	17	**983**
	T_v	0	133	793	0	155	830	0	9	930	0	14	976
	CCP	.	0.806	***0.950***	.	0.939	***0.994***	.	0.529	***0.946***	.	0.824	***0.993***
N=500	***T_s***	0	141	**859**	0	141	**859**	0	9	**991**	0	9	**991**
	T_v	0	122	817	0	133	849	0	5	941	0	8	980
	CCP	.	0.865	***0.951***	.	0.943	***0.988***	.	0.556	***0.950***	.	0.889	***0.989***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 25: Count on Model Selection and Conditional Coverage Probability (the correct model = M_3, Corr=0.7, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	45	160	**795**	45	160	**795**	60	55	**885**	60	55	**885**
	T_v	25	121	748	40	141	784	30	33	827	52	42	871
	CCP	0.556	0.756	***0.941***	0.889	0.881	***0.986***	0.500	0.600	***0.934***	0.867	0.764	***0.984***
N=100	T_s	3	182	**815**	3	182	**815**	13	39	**948**	13	39	**948**
	T_v	2	140	772	3	170	805	3	26	895	6	34	935
	CCP	0.667	0.769	***0.947***	1.000	0.934	***0.988***	0.231	0.667	***0.944***	0.462	0.872	***0.986***
N=200	T_s	0	160	**840**	0	160	**840**	0	23	**977**	0	23	**977**
	T_v	0	131	793	0	155	837	0	17	920	0	22	971
	CCP	.	0.819	***0.944***	.	0.969	***0.996***	.	0.739	***0.942***	.	0.957	***0.994***
N=350	T_s	0	161	**839**	0	161	**839**	0	8	**992**	0	8	**992**
	T_v	0	123	798	0	151	824	0	6	945	0	6	976
	CCP	.	0.764	***0.951***	.	0.938	***0.982***	.	0.750	***0.953***	.	0.750	***0.984***
N=500	T_s	0	181	**819**	0	181	**819**	0	9	**991**	0	9	**991**
	T_v	0	146	784	0	173	812	0	5	952	0	7	983
	CCP	.	0.807	***0.957***	.	0.956	***0.991***	.	0.556	***0.961***	.	0.778	***0.992***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 26: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.9, Offset=0)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	179	107	**714**	179	107	**714**	185	41	**774**	185	41	**774**
	T_v	64	68	672	110	92	702	64	19	729	110	32	761
	CCP	0.358	0.636	***0.941***	0.615	0.860	***0.983***	0.346	0.463	***0.942***	0.595	0.780	***0.983***
N=100	T_s	89	97	**814**	89	97	**814**	97	15	**888**	97	15	**888**
	T_v	9	65	776	23	87	807	9	6	845	23	9	878
	CCP	0.101	0.670	***0.953***	0.258	0.897	***0.991***	0.093	0.400	***0.952***	0.237	0.600	***0.989***
N=200	T_s	14	163	**823**	14	163	**823**	40	13	**947**	40	13	**947**
	T_v	0	113	782	0	152	813	0	4	895	1	9	935
	CCP	0.000	0.693	***0.950***	0.000	0.933	***0.988***	0.000	0.308	***0.945***	0.025	0.692	***0.987***
N=350	T_s	0	160	**840**	0	160	**840**	2	13	**985**	2	13	**985**
	T_v	0	121	803	0	149	831	0	1	940	0	7	976
	CCP	.	0.756	***0.956***	.	0.931	***0.989***	0.000	0.077	***0.954***	0.000	0.538	***0.991***
N=500	T_s	0	173	**827**	0	173	**827**	1	10	**989**	1	10	**989**
	T_v	0	117	777	0	164	819	0	0	932	0	5	981
	CCP	.	0.676	***0.940***	.	0.948	***0.990***	0.000	0.000	***0.942***	0.000	0.500	***0.992***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 27: Count on Model Selection and Conditional Coverage Probability (the correct model = M_1, Corr=0.1, Offset=0.8473)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***
N=50	T_s	155	**479**	366	155	**479**	366	278	**182**	540	278	**182**	540
	T_v	150	419	294	153	467	353	271	130	443	275	170	520
	CCP	0.968	***0.875***	0.803	0.987	***0.975***	0.964	0.975	***0.714***	0.820	0.989	***0.934***	0.963
N=100	T_s	20	**828**	152	20	**828**	152	78	**518**	404	78	**518**	404
	T_v	15	790	95	19	822	121	70	488	257	76	512	328
	CCP	0.750	***0.954***	0.625	0.950	***0.993***	0.796	0.897	***0.942***	0.636	0.974	***0.988***	0.812
N=200	T_s	0	**973**	27	0	**973**	27	5	**842**	153	5	**842**	153
	T_v	0	932	6	0	964	13	5	810	58	5	835	90
	CCP	.	***0.958***	0.222	.	***0.991***	0.481	1.000	***0.962***	0.379	1.000	***0.992***	0.588
N=350	T_s	0	**1000**	0	0	**1000**	0	0	**988**	12	0	**988**	12
	T_v	0	939	0	0	988	0	0	932	1	0	976	3
	CCP	.	***0.939***	.	.	***0.988***	.	.	***0.943***	0.083	.	***0.988***	0.250
N=500	T_s	0	**1000**	0	0	**1000**	0	0	**998**	2	0	**998**	2
	T_v	0	945	0	0	987	0	0	944	0	0	986	0
	CCP	.	***0.945***	.	.	***0.987***	.	.	***0.946***	0.000	.	***0.988***	0.000

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 28: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.3, Offset=0.8473)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	182	**361**	457	182	**361**	457	292	**109**	599	292	**109**	599
	T_v	162	325	367	175	344	426	266	83	493	284	95	566
	CCP	0.890	***0.900***	0.803	0.962	***0.953***	0.932	0.911	***0.761***	0.823	0.973	***0.872***	0.945
N=100	T_s	34	**701**	265	34	**701**	265	122	**357**	521	122	**357**	521
	T_v	30	659	173	33	692	227	112	321	335	118	349	443
	CCP	0.882	***0.940***	0.653	0.971	***0.987***	0.857	0.918	***0.899***	0.643	0.967	***0.978***	0.850
N=200	T_s	0	**936**	64	0	**936**	64	4	**730**	266	4	**730**	266
	T_v	0	899	25	0	932	42	4	701	102	4	726	163
	CCP	.	***0.960***	0.391	.	***0.996***	0.656	1.000	***0.960***	0.383	1.000	***0.995***	0.613
N=350	T_s	0	**992**	8	0	**992**	8	0	**931**	69	0	**931**	69
	T_v	0	944	1	0	983	2	0	890	8	0	923	27
	CCP	.	***0.952***	0.125	.	***0.991***	0.250	.	***0.956***	0.116	.	***0.991***	0.391
N=500	T_s	0	**1000**	0	0	**1000**	0	0	**982**	18	0	**982**	18
	T_v	0	953	0	0	988	0	0	939	3	0	972	5
	CCP	.	***0.953***	.	.	***0.988***	.	.	***0.956***	0.167	.	***0.990***	0.278

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 29: Count on Model Selection and Conditional Coverage Probability (the correct model = M_1, Corr=0.5, Offset=0.8473)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	215	**223**	562	215	**223**	562	281	**46**	673	281	**46**	673
	T_v	194	187	465	209	211	527	245	26	568	270	38	637
	CCP	0.902	***0.839***	0.827	0.972	***0.946***	0.938	0.872	***0.565***	0.844	0.961	***0.826***	0.947
N=100	T_s	45	**572**	383	45	**572**	383	158	**191**	651	158	**191**	651
	T_v	35	542	265	41	563	339	134	169	465	150	184	579
	CCP	0.778	***0.948***	0.692	0.911	***0.984***	0.885	0.848	***0.885***	0.714	0.949	***0.963***	0.889
N=200	T_s	5	**842**	153	5	**842**	153	31	**532**	437	31	**532**	437
	T_v	3	804	68	4	832	101	23	499	181	30	522	276
	CCP	0.600	***0.955***	0.444	0.800	***0.988***	0.660	0.742	***0.938***	0.414	0.968	***0.981***	0.632
N=350	T_s	0	**955**	45	0	**955**	45	2	**781**	217	2	**781**	217
	T_v	0	905	13	0	946	19	2	745	57	2	773	90
	CCP	.	***0.948***	0.289	.	***0.991***	0.422	1.000	***0.954***	0.263	1.000	***0.990***	0.415
N=500	T_s	0	**996**	4	0	**996**	4	0	**915**	85	0	**915**	85
	T_v	0	945	0	0	987	0	0	873	6	0	907	23
	CCP	.	***0.949***	0.000	.	***0.991***	0.000	.	***0.954***	0.071	.	***0.991***	0.271

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 30: Count on Model Selection and Conditional Coverage Probability (the correct model = M_1, Corr=0.7, Offset=0.8473)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	311	**90**	599	311	**90**	599	343	**7**	650	343	**7**	650
	T_v	235	66	519	291	82	572	250	1	568	318	3	622
	CCP	0.756	***0.733***	0.866	0.936	***0.911***	0.955	0.729	***0.143***	0.874	0.927	***0.429***	0.957
N=100	T_s	165	**306**	529	165	**306**	529	285	**21**	694	285	**21**	694
	T_v	118	278	372	148	299	454	188	14	501	244	19	612
	CCP	0.715	***0.908***	0.703	0.897	***0.977***	0.858	0.660	***0.667***	0.722	0.856	***0.905***	0.882
N=200	T_s	23	**642**	335	23	**642**	335	144	**200**	656	144	**200**	656
	T_v	12	610	156	17	633	231	73	181	311	117	196	453
	CCP	0.522	***0.950***	0.466	0.739	***0.986***	0.690	0.507	***0.905***	0.474	0.813	***0.980***	0.691
N=350	T_s	1	**852**	147	1	**852**	147	15	**504**	481	15	**504**	481
	T_v	0	810	34	1	847	66	0	473	97	6	500	188
	CCP	0.000	***0.951***	0.231	1.000	***0.994***	0.449	0.000	***0.938***	0.202	0.400	***0.992***	0.391
N=500	T_s	0	**938**	62	0	**938**	62	2	**653**	345	2	**653**	345
	T_v	0	891	6	0	930	14	1	615	31	2	645	67
	CCP	.	***0.950***	0.097	.	***0.991***	0.226	0.500	***0.942***	0.090	1.000	***0.988***	0.194

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 31: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.9, Offset=0.8473)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	392	**56**	552	392	**56**	552	401	**10**	589	401	**10**	589
	T_v	266	27	469	344	47	518	269	3	500	348	7	551
	CCP	0.679	***0.482***	0.850	0.878	***0.839***	0.938	0.671	***0.300***	0.849	0.868	***0.700***	0.935
N=100	T_s	337	**29**	634	337	**29**	634	343	**2**	655	343	**2**	655
	T_v	178	13	455	270	21	539	181	0	470	273	0	555
	CCP	0.528	***0.448***	0.718	0.801	***0.724***	0.850	0.528	***0.000***	0.718	0.796	***0.000***	0.847
N=200	T_s	273	**125**	602	273	**125**	602	333	**0**	667	333	**0**	667
	T_v	60	116	293	131	122	412	65	0	335	147	0	469
	CCP	0.220	***0.928***	0.487	0.480	***0.976***	0.684	0.195	***.***	0.502	0.441	***.***	0.703
N=350	T_s	101	**416**	483	101	**416**	483	248	**2**	750	248	**2**	750
	T_v	10	396	125	27	409	212	17	0	202	49	2	351
	CCP	0.099	***0.952***	0.259	0.267	***0.983***	0.439	0.069	***0.000***	0.269	0.198	***1.000***	0.468
N=500	T_s	41	**568**	391	41	**568**	391	185	**25**	790	185	**25**	790
	T_v	2	550	44	10	566	97	3	20	84	20	25	208
	CCP	0.049	***0.968***	0.113	0.244	***0.996***	0.248	0.016	***0.800***	0.106	0.108	***1.000***	0.263

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 32: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_2$, Corr=0.1, Offset=0.8473)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	***T_s***	**833**	159	8	**833**	159	8	**942**	40	18	**942**	40	18
	T_v	786	146	5	825	155	5	892	34	12	934	38	14
	CCP	***0.944***	0.918	0.625	***0.990***	0.975	0.625	***0.947***	0.850	0.667	***0.992***	0.950	0.778
N=100	***T_s***	**825**	175	0	**825**	175	0	**964**	35	1	**964**	35	1
	T_v	777	160	0	815	170	0	907	31	0	952	33	0
	CCP	***0.942***	0.914	.	***0.988***	0.971	.	***0.941***	0.886	0.000	***0.988***	0.943	0.000
N=200	***T_s***	**848**	152	0	**848**	152	0	**984**	16	0	**984**	16	0
	T_v	804	141	0	837	150	0	943	14	0	974	16	0
	CCP	***0.948***	0.928	.	***0.987***	0.987	.	***0.958***	0.875	.	***0.990***	1.000	.
N=350	***T_s***	**832**	168	0	**832**	168	0	**984**	16	0	**984**	16	0
	T_v	796	160	0	822	168	0	943	14	0	974	16	0
	CCP	***0.957***	0.952	.	***0.988***	1.000	.	***0.958***	0.875	.	***0.990***	1.000	.
N=500	***T_s***	**850**	150	0	**850**	150	0	**989**	11	0	**989**	11	0
	T_v	803	143	0	836	148	0	936	11	0	973	11	0
	CCP	***0.945***	0.953	.	***0.984***	0.987	.	***0.946***	1.000	.	***0.984***	1.000	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 33: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_2$, Corr=0.3, Offset=0.8473)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***
N=50	T_s	**805**	154	41	**805**	154	41	**910**	35	55	**910**	35	55
	T_v	764	138	21	795	149	26	862	29	32	899	32	39
	CCP	***0.949***	0.896	0.512	***0.988***	0.968	0.634	***0.947***	0.829	0.582	***0.988***	0.914	0.709
N=100	T_s	**817**	182	1	**817**	182	1	**964**	24	12	**964**	24	12
	T_v	777	167	0	808	176	0	914	20	3	952	21	6
	CCP	***0.951***	0.918	0.000	***0.989***	0.967	0.000	***0.948***	0.833	0.250	***0.988***	0.875	0.500
N=200	T_s	**847**	153	0	**847**	153	0	**984**	16	0	**984**	16	0
	T_v	806	145	0	838	151	0	937	15	0	974	15	0
	CCP	***0.952***	0.948	.	***0.989***	0.987	.	***0.952***	0.938	.	***0.990***	0.938	.
N=350	T_s	**836**	164	0	**836**	164	0	**984**	16	0	**984**	16	0
	T_v	788	149	0	826	160	0	922	15	0	973	16	0
	CCP	***0.943***	0.909	.	***0.988***	0.976	.	***0.937***	0.938	.	***0.989***	1.000	.
N=500	T_s	**831**	169	0	**831**	169	0	**985**	15	0	**985**	15	0
	T_v	789	160	0	821	163	0	935	13	0	974	13	0
	CCP	***0.949***	0.947	.	***0.988***	0.964	.	***0.949***	0.867	.	***0.989***	0.867	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 34: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_2$, Corr=0.5, Offset=0.8473)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**835**	98	67	**835**	98	67	**894**	33	73	**894**	33	73
	T_v	781	84	31	820	91	46	836	27	35	878	29	50
	CCP	***0.935***	0.857	0.463	***0.982***	0.929	0.687	***0.935***	0.818	0.479	***0.982***	0.879	0.685
N=100	T_s	**847**	139	14	**847**	139	14	**949**	22	29	**949**	22	29
	T_v	796	124	0	840	133	2	894	15	4	942	20	7
	CCP	***0.940***	0.892	0.000	***0.992***	0.957	0.143	***0.942***	0.682	0.138	***0.993***	0.909	0.241
N=200	T_s	**818**	182	0	**818**	182	0	**967**	32	1	**967**	32	1
	T_v	768	161	0	811	177	0	907	30	0	956	32	0
	CCP	***0.939***	0.885	.	***0.991***	0.973	.	***0.938***	0.938	0.000	***0.989***	1.000	0.000
N=350	T_s	**835**	165	0	**835**	165	0	**983**	17	0	**983**	17	0
	T_v	775	155	0	828	163	0	920	17	0	976	17	0
	CCP	***0.928***	0.939	.	***0.992***	0.988	.	***0.936***	1.000	.	***0.993***	1.000	.
N=500	T_s	**858**	142	0	**858**	142	0	**989**	11	0	**989**	11	0
	T_v	814	132	0	846	139	0	939	8	0	977	10	0
	CCP	***0.949***	0.930	.	***0.986***	0.979	.	***0.949***	0.727	.	***0.988***	0.909	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 35: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.7, Offset=0.8473)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***
N=50	T_s	**754**	103	143	**754**	103	143	**811**	43	146	**811**	43	146
	T_v	715	74	59	747	93	86	770	21	62	802	35	89
	CCP	***0.948***	0.718	0.413	***0.991***	0.903	0.601	***0.949***	0.488	0.425	***0.989***	0.814	0.610
N=100	T_s	**815**	129	56	**815**	129	56	**911**	18	71	**911**	18	71
	T_v	760	118	7	806	123	10	850	14	13	901	15	18
	CCP	***0.933***	0.915	0.125	***0.989***	0.953	0.179	***0.933***	0.778	0.183	***0.989***	0.833	0.254
N=200	T_s	**847**	148	5	**847**	148	5	**969**	12	19	**969**	12	19
	T_v	801	131	0	836	144	0	917	7	0	956	11	0
	CCP	***0.946***	0.885	0.000	***0.987***	0.973	0.000	***0.946***	0.583	0.000	***0.987***	0.917	0.000
N=350	T_s	**830**	170	0	**830**	170	0	**985**	14	1	**985**	14	1
	T_v	795	146	0	827	165	0	943	10	0	982	14	0
	CCP	***0.958***	0.859	.	***0.996***	0.971	.	***0.957***	0.714	0.000	***0.997***	1.000	0.000
N=500	T_s	**845**	155	0	**845**	155	0	**986**	14	0	**986**	14	0
	T_v	805	139	0	838	153	0	942	10	0	978	13	0
	CCP	***0.953***	0.897	.	***0.992***	0.987	.	***0.955***	0.714	.	***0.992***	0.929	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 36: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.9, Offset=0.8473)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**617**	105	278	**617**	105	278	**688**	29	283	**688**	29	283
	T_v	580	72	119	610	99	174	648	11	122	680	24	179
	CCP	***0.940***	0.686	0.428	***0.989***	0.943	0.626	***0.942***	0.379	0.431	***0.988***	0.828	0.633
N=100	T_s	**696**	93	211	**696**	93	211	**772**	16	212	**772**	16	212
	T_v	661	62	30	687	84	48	733	5	30	762	13	48
	CCP	***0.950***	0.667	0.142	***0.987***	0.903	0.227	***0.949***	0.313	0.142	***0.987***	0.813	0.226
N=200	T_s	**803**	82	115	**803**	82	115	**877**	7	116	**877**	7	116
	T_v	765	62	0	793	76	1	833	2	0	865	4	1
	CCP	***0.953***	0.756	0.000	***0.988***	0.927	0.009	***0.950***	0.286	0.000	***0.986***	0.571	0.009
N=350	T_s	**848**	110	42	**848**	110	42	**937**	8	55	**937**	8	55
	T_v	807	90	0	844	107	0	891	3	0	932	5	0
	CCP	***0.952***	0.818	0.000	***0.995***	0.973	0.000	***0.951***	0.375	0.000	***0.995***	0.625	0.000
N=500	T_s	**826**	164	10	**826**	164	10	**968**	6	26	**968**	6	26
	T_v	783	127	0	815	157	0	914	1	0	955	4	0
	CCP	***0.948***	0.774	0.000	***0.987***	0.957	0.000	***0.944***	0.167	0.000	***0.987***	0.667	0.000

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 37: Count on Model Selection and Conditional Coverage Probability (the correct model = M_3, Corr=0.1, Offset=0.8473)

		AIC						**BIC**					
		1-a=.95			**1-a=.99**			**1-a=.95**			**1-a=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	1	170	**829**	1	170	**829**	3	48	**949**	3	48	**949**
	T_v	0	135	779	1	157	817	2	32	892	3	41	933
	CCP	0.000	0.794	***0.940***	1.000	0.924	***0.986***	0.667	0.667	***0.940***	1.000	0.854	***0.983***
N=100	T_s	0	165	**835**	0	165	**835**	0	38	**962**	0	38	**962**
	T_v	0	136	780	0	156	825	0	29	896	0	34	948
	CCP	.	0.824	***0.934***	.	0.945	***0.988***	.	0.763	***0.931***	.	0.895	***0.985***
N=200	T_s	0	172	**828**	0	172	**828**	0	25	**975**	0	25	**975**
	T_v	0	149	787	0	169	819	0	21	928	0	24	965
	CCP	.	0.866	***0.950***	.	0.983	***0.989***	.	0.840	***0.952***	.	0.960	***0.990***
N=350	T_s	0	156	**844**	0	156	**844**	0	16	**984**	0	16	**984**
	T_v	0	142	814	0	150	839	0	13	945	0	16	977
	CCP	.	0.910	***0.964***	.	0.962	***0.994***	.	0.813	***0.960***	.	1.000	***0.993***
N=500	T_s	0	156	**844**	0	156	**844**	0	19	**981**	0	19	**981**
	T_v	0	134	799	0	150	834	0	16	927	0	18	968
	CCP	.	0.859	***0.947***	.	0.962	***0.988***	.	0.842	***0.945***	.	0.947	***0.987***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 38: Count on Model Selection and Conditional Coverage Probability (the correct model = M_3, Corr=0.3, Offset=0.8473)

		AIC						**BIC**					
		1-a=.95			**1-a=.99**			**1-a=.95**			**1-a=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	2	193	**805**	2	192	**806**	11	60	**929**	11	59	**930**
	T_v	1	152	764	2	181	798	8	38	884	11	50	921
	CCP	0.500	0.788	***0.949***	1.000	0.943	***0.990***	0.727	0.633	***0.952***	1.000	0.847	***0.990***
N=100	T_s	0	173	**827**	0	173	**827**	0	44	**956**	0	44	**956**
	T_v	0	145	773	0	163	812	0	29	899	0	37	939
	CCP	.	0.838	***0.935***	.	0.942	***0.982***	.	0.659	***0.940***	.	0.841	***0.982***
N=200	T_s	0	156	**844**	0	156	**844**	0	21	**979**	0	21	**979**
	T_v	0	129	790	0	149	830	0	13	919	0	18	964
	CCP	.	0.827	***0.936***	.	0.955	***0.983***	.	0.619	***0.939***	.	0.857	***0.985***
N=350	T_s	0	162	**838**	0	162	**838**	0	12	**988**	0	12	**988**
	T_v	0	136	794	0	153	830	0	6	936	0	8	978
	CCP	.	0.840	***0.947***	.	0.944	***0.990***	.	0.500	***0.947***	.	0.667	***0.990***
N=500	T_s	0	158	**842**	0	158	**842**	0	14	**986**	0	14	**986**
	T_v	0	134	784	0	151	828	0	8	921	0	13	970
	CCP	.	0.848	***0.931***	.	0.956	***0.983***	.	0.571	***0.934***	.	0.929	***0.984***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 39: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.5, Offset=0.8473)

		AIC						BIC					
		1-a=.95			1-a=.99			1-a=.95			1-a=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	10	169	**821**	10	169	**821**	22	52	**926**	22	52	**926**
	T_v	9	135	760	10	161	806	17	36	858	22	47	909
	CCP	0.900	0.799	***0.926***	1.000	0.953	***0.982***	0.773	0.692	***0.927***	1.000	0.904	***0.982***
N=100	T_s	0	148	**852**	0	148	**852**	1	39	**960**	1	39	**960**
	T_v	0	119	801	0	140	840	1	26	902	1	35	948
	CCP	.	0.804	***0.940***	.	0.946	***0.986***	1.000	0.667	***0.940***	1.000	0.897	***0.988***
N=200	T_s	0	156	**844**	0	156	**844**	0	16	**984**	0	16	**984**
	T_v	0	136	807	0	151	838	0	11	940	0	14	975
	CCP	.	0.872	***0.956***	.	0.968	***0.993***	.	0.688	***0.955***	.	0.875	***0.991***
N=350	T_s	0	169	**831**	0	169	**831**	0	12	**988**	0	12	**988**
	T_v	0	129	795	0	159	825	0	7	943	0	8	980
	CCP	.	0.763	***0.957***	.	0.941	***0.993***	.	0.583	***0.954***	.	0.667	***0.992***
N=500	T_s	0	146	**854**	0	146	**854**	0	8	**992**	0	8	**992**
	T_v	0	126	804	0	142	841	0	5	932	0	8	977
	CCP	.	0.863	***0.941***	.	0.973	***0.985***	.	0.625	***0.940***	.	1.000	***0.985***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 40: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.7, Offset=0.8473)

		AIC						BIC					
		1-a=.95			1-a=.99			1-a=.95			1-a=.99		
	Model	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***
N=50	T_s	51	159	**790**	51	159	**790**	63	52	**885**	63	52	**885**
	T_v	23	110	744	40	139	779	27	27	828	49	41	869
	CCP	0.451	0.692	***0.942***	0.784	0.874	***0.986***	0.429	0.519	***0.936***	0.778	0.788	***0.982***
N=100	T_s	4	189	**807**	4	189	**807**	15	24	**961**	15	24	**961**
	T_v	3	161	769	4	183	796	6	16	921	13	21	949
	CCP	0.750	0.852	***0.953***	1.000	0.968	***0.986***	0.400	0.667	***0.958***	0.867	0.875	***0.988***
N=200	T_s	0	167	**833**	0	167	**833**	1	19	**980**	1	19	**980**
	T_v	0	130	785	0	157	820	0	12	926	0	16	967
	CCP	.	0.778	***0.942***	.	0.940	***0.984***	0.000	0.632	***0.945***	0.000	0.842	***0.987***
N=350	T_s	0	145	**855**	0	145	**855**	0	16	**984**	0	16	**984**
	T_v	0	111	803	0	136	841	0	6	925	0	14	967
	CCP	.	0.766	***0.939***	.	0.938	***0.984***	.	0.375	***0.940***	.	0.875	***0.983***
N=500	T_s	0	154	**846**	0	154	**846**	0	13	**987**	0	13	**987**
	T_v	0	129	798	0	144	835	0	4	930	0	7	976
	CCP	.	0.838	***0.943***	.	0.935	***0.987***	.	0.308	***0.942***	.	0.538	***0.989***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 41: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.9, Offset=0.8473)

		AIC						BIC					
		1-a=.95			1-a=.99			1-a=.95			1-a=.99		
	Model	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***
N=50	T_s	183	116	**701**	183	116	**701**	185	42	**773**	185	42	**773**
	T_v	62	84	644	112	104	682	62	20	712	112	33	754
	CCP	0.339	0.724	***0.919***	0.612	0.897	***0.973***	0.335	0.476	***0.921***	0.605	0.786	***0.975***
N=100	T_s	105	85	**810**	105	85	**810**	110	22	**868**	110	22	**868**
	T_v	15	63	758	39	79	801	15	11	810	39	17	858
	CCP	0.143	0.741	***0.936***	0.371	0.929	***0.989***	0.136	0.500	***0.933***	0.355	0.773	***0.988***
N=200	T_s	15	142	**843**	15	142	**843**	34	9	**957**	34	9	**957**
	T_v	0	112	796	0	135	834	0	3	903	1	7	947
	CCP	0.000	0.789	***0.944***	0.000	0.951	***0.989***	0.000	0.333	***0.944***	0.029	0.778	***0.990***
N=350	T_s	0	159	**841**	0	159	**841**	8	6	**986**	8	6	**986**
	T_v	0	112	806	0	150	836	0	0	944	0	4	981
	CCP	.	0.704	***0.958***	.	0.943	***0.994***	0.000	0.000	***0.957***	0.000	0.667	***0.995***
N=500	T_s	0	161	**839**	0	161	**839**	0	10	**990**	0	10	**990**
	T_v	0	120	796	0	155	826	0	2	939	0	7	977
	CCP	.	0.745	***0.949***	.	0.963	***0.985***	.	0.200	***0.948***	.	0.700	***0.987***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 42: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.1, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	***T_s***	206	**362**	432	206	**362**	432	294	**128**	578	294	**128**	578
	T_v	192	303	368	203	338	412	277	83	497	291	106	556
	CCP	0.932	***0.837***	0.852	0.985	***0.934***	0.954	0.942	***0.648***	0.860	0.990	***0.828***	0.962
N=100	***T_s***	40	**732**	228	40	**732**	228	156	**365**	479	156	**365**	479
	T_v	36	685	158	38	720	203	151	325	331	154	354	424
	CCP	0.900	***0.936***	0.693	0.950	***0.984***	0.890	0.968	***0.890***	0.691	0.987	***0.970***	0.885
N=200	***T_s***	2	**953**	45	2	**953**	45	17	**756**	227	17	**756**	227
	T_v	2	899	18	2	940	31	17	716	94	17	745	147
	CCP	1.000	***0.943***	0.400	1.000	***0.986***	0.689	1.000	***0.947***	0.414	1.000	***0.985***	0.648
N=350	***T_s***	0	**993**	7	0	**993**	7	0	**933**	67	0	**933**	67
	T_v	0	945	0	0	987	4	0	893	16	0	929	27
	CCP	.	***0.952***	0.000	.	***0.994***	0.571	.	***0.957***	0.239	.	***0.996***	0.403
N=500	***T_s***	0	**999**	1	0	**999**	1	0	**984**	16	0	**984**	16
	T_v	0	929	1	0	977	1	0	920	3	0	964	5
	CCP	.	***0.930***	1.000	.	***0.978***	1.000	.	***0.935***	0.188	.	***0.980***	0.313

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 43: Count on Model Selection and Conditional Coverage Probability (the correct model = M_1, Corr=0.3, Offset=2.1972)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	T_s	221	**249**	530	221	**249**	530	298	**63**	639	298	**63**	639
	T_v	200	203	458	213	236	510	269	34	560	290	55	617
	CCP	0.905	***0.815***	0.864	0.964	***0.948***	0.962	0.903	***0.540***	0.876	0.973	***0.873***	0.966
N=100	T_s	61	**628**	311	61	**628**	311	189	**226**	585	189	**226**	585
	T_v	55	573	209	60	616	262	174	185	393	187	216	499
	CCP	0.902	***0.912***	0.672	0.984	***0.981***	0.842	0.921	***0.819***	0.672	0.989	***0.956***	0.853
N=200	T_s	2	**888**	110	2	**888**	110	33	**604**	363	33	**604**	363
	T_v	2	837	70	2	880	83	32	560	182	33	598	257
	CCP	1.000	***0.943***	0.636	1.000	***0.991***	0.755	0.970	***0.927***	0.501	1.000	***0.990***	0.708
N=350	T_s	0	**984**	16	0	**984**	16	1	**851**	148	1	**851**	148
	T_v	0	938	2	0	978	4	1	818	29	1	847	61
	CCP	.	***0.953***	0.125	.	***0.994***	0.250	1.000	***0.961***	0.196	1.000	***0.995***	0.412
N=500	T_s	0	**998**	2	0	**998**	2	0	**945**	55	0	**945**	55
	T_v	0	942	0	0	986	1	0	897	7	0	936	15
	CCP	.	***0.944***	0.000	.	***0.988***	0.500	.	***0.949***	0.127	.	***0.990***	0.273

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 44: Count on Model Selection and Conditional Coverage Probability (the correct model = M_1, Corr=0.5, Offset=2.1972)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	265	**150**	585	265	**150**	585	309	**27**	664	309	**27**	664
	T_v	234	125	500	253	140	556	266	8	577	296	18	635
	CCP	0.883	***0.833***	0.855	0.955	***0.933***	0.950	0.861	***0.296***	0.869	0.958	***0.667***	0.956
N=100	T_s	120	**447**	433	120	**447**	433	254	**87**	659	254	**87**	659
	T_v	102	418	316	112	438	372	220	71	498	242	80	581
	CCP	0.850	***0.935***	0.730	0.933	***0.980***	0.859	0.866	***0.816***	0.756	0.953	***0.920***	0.882
N=200	T_s	12	**786**	202	12	**786**	202	73	**408**	519	73	**408**	519
	T_v	12	744	122	12	778	154	57	371	276	69	402	368
	CCP	1.000	***0.947***	0.604	1.000	***0.990***	0.762	0.781	***0.909***	0.532	0.945	***0.985***	0.709
N=350	T_s	0	**923**	77	0	**923**	77	4	**683**	313	4	**683**	313
	T_v	0	896	18	0	918	34	2	662	80	4	678	148
	CCP	.	***0.971***	0.234	.	***0.995***	0.442	0.500	***0.969***	0.256	1.000	***0.993***	0.473
N=500	T_s	0	**982**	18	0	**982**	18	1	**818**	181	1	**818**	181
	T_v	0	933	4	0	980	6	0	782	27	0	817	53
	CCP	.	***0.950***	0.222	.	***0.998***	0.333	0.000	***0.956***	0.149	0.000	***0.999***	0.293

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 45: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_1$, Corr=0.7, Offset=2.1972)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	325	**69**	606	325	**69**	606	350	**6**	644	350	**6**	644
	T_v	270	49	534	301	61	583	282	1	571	319	3	621
	CCP	0.831	***0.710***	0.881	0.926	***0.884***	0.962	0.806	***0.167***	0.887	0.911	***0.500***	0.964
N=100	T_s	215	**208**	577	215	**208**	577	291	**15**	694	291	**15**	694
	T_v	165	180	432	197	203	507	209	8	535	253	14	621
	CCP	0.767	***0.865***	0.749	0.916	***0.976***	0.879	0.718	***0.533***	0.771	0.869	***0.933***	0.895
N=200	T_s	53	**549**	398	53	**549**	398	190	**95**	715	190	**95**	715
	T_v	36	521	219	46	541	293	102	81	378	151	92	526
	CCP	0.679	***0.949***	0.550	0.868	***0.985***	0.736	0.537	***0.853***	0.529	0.795	***0.968***	0.736
N=350	T_s	5	**784**	211	5	**784**	211	51	**386**	563	51	**386**	563
	T_v	3	762	62	5	778	110	26	372	157	38	383	278
	CCP	0.600	***0.972***	0.294	1.000	***0.992***	0.521	0.510	***0.964***	0.279	0.745	***0.992***	0.494
N=500	T_s	1	**901**	98	1	**901**	98	10	**558**	432	10	**558**	432
	T_v	0	865	16	0	891	35	3	529	59	5	551	137
	CCP	0.000	***0.960***	0.163	0.000	***0.989***	0.357	0.300	***0.948***	0.137	0.500	***0.987***	0.317

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 46: Count on Model Selection and Conditional Coverage Probability (the correct model = M_1, Corr=0.9, Offset=2.1972)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***	***M₂***	***M₁***	***M₃***
N=50	T_s	399	**71**	530	399	**71**	530	413	**25**	562	413	**25**	562
	T_v	283	30	452	353	51	500	291	5	480	363	12	530
	CCP	0.709	***0.423***	0.853	0.885	***0.718***	0.943	0.705	***0.200***	0.854	0.879	***0.480***	0.943
N=100	T_s	354	**42**	604	354	**42**	604	366	**5**	629	366	**5**	629
	T_v	183	12	473	278	27	546	185	1	488	284	4	567
	CCP	0.517	***0.286***	0.783	0.785	***0.643***	0.904	0.505	***0.200***	0.776	0.776	***0.800***	0.901
N=200	T_s	297	**73**	630	297	**73**	630	320	**0**	680	320	**0**	680
	T_v	96	62	330	185	71	460	96	0	372	188	0	507
	CCP	0.323	***0.849***	0.524	0.623	***0.973***	0.730	0.300	***.***	0.547	0.588	***.***	0.746
N=350	T_s	158	**310**	532	158	**310**	532	286	**0**	714	286	**0**	714
	T_v	26	301	160	60	307	275	34	0	235	83	0	394
	CCP	0.165	***0.971***	0.301	0.380	***0.990***	0.517	0.119	***.***	0.329	0.290	***.***	0.552
N=500	T_s	53	**499**	448	53	**499**	448	237	**6**	757	237	**6**	757
	T_v	1	480	81	8	496	162	8	3	126	31	6	251
	CCP	0.019	***0.962***	0.181	0.151	***0.994***	0.362	0.034	***0.500***	0.166	0.131	***1.000***	0.332

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 47: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_2$, Corr=0.1, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**794**	162	44	**794**	162	44	**885**	52	63	**885**	52	63
	T_v	746	140	26	785	152	35	832	41	42	875	45	53
	CCP	***0.940***	0.864	0.591	***0.989***	0.938	0.795	***0.940***	0.788	0.667	***0.989***	0.865	0.841
N=100	T_s	**831**	167	2	**831**	167	2	**963**	29	8	**963**	29	8
	T_v	770	161	0	817	167	1	899	28	2	949	29	6
	CCP	***0.927***	0.964	0.000	***0.983***	1.000	0.500	***0.934***	0.966	0.250	***0.985***	1.000	0.750
N=200	T_s	**832**	168	0	**832**	168	0	**972**	28	0	**972**	28	0
	T_v	789	157	0	825	165	0	922	25	0	964	26	0
	CCP	***0.948***	0.935	.	***0.992***	0.982	.	***0.949***	0.893	.	***0.992***	0.929	.
N=350	T_s	**842**	158	0	**842**	158	0	**988**	12	0	**988**	12	0
	T_v	791	151	0	834	157	0	932	12	0	979	12	0
	CCP	***0.939***	0.956	.	***0.990***	0.994	.	***0.943***	1.000	.	***0.991***	1.000	.
N=500	T_s	**848**	152	0	**848**	152	0	**995**	5	0	**995**	5	0
	T_v	799	146	0	837	152	0	940	5	0	983	5	0
	CCP	***0.942***	0.961	.	***0.987***	1.000	.	***0.945***	1.000	.	***0.988***	1.000	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 48: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_2$, Corr=0.3, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**789**	134	77	**789**	134	77	**869**	47	84	**869**	47	84
	T_v	751	106	43	779	121	58	822	31	50	856	39	65
	CCP	***0.952***	0.791	0.558	***0.987***	0.903	0.753	***0.946***	0.660	0.595	***0.985***	0.830	0.774
N=100	T_s	**828**	160	12	**828**	160	12	**953**	24	23	**953**	24	23
	T_v	783	142	2	819	159	5	902	18	6	944	23	10
	CCP	***0.946***	0.888	0.167	***0.989***	0.994	0.417	***0.946***	0.750	0.261	***0.991***	0.958	0.435
N=200	T_s	**842**	158	0	**842**	158	0	**982**	16	2	**982**	16	2
	T_v	796	145	0	832	152	0	928	13	0	969	14	0
	CCP	***0.945***	0.918	.	***0.988***	0.962	.	***0.945***	0.813	0.000	***0.987***	0.875	0.000
N=350	T_s	**844**	156	0	**844**	156	0	**991**	9	0	**991**	9	0
	T_v	806	147	0	834	155	0	945	8	0	977	9	0
	CCP	***0.955***	0.942	.	***0.988***	0.994	.	***0.954***	0.889	.	***0.986***	1.000	.
N=500	T_s	**821**	179	0	**821**	179	0	**988**	12	0	**988**	12	0
	T_v	787	172	0	816	176	0	949	12	0	981	12	0
	CCP	***0.959***	0.961	.	***0.994***	0.983	.	***0.961***	1.000	.	***0.993***	1.000	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 49: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.5, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**759**	128	113	**759**	128	113	**838**	40	122	**838**	40	122
	T_v	700	99	60	749	115	87	767	28	69	825	32	96
	CCP	***0.922***	0.773	0.531	***0.987***	0.898	0.770	***0.915***	0.700	0.566	***0.984***	0.800	0.787
N=100	T_s	**818**	143	39	**818**	143	39	**914**	24	62	**914**	24	62
	T_v	776	128	6	809	138	13	866	17	16	901	22	29
	CCP	***0.949***	0.895	0.154	***0.989***	0.965	0.333	***0.947***	0.708	0.258	***0.986***	0.917	0.468
N=200	T_s	**850**	147	3	**850**	147	3	**978**	13	9	**978**	13	9
	T_v	800	131	0	840	143	0	923	10	0	967	12	1
	CCP	***0.941***	0.891	0.000	***0.988***	0.973	0.000	***0.944***	0.769	0.000	***0.989***	0.923	0.111
N=350	T_s	**837**	163	0	**837**	163	0	**986**	14	0	**986**	14	0
	T_v	800	147	0	832	159	0	940	13	0	980	14	0
	CCP	***0.956***	0.902	.	***0.994***	0.975	.	***0.953***	0.929	.	***0.994***	1.000	.
N=500	T_s	**828**	172	0	**828**	172	0	**988**	12	0	**988**	12	0
	T_v	783	153	0	819	168	0	932	9	0	977	12	0
	CCP	***0.946***	0.890	.	***0.989***	0.977	.	***0.943***	0.750	.	***0.989***	1.000	.

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 50: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.7, Offset=2.1972)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***	***M_2***	***M_1***	***M_3***
N=50	***T_s***	**721**	93	186	**721**	93	186	**784**	29	187	**784**	29	187
	T_v	678	60	87	710	83	125	734	15	88	770	23	126
	CCP	***0.940***	0.645	0.468	***0.985***	0.892	0.672	***0.936***	0.517	0.471	***0.982***	0.793	0.674
N=100	***T_s***	**776**	112	112	**776**	112	112	**858**	21	121	**858**	21	121
	T_v	738	88	19	765	107	35	812	10	26	846	18	43
	CCP	***0.951***	0.786	0.170	***0.986***	0.955	0.313	***0.946***	0.476	0.215	***0.986***	0.857	0.355
N=200	***T_s***	**843**	142	15	**843**	142	15	**949**	13	38	**949**	13	38
	T_v	808	123	0	837	138	0	911	8	1	942	11	2
	CCP	***0.958***	0.866	0.000	***0.993***	0.972	0.000	***0.960***	0.615	0.026	***0.993***	0.846	0.053
N=350	***T_s***	**838**	162	0	**838**	162	0	**987**	9	4	**987**	9	4
	T_v	801	142	0	832	159	0	944	8	0	980	9	0
	CCP	***0.956***	0.877	.	***0.993***	0.981	.	***0.956***	0.889	0.000	***0.993***	1.000	0.000
N=500	***T_s***	**864**	136	0	**864**	136	0	**993**	6	1	**993**	6	1
	T_v	815	117	0	856	130	0	938	5	0	984	5	0
	CCP	***0.943***	0.860	.	***0.991***	0.956	.	***0.945***	0.833	0.000	***0.991***	0.833	0.000

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 51: Count on Model Selection and Conditional Coverage Probability (the correct model = M_2, Corr=0.9, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	**602**	106	292	**602**	106	292	**668**	35	297	**668**	35	297
	T_v	567	68	153	593	97	196	627	17	156	657	28	200
	CCP	***0.942***	0.642	0.524	***0.985***	0.915	0.671	***0.939***	0.486	0.525	***0.984***	0.800	0.673
N=100	T_s	**668**	101	231	**668**	101	231	**738**	25	237	**738**	25	237
	T_v	629	57	49	658	90	81	694	10	51	725	20	85
	CCP	***0.942***	0.564	0.212	***0.985***	0.891	0.351	***0.940***	0.400	0.215	***0.982***	0.800	0.359
N=200	T_s	**757**	85	158	**757**	85	158	**826**	15	159	**826**	15	159
	T_v	713	62	2	746	78	9	778	6	3	814	10	10
	CCP	***0.942***	0.729	0.013	***0.985***	0.918	0.057	***0.942***	0.400	0.019	***0.985***	0.667	0.063
N=350	T_s	**840**	78	82	**840**	78	82	**910**	7	83	**910**	7	83
	T_v	787	58	0	830	73	1	849	3	0	898	6	1
	CCP	***0.937***	0.744	0.000	***0.988***	0.936	0.012	***0.933***	0.429	0.000	***0.987***	0.857	0.012
N=500	T_s	**852**	128	20	**852**	128	20	**949**	4	47	**949**	4	47
	T_v	814	106	0	839	125	0	906	2	0	933	4	0
	CCP	***0.955***	0.828	0.000	***0.985***	0.977	0.000	***0.955***	0.500	0.000	***0.983***	1.000	0.000

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 52: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.1, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	3	193	**804**	3	193	**804**	7	63	**930**	7	63	**930**
	T_v	1	156	751	1	176	790	5	43	867	5	51	915
	CCP	0.333	0.808	***0.934***	0.333	0.912	***0.983***	0.714	0.683	***0.932***	0.714	0.810	***0.984***
N=100	T_s	0	161	**839**	0	161	**839**	0	41	**959**	0	41	**959**
	T_v	0	137	798	0	154	826	0	30	910	0	38	942
	CCP	.	0.851	***0.951***	.	0.957	***0.985***	.	0.732	***0.949***	.	0.927	***0.982***
N=200	T_s	0	147	**853**	0	147	**853**	0	16	**984**	0	16	**984**
	T_v	0	133	811	0	146	844	0	16	937	0	16	975
	CCP	.	0.905	***0.951***	.	0.993	***0.989***	.	1.000	***0.952***	.	1.000	***0.991***
N=350	T_s	0	164	**836**	0	164	**836**	0	21	**979**	0	21	**979**
	T_v	0	138	802	0	158	825	0	17	939	0	20	966
	CCP	.	0.841	***0.959***	.	0.963	***0.987***	.	0.810	***0.959***	.	0.952	***0.987***
N=500	T_s	0	158	**842**	0	158	**842**	0	11	**989**	0	11	**989**
	T_v	0	134	804	0	151	835	0	7	940	0	10	978
	CCP	.	0.848	***0.955***	.	0.956	***0.992***	.	0.636	***0.950***	.	0.909	***0.989***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 53: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.3, Offset=2.1972)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	7	182	**811**	7	182	**811**	17	59	**924**	17	59	**924**
	T_v	5	136	770	6	164	801	14	31	874	15	48	910
	CCP	0.714	0.747	***0.949***	0.857	0.901	***0.988***	0.824	0.525	***0.946***	0.882	0.814	***0.985***
N=100	T_s	0	193	**807**	0	193	**807**	0	40	**960**	0	40	**960**
	T_v	0	158	759	0	182	799	0	30	906	0	34	949
	CCP	.	0.819	***0.941***	.	0.943	***0.990***	.	0.750	***0.944***	.	0.850	***0.989***
N=200	T_s	0	178	**822**	0	178	**822**	0	26	**974**	0	26	**974**
	T_v	0	146	780	0	166	811	0	15	922	0	20	962
	CCP	.	0.820	***0.949***	.	0.933	***0.987***	.	0.577	***0.947***	.	0.769	***0.988***
N=350	T_s	0	159	**841**	0	159	**841**	0	19	**981**	0	19	**981**
	T_v	0	131	792	0	153	835	0	14	925	0	16	972
	CCP	.	0.824	***0.942***	.	0.962	***0.993***	.	0.737	***0.943***	.	0.842	***0.991***
N=500	T_s	0	172	**828**	0	172	**828**	0	12	**988**	0	12	**988**
	T_v	0	146	786	0	163	818	0	10	936	0	11	976
	CCP	.	0.849	***0.949***	.	0.948	***0.988***	.	0.833	***0.947***	.	0.917	***0.988***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 54: Count on Model Selection and Conditional Coverage Probability (the correct model $= M_3$, Corr=0.5, Offset=2.1972)

		AIC						BIC					
		1-α=.95			1-α=.99			1-α=.95			1-α=.99		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	25	166	**809**	25	166	**809**	32	60	**908**	32	60	**908**
	T_v	19	115	766	24	139	801	23	30	854	29	43	897
	CCP	0.760	0.693	***0.947***	0.960	0.837	***0.990***	0.719	0.500	***0.941***	0.906	0.717	***0.988***
N=100	T_s	1	176	**823**	1	176	**823**	5	34	**961**	5	34	**961**
	T_v	0	141	781	0	165	820	2	22	913	2	28	955
	CCP	0.000	0.801	***0.949***	0.000	0.938	***0.996***	0.400	0.647	***0.950***	0.400	0.824	***0.994***
N=200	T_s	0	163	**837**	0	163	**837**	0	16	**984**	0	16	**984**
	T_v	0	140	798	0	154	829	0	12	934	0	13	974
	CCP	.	0.859	***0.953***	.	0.945	***0.990***	.	0.750	***0.949***	.	0.813	***0.990***
N=350	T_s	0	166	**834**	0	166	**834**	0	17	**983**	0	17	**983**
	T_v	0	142	795	0	160	828	0	11	939	0	14	977
	CCP	.	0.855	***0.953***	.	0.964	***0.993***	.	0.647	***0.955***	.	0.824	***0.994***
N=500	T_s	0	160	**840**	0	160	**840**	0	11	**989**	0	11	**989**
	T_v	0	135	797	0	153	832	0	8	940	0	10	980
	CCP	.	0.844	***0.949***	.	0.956	***0.990***	.	0.727	***0.950***	.	0.909	***0.991***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 55: Count on Model Selection and Conditional Coverage Probability (the correct model = M_3, Corr=0.7, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	60	141	**799**	60	141	**799**	69	46	**885**	69	46	**885**
	T_v	36	106	751	48	123	788	38	25	834	53	31	873
	CCP	0.600	0.752	***0.940***	0.800	0.872	***0.986***	0.551	0.543	***0.942***	0.768	0.674	***0.986***
N=100	T_s	4	181	**815**	4	181	**815**	12	38	**950**	12	38	**950**
	T_v	1	130	769	2	166	806	3	14	895	5	27	937
	CCP	0.250	0.718	***0.944***	0.500	0.917	***0.989***	0.250	0.368	***0.942***	0.417	0.711	***0.986***
N=200	T_s	0	170	**830**	0	170	**830**	0	26	**974**	0	26	**974**
	T_v	0	136	788	0	160	819	0	11	922	0	20	959
	CCP	.	0.800	***0.949***	.	0.941	***0.987***	.	0.423	***0.947***	.	0.769	***0.985***
N=350	T_s	0	152	**848**	0	152	**848**	0	13	**987**	0	13	**987**
	T_v	0	122	807	0	144	843	0	6	943	0	9	981
	CCP	.	0.803	***0.952***	.	0.947	***0.994***	.	0.462	***0.955***	.	0.692	***0.994***
N=500	T_s	0	144	**856**	0	144	**856**	0	14	**986**	0	14	**986**
	T_v	0	115	819	0	139	851	0	8	943	0	11	980
	CCP	.	0.799	***0.957***	.	0.965	***0.994***	.	0.571	***0.956***	.	0.786	***0.994***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Table 56: Count on Model Selection and Conditional Coverage Probability (the correct model = M_3, Corr=0.9, Offset=2.1972)

		AIC						**BIC**					
		1-α=.95			**1-α=.99**			**1-α=.95**			**1-α=.99**		
	Model	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3	M_2	M_1	M_3
N=50	T_s	190	121	**689**	190	121	**689**	194	52	**754**	194	52	**754**
	T_v	57	86	636	123	104	672	57	31	692	124	41	736
	CCP	0.300	0.711	***0.923***	0.647	0.860	***0.975***	0.294	0.596	***0.918***	0.639	0.788	***0.976***
N=100	T_s	108	101	**791**	108	101	**791**	110	22	**868**	110	22	**868**
	T_v	12	82	749	34	98	782	12	16	822	34	21	857
	CCP	0.111	0.812	***0.947***	0.315	0.970	***0.989***	0.109	0.727	***0.947***	0.309	0.955	***0.987***
N=200	T_s	20	153	**827**	20	153	**827**	47	6	**947**	47	6	**947**
	T_v	0	113	787	4	144	820	0	2	897	4	4	939
	CCP	0.000	0.739	***0.952***	0.200	0.941	***0.992***	0.000	0.333	***0.947***	0.085	0.667	***0.992***
N=350	T_s	1	151	**848**	1	151	**848**	7	11	**982**	7	11	**982**
	T_v	0	112	799	0	145	841	0	2	927	0	8	973
	CCP	0.000	0.742	***0.942***	0.000	0.960	***0.992***	0.000	0.182	***0.944***	0.000	0.727	***0.991***
N=500	T_s	0	164	**836**	0	164	**836**	1	14	**985**	1	14	**985**
	T_v	0	110	786	0	150	828	0	1	926	0	7	976
	CCP	.	0.671	***0.940***	.	0.915	***0.990***	0.000	0.071	***0.940***	0.000	0.500	***0.991***

Note:

(1) T_s= Count on Model Selection. T_v = Count on the coverage of confidence intervals conditional on model selection. *CCP*= Conditional Coverage Probability. *N*=sample size.

(2) **Bold** values (other than the Row and Column Titles) indicate Count on selecting the correct model; ***bold italic*** values indicate the Coverage Probability conditional on selecting the correct model.

(3) The count is based upon 1000 iterations.

Appendix B: SAS Code for Simulation

```
OPTIONS NODATE PAGENO=1 LINESIZE=80 PAGESIZE=60;

***********************************************************************
SAS program to simulate the effects of model selection on the coverage
probability of confidence intervals for binary logistic regression
models. Factors manipulated include covariate correlation, true model,
sample size, automatic model selectors and an offset term for
dichotomization threshold. Following is the sample code for the true
model of M1 with covariate correlation 0.1 and offset term 0;
***********************************************************************;

LIBNAME LOGIT1 'C:\DIS TOPIC\SAS';
LIBNAME LOGIT2 'C:\DIS TOPIC\SAS';

* define the true model from which data are generated;
%LET COVARIATE=ACTSAT GPACORE;
* underlying parameter for ACTSAT;
%LET BETA1=2;
* underlying parameter for GPACORE;
%LET BETA2=1;
* offset for dichotomization;
%LET OFFSET=0;

*to direct the sas log to an external file;
PROC PRINTTO LOG='C:\DIS TOPIC\SAS\LOGFILE.TMP';
RUN;

* Prepare correlated covariates with the matrix decomposition procedure
(Kaiser and Dickman, 1962).The correlation is varied as .1, .3, .5, .7
and .9 in accordance with the design conditions;
***********************************************************************;
*define population correlation matrix;
DATA R (TYPE=CORR);_TYPE='CORR';
INPUT _TYPE_ $ _NAME_ $ ACTSAT GPACORE;
CARDS;

CORR ACTSAT   1.000  .
CORR GPACORE   .100  1.000
;
*apply principal component factorization to the population correlations
matrix and obtain factor pattern matrix for later data generation;

PROC FACTOR N=2 OUTSTAT=FACOUT;

DATA PATTERN; SET FACOUT;                *obtain factor pattern matrix;
IF _TYPE_='PATTERN';
DROP _TYPE_ _NAME_;
RUN;
***********************************************************************;

*create a macro for data generation;
```

```
%MACRO DATA;

PROC IML;
      USE PATTERN;
      READ ALL VAR _NUM_ INTO F;
F=F`;      *F contains principal component factor pattern coefficients;

*generate two random uncorrelated normal variables;

DATA=RANNOR (J(&N,2,0));
DATA=DATA`;             *transpose for multiplication;
Z=F*DATA;              *transform uncorrelated variables to correlated;
Z=Z`;                   *transpose sample data matrix;
ACTSAT=Z[,1];
GPACORE=Z[,2];
Z=ACTSAT ||GPACORE;
CREATE B FROM Z [COLNAME={ACTSAT GPACORE }];
APPEND FROM Z;
DATA RESPONSE (DROP=I J Y PROB); SET B;

*for model M3, create an academic variable as sum of the specified
covariates;

ACADE=GPACORE+ACTSAT;

* generate the response data from the 'true' model;
      DO I=1 TO &N;
            PROB=EXP(&OFFSET+&BETA1*ACTSAT+&BETA2*GPACORE)/
                        (1+EXP(&OFFSET+&BETA1*ACTSAT+&BETA2*GPACORE));
      END;
      DO J=1 TO &N;
            Y=RANUNI(0);
            IF Y<=PROB THEN GRADRATE=1;   * Dichotomize the variable;
            ELSE GRADRATE=0;
      END;
RUN;
%MEND DATA;

*to direct the output to an external file;

PROC PRINTTO PRINT='C:\DIS TOPIC\SAS\M&C.TMP';

*create a macro for model selection, parameter estimates and naive
confidence intervals construction;

%MACRO SELECTION;

*specify conditions of sample size;
      %DO S=1 %TO 5;

            %IF &S=1 %THEN %DO; %LET N=50; %END;
            %IF &S=2 %THEN %DO; %LET N=100; %END;
            %IF &S=3 %THEN %DO; %LET N=200; %END;
            %IF &S=4 %THEN %DO; %LET N=350; %END;
            %IF &S=5 %THEN %DO; %LET N=500; %END;
```

```
* specify the number of replications;
            %DO NITER=1 %TO 1000;

* run the data generation macro;
      %DATA;

*define alpha levels;
            %DO A=1 %TO 2;
                  %IF &A=1 %THEN %DO; %LET ALPHA=.05;%END;
                  %IF &A=2 %THEN %DO; %LET ALPHA=.01;%END;

* specify the models available for selection;
                  %DO M=1 %TO 3;
                        %LET COVER=COVER&M;
                        %LET AIC=AIC&M;
                        %LET BIC=BIC&M;
                        %LET DATA=DATA&M;
                              %IF &M=1 %THEN %DO;
                                    %LET VARIABLES=ACTSAT;
                              %END;
                              %IF &M=2 %THEN %DO;
                                    %LET VARIABLES=ACTSAT GPACORE;
                              %END;
                              %IF &M=3 %THEN %DO;
                                    %LET VARIABLES=ACADE;
                              %END;

*Direct output to an external file;
      FILENAME NEWOUT 'C:\DIS TOPIC\SAS\SAMPLEOUT.TMP';
      PROC PRINTTO PRINT=NEWOUT NEW;

*run logistic regression with the three models to fit the same sample
data;
      PROC LOGISTIC DATA=RESPONSE;
      MODEL GRADRATE (EVENT='1')= &VARIABLES/ALPHA=&ALPHA CLPARM=PL;
      RUN;

*direct sas output back to sas output window;
      PROC PRINTTO PRINT=PRINT;
      RUN;

* extract aic from the fit results;
      DATA &AIC; INFILE NEWOUT;
      INPUT PRIOR1 $ @;
      IF PRIOR1='AIC' THEN DO;
            INPUT V1 &AIC ;
            KEEP &AIC;
            OUTPUT;
      END;

* extract bic from the fit results;
      DATA &BIC; INFILE NEWOUT;
      INPUT PRIOR2 $ @;
      IF PRIOR2='SC' THEN DO;
            INPUT V2 &BIC ;
            KEEP &BIC;
            OUTPUT;
```

```
      END;

* extract lower and higher bound of the naive confidence intervals;
      DATA CI; INFILE NEWOUT;
      INPUT PAR $ @;
      IF PAR='ACTSAT'OR PAR='ACADE' THEN DO;
            INPUT EST LOW HIGH ;
            KEEP LOW HIGH;
            OUTPUT;
      END;

      PROC IML;
            USE CI;
      READ ALL VAR _NUM_ INTO I;
      LOW=I[3,1];HIGH=I[3,2];
      IL=LOW||&BETA1||HIGH;
      CREATE IV FROM IL[COLNAME={LOW BETA1 HIGH}];
      APPEND FROM IL;

*examine if the nominal confidence intervals cover the parameter for
ACTSAT;
      DATA &COVER; SET IV;
      IF LOW<=BETA1<=HIGH THEN &COVER=1;
            ELSE &COVER=0;
            KEEP &COVER;
            OUTPUT;
      RUN;

*merge aic, bic and coverage results;
      DATA &DATA;
      MERGE &AIC &BIC &COVER;
      RUN;

                  %END; *MODEL1-3;

*comebine results of the three models;
      DATA LOGIT&A;
      MERGE DATA1 DATA2 DATA3;
      MINAIC=MIN(OF AIC1-AIC3);
      MINBIC=MIN(OF BIC1-BIC3);
      N=&N;

      IF MINAIC=AIC1 THEN AMODEL=1;
      ELSE IF MINAIC=AIC2 THEN AMODEL=2;
      ELSE AMODEL=3;

      IF MINBIC=BIC1 THEN BMODEL=1;
      ELSE IF MINBIC=BIC2 THEN BMODEL=2;
      ELSE BMODEL=3;

      PROC APPEND BASE=LOGIT&A.RESULTS FORCE;
      RUN;

            %END; *ALPHA1-2;
            %END; *ITERATION;
      %END; *SAMPLE SIZE 1-5;
%MEND SELECTION;  *close the macro;
```

```
*run the selection macro;
%SELECTION;

*create macro to export the result;
%MACRO LOGITRESULTS;

      %DO A=1 %TO 2;

            DATA MODEL&A; SET LOGIT&A.RESULTS;
            IF AMODEL=1 THEN AMODEL1=1; ELSE AMODEL1=0;
            IF AMODEL=2 THEN AMODEL2=1; ELSE AMODEL2=0;
            IF AMODEL=3 THEN AMODEL3=1; ELSE AMODEL3=0;

            IF BMODEL=1 THEN BMODEL1=1; ELSE BMODEL1=0;
            IF BMODEL=2 THEN BMODEL2=1; ELSE BMODEL2=0;
            IF BMODEL=3 THEN BMODEL3=1; ELSE BMODEL3=0;
            RUN;

            PROC SORT; BY N;

            DATA CVRATE&A (DROP=COVER1-COVER3 AMODEL BMODEL AIC1-AIC3
            MINAIC BIC1-BIC3 MINBIC);
                  SET MODEL&A;
            IF AMODEL=1 THEN ACV=COVER1;
            ELSE IF AMODEL=2 THEN ACV=COVER2;
            ELSE ACV=COVER3;

            IF AMODEL=1 AND COVER1=1 THEN ACONCV1=1; ELSE IF AMODEL=1
            AND COVER1=0 THEN ACONCV1=0;ELSE ACONCV1=.;
            IF AMODEL=2 AND COVER2=1 THEN ACONCV2=1; ELSE IF AMODEL=2
            AND COVER2=0 THEN ACONCV2=0;ELSE ACONCV2=.;
            IF AMODEL=3 AND COVER3=1 THEN ACONCV3=1; ELSE IF AMODEL=3
            AND COVER3=0 THEN ACONCV3=0;ELSE ACONCV3=.;

            IF BMODEL=1 THEN BCV=COVER1;
            ELSE IF BMODEL=2 THEN BCV=COVER2;
            ELSE BCV=COVER3;
            IF BMODEL=1 AND COVER1=1 THEN BCONCV1=1; ELSE IF BMODEL=1
            AND COVER1=0 THEN BCONCV1=0;ELSE BCONCV1=.;
            IF BMODEL=2 AND COVER2=1 THEN BCONCV2=1; ELSE IF BMODEL=2
            AND COVER2=0 THEN BCONCV2=0;ELSE BCONCV2=.;
            IF BMODEL=3 AND COVER3=1 THEN BCONCV3=1; ELSE IF BMODEL=3
            AND COVER3=0 THEN BCONCV3=0;ELSE BCONCV3=.;
            RUN;
      %END;
%MEND;

*export results;
%LOGITRESULTS;

PROC PRINTTO PRINT='C:\DIS TOPIC\SAS\FINAL.TMP';

*create a macro to tabulate and report results;
%MACRO RESULT;

      %LET NITER=1000;
```

```
      %DO A=1 %TO 2;
            %IF &A=1 %THEN %DO; %LET RESULT=CVRATE1; %END;
            %IF &A=2 %THEN %DO; %LET RESULT=CVRATE2; %END;

            PROC IML;
                  USE &RESULT;
            READ ALL VAR {AMODEL1 AMODEL2 AMODEL3} INTO AMM;
            READ ALL VAR {ACONCV1 ACONCV2 ACONCV3} INTO ACC;
            READ ALL VAR {ACV} INTO ATC;

            READ ALL VAR {BMODEL1 BMODEL2 BMODEL3} INTO BMM;
            READ ALL VAR {BCONCV1 BCONCV2 BCONCV3} INTO BCC;
            READ ALL VAR {BCV} INTO BTC;

            %DO S=1 %TO 5;
**********************************************************************;
*AIC;
                  AMM&S=AMM[(&S-1)*&NITER+1:&S*&NITER,];
                  ACC&S=ACC[(&S-1)*&NITER+1:&S*&NITER,];
                  ATC&S=ATC[(&S-1)*&NITER+1:&S*&NITER,];
                  ASELECT&S=AMM&S[+,];
                  ABCOV&S=ACC&S[+,];
                  ACONCOV&S=ABCOV&S/ASELECT&S;
                  ATOTALCONCOV&S=ATC&S[+,]/&NITER;

                  ACI_COVER_RATE&S=ASELECT&S//ABCOV&S//ACONCOV&S;
*BIC;
                  BMM&S=BMM[(&S-1)*&NITER+1:&S*&NITER,];
                  BCC&S=BCC[(&S-1)*&NITER+1:&S*&NITER,];
                  BTC&S=BTC[(&S-1)*&NITER+1:&S*&NITER,];
                  BSELECT&S=BMM&S[+,];
                  BBCOV&S=BCC&S[+,];
                  BCONCOV&S=BBCOV&S/BSELECT&S;
                  BTOTALCONCOV&S=BTC&S[+,]/&NITER;

                  BCI_COVER_RATE&S=BSELECT&S//BBCOV&S//BCONCOV&S;
**********************************************************************;
*TABULATE THE RESULTS;

                  CI_COVER_RATE&S=ACI_COVER_RATE&S||BCI_COVER_RATE&S;
                  TOTALCONCOV&S=ATOTALCONCOV&S||BTOTALCONCOV&S;

                  TITLE 'TRUE MODEL=M1, COVARIATE CORR=.1, OFFSET=0 ';

                  PRINT CI_COVER_RATE&S [ROWNAME={TIMES_MODEL_SELECTED
                  TIMES_BETA_COVERED BETA_COVERRATE}
                  COLNAME={AICMODEL1 AICMODEL2 AICMODEL3 BICMODEL1
                  BICMODEL2 BICMODEL3}]
                  TOTALCONCOV&S [COLNAME={AIC_TOTAL_COVERRATE
                  BIC_TOTAL_COVERRATE}];
            %END;
      QUIT;
      %END;
%MEND;
%RESULT;
```

References

Arabatzis, A. A., Gregoire, T. G. & Reynolds, M. R. (1989). Conditional interval estimation of the mean following rejection of a two sided test. *Communications in Statistics-Theory and Methods*, 18, 4359–4373.

Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. in *2nd international Symposium on Information Theory*, eds. B. N . Petrov and F. Csaki, Budapest: Akademia Kiado, 267-281.

Akaike, H. (1978). A Bayesian Analysis of the Minimum ACI Procedure. *Annals of the Institute of Statistical Mathematics, Series A*, 30, 9-14.

Aranda-Ordaz, F. J. (1981). On Two Families of Transformations to Additivity for Binary Response Data. *Biometrika*, 68, 357-364.

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.

Bickel, P. J., & Doksum, K. A. (1981). An Analysis of Transformation Revisited. *Journal of the American Statistical Association*, 76, 296-311.

Box, G. E. P. & Cox, D.R. (1964). An analysis of transformation. *Journal of Royal Statistical Society*, *Series B*, 16, 211-252.

Box, G. E. P. & Cox, D.R. (1982). An analysis of transformation revisited, rebutted. *Journal of American Statistical Association*, 77, 209-210.

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 American Statistical Association*, 87, 738-354.

Buckland, S.T., Burnham, K.P.& Augustin, N.H. (1997). Model Selection: An Integral Part of Inference, *Biometrics,* 53, No. 2. 603-618.

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

Carroll, R. J. & Ruppert, D. (1981). Prediction and the Power Transformation Family. *Biometrika*, 68, 609-616.

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.

Conger, A.J. (1974). A Revised Definition for Suppressor Variables: a Guide To Their Identification and Interpretation. *Educational and Psychological Measurement*, 34, 35-46.

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

Dijkstra, T. K. (ed.) (1988). *On Model Uncertainty and Its Statistical Implications*. Berlin: Springer.

Draper, D.(1995). Assessment and Propagation of Model Uncertainty. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57, 45-97.

Fan, X., FelsOvályi, A., Sivo, S.A. & Keenan, S. C. (2001). *SAS for Monte Carlo studies: a guide for quantitative researchers*. Cary, NC: SAS Institute Inc.

Furnival, G.M. & Wilson, R.W. (1974). Regression by leaps and bounds. *Technometrics,* 16, 499–511.

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.

Guerrero, V. M., & Johnson, R. A. (1982). Use of the Box-Cox Transformation With Binary Response Models. *Biometrika*, 69,309-314.

Han, C. P. (1998). Conditional confidence intervals of regression coefficients following rejection of a preliminary test. in *Applied Statistical Science*, ed. by S. E. Ahmed, M. Ahsanullah, and B. K. Sinha, vol. III, pp. 193–202. Nova Science.

Hinkley, D. V. & Runger, G. (1984). The Analysis of Transformed Data. *Journal of the American Statistical Association,* 79, 302-309.

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

Hjorth, U. & Holmqvist, L. (1981). On Model Selection Based on Validation with Applications to Pressure and Temperature Prognosis. *Applied Statistics*, 30 264-274.

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

Hjorth, U. (1987). On model selection in the computer age. *Technical Report LiTH-MAT-R-87-08.* Linköping University, Linköping.

Hjorth, U. (1989) On model selection in the computer age. *Journal of Statistical planning and inference*, 23, 101-115.

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.

Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics*, 32, l-49.

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.

Hook, E. B. & Regal, R.R. (1997). Validity of methods for model selection, weighting for model uncertainty, and small sample adjustment in capture-recapture estimation. *American Journal of Epidemiology,* 145, 1138–1144.

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.

Horst, P. (1966). *Psychological measurement and prediction.* Belmont, CA: Wadsworth.

Hosmer, D. W. & Lemeshow, S. (1989). *Applied logistic regression*, New York: John Wiley.

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.

Judge, G.G. & Bock, M.E. (1978). *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics.* North-Holland, Amsterdam.

Judge, G.G. & Bock, M.E. (1983). Biased estimation. In: Griliches, Z., Intriligator, M.D. (Eds.), *Handbook of Econometrics,* Vol. 1. North-Holland, Amsterdam (Chapter 10).

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(4), 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,

Kass, R. E. & Raftery, A. E. (1995) Bayes factors. *Journal of the American Statistical Association*, 90, 773-795.

Le, N. D., Raftery, A. E. & Martin, R. D. (1993). Robust model comparison for autoregressive processes with robust Bayes factors. *Technical Report 123,* Department of Statistics, University of British Columbia, Vancouver.

Leamer, E.E. (1978). Regression Selection Strategies and Revealed Priors. *Journal of the American Statistical Association*, 73, 580-587.

Leeb, H. (2002): The distribution of a linear predictor after model selection: unconditional finite-sample distributions and asymptotic approximations. *Technical Report 2002-02*, Department of Statistics, University of Vienna.

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. (2001). The variance of an integrated process need not diverge to infinity and related results on partial sums of stationary processes. *Econometric Theory*, 17, 671–685.

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. (2004). Sparse estimators and the oracle property, or the return of Hodge's estimator. *Technical Report 2004dec-1*, Department of Statistics, Yale University.

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.

Magnus, J.R. (1999). The traditional pretest estimator. *Theory of Probability and its Applications*, 44, 293-308.

Magnus, J.R. & Durbin, J.(1999). Estimation of regression coefficients of interest when other regression coefficients are of no interest. *Econometrics,* 67, 639-643.

Madigan, D. & York, J. (1995) Bayesian graphical models for discrete data. *International Statistical Review*, 63, 215-232.

McArdle, J.J. & Hamagami, F. (1994). Logit and multilevel logit modeling of college graduation for 1984-1985 freshman student-athletes. *Journal of the American Statistical Association.* 89, 1107-1123.

McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models* (2nd ed.), New York: Chapman and Hall.

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.

Mittelhammer, R.C. (1984). Restricted least squares, pre-test, OLS and Stein-rule estimators: Risk comparisons under model misspecification. *Journal of Econometrics,* 25, 151-164.

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

Peers, H.W. (1965). On Confidence Points and Bayesian Probability Points in the Case of Several Parameters. *Journal of the Royal Statistical Society. Series B (Methodological)*, 27, 9-16.

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., 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.

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.

Roehrig, C.S. (1984). Optimal critical regions for pre-test estimators using a Bayes risk criterion. *Journal of Econometrics,* 25, 3-14.

Saleh, M.A.K.E. & Sen, P.K.(1983). Nonparametric tests of location after a preliminary test on regression in the multivariate case. *Communications in Statistics-Theory and Methods*, 12, 1855-1872

SAS OnlineDoc 9.1.3. (2007). SAS Institute, Inc. Retrieved Dec. 10, 2007, from http://support.sas.com/onlinedoc/913/docMainpage.jsp

Schervish, M. J. & Tsay, R. S. (1988). Bayesian modeling and forecasting in autoregressive models. In *Bayesian Analysis of Time Series and Dynamic Models* (ed. J. C. Spall), 23-52. New York: Dekker.

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

Sclove, S.L., Morris, C. & Radhakrishnan, R. (1972). Non-optimality of preliminary- test estimators for the mean of a multivariate normal distribution. *The Annals of Mathematical Statistics,* 43, 1481-1490.

Sen, P.K. (1979). Asymptotic properties of maximum likelihood estimators based on conditional specification. *Annals of Statistics,* 7, 1019-1033.

Shibata, R. (1984). Approximate efficiency of a selection procedure for the number of regression variables. *Biometrika,* 71, 43-9.

Siqueira, A.L. & Taylor, J.M.G. (1999). Treatment Effects in a Logistic Model Involving the Box-Cox Transformation. *Journal of the American Statistical Association*, 94, 240-246.

Taylor, J. M. G. (1986). The Retransformed Mean after a Fitted Power Transformation. *Journal of the American Statistical Association*, 81, 114-118.

Taylor, J. M. G. (1988). The Cost of Generalizing Logistic Regression. *Journal of the American Statistical Association*, 83, 1078-1083.

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.

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