ERRATA

Page 28, line 7. The sentence should read as follows: "But among these Bayesian criteria, BIC has the most general asymptotic properties in the sense that it satisfies the properties of a consistent criterion."

Page 31, line 1. The sentence should read as follows: "One major limitation of Hocking's criterion is that it is not applicable for nonstochastic regressors."

Page 46, line 8. Read $f' \leq f_{opt}$ for $f' \leq f_t$.

Page 47, line 12. The sentence should read as follows: "Once an equilibrium state has been achieved for a given temperature, the temperature is reduced as defined in step 5 and the process started again taking values of the last iteration of the algorithm as the initial values."

Page 182, line 5. The sentence should read as follows: "Our goal is to choose the penalty in such a way that none of the competing models is unknowingly favoured over the others."

Page 241, lines 4-3 from bottom should read as follows: "This implies that the optimal penalty depends not only on the sample size and the number of free parameters, but also on the competing data generating processes."

Page 242, last dot point should read as follows: "The difference between the largest and the smallest MAPCS obtained using the SAO technique is very small, which implies that for equi-dimensional competing alternative models, the MAPCS is insensitive to the starting parameter values of the SAO technique."

Page 244, line 4. Read MAPCS for MACPS.

MONASH UNIVERSITY

THESIS ACCEPTED IN SATISFACTION OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ON............... [redacted] ....

for Sec. Research Graduate School Committee

MODEL SELECTION: AN OPTIMAL APPROACH TO CONSTRUCTING A PENALTY FUNCTION IN SMALL SAMPLES

A thesis submitted for the degree of Doctor of Philosophy in Econometrics

by

GOPAL KRISHNA BOSE
BSc(Hons) and MSc in Statistics (Dhaka, Bangladesh)
MSc in Computation (Manchester, The United Kingdom)

Department of Econometrics and Business Statistics
Faculty of Business and Economics
Monash University
Australia

May 2002

CONTENTS

ACKNOWLEDGMENTS

I wish to acknowledge my deep gratitude to my supervisor, Professor Maxwell L. King, for his guidance, consistent instructions, insightful comments and suggestions throughout the development of this PhD thesis. Without his help and encouragement, this thesis could not have been written. I deeply appreciate his special attention to thesis writing, which improved my writing skills as well as my presentation style of the research results.

I would like to thank Associate Professor Ralph D. Snyder for allowing me to work with him as a part-time Research Assistant for more than two and a half years of my PhD work.

I very much appreciate the Monash Graduate Scholarship (MGS), which was awarded to me by Monash University and the Department of Econometrics and Business Statistics for employing me as a part-time Research Assistant and part-time Assistant Lecturer. The Department of Education, Training and Youth Affairs deserve thanks for their policy of HECS exemption for postgraduate research study.

My thanks are due to Ms. Inge Meldgaard for her technical assistance in various aspects of my computer needs during the period of this PhD work. I wish to thank

Mrs. Mary Englefield, Mrs. Philippa Geurens, and Ms. Julie Larsen for their kind cooperation and administrative help at all stages of the development of this thesis.

I would like to thank my friends, relatives and well wishers for their moral support and encouragement throughout the development of this PhD thesis.

I must appreciate the patience, moral support and understanding of my wife, Sharmila, who practically took care of most of the important aspects of the well being of the whole family while I was involved in the PhD program. My thanks are due to my son, Shuvagata, and daughter, Shuktika, who subsisted on less than their rightful share of attention for the period of my PhD research. Their sentimental attachment strengthened my determination to complete my PhD thesis.

Finally, I dedicate this thesis to the memory of my late father Shama Kanta Bose, who left us long ago, and my mother, Santi Rani Bose, whose profound love and blessing is the corner stone to all achievements in my life.

Above all, I am grateful to God Ramakrishna, without whose blessing, mercy and unceasing help, nothing could have been done by my own humble self.

DECLARATION

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution; and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

[redacted]

Gopal Krishna Bose

ABSTRACT

This thesis is concerned with model selection in the context of the linear regression model and the construction of an appropriate procedure to find the best model from a set of alternative competing models. Of the many available techniques, the information criteria (IC) based model selection technique is the most widely used. There are many IC available in the literature. But from the literature, it is observed that none of the existing criteria perform well in all situations. So from the user's point of view, which criteria one should use to select the best model for a particular data set is a question that is unresolved. The aim of this thesis is to answer this question by developing an IC procedure, which performs better on average in all situations than the best of the existing criteria.

In Chapter 2, we review the relevant literature on model selection. We begin with the literature on model selection based on error sum of squares rules and then review commonly used IC based model selection procedures. We discuss the advantages and drawbacks of the existing criteria, which motivate us to develop a new technique, the performance of which is better on average in all situations than the existing criteria. We also review the simulated annealing optimisation (SAO) technique and its use in econometrics.

A new and more efficient technique for selecting the best combination of parameter drawings and number of replications for estimating the average probability of correct selection (APCS) of the true model via simulation is presented in Chapter 3. A generalised form of penalty functions of six existing IC is also presented. In this chapter, we also propose a new method of model selection. Simulation results show that the mean APCS (MAPCS) obtained from this proposed method is always higher than that of the best of the existing IC.

The application of the SAO technique to maximise the MAPCS for an additive penalty and the maximised log-likelihood; and for a multiplicative penalty and mean squared error, is discussed in Chapter 4. From the simulation results, we observe that the MAPCS obtained from the SAO technique with both types of penalties are always higher than those of the best of the existing IC. We also see that for the same model, the relative penalty varies from data set to data set; and for a particular data set, the relative penalties are different for those competing models with the same number of parameters. This implies that the penalties are not only a function of n, the sample size, and k, the number of free parameters, but also the data generating process, this is in contrast to existing penalty functions that are a function of n and k only. Another interesting outcome of these simulation experiments is that exactly the same MAPCS is obtained from different sets of relative penalties, which implies that there may not be a unique form of penalty function.

In Chapter 5 we investigate the use of the additive and multiplicative penalty with the SAO technique described in Chapter 4 for the special case of equi-dimensional alternative competing models. From the results of the simulation experiments, it is apparent that in all situations the MAPCS obtained using the SAO technique is higher than those of the existing criteria. Another notable outcome is that although the relative penalties are zero for the existing criteria when models are equi-dimensional, the relative penalties that maximise the MAPCS using the SAO technique are different from zero.

In Chapter 6 we develop a method based on the SAO to select the best model in such a way that all the competing models have an equal chance on average of being selected. We apply the SAO technique to minimise the standard deviation (SD) among the APCS for selecting the best model. The simulation results show that for most of the cases with competing models of different dimensions, the variation among the APCS is zero or close to zero, which indicates that the APCS are equal or nearly equal. But, generally the MAPCS is lower than that of the best of the existing criteria. For equi-dimensional competing models, the MAPCS obtained using the SAO technique is generally higher than those of the existing criteria with zero or close to zero variation among the APCS. Therefore, for equi-dimensional competing alternative models, the application of the SAO technique to minimise the variation among the APCS seems to be the best way of selecting the best model without favouring one model over another.

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND OF MODELLING AND MODEL SELECTION

In any modelling, a real-world problem is analysed and, as a result of the analysis, a model that approximates the real situation is developed. According to Pearce (1992), "A model is a formal or informal framework of analysis which seeks to abstract from the complexities of the real world those characteristics of an economic system which are crucial for an understanding of the behavioural, institutional, and technical relationships which underlie the system". Harvey (1981) asserted five characteristics of a good model, namely parsimony, identifiability, goodness of fit, theoretical consistency, and predictive power. Proper specification of the econometric model plays an important role in the selection of the best model from a set of alternative models. Several researchers have stressed the importance of misspecification tests in econometric model building process. For example, Malinvaud (1981) contended that econometricians should place special emphasis on the testing of model specification. In formulating a regression model, misspecification error may arise from (i) omitted variables, (ii) incorrect functional form, (iii) autocorrelation, (iv) heteroscedasticity,

(v) lack of regression parameter consistency, (vi) non-normality of disturbances and (viii) invalid assumptions about the exogeneity of one or more regressors. There are a large number of tests designed to detect these misspecifications and many of these tests are available in modern computer packages which are very comprehensive, for example see Microfit (Pesaran and Pesaran 1987) and SHAZAM (White, 1978). There is a lot of literature on the topic of misspecification tests. Examples are, Eastwood and Godfrey (1992), Ramsey (1969), Durbin and Watson (1951), O'Hagan and McCabe (1975). Discussions of these are beyond the scope of this thesis.

The level of sophistication of a model depends on a number of things, including the mathematical background of the modellers, the nature of the problem, the available information including reliable data and so on. In econometrics and applied statistics, we generally deal with sample data and usually want to draw inferences about the relationship of the variables in the population from which the sample data have been drawn. Because modelling is concerned with the population, we should expect that any model we develop on the basis of a sample will provide an approximation of the relationship for the population. Generally, it is possible to develop many different models for a particular data set and set of variables, and we need to make a choice as to which one of these models is the best approximation of the population, substantively interpretable and as simple (parsimonious) as possible. For example, if the models with and without higher-order interaction terms fit the data well, the simpler models are usually preferable, because higher-order interaction terms are generally difficult to interpret.

Data play a vital role in the econometric model building process. Without good, consistent and reliable data, even the most sophisticated and mathematically sound models fail to represent the phenomenon of interest. In selecting the best representative model for a particular data set, one must keep in mind the purpose of the model, the availability and accuracy of data, ease of model application, and accuracy of the selected model. Models with different subsets of variables produce very different results, raising questions about which one is the best representative model for a particular set of variables. The selection of the best model for a particular data set and set of variables is an important issue in econometrics and statistics for the purpose of valid estimation, inference and prediction. The process of choosing a model from a set of alternative models using the available data and set of variables, is known as model selection. In regression analysis, model selection is the process of selecting a subset of independent variables which best explain changes in the dependent variable.

Econometric modelling usually involves the estimation of a range of models, then the choice of a model that best fits the available data. There are several selection techniques available in the literature to choose the best desirable model using the available data. These selection techniques can be grouped into the following four major classes: (i) hypothesis testing based procedures, (ii) Bayesian criteria, (iii) the use of residual sum of squares, and (iv) information criteria (IC). Hypothesis testing based procedures involve the use of a series of pairwise hypothesis tests to select the final model. Several researchers have proposed various testing procedures for model selection, see for example, Gaver and Geisel (1974), Atkinson and Fedorov (1975), Leamer (1978), White (1982a, 1982b, 1983, 1990), MacKinnon (1983), Davidson and MacKinnon (1984), Bunke and Droge (1985), Linhart and Zuchini (1986), McAleer (1987), Grassa (1989), Brownstone (1990), Pötscher (1991), and Maddala (1992). Unfortunately, this method of model selection has many limitations as mentioned by Granger *et al.* (1995). At each step, one model is chosen as the null hypothesis and can be unfairly favoured because the probability of wrongly rejecting it is set to a small value like 0.05 or 0.01. This can be particularly troublesome in situations where the test being used is not particularly powerful and therefore a choice of the null hypothesis model is the most likely outcome. There is also the well known problem of pre-test bias (see for example Wallace, 1977, and Giles and Giles, 1993). Finally different researchers working on the same model selection problem could easily end up with different final models purely because they performed the tests in a different order or used different significance levels. In Bayesian criteria the penalised posterior probability is maximised using Bayes theorem. It is a well established concept in model selection and it uses posterior odds ratios for the comparison of the model. DeGroot (1970) gave a detailed description of Bayesian criteria.

The available methods that use residual sum of squares for regression model selection include the use of the coefficient of multiple determination (R^2) and adjusted coefficient of determination ($\bar{R}^2$) proposed by Theil (1961), Mallows' (1964) C_p criterion, Amemiya's (1980) PC criterion (see for example, Judge *et al.* (1985, 1988)). Nowadays C_p and PC are regarded as IC and these IC are related to the IC proposed by Rothman's (1968) J_p and Hocking's (1976) S_p and are based on minimising the mean square error of prediction. Zhang's (1992) final prediction error (FPE) criterion and Rahman's (1998) generalised model selection criterion for linear regression are also based on residual sum of squares and used as IC. R^2 is a nondecreasing function of the number of explanatory variables and generally is inadequate to pick out the best model. Thus, $\bar{R}^2$, which is R^2 adjusted for the residual degrees of freedom was defined to overcome this problem. Unfortunately, it does not penalise the loss of degrees of freedom sufficiently. Dhrymes (1970) mentioned that if the purpose of the selected model is prediction, then this technique is not suitable as $\bar{R}^2$ produces an unnecessarily large prediction error. As there are several problems with using hypothesis testing and residual sum of squares based model selection techniques, nowadays IC based model selection procedures are frequently used by researchers to select the best model for a particular data set.

An IC based technique is the most widely accepted class of model selection procedures and is based on choosing the model with the largest maximised log-likelihood function minus a penalty which is an increasing function of the number of free parameters included in the model. Examples of these procedures include Akaike's (1973, 1974) IC (AIC), Schwarz's (1978) Bayesian IC (BIC) and Hannan and Quinn's (1979) procedure denoted HQ. Fox (1995) has also demonstrated how a number of other model selection procedures, particularly those developed for choosing from a set of linear regression models, can be thought of in the IC framework. Those of interest in this thesis include Theil's (1961) $\bar{R}^2$ criterion denoted by RBAR, Mallows' (1964) C_p procedure denoted MCP, Schmidt's (1975) generalised cross-validation (GCV) procedure and Hocking's (1976) S_p, criterion denoted by HOC. As Granger *et al.* (1995) and others have noted, the IC approach has the advantages that (i) no particular model is favoured because it has been chosen to be a null hypothesis, (ii) the order of computation is irrelevant, (iii) pretesting bias is not an issue, (iv) if the IC procedure is asymptotically consistent, the correct model is chosen with probability one asymptotically, and (v) there is no need to choose an arbitrary level of significance although there is the related issue of which penalty function is appropriate.

The latter is a major issue, which has featured in the literature. There is currently little agreement about what the form of the penalty function should be. The early literature focused on asymptotic arguments to justify various choices of penalty functions, see for example Akaike (1973), Schwarz (1978), and Hannan and Quinn (1979). Since then we have seen a number of Monte Carlo studies of the small sample properties of different model selection procedures which have shown a number of problems. They suggest that asymptotic properties are no guarantee of acceptable small sample properties. For example, Grose and King (1994), in the context of choosing between first-order autoregressive and first-order moving average disturbances in the linear regression model, have shown that a particular model can be unfairly favoured because of the functional form of its log-likelihood function. They also found that the presence of nuisance parameters can adversely affect the probabilities of correct selection.

There are clear parallels between the hypothesis testing literature and the model selection literature, although it does appear that the latter lags the former. The computer revolution has meant that we can now ask what kind of testing procedures would we like to use rather than what kind of testing procedure is convenient to use (see for example King, 1987). We should be asking similar questions for model selection, in the context of finite samples. We now regularly use simulation, a numerical technique for conducting experiments on a digital computer, to find critical values for hypothesis tests. Can we use similar methods to find penalty values for IC procedures that are in some sense optimal?

The penalty function of almost all existing IC is a function of n, the number of observations and k, the number of free parameters. Thus a change of data set and sets of independent variables do not have any impact on the penalty function provided n

and k remain unchanged. Another problem of existing IC is that none of these performs well in all situations. For example, BIC generally favour the model with the smallest number of parameters, while $\overline{R}^2$ favours the model with the largest number of parameters. The performance of HQ is in between BIC and $\overline{R}^2$. Thus, from the user's point of view, it is very difficult to choose an IC for selecting the best model for a particular data set and set of independent variables.

In this thesis we investigate a new approach to IC model selection in the context of the classical problem of choosing between different linear regression models. We use the idea of a data-oriented penalty function, which was first introduced by Rao and Wu (1989), for the model selection problem in the linear regression models. For the selection of AR time series models, Chen *et al.* (1993) used the same idea for finding penalties. For the model selection problem in the context of classical regression model using the general information criteria (GIC), Bai *et al.* (1999) also applied the idea of data-oriented penalties. Here, we also use the same idea, but our approach involves the use of a simulation method to estimate probabilities of correct selection and choosing penalties that optimise these probabilities on average. Maximising simulated probabilities can be a difficult optimisation problem. We use a relatively new optimisation algorithm called simulated annealing to overcome any difficulties in this regard. The main feature of this optimisation algorithm is that it can find the global maximum/minimum in the presence of local maximum/minimum, and it is a very robust algorithm (Goffe *et al.* (1994)). A detailed description of this optimisation technique is given in Section 2.4 of Chapter 2.

A similar approach to model selection has been investigated by Kwek and King (1997a, 1997b) and Kwek (2000) in the context of selecting between different conditional heteroscedastic processes. Billah and King (2000a, 2000b) have also considered a similar method for choosing between different time-series processes for linear regression disturbances. Each of these studies has involved choosing between variance-covariance matrix functions with restricted parameter spaces. In this thesis we consider model selection of different mean processes with unrestricted parameter spaces.

The specific aims of this thesis are to:

(i) develop a generalised form of some well used IC (AIC, BIC, HOC, HQ, GCV),

(ii) derive a method of model selection which performs better than all the existing criteria in all situations,

(iii) to investigate the use of multiplicative penalties with mean squared error to select the best model,

(iv) investigate the performance of the simulated annealing optimisation (SAO) technique to find the penalty for selecting the best model using additive and multiplicative penalties,

(v) study the performance of additive and multiplicative penalties with the SAO technique using the idea of controlling the probability of correct selection, and

(vi) make recommendations for the use of appropriate methods of model selection,

all in the context of the linear regression model.

1.2 OUTLINE OF THE THESIS

In Chapter 2, we present a survey of the relevant literature in the field of model selection. The use of sum of squared errors for selecting the best model is discussed in Section 2.2. The idea of multiplicative penalties is introduced and the technique of using multiplicative penalties for selecting the best model is also discussed. In Section 2.3, we review widely used IC procedures. The major advantages and disadvantages of using IC for model selection purposes are also discussed. In addition, this section contains some directions for overcoming these problems of model selection. SAO is introduced in Section 2.4, which also contains a literature survey of the use of SAO in econometrics.

A generalised form of six widely used IC is given in Chapter 3. A new method of selecting the true model on the basis of the maximum average probability of correct selection of models is introduced in Section 3.2. In Section 3.3.1, we propose a new method of data generating processes for simulation experiments for selecting a model from a set of competing alternative models in linear regression settings. In Section 3.3.2, it is shown through the Monte Carlo technique that for a fixed number of simulations qN, where q is the number of parameter drawings and N is the number of replications, greatest accuracy is obtained by setting N=1. In this chapter, we propose five new IC and the performance of these IC are evaluated in Section 3.3.3 via Monte Carlo experiments.

In Chapter 4, we introduce the application of the SAO technique to selecting the best model from a set of competing alternative models. The idea of an additive penalty for the maximised log-likelihood and a multiplicative penalty for the mean squared error is introduced in Sections 4.2 and 4.3, respectively. Simulation experiments were conducted to evaluate the performance of these penalties compared to those of existing criteria. The experimental design of these simulations is outlined in Section 4.4 and results are presented in Section 4.5.

In Chapter 5 we investigate selection between competing models with an equal number of parameters, which is a special case of Chapter 4, where the models have different numbers of parameters. For the existing IC, the effect of the penalty function is cancelled out in selecting the best model from a set of equi-dimensional competing models, because the penalty functions of the existing criteria are functions of k, the number of free parameters and n the sample size. In this situation, selection of the best model depends only on the maximised log-likelihood. The objective of this chapter is to investigate the performance of the SAO technique for finding optimum penalties, when the competing models are equi-dimensional. Section 5.2

contains a description of the Monte Carlo simulation experiments and Section 5.3 presents the results of these simulation experiments.

In Chapter 6, we discuss the issue of finding penalties that make the average probability of correct selection (APCS) equal. We propose a method of minimisation of variation among the APCS in Section 6.2. We apply the SAO technique to implement our proposed method to find the penalties for selecting the models with the constraint that the variation among the APCS is as minimal as possible. The standard deviation among the APCS is used as a measure of variation. Section 6.3 gives a description of the Monte Carlo experiments and discussion of the results, when the competing models have an unequal number of parameters. Section 6.4 is similar to Section 6.3, but the competing models have an equal number of parameters.

Chapter 7 is the final chapter of this thesis. In this chapter we summarise the findings of our research work, make some concluding remarks and make some suggestions for future research in the area of model selection.

1.3 COMPUTATIONS

We used GAUSS System Version 3.2.18 (Aptech Systems Inc.) on a Pentium II personal computer with 64MB RAM for all simulation experiments used for generating results for this thesis. We also used SPSS 9.0 for Windows and Microsoft Excel 97 for some of our computations.

CHAPTER 2

LITERATURE SURVEY OF MODEL SELECTION AND THE SIMULATED ANNEALING OPTIMISATION TECHNIQUE

2.1 INTRODUCTION

In econometrics we are often forced to ask the data to choose a model for us from a set of alternative models. There is a lot of literature on this problem with a range of methods or strategies being suggested as possible solutions. A sequence of pairwise tests is one of the techniques often used to select the best model. Unfortunately this technique has several drawbacks, which were discussed in Chapter 1. Akaike (1974) pointed out that the use of hypothesis testing is not a proper method of statistical model selection. Granger *et al.* (1995) contend that model selection should be based on well-thought-out model selection procedures rather than a series of classical pairwise tests and information criteria (IC) based model selection procedures provide a good framework. These procedures overcome most of the disadvantages of pairwise tests of hypotheses, which were discussed in the previous chapter. The purpose of the first two sections of this chapter is to survey the relevant literature on model selection based on residual sum of squares and IC procedures.

An IC based model selection procedure depends on maximised log-likelihood functions and a penalty function, and the model with the largest penalised log-likelihood value among the competing models is selected as the best model. Unfortunately there is little agreement about the form of penalty function and none of the existing IC procedures perform well in all situations. In this thesis we propose a simulation based technique to estimate probabilities of correct selection and choosing penalties that optimise these probabilities on average. Our objective function is made up of averaged probabilities which in turn have been estimated by the Monte Carlo method; it is therefore a step function (details are discussed in Section 3.2 of Chapter 3). These types of functions are difficult to optimise using standard numerical iterative methods. They have many plateaux, which will cause standard optimisation techniques great difficulties. Most of the standard iterative algorithms fail to find the global maximum or minimum for these types of objective functions. From the work of Kirkpatrick *et al.* (1983), Romeo *et al.* (1984), White (1984) and Goffe *et al.* (1994), it does seem that the simulated annealing optimisation (SAO) algorithm performs well at finding the global maxima in the presence of local maxima and for functions like ours which have plateaux and other ill-behaviour. The algorithm works well because it accepts both uphill and downhill moves in a random but systematic manner thus allowing the algorithm to by-pass local maxima/minima and plateaux. In Section 2.3 we will provide a brief literature review of this optimisation technique used in econometrics. In the final section we will make some concluding remarks.

2.2 USE OF RESIDUAL SUM OF SQUARES FOR MODEL SELECTION

The purpose of this section is to survey the relevant literature on model selection, which use residual sum of squares. The available methods in the literature that use residual sum of squares for model selection are the coefficient of multiple determination and the adjusted coefficient of determination suggested by Theil (1961), Mallows' (1964) C_p criterion and Amemiya's (1980) PC criterion. Nowadays C_p and PC are regarded as IC and these IC are related to the IC proposed by Rothman (1968) denoted J_p and Hocking (1976) known as S_p. Zhang's (1992) final prediction error (FPE) criterion and Rahman's (1998) generalised model selection criterion are also based on the residual sum of squares.

The coefficient of multiple determination denoted by R^2 was the first criterion used in econometrices and other areas for model selection purposes. It is also used as a goodness of fit statistic for selecting a model. The interpretation of R^2 is, the proportion of the variation in the dependent variable that is explained by the independent variables in the model and its value lies between zero and one. In the linear regression model

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}, \tag{2.1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations on the dependent variable, $\mathbf{X}$ is an $n \times k^*$ matrix with a column vector of ones in its first column and in the remaining $(k^* - 1)$

columns (k^*-1) non-stochastic variables (i.e., there are (k^*-1) non-constant regressors), $\mathbf{u}$ is an $n \times 1$ vector of random disturbances following $N(0, \sigma^2 \mathbf{I})$ and β is a vector of k^* parameters. R^2 is defined as

$$R^2 = \frac{\sum_{t=1}^{n} (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^{n} (y_t - \bar{y})^2} = \frac{SSR}{TSS}, \tag{2.2}$$

where SSR is the sum of squares due to the regression, TSS is the total sum of squares, $\hat{y}_t$ is the estimated value of y_t based on the ordinary least square method, and $\bar{y}$ is the mean value of y_t, $(t = 1, 2, \ldots, n)$. It is clear that by simply adding new independent variables to (2.1), the value of R^2 will increase and never decrease, and $R^2 \to 1$ as $k^* \to n$. This can be demonstrated in the following way. Suppose we have a model which regresses a variable y, on the k° variables $x_1, x_2, \ldots, x_{k^\circ}$. This model is equivalent to the model which regresses y on the $k^\circ + 1$ variables $x_1, x_2, \ldots, x_{k^\circ+1}$, subject to the restriction that the coefficient on $x_{k^\circ+1}$ is zero. So the R^2 value from the latter model is either greater than or equal to (if and only if the estimated coefficient of $x_{k^\circ+1}$ is identically zero) the earlier model. Hence the R^2 criterion for selecting the true model is inadequate as the model with a larger number of unnecessary independent variables will produce a larger R^2 value. To overcome this deficiency of R^2, Theil (1961) suggested an adjusted R^2 denoted by $\bar{R}^2$, defined by

$$\bar{R}^2 = 1 - \frac{\frac{RSS}{n-k^*}}{\frac{TSS}{n-1}}. \tag{2.3}$$

This adjusted coefficient of multiple determination is adjusted for the residual degrees of freedom. Since $R^2 = 1 - \frac{RSS}{TSS}$, so $\bar{R}^2$ can be expressed as a function of R^2 as follows:

$$\bar{R}^2 = 1 - \frac{n-1}{n-k^*}(1-R^2). \tag{2.4}$$

As mentioned earlier, with the inclusion of an additional variable, R^2 cannot fall, but $\bar{R}^2$ may fall as it takes account of the residual degrees of freedom. Actually $\bar{R}^2$ is R^2 with a penalty for the additional regressors. Dhrymes (1970) showed that the value of $\bar{R}^2$ will increase by adding an additional regressor if the t-value of the coefficient of this added regressor is greater than unity in absolute value. He analytically showed that

$$\frac{1-\bar{R}^2_{k^\diamond}}{1-\bar{R}^2_{k^\diamond-1}} = \frac{n-k^\diamond}{(n-k^\diamond-1)+t^2_{k^\diamond}}, \tag{2.5}$$

where $\bar{R}^2_{k^\diamond}$ and $\bar{R}^2_{k^\diamond-1}$ are the values of the coefficient of determination of the models with $k^\diamond$ and $(k^\diamond-1)$ non-constant regressors, respectively and $t^2_{k^\diamond}$ is the value of the square of the t-statistics for the $k^{\diamond th}$ regression coefficient of the model with $k^\diamond$ non-constant regressors.

From this equation it is clear that $\overline{R}_k^2$ will increase for the added k^{th} non-constant regressor if $\frac{n-k^*}{(n-k^*-1)+t_k^2}<1$, which is only true when $t_k^2>1$. According to Theil (1971), the model selected on the basis of largest $\overline{R}^2$ value 'on average' is the correct model.

One of the main problems of $\overline{R}^2$ is that it does not penalise the loss of degrees of freedom sufficiently. For example, suppose the number of regressors in the model is large with all relevant regressors along with some irrelevant regressors in the model. In this situation the estimated value of the residual variance is unbiased, but $\overline{R}^2$ tends to select the model with a large number of regressors. Another problem with $\overline{R}^2$ is its poor predictive performance. According to Schmidt (1973, 1975), if the model contains variables of the true model with some irrelevant variables, then the $\overline{R}^2$ criterion tends to fail to identify the true model. Dhrymes (1970) also mentioned that $\overline{R}^2$ produces unnecessarily large prediction errors, so this technique may not be appropriate if the selected model is used for prediction purposes.

If it can be assumed that the behavior of the regressors in the future is the same as in the sample, then the mean squared error may be shown to be approximately equal to

$$\frac{2(k_j^*-1)\,\sigma^2}{n}+\frac{RSS_j}{n}, \tag{2.6}$$

where RSS_j is the residual sum of squares. $(k_j^* - 1)$ is the number of non-constant regressors included in the j^{th} model and σ^2 is the unknown population error variance. Mallows' C_p criterion and Amemiya's PC criterion differ only in the estimated value of this σ^2. Mallows' C_p criterion uses $\dfrac{RSS}{n-k^*+1}$ as an estimate of σ^2, where RSS is the residual sum of squares obtained from the complete model using all $(k^* - 1)$ non-constant regressors, while the estimated value of σ^2 in Amemiya's PC criterion is $\dfrac{RSS_j}{n-k_j^*+1}$, where RSS_j is the residual sum of squares obtained from the model using only $(k_j^* - 1)$ regressors and $k_j^* < k^*$. So these two criteria can be written as follows:

$$C_p = \frac{2(k^* - 1)\,RSS}{n - k^* + 1} + RSS_j\,. \tag{2.7}$$

$$PC = \frac{2(k_j^* - 1)RSS_j}{n - k_j^* + 1} + RSS_j = \frac{n + k_j^* - 1}{n - k_j^* + 1}RSS_j\,. \tag{2.8}$$

In Amemiya's PC criterion it is assumed that the model with $(k_j^* - 1)$ non-constant regressors is the correct model and the model with $(k^* - 1)$ non-constant regressors includes a number of irrelevant regressors. Maddala (1992) argued that it is not a reasonable assumption that every one of the models is the true model to estimate σ^2, rather the asymptotic estimate of σ^2 used in C_p is more reasonable.

According to Gorman and Toman (1966), Mallows' C_p for the multiple linear regression model with known variance σ^2 can be written as follows:

$$C_p = \frac{RSS_j}{\sigma^2} - (n - 2(k_j^* - 1)), \tag{2.9}$$

where RSS_j is the residual sum of squares, n is the sample size and $(k_j^* - 1)$ is the number of non-constant regressors for the j^{th} model.

Under the assumption that the regressors follow a multivariate normal distribution, Hocking (1976) derived the S_p criterion by minimising the conditional mean squared error of prediction, which can be written as:

$$S_p = \frac{n - k_j^* + 1}{n - k_j^*} RSS_j. \tag{2.10}$$

The basic difference between Mallows' C_p, Amemiya's PC and Hocking's S_p criterion is that the former two depend on the assumption of non-stochastic regressors, whereas the latter depends on the assumption of stochastic regressors. Both Mallows' C_p and Amemiya's PC reduce to Hocking's S_p criterion for stochastic regressors (Kinal and Lahiri (1984)).

Zhang (1992) mentioned that all the existing criteria can be shown to be asymptotically equivalent to minimising (with respect to k_j^0)

$$C(k_j^{\xi}, \eta) = RSS(k_j^{\xi}) + \eta k_j^{\xi}, \; 0 < k_j^{\xi} \leq (k_j^{*} - 1), \tag{2.11}$$

where $RSS(k_j^{\xi})$ is the residual sum of squares when only the first k_j^{ξ} regressors are entered in the j^{th} model and η represent the penalty for over fitting. This criterion is called the final prediction error (FPE) criterion. According to Zhang, AIC, C_p and BIC are special cases of the above criterion. When $\eta = 2$, (2.11) corresponds to AIC and C_p; and (2.11) corresponds to BIC when $\eta = \ln n$.

It was mentioned earlier that with the inclusion of an additional variable into an existing model, the value of R^2 increases, which means that the value of the residual sum of squares decreases. So there is a tendency to select the model with an unnecessarily large number of independent variables if we use the smallest residual sum of squares as the criteria for model selection. Hence we need some adjustment to the residual sum of squares, which can be done with the help of a penalty function. Let us assume q_j (multiplicative penalty) is the penalty function for the j^{th} model. Then the model with smallest J_j (multiplicative information criteria for the j^{th} model) will be selected, where Rahman (1998) defined J_j as follows:

$$J_j = RSS_j q_j. \tag{2.12}$$

J_j is called the penalised sum of squares error and model j will be accepted over all other models i if

$$J_j < J_i, \quad \forall i = 1, 2, \ldots, (j-1), (j+1), \ldots, m. \tag{2.13}$$

Rahman (1998) gave the functional form of a possible q_j as follows:

$$q_j = (n - k_j)^{a_1} (a_2)^{\frac{k_j}{n}}. \tag{2.14}$$

where a_1 and a_2 are arbitrary constants, and k_j is the number of free parameters included in the model under consideration.

Analytically he showed that all the existing criteria are a special case of this new criterion in the linear regression setting. For example, if $q_j = e^{\frac{2k_j}{n}}$ then J_j is approximately equivalent to Akaike's information criterion (AIC) and if $q_j = \frac{k_j}{n}$ then J_j is approximately equivalent to Schwartz's Bayesian information criterion (BIC).

He also mentioned that by choosing appropriate values of a_1 and a_2, it is possible to develop an infinite number of new criteria which will perform well in a range of situations. But the problem with this penalty function is setting the values of a_1 and a_2 for a particular data set. Also the penalty function is a function of n, the sample size, and k_j the number of free parameters; and independent of data values, i.e. for the same set of competing models, a change of data sets or regressors does not have any impact on the penalty function. To overcome these problems, we redefine the multiplicative information criteria using mean squared error, which is discussed in Chapter 4.

2.3 INFORMATION CRITERIA BASED MODEL SELECTION

The purpose of this section is to review the relevant literature on information criteria (IC) based model selection. An IC model selection procedure is based on choosing the model with the largest maximised log-likelihood minus a penalty term. There are a number of penalty functions available in the literature. All of them are a function of the number of free parameters in the model and many include the sample size. But from the user's point of view, which one is the best for a particular data and set of models is a question to be answered as there is little agreement about what the correct answer to this question is. Among the available IC, AIC, proposed by Akaike (1973), is the most widely used and popular criterion in economics and econometrics (see for example, Hurvich and Tsai (1991), Mills and Prasad (1992), Fox (1995) and Hughes (1997)). Therefore, we begin our review with AIC, followed by the BIC model selection criterion proposed by Schwarz (1978) which assumes a prior distribution of the parameter of the proposed model. In addition to these two widely used criteria, we will look at the literature on some other criteria, which include Schmidt's (1975) Generalized Cross Validation (GCV) criterion, Hannan and Quinn's (1979) criterion, (HQ), Hocking's (1976) criterion (HOC), and JIC, recently proposed by Rahman and King (1999).

Akaike's Information Criterion (AIC) grew out of Akaike's (1973, 1974) research on selecting the best order of an autoregressive process. It gives a measure of the distance between estimated model and the true data generating process using Kullback-Leibler's (1951) information theory. AIC was derived under the assumption that the true distribution can be described by the given model when its parameters are suitably adjusted. Several authors define AIC in different forms, but the most popular one for any general model selection purpose can be written as the penalised log-likelihood form

$$\mathrm{AIC}_j = L_j(\hat{\theta}_j) - k_j, \tag{2.15}$$

where $L_j(\hat{\theta}_j)$ is the maximised log-likelihood function and $\hat{\theta}_j$ is the maximum likelihood estimator of θ_j, the vector of k_j free parameters included in the model M_j, $j = 1, 2, \ldots, m$. AIC selects the model for which AIC_j is the maximum among the m models.

The mean expected log-likelihood can be used as a measure of the goodness of fit of a model, which is defined as the mean of the expected log-likelihood of the maximum likelihood. The larger the mean expected log-likelihood, the better the fit of the model. When there are several models whose values of the maximum log-likelihood function are the same or approximately the same then one should choose the model with the smallest number of parameters, which is called the principle of parsimony. AIC was developed to measure these two properties of the model, though it is successful at measuring the first property rather then the second. Hurvich and Tsai (1989) concluded from their Monte Carlo study that AIC tends to overfit the data for very small samples. Sugiuro (1978) proposed a finite sample correction of AIC for the case of data generated from a normal distribution. To overcome the overfitting problem of AIC, Hurvich and Tsai (1989) derived a bias corrected version of AIC for regression and autoregressive time series models which they called AIC_c and expressed it as:

$$\mathrm{AIC}_c = \mathrm{AIC}_j + \frac{(k_j+1)(k_j+2)}{n-k_j-2}. \tag{2.16}$$

This modification of AIC is useful when the sample size is small relative to the dimension of the model and is asymptotically efficient if the true model has infinite dimensions. The simulation results of Hurvich and Tsai (1993) demonstrate that in small samples, AIC_c is superior in terms of bias and strongly outperforms AIC for the models that contain more unknown parameters than the univariate AR models.

A generalised form of AIC was suggested by Bhansali and Downham (1977) and is

$$\mathrm{AIC}(\gamma)_j = L_j(\hat{\theta}_j) - \gamma k_j, \tag{2.17}$$

where γ is a constant and greater than one. They argued that this generalisation will help to improve the problem of overfitting and if γ increases gradually as n increases, then it will be a consistent criterion, which is discussed below.

Consistency is an important asymptotic property of a model selection criterion. It requires the model selection criterion to be able to select the true model from a set of models with probability tending to one as the sample size tends to infinity, assuming that true model is contained in the choice set of models. Shibata (1976), who introduced the idea of consistency, recognised that AIC is not a consistent criterion. Nishii (1988) developed a technique to test the consistency of a criterion, which is briefly discussed below.

Let IC_j be the information criterion for the j^{th} model and defined as

$$\mathrm{IC}_j = L_j(\hat{\theta}_j) - p_j = L_j(\hat{\theta}_j) - k_j D_n, \tag{2.18}$$

where p_j is the penalty for the j^{th} model and D_n is a function of n.

By Theorem 4 of Nishii (1988), IC_j will be strongly consistent if D_n satisfies the following two conditions:

$$\text{(i)} \quad c_1 = \lim_{n \to \infty} \frac{D_n}{n} = 0 \text{ and (ii) } c_2 = \lim_{n \to \infty} \frac{D_n}{\ln \ln n} = +\infty;$$

and IC_j will be weakly consistent if D_n satisfies the conditions

$$\text{(i)} \quad c_1 = \lim_{n \to \infty} \frac{D_n}{n} = 0 \text{ and (ii) } c_2 = \lim_{n \to \infty} D_n = +\infty.$$

If we apply this to AIC, it is transparent that AIC is not consistent as the penalty function for AIC is independent of n, the sample size. Some other researchers have also shown that AIC is not consistent (see for example, Atkinson (1980), Hannan (1982), Shibata, (1986), and Koehler and Murphree (1988)).

Akaike's 1973 and 1974 research work was extended by Schwarz (1978), who gave a Bayesian solution to the problem. As it is based on the Bayesian approach, it has become known as the Bayesian Information Criterion (BIC). During 1978 two more Bayesian criterion were proposed by Sawa (1978) and Leamer (1978). But among these Bayesian criteria, BIC has the most general asymptotic properties. Rissanen (1978) also proposed an IC based on the Bayesian approach, which is the same as BIC. In a similar way to AIC, BIC can be expressed in the penalised log-likelihood form as follows:

$$\mathrm{BIC}_j = L_j(\hat{\theta}_j) - \frac{k_j}{2}\ln n \,. \tag{2.19}$$

BIC assumes a proper prior distribution of the parameters of the proposed model and selects the model with the highest asymptotic posterior probability. So the use of Schwarz's criterion may be difficult if the proper prior distribution of the parameters is not clearly defined (Akaike, 1981). The criterion developed by Rissanen (1978, 1986, 1987, 1988) on the basis of the minimum description length (MDL) overcomes this problem of BIC.

Hannan and Quinn (1979) derived a criterion for selection of the order of an autoregressive model, i.e. to determine the most desirable lag-length, which is known as the Hannan and Quinn (HQ) criterion. This criterion is less commonly used, but is ideal for the comparison of AIC and BIC, because it shares the property of both AIC and BIC. The criterion can be written in penalised log-likelihood form as

$$\mathrm{HQ}_j = L_j(\hat{\theta}_j) - k_j \ln \ln n . \tag{2.20}$$

They recognised that BIC is a consistent criterion and the characteristics of HQ are similar to those of 3IC with the exception that in small samples, the two criteria are likely to select different models. Fox (1995) noted that this criterion shares a property of both AIC and BIC in that its marginal penalty is constant as k_j increases for fixed n. Many researchers showed that this is a consistent criterion, see for example Nishii (1988), Hannan and Quinn (1979), and Atkinson (1981).

The variable selection criterion proposed by Mallows (1964) has been widely used in many social sciences including economics and econometrics. As new and more efficient methods became available for variable and model selection purposes, the acceptance of this criterion has decreased. The penalised log-likelihood form of this criterion henceforth will be denoted as MCP to distinguish it from C_p defined in Section 2.2. It has the following form:

$$\mathrm{MCP}_j = L_j(\hat{\theta}_j) - \frac{n}{2} \ln\left(1 + \frac{2k_j}{n - k^*}\right), \tag{2.21}$$

where k^* is the dimension of the largest model which nests all possible models. This is a consistent criterion (see. Atkinson (1981) and Nishii (1988)). In Mallows' criterion, the variance estimate is obtained from a regression model which includes the entire set of regressors. An alternative estimate of the variance has been suggested using only the number of regressors in the model under consideration, which results in a new criterion proposed by Rothman (1968) and can be expressed in the following penalised log-likelihood form.

$$J_{P_i} = L_i(\hat{\theta}_i) - \frac{n}{2}\ln(n + k_i) + \frac{n}{2}\ln(n - k_i) . \tag{2.22}$$

The same criterion was suggested by Akaike (1969) and Amemiya (1972, 1980). Akaike called it the Final Prediction Criterion (FPC), while Amemiya called it the Prediction Criterion (PC).

Hocking (1976) suggested a model selection criterion denoted by S_p. The penalised log-likelihood form of this criterion can be given by

$$\mathrm{HOC}_i = L_i(\hat{\theta}_i) + \frac{n}{2}\ln(n - k_i) + \frac{n}{2}\ln(n - k_i - 1) . \tag{2.23}$$

This criterion was thoroughly reviewed by Thompson (1978), and was given an alternative justification by Breiman and Freedman (1983). In his paper, Thompson (1978) gave the derivation and justification of this criterion along with C_p and J_p.

One major limitation of Hocking's criterion is it is not applicable for nonstochastic regressors.

Cross-validation is a statistical tool, which uses the technique of splitting data into two sets. The first set is used to develop the model called the model building set, while the second set is used to evaluate the reasonableness and predictive ability of the selected model. This technique was suggested by Schmidt (1971, 1974, 1975), Allen (1971, 1974), Stone (1974) and Geisser (1974, 1975). Schmidt called it the cross-validation score SSPE (Sum of Squared Predictive Error), while Allen called it PRESS (Predictive Sum of Squares). Typically this technique involves deleting an observation from the sample, then fitting the model using the reduced sample, which is used to predict the deleted observation. This is repeated for each observation in the sample and the model with the smallest mean squared error of prediction is selected as the best model. The direct computation of the CV is burdensome. Schmidt (1971) and Allen (1971) gave a formula to compute CV directly, but this criterion is different in nature from the other model selection criterion. Craven and Wahba (1979) gave a generalised form of CV and called it Generalised Cross Validation (GCV), which is an approximation of CV and is a comparable form to the other criteria. Nishii (1988) showed that GCV is not a consistent criterion. According to Fox (1995), the penalised log-likelihood form of the GCV criterion can be written as

$$\mathrm{GCV}_j = L_j(\hat{\theta}_j) + n\ln\left(1 - \frac{k_j}{2}\right). \tag{2.24}$$

Fox (1995) expressed Theil's (1961) $\bar{R}^2$ criterion in the following penalised log-likelihood form for model selection purpose and we will henceforth denote it as RBAR to distinguish it from R^2 and $\bar{R}^2$ defined in Section 2.2.

$$\mathrm{RBAR}_j = L_j(\hat{\theta}_j) + \frac{n}{2}\ln(n - k_j). \tag{2.25}$$

Rahman and King (1997) derived an analytical formula for finding the probability of correct selection of the true model from a set of competing alternative regression models. They observed that for samples of size at least 8 and when the model with the lowest number of regressors is true then the probabilities of correct section obtained from AIC, BIC and RBAR satisfy the inequality BIC>AIC>RBAR. On the other hand, when the model with the highest number of regressors is true then the picture is exactly reverse, i.e. the probabilities of correct selection are RBAR>AIC>BIC. This behaviour of BIC and RBAR motivated them to develop a new criterion on the basis of the simple average of the penalty functions of BIC and RBAR. They called it the joint information criterion (JIC) (Rahman and King (1999)) which can be expressed as the following penalised log-likelihood form.

$$\mathrm{JIC}_j = L_j(\hat{\theta}_j) - \frac{1}{4}\left(k_j \ln n - n\ln\left(1 - \frac{k_j}{n}\right)\right). \tag{2.26}$$

They showed that it is a strongly consistent criterion, which performs well in a range of situations and therefore is a very competitive model selection criterion. In their paper, they mentioned that the probability of correct selection is directly related to the sample size but inversely related to the error variance.

The small sample performance of AIC and BIC was studied by Lütkephol (1984). He concluded that for both small and large samples, the performance of AIC is better than BIC in one-step-ahead forecasting for selecting AR(1) and MA(1) models, but worse in five-step-ahead forecasting. Meese and Geweke (1984) also compared the forecasting ability of AIC and BIC along with three other criteria and concluded that AIC performed the best in most of the cases. Schwarz (1978) reported that for large samples, the performance of BIC differs markedly from that of AIC with respect to selecting the correct model. For selecting the best model from a large set of models, Kohn (1983) found that BIC consistently chooses a model with the smaller dimension. From the research results of Hurvich and Tsai (1991), it is clear that for small samples the modified version of AIC, AIC_c, performs significantly better than AIC and BIC and marginally better than these criteria for moderately small samples. In another research paper, Hurvich and Tsai (1990) reported that under certain conditions, AIC is likely to perform better than BIC in small samples. Crato and Ray (1996) conducted a large-scale simulation study to compare the performance of AIC, AIC_c and BIC, and concluded that for pure fractional noise, the performance of BIC is better than the other two criteria. In order to compare the performance of AIC, AIC_c, BIC, HQ, MLD and RBAR, Mills and Prasad (1992) conducted simulation experiments in determining the correct data generating process in autoregressive and linear regression models. From their simulation results, they discovered that the performance of AIC_c is better in small samples and with an increase in sample size, the relative performance of AIC_c gradually becomes worse. AIC tends to over parameterise the model. They concluded that from a user's point of view, BIC should be the criterion to be used for selecting the correct model. From the simulation study of Schmidt and Tscherning (1993), it was concluded that among AIC_c, BIC and HQ, the performance of AIC_c is the best and HQ is the worst in selecting time-series ARIMA models. More detailed discussion of the above mentioned information criteria can also be found in Hughes (1997), who derived an AIC-type criteria using KL information in the presence of one-sided information on the parameters under dispute. He pointed out that AIC out performs BIC if the larger model is the true model. On the other hand, the reverse picture is observed if the model with smaller number of parameters is the true model. From a simulation study, Kwek (2000) found that in small samples, the performance of RBAR is the best and BIC is the worst procedure for selecting the correct model among autoregressive conditional heteroscedastic (ARCH) models and generalised ARCH (GARCH) models.

Stone (1977) and Nishii (1986) noted that CV is asymptotically equivalent to AIC. From the simulation results of Holmes and Hutton (1989), it is apparent that if there is a weak relationship between the dependent and independent variables, then the RBAR criterion performs better in terms of selecting the true model. But if the relationship is strong then AIC, BIC, HQ and PC choose the correct model with high probabilities and in this situation the performance of BIC is the best, while the performance of RBAR is the worst. For selecting the best predictive model from a class of linear regression models, Shao (1993) showed that CV is asymptotically equivalent to AIC and MCP in terms of consistency. Shibata (1981) showed that AIC is asymptotically equivalent to MCP. For the variable selection problem in the linear regression model, Thompson (1978) recommended Mallows' MCP criterion for fixed regressors, while for random regressors his preferred criterion is Hocking's HOC.

The penalty function of all the above mentioned criteria in the penalised log-likelihood form is a function of the number of free parameters involved in the model under consideration and often the sample size. It implies that for a fixed sample size and a particular set of models, changes of the data set do not have any impact on the penalty function. This motivates researchers to find data oriented penalty functions, so that a change of data set will also change the penalty value in numerical terms. As far as we know, for model selection in linear regression settings, Rao and Wu (1989) first introduced the idea of a flexible penalty function based on data in linear regression models. They analytically showed that the criterion proposed by them on the basis of a flexible penalty function is strongly consistent without making any distributional assumptions. They argued that the performance of this type of criteria is better than the criteria based on the fixed penalties. Chen *et al.* (1993) used a data oriented penalty function for selecting AR models for time series. Bai *et al.* (1999) extended the work of Rao and Wu (1989) for selecting linear regression models using

General Information Criterion (GIC). They argued that the fixed choice of penalties may not be good in all situations. They showed that the data-oriented penalties guarantee strong consistency and have some advantages over fixed penalties.

It is well known that the performance of the existing criteria varies from situation to situation and none of the criteria performs well in all situations. Some criteria favour the model with the smallest number of parameters, while others favour the model with the largest number of parameters. For example, BIC always favours the model with the smallest number of parameters among competing models, while RBAR favours the model with the largest number of parameters. As a result, the average probabilities of correct selection vary from model to model. Ideally a model selection criterion should select the true model without favouring one model over others. So the idea of making the probabilities of correct selection equal has evolved. This is called controlling the probability of correct selection.

King *et al.* (1995) proposed an algorithm, which makes the probability of correct selection equal. They proposed two approaches. Their first approach is based on the idea of a common model. But the problem with this approach is that there may not always be a common model and there is no fixed rule for finding the probabilities when the competing models are nested. Their second approach is based on representative fixed points and they proposed two techniques for selecting representative fixed points. The first entirely depends on the judgment of the user so for the same data set and models, the conclusions of different researchers may vary.

Forbes *et al.* (1995) proposed three techniques for controlling the probability of correct selection of one model over the other. Their first two methods are for the variable selection purpose and the final one is for general model selection, but the problem with this approach is that the penalties that control the probability of correct selection are approximate penalties. Hossain (1998) proposed an empirical based information criterion called CIC, which is based on King *et al.*'s (1995) algorithm for controlling probabilities of correct selection and the bootstrap sampling method. He applied CIC to selecting the correct model from a set of linear, log-linear and Box-Cox transformation models. From his study, it is observed that the performance of CIC is better than the existing IC for selecting the correct model.

On the topic of finding empirical penalty functions, related work has been done by Kwek (2000) for selecting conditional heteroscedastic (CH) and autoregressive conditional heteroscedastic (ARCH) models, while Billah (2001) used the same technique for selecting appropriate time series models including exponential smoothing models. In her study, Kwek concluded that CH information criterion and the optimal small sample procedures proposed by her outperform all other existing IC for selecting ARCH and GARCH models in small samples. She suggested that BIC should not be used for selecting CH models, because it has been built without considering the one sided information used in the estimation of CH models. From the simulation results, she concluded that among the existing criteria, in small samples the performance of RBAR is the best for selecting ARCH and GARCH models followed by AIC, and BIC is the worst performing criterion. But in large samples, AIC performed relatively better than RBAR. The well known tendency of BIC to under fit and the tendency of AIC to over fit in linear regression models was also observed in the results of her experiments. She mentioned that although BIC has the strongest consistency property, it under fits the model in small samples and as a result, the performance of BIC is the worst in selecting both ARCH and GARCH models. This indicates that the efficiency of an IC is not guaranteed by its consistency property.

In order to see the performance of the widely used existing IC procedures for selecting exponential smoothing models on the basis of mean average probability of correct selection (MAPCS), Billah (2001) conducted a Monte Carlo simulation study. From his study he found that the performance of BIC is the best followed by HQ and MCP is the worst at selecting exponential smoothing models. He proposed two penalty estimation methods (PEM), PEM-GS and PEM-SA on the basis of maximising the MAPCS for the model selection problem for the linear regression model with different ARMA error processes. From his simulation study, he concluded that for small samples the newly proposed methods consistently perform better than the existing IC. He also did simulation experiments for selecting models on the basis of the model's forecasting accuracy, and concluded that the model selection methods proposed by him were better than available IC methods.

2.4 SIMULATED ANNEALING OPTIMISATION TECHNIQUE

In any model building process, the modeller needs to estimate the model parameters. There are some methods of estimation, for example the least squares method, where the parameters can be estimated directly. But, there are many methods, for example non-linear least squares, maximum likelihood and the generalised method of moments, where the estimation process is totally dependent on a numerical optimisation technique, which attempts to iterate to the desired solution. Several iterative search algorithms are available in the literature. For example, the Gauss-Newton method and Newton-Rapson method. All iterative methods use the following four steps to find the best estimates of the parameters (SAS (1992)):

(i) The modeller has to provide initial starting values for the parameter estimates,

(ii) the algorithm selects updated values for the parameter estimates such that the error sum of squares/log-likelihood for the updated values is less/greater than the error sum of squares/log-likelihood for the initial starting values,

(iii) the algorithm continues to select updated values for the parameter estimates that reduce the error sum of squares/increase the log-likelihood, and

(iv) the algorithm stops when a convergence criterion is met.

For all iterative methods, providing good initial values can be important, because the solution found will depend on the initial starting value. Unfortunately, for many problems there is no way to find a starting value that guarantees the global optimum. Bad starting values can increase computer time and could prevent the procedure from finding the correct estimates of the parameters, which is one of the drawbacks of these methods. Another problem is that for some of these methods, e.g. the Newton-Rapson method, both first and second derivatives are necessary and in some situations these derivatives do not exist. Most of the existing iterative algorithms assume that the functional form of the objective function is approximately quadratic and the function has one optimum. Unfortunately, some functions violate these assumptions also. For multiextrema functions, these iterative methods cannot get away from a local extrema and converge only to one of the local extrema, subject to the starting value of the parameter. This is the major drawback of these iterative methods and in this situation, researchers typically try to find the global optimum by using different arbitrary starting values (see Cramer (1986) and Finch *et al.* (1989)). Even if the algorithm of these methods converges, it does not guarantee that the estimated value is the global optimum. Most of the popular packages, for example SAS, RATS and TSP, used for econometric and statistical analysis use these methods. To overcome these problems, one solution would be to introduce a global optimisation method that can avoid local maximum or minimum, and can find the optimum value of the parameters from the entire parameter space. The task of a global optimisation method is to find absolutely the best set of parameter values to optimise an objective function.

There are a number of different ways of finding the global optimum of a function. The earliest methods were associated with the grid search technique and the function is evaluated at equi-spaced points throughout the parameter space. Although this method is typically successful at finding the global optimum of a complicated function, the computational time of this method is too high for a refined search (Billah and King (2000b)). Global optimisation techniques were developed to overcome the problems of grid search and the conventional existing iterative methods. The numerical optimisation techniques have several advantages over the grid search method. There are several numerical global optimisation techniques available in the literature and specific optimisation methods have been developed for many classes of optimisation problems. A comprehensive list includes, i) Mixed Integer Programming, ii) Interval Methods, iii) Clustering Methods, iv) Evolutionary Algorithms, v) Hybrid Methods, vi) Statistical Methods, vii) Tabu Search and viii) Simulated Annealing. Gray *et al.* (1997) did a comprehensive survey of these methods and the description of them is available on the internet site http://www.cs.sadia.gov/opt/survey/main.html. There are several global optimisation programs available to solve different types of optimisation problems. For example, ASA-CalTech Adaptive Simulated Annealing for finding the global optimum of a continuous non-convex function over a multidimensional interval, CURVI-Bound

Constrained Global Optimisation, for solving constrained and unconstrained nonlinear optimisation problems; and SIMANN-Simulated Annealing, which implements the continuous simulated annealing optimisation algorithm described in Corona *et al.* (1987). A review of the available global optimisation programs was done by Pinter (1996) and is available on the internet site http://mat.gsia.cum.edu/pinter.fil.

Selection of a global optimisation method depends on the nature of the problem. The objective function we have to optimise is a step function. From the literature it is apparent that SAO performs well at finding the global maxima in the presence of local maxima and for functions like ours which have plateaux and other ill-behaviour (see Kirkpatrick *et al.* (1983), Romeo *et al.* (1984), White (1984) and Goffe *et al.* (1994)). A comprehensive discussion of the theoretical and practical details of SAO is given in Aarts *et al.* (1997). Simulated annealing is one of the numerical optimisation techniques, which is a probabilistic method for finding the global maximum or minimum of a function that may possess several local maxima or minima (Kirkpatrick *et al.* (1983) and Cerny (1985)). Annealing is a heat-treating process that is applied to glass, metals or materials and involves slowly cooling them to obtain a strong crystalline structure. The basic idea of SAO comes from the theory of thermodynamics and it is a numerical optimisation technique based on a Monte Carlo approach for finding the global optimum of an objective function in the presence of several local optimum. The main advantage of SAO is its ability to move from the region of a local optima to the region of a global optimum. Another advantage of this optimisation technique is, like the conventional algorithms, it assumes very little about the function and can handle the optimisation problem very efficiently and it is explicitly designed for functions with multiple optima (Corana *et al.* (1987); Goffe *et al.* (1994)).

There are several SAO algorithms available in the literature. An early one was introduced for the search of optima of discrete variables called combinatorial SAO (Lawler (1976) and Papadimitriou and Steiglitz (1982)). The combinatorial SAO algorithm has been used successfully in computer and circuit design (Kirkpatrick *et al.* 1983 and Wang *et al.* (1988)), image processing (Carnevali *et al.* (1985)), re-construction of pollycrystalline structures (Telly *et al.* (1987)), neural networks (Wasserman and Schwarz (1988)), and pollution control (Derwent (1988)). Other SAO algorithms are for example, adaptive random search (Pronzato *et al.* (1984)), fast SAO (Szu and Hartly (1987)), down hill simplex with annealing (Vetterling *et al.* (1994)) and direct search SAO (Ali *et al.* (1997)). The implementation of a SAO algorithm involves the application of the Metropolis algorithm (Metropolis *et al.* (1953)), which is the heart of the SAO technique.

The combinatorial SAO algorithm was modified by Vanderbilt and Louie (1984), Bohachevsky *et al.* (1986) and Corana *et al.* (1987) to optimise functions of continuous variables. Among these, the Corana *et al.* (1987) implementation of SAO appears to be the best in terms of combination of ease of use and robustness (Goffe *et al.* (1994)). Goffe *et al.* (1994) tested the effectiveness of the SAO technique against some well known optimisation algorithms, namely, the Nelder and Mead simplex method and a general purpose global optimiser using Adaptive Random Search (ARS). From their test results it was observed that SAO never failed to reach the minimum of the function and is the best of the three methods with respect to reliability followed by simplex method. ARS is the least reliable of the three methods.

Goffe *et al.* (1994) compared the Corana *et al.* (1987) implementation of SAO with three multivariable optimisation algorithms in the IMSL library on four econometric models of four different natures. The first model was an example of multiple minima, which contains only two parameters (Judge *et al.* (1985, pp.956-957)) and the second was a rational expectations exchange rate model with 14 parameters. The third model was an efficiency study of the banking industry using a translog cost frontier system with 62 parameters and the fourth fits a neural network to a chaotic time series with 35 parameters. On the Judge *et al.* function, SAO correctly differentiated between the local and the global minima and finds the global optima, while the conventional algorithms failed. All conventional algorithms failed to find the optimum of the second model and did not offer any reason for their failure, but SAO was able to identify the reason. After correcting the problem, SAO found the optimum easily, but conventional algorithms were successful only 21% of the time. None of the conventional algorithms was able to find the optimum of the third model, while SAO did it easily. In the case of the fourth model, SAO found a much better optimum than did any of the conventional algorithms. They concluded that if SAO was not able to find the global optimum, then the nature of the function makes it impossible for any other method to find the global optimum.

Let $\varphi = (\varphi_1, \varphi_2, \ldots \varphi_n)'$ be an $n \times 1$ parameter vector to be estimated each ranging in a finite continuous interval and $f(\varphi)$ be the bounded function to be maximised. The implementation of the Corana *et al.* (1987) SAO algorithm requires the step length vector for φ say $\mathbf{v}$ and the temperature T. For maximising a function the following steps are required for implementing Corana *et al.* (1987) algorithm.

Step 1 (Initialisation)

Let the initial values for φ, $\mathbf{v}$ and T be φ_0, $\mathbf{v}_0$ and T_0, respectively. Let the value of the objective function at the initial parameter vector φ_0 be f_0. Also set $\varphi_{opt} = \varphi_0$ and $f_{opt} = f_0$, where *opt* stands for optimum.

Step 2 (Selection of new point)

Randomly select another point φ' in the parameter space within a neighbourhood of the original parameter value using the following equation.

$$\varphi_i' = \varphi_i + rv_i \,. \tag{2.27}$$

where r is a random number generated in the range [-1,1] by a pseudorandom number generator and v_i is the i^{th} component of the step vector $\mathbf{v}$. If φ' lies outside the definition domain of $f(\varphi)$ then a new point is randomly generated until a point is found inside the definition domain.

Step 3 (Accept or reject the point and use of Metropolis criterion)

Let the value of the objective function at this new point be $f' = f(\varphi')$. If $f' > f_{opt}$ (uphill move) then accept the new point and set $\varphi_{opt} = \varphi'$ and $f_{opt} = f'$. If $f' \leq f_i$ then accept or reject the point with acceptance probability p, computed using the Metropolis criterion as follows:

$$p = \exp\left(\frac{f' - f_{opt}}{T}\right). \tag{2.28}$$

This computed value of p is then compared with a pseudorandom number p', which is generated from the uniform distribution in the range [0,1]. If $p' < p$, the point is accepted (downhill move) otherwise it is rejected. In the case of acceptance the values of φ_{opt} and f_{opt} are updated by φ' and f', respectively, and in case of rejection, there is no change in φ_{opt} and f_{opt}. Lower temperatures and larger differences in the function values are the two factors that decrease the probability of a down hill move.

Step 4 (Adjustment of step length vector $\mathbf{v}$ *)*

As mentioned by Corana *et al.* (1987), both a higher number of acceptances or higher number of rejections wastes computational effort. On the contrary, 50% of moves accepted and 50% rejected indicates that the algorithm is running well. To make this the case, after N_s steps through all elements of φ, the step length vector $\mathbf{v}$ is adjusted so that 50% of all moves are accepted. The objective of doing so is to sample the function widely. If more than 60% of the points are accepted for φ_i, then the relevant elements of $\mathbf{v}$ are enlarged by the factor $1+2.5c_i(a_i / N_s - 0.6)$, where a_i is the number of points accepted and c_i is the i^{th} element of the vector that controls step variation. The element is declined by, $1+2.5c_i(0.4 - a_i / N_s)$, if less than 40% of the points are accepted. Once an equilibrium state has been achieved for a given temperature, the temperature is reduced to a new temperature as defined in step 5 and the process started again taking values of the last iteration of the algorithm as the initial values.

Step 5 (Temperature reduction)

After N_T times (N_T, is the number set by the user for temperature reduction test) through the steps 1 to 4, the temperature, T, is reduced. The new temperature is given by

$$T' = r_T T, \tag{2.29}$$

where r_T is the temperature reduction coefficient which lies between 0 and 1.

Step 6 (Terminating criterion)

The last N_ε (N_ε is a number set by the user for successive temperature reductions to test for termination) values of the largest function values from the end of each temperature reduction are recorded and compared with the most recent value. The algorithm will terminate if all of these N_ε differences from the most recent value are less than the terminating criterion ε, a very small number.

Goffe *et al.* (1994) introduced four extensions of the Corana *et al.* algorithm. The first modification allows the researcher to test if SAO has indeed found the global optima; the second modification allows the researcher to restrict the search area to a subset of the parameter space. The third extension permits the researcher to determine a critical initial parameter for the algorithm and the final one directs the selection of the initial temperature, an essential parameter that controls the robustness of the algorithm. This allows the researcher to minimise the execution time of the algorithm.

The simulated annealing algorithm has several potential advantages over conventional optimisation algorithms. First, it can distinguish between different local maxima and can escape from local maxima by moving both uphill and downhill. The algorithm makes very few assumptions regarding the function to be maximised. It is robust with respect to non-quadratic surfaces so the function need not be approximately quadratic; and even need not be differentiable (see Corana *et al.*

(1987)). Second, a very large element in the step length vector of SAO indicates that the function is very flat in that parameter, which is valuable information for the researcher about the function. Third, SAO can identify corner solutions because it can snuggle up to a corner for functions that do not exist at that corner. The final and the most important advantage of SAO is that it can optimise functions that are very complex or impossible to optimise. Simulated annealing requires high computational power, which is the only drawback of this method. However, recent developments with respect to computer power largely eliminates this problem. As compared to other global optimisation methods, SAO has many advantages. These include the relative ease of implementation, applicability to almost any problem and the ability to provide reasonably good solutions for most problems. Depending on the problem to which it is applied, SAO appears to be competitive with many of the best heuristic methods (Johnson *et al.* (1997)). In this thesis we use SAO to estimate penalties that maximise the average probability of correct selection of linear regression models in small samples.

2.5. CONCLUDING REMARKS

The main purpose of this chapter was to review different model selection procedures. This chapter also contains a review of relevant literature on the global optimisation technique SAO and its applications in econometrics. Several researchers have argued that a sequence of pairwise tests has several drawbacks, so may not be appropriate for the purpose of model selection. We reviewed the literature of two alternative methods, model selection based on residual sum of squares and model selection based on IC.

There are several model selection procedures available in the literature which are based on residual sum of squares. We discussed the following: coefficient of multiple determination and adjusted coefficient of determination by Theil (1961), Mallows' (1964) C_p criterion, Amemiya's (1980) PC criterion, Rothman's (1968) J_p and Hocking's (1976) S_p, Zhang's (1992) final prediction error (FPE) criterion and Rahman's (1998) generalised model selection criteria are also based on residual sum of squares. The advantages and disadvantages of residual sum of squares based model selection was also discussed. In Chapter 4 a new information criterion based on residual mean square will be proposed for model selection.

Nowadays IC based model selection procedures are widely accepted for model selection purposes in different areas of research. There are several IC procedures available in the literature. We selected some of the widely used IC procedures, for example, AIC, proposed by Akaike (1973), BIC proposed by Schwarz (1978), Schmidt's (1975) Generalized Cross Validation criterion (GCV), Hannan and Quinn's (1979) criterion (HQ), Hocking's (1976) criterion (HOC), and JIC, proposed by Rahman and King (1999). We discussed some of the main features of these procedures mainly in relation to selection of linear regression models.

The final section of this chapter contains some background on global optimisation techniques with particular emphasis on the SAO technique. This section also contains a review of the literature on SAO and its application in econometrics. From the literature, it is apparent that for finding the global optimum of any function which has several local optima, SAO is nearly always successful in contrast to the outcomes from standard iterative methods.

CHAPTER 3

IMPROVED PENALTY FUNCTIONS FOR INFORMATION CRITERIA BASED MODEL SELECTION[1]

3.1 INTRODUCTION

One of the principal decision making problems faced by applied statisticians and econometricians is that of choosing an appropriate model from a number of competing models for a particular data set. This problem can be solved in several ways. One, and probably the most popular way, is to use an information criterion (IC) to make the choice. In general, an IC model selection procedure is based on choosing the model with the largest maximised log-likelihood function minus a penalty function which depends on the number of parameters and in most cases the sample size. At present, a number of model selection criteria that fall into this category are available. They include AIC, BIC, GCV, HQ, RBAR and HOC. These criteria were discussed in Chapter 2. Many of these procedures were originally developed with particular types of models in mind, and not necessarily as IC procedures. Fox (1995)

[1] A paper based on some of the findings reported in this chapter and Chapter 4 has been accepted for publication in *Computer Aided Econometrics*, edited by D.E.A Giles, see King and Bose (2002).

expressed these procedures in this framework by finding their implied penalty functions in the context of choosing between different linear regression models.

A small number of researchers have conducted Monte Carlo studies comparing various subsets of the above criteria in a variety of different settings; see for example Bora-Senta and Kounias (1986), Hurvich and Tsai (1989, 1991), Mills and Prasad (1992) and Hughes (1997). The general conclusion one draws from these and other studies is that no one procedure dominates: for models with fewer parameters, BIC does well but at the expense of selecting larger models when these are indeed the true model. AIC and RBAR, on the other hand, favour larger models with lower relative probabilities of selecting smaller models when these models are true. HQ is somewhere between these two extremes, not being as harsh as BIC nor as generous as AIC and RBAR on larger models when they are true.

As noted by Pötscher (1991), maximizing an IC is equivalent to testing each model against all other models by means of a standard likelihood ratio test and selecting that model which is accepted against all others. The choice of penalty function determines the values of the critical values of the tests. In hypothesis testing, when critical values are changed, this either increases the probability of a Type I error while decreasing the probability of a Type II error, or vice versa. Clearly, in the context of model selection, a change in penalty function induces similar changes in the probabilities of various types of errors. Therefore it is not surprising that no one IC procedure dominates all other procedures.

We are then left with the question of which IC procedure to use in practice. Given that many of the above procedures have an asymptotic justification, it is far from clear which procedure should be favoured in small samples. In this chapter, we outline an alternative approach for calculating a penalty function based on (i) treating all models equally; (ii) calculating the average probability of correct selection for a given model using a Bayesian prior distribution to weigh different parameter values; and (iii) optimizing the mean of these average probabilities of correct selection. In related work, similar ideas have been proposed for IC model selection procedures in the context of selecting ARCH and ARCH type models by Kwek and King (1998), selecting a structural break in a linear regression by Azam and King (1998) and selecting an ARMA time series model by Billah and King (1998). A major problem in this work has been the high computational cost involved in finding the penalties. The approach outlined in this chapter is much more manageable in this regard.

The plan of this chapter is as follows. A new model selection technique is outlined in Section 2. In Section 3, we outline and discuss two sets of Monte Carlo experiments. The purpose of the first set of experiments is to investigate what is the most appropriate combination of the number of parameter drawings (q) and the number of replications (N) for a fixed total qN when estimating the average probability of correct selection of the true model using simulation methods. The second set of experiments is conducted to compare the performance of our new approach with selected existing criteria in the context of selecting the true model in the classical linear regression setting. In Section 3.3.1, we discuss the data generating process used to conduct Monte Carlo simulation experiments. Section 3.3.2 contains an outline of the experiments and discussions of results for the first set of experiments. The outline of the experiments, results and discussions of the second set of experiments are presented in Section 3.3.3. Section 3.4, the final section, contains some concluding remarks.

3.2 PROPOSED TECHNIQUE

We are interested in selecting a model from m alternative models, M_1, M_2, . . ., M_m, for a given data set. Let the model M_j, $j = 1, 2, \ldots, m$, be represented by

$$\mathbf{y} = f(\mathbf{X}_j, \theta_j, \mathbf{u}_j), \tag{3.1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations on the dependent variable, θ_j is a vector of k_j free parameters, $\mathbf{X}_j$ is an $n \times k_j^*$ matrix, $k_j^* = (k_j - 1)$, and $\mathbf{u}_j$ is an $n \times 1$ vector of random disturbances distributed as $\mathrm{N}(0, \sigma_j^2 \mathbf{I})$. $\mathbf{X}_j$ contains a column vector of ones in its first column and $(k_j^* - 1)$ vector of observations on non-stochastic variables (e.g., in linear regression model there are $(k_j^* - 1)$ non-constant regressors) in the remaining $(k_j^* - 1)$ columns. Let the log-likelihood function for the model M_j be

$L_j(\theta_j)$ and the maximised value of $L_j(\theta_j)$ be $L_j(\hat{\theta}_j)$, where $\hat{\theta}_j$ is the maximum likelihood estimator of θ_j. Let p_j denote the penalty for model M_j. In almost all IC based model selection procedures, the model with the largest I_j is selected, where I_j is given by

$$I_j = L_j(\hat{\theta}_j) - p_j \tag{3.2}$$

This I_j is called the penalised log-likelihood. Following Fox (1995), the penalised log-likelihood forms of AIC, BIC, HQ, RBAR, GCV and HOC are given by,

$$\mathrm{AIC}_j = L_j(\hat{\theta}_j) - k_j, \tag{3.3}$$

$$\mathrm{BIC}_j = L_j(\hat{\theta}_j) - \frac{k_j}{2}\ln(n), \tag{3.4}$$

$$\mathrm{HQ}_j = L_j(\hat{\theta}_j) - k_j\ln(\ln(n)), \tag{3.5}$$

$$\mathrm{RBAR}_j = L_j(\hat{\theta}_j) + \frac{n}{2}\ln(n - k_j), \tag{3.6}$$

$$\mathrm{GCV}_j = L_j(\hat{\theta}_j) + n\ln(\frac{n - k_j}{n}), \tag{3.7}$$

and

$$\mathrm{HOC}_j = L_j(\hat{\theta}_j) + \frac{n}{2}\ln\left\{(n - k_j)(n - k_j - 1)\right\}. \tag{3.8}$$

The penalty functions of AIC, BIC and HQ for the j^{th} model can be written in the general form

$$p_j = \lambda_1 k_j, \tag{3.9}$$

where λ_1 is a known function of n and

$\lambda_1 = 1$ for AIC,

$= \frac{1}{2}\ln n$ for BIC, and

$= \ln(\ln n)$ for HQ.

Also the penalty function of GCV for the j^{th} model can be written as $n\ln n - n\ln(n-k_j)$. Since $n\ln n$ is constant for a particular selection problem, a criterion with penalty function $n\ln n - n\ln(n-k_j)$ is equivalent to one with penalty function $-n\ln(n-k_j)$. The penalty function of HOC can be written as $-\frac{n}{2}\ln(n-k_j)-\frac{n}{2}\ln(n-k_j-1)$ which approaches $-n\ln(n-k_j)$ as $n\to\infty$. Hence the penalty functions of RBAR, GCV and HOC (asymptotically) for the j^{th} model can be written in the common form

$$p_j = \lambda_2 \ln(n-k_j), \tag{3.10}$$

where λ_2 is a known constant depending on the criterion and

$\lambda_2 = -\frac{n}{2}$ for RBAR,

$= -n$ for GCV, and

$= -n$ for HOC (asymptotically).

Thus the penalty function of any of these six listed criteria for the j^{th} model is either of the form $\lambda_1 k_j$ or $\lambda_2 \ln(n-k_j)$. These can be generalised to $\lambda_1 k_j + \lambda_2 \ln(n-k_j)$. Here λ_1 and λ_2 may be functions of sample size n.

We know that the performances of different model selection criteria vary from situation to situation. For example, in the context of selecting the true model from a set of linear regression models, Rahman and King (1997) observed that for a sample size of at least 13, and when the model with the lowest number of regressors is true, then the performance of BIC is better than that of GCV, which is better than that of AIC, and which in turn is better than that of the RBAR criterion in terms of probability of correctly selecting the model. On the other hand, when the sample size is at least 13, and the model with the highest number of regressors is true, then the performances of BIC, GCV, AIC and the RBAR criteria are exactly reversed. The performances of different model selection criteria also vary from data set to data set. Therefore, the question arises as to which IC procedure should be used for a particular data set and group of competing models. Is there one we can have confidence in for all situations?

Given that we have seen that the penalty functions of six of the main IC procedures can be generalised to

$$p_j = \lambda_1 k_j + \lambda_2 \ln(n-k_j), \tag{3.11}$$

then the choice of procedure (at least out of these six procedures) just involves a choice of λ_1 and λ_2 values for this penalty function. Clearly we would like to make the best choice of λ_1 and λ_2 values for a particular sample size and set of models to be chosen from. Also we should not restrict λ_1 and λ_2 to the six sets of values implied by the above six procedures, but allow λ_1 and λ_2 to be chosen to suit our particular circumstances.

The penalty function (3.11) is obtained by adding the penalty functions (3.9) and (3.10). We can define another penalty function by multiplying the penalty functions (3.9) and (3.10) as,

$$p_j = \lambda_1 k_j \ln(n - k_j). \tag{3.12}$$

In equation (3.9), λ_1 is a function of n the sample size and k_j is the number of free parameters in the j^{th} model. Instead of k_j, if we consider a fractional power of k_j, then we can define a new penalty function as,

$$p_j = \lambda_1 k_j^{\lambda_2}. \tag{3.13}$$

The question then is, how do we find optimal values of λ_1 and λ_2 for our proposed penalty functions? Our suggestion is as follows. For each model under consideration and for a given choice of penalty values, we estimate the average probability of correctly selecting this model when it is indeed the true model. For the same penalty set, we then average these probabilities of correct selection (thus treating all models equally). λ_1 and λ_2 are then chosen to maximise this average.

In order to understand more closely what is involved, let CS denote the event of correct selection and $P(CS|M_j, \theta_j, \lambda_1, \lambda_2)$ denote the probability of correct selection when the model M_j is true with parameter vector θ_j, and λ_1 and λ_2 are used in (3.11).

Therefore,

$$P(CS|M_j, \theta_j, \lambda_1, \lambda_2) = P\left[(I_j - I_i) > 0 | M_j, \theta_j, \lambda_1, \lambda_2; i = 1, 2, \ldots, (j-1), (j+1), \ldots, m\right]$$

$$= P\left[\left(L_j(\hat{\theta}_j) - L_i(\hat{\theta}_i)\right) > (p_j - p_i) | M_j, \theta_j, \lambda_1, \lambda_2; i = 1, 2, \ldots, (j-1), (j+1), \ldots, m\right]$$

$$= \psi\left(\theta_j, \lambda_1, \lambda_2\right) \text{ (say)}.$$

then the average probability of correct selection when the model M_j is true will be

$$\underset{\theta_j}{E}\left[P\left(CS|M_j, \theta_j, \lambda_1, \lambda_2\right)\right] = \int \psi\left(\theta_j, \lambda_1, \lambda_2\right) g\left(\theta_j\right) d\theta_j, \qquad (3.14)$$

where $g\left(\theta_j\right)$ is the prior density function of the vector of parameters θ_j. Thus the mean of average probability of correct selection will be

$$\frac{\sum_{j=1}^{m} \underset{\theta_j}{E}\left[P\left(CS|M_j, \theta_j, \lambda_1, \lambda_2\right)\right]}{m}$$

which involves the unknown constants λ_1 and λ_2.

For our approach to be operational, we need a method of estimating (3.14). Given that $g(\theta_j)$ is a joint density function, the Monte Carlo estimate of (3.14) can be found by first taking a large sample of drawings from $g(\theta_j)$ which we will denote by $\theta_j(i)$, $i = 1, 2, \ldots, q$, and then calculating $\frac{1}{q}\sum_{i=1}^{q} \mathrm{P}(\mathrm{CS}|M_j, \theta_j(i), \lambda_1, \lambda_2)$.

This then requires us to estimate $\mathrm{P}(\mathrm{CS}|M_j, \theta_j(i), \lambda_1, \lambda_2)$ for given M_j, $\theta_j(i)$, λ_1, λ_2, which can be achieved by a straightforward Monte Carlo simulation of N replications. After some experimentation with a range of settings, we find that for a fixed total number of simulations qN, good results are achieved by using only one replication ($N = 1$) in the estimation of $\mathrm{P}(\mathrm{CS}|M_j, \theta_j(i), \lambda_1, \lambda_2)$, and the maximum number of drawings of θ_j from $g(\theta_j)$. Results of these simulation experiments are reported in the Section 3.3.2.

Unfortunately, the problem of maximizing our estimate of (3.14) with respect to λ_1 and λ_2 requires considerable computational effort. We suggest that for each of the m models, q random drawings of θ_j are obtained from $g(\theta_j)$ and then model (3.1) is used to generate q $\mathbf{y}$ vectors. For each $\mathbf{y}$ vector the likelihood functions of each of the models are then maximised, and the maximised values are stored. This is repeated for each model so that in total a file of m^2q maximised likelihoods is generated. This can then be used to estimate (3.14) for different values of λ_1 and λ_2. In our stud[illegible]d q = 2000 and N=1.

3.3 THE MONTE CARLO STUDIES

In this section we describe two sets of simulation experiments. The purpose of the first set of experiments is to find the best combination of N, the number of replications, and q, the number of parameter drawings, for a simulation experiment. The second set of experiments was conducted to evaluate the performance of the newly proposed criteria against the existing IC listed above. The plan of this section is as follows. In Section 3.3.1, we describe the data generating process used for the simulation experiment. Section 3.3.2 is devoted to the first set of experiments while in Section 3.3.3, we describe the second set of experiments.

3.3.1 DATA GENERATING PROCESS

Suppose we have n observations on each variable, then the j^{th} model can be written in matrix notation as

$$\mathbf{y} = \mathbf{X}_j\beta_j + \mathbf{u}_j, \tag{3.15}$$

where $\mathbf{y}$ is $n\times 1$, X_j is $n\times k_j^*$, β_j is $k_j^*\times 1$ vector of coefficients, and $\mathbf{u}_j$ is an $n\times 1$ disturbance vector distributed as $\mathrm{N}(0,\sigma_j^2\mathbf{I})$.

We used different designs for the simulation experiments. In each case, the random vector $\mathbf{y}$ is generated for a set of randomly selected values of β_j and σ_j^2 obtained from the prior distribution suggested by Zellner (1971), namely the inverted gamma distribution for σ_j^2 and multivariate normal distribution for β_j. That is, the values of β_j and σ_j^2 are randomly and independently chosen from independent multivariate normal and inverted gamma distributions, respectively. Therefore, to generate random vectors $\mathbf{y}$ and for estimating (3.14), the following steps were followed.

Step 1: We randomly selected a value of σ_j^2 for the j^{th} model from an inverted gamma distribution, as follows.

A random sample of size $(n-k_j^*)$ is drawn from the N(0,1) distribution, where k_j^* is the number of regressors including the constant in the j^{th} model. Let the sample values be $z_1, z_2, \ldots, z_{n-k_j^*}$. Then we compute $\chi^2_{n-k_j^*} = \sum_{i=1}^{n-k_j^*} z_i^2$, which is distributed as a chi-squared variable with $(n-k_j^*)$ degrees of freedom. Finally a value of σ_j^2 is obtained by using the formula $\sigma_j^2 = (n-k_j^*)s_r^2 \dfrac{1}{\chi^2_{n-k_j^*}}$, where s_r^2 is an arbitrary positive value held constant for the experiment.

Step 2: We randomly selected a vector of parameters $\beta_j = (\beta_0, \beta_1, \ldots, \beta_{k-1})'$ for the j^{th} model in the following way.

We drew a random sample of size k_j^* from the N(0,1) distribution. Let the sample values be $w_1, w_2, \ldots, w_{k_j^*}$. Then

$$\beta_j = \mathbf{W}_j \tilde{\sigma}_j,$$

where $\mathbf{W}_j = (w_1, w_2, \ldots, w_{k_j^*})'$ and $\tilde{\sigma}_j$ is a randomly selected value of the error standard deviation obtained by applying step 1 again for another arbitrary positive value say s_{β}^2 and $\tilde{\sigma}_j$ is independent of σ_j^2.

The arbitrary values of s_ε^2 in step 1 and s_β^2 are used to generate data. These values are chosen by a trial and error method, so that the expected probability of correct selection is a middle level of probability (say 0.5) for small sample for a particular simulation experiment. (Low average probabilities and high average probabilities are of less interest.)

Step 3: To generate the dependent random vector $\mathbf{y}$ for the j^{th} model, we drew a random sample of size n from the N(0,1) distribution and let the sample values be $v_{j1}, v_{j2}, \ldots, v_{jn}$. Then we obtained $\mathbf{y}$ by using the following formula

$$\mathbf{y} = \mathbf{X}_j\beta_j + \mathbf{v}_j\sigma_j,$$

where $\mathbf{v}_j = (v_{j1}, v_{j2}, \ldots, v_{jn})'$.

Step 4: Using the generated random vector $\mathbf{y}$, we fit all the plausible models in the set and found the maximised log-likelihood value say $L_j(\hat{\beta}_j, \hat{\sigma}_j)$ for the j^{th} model, where $\hat{\beta}_j = (\mathbf{X}_j'\mathbf{X}_j)^{-1}\mathbf{X}_j'\mathbf{y}$, $\hat{\sigma}_j^2 = \dfrac{\mathbf{u}_j'\mathbf{u}_j}{n}$ and $\mathbf{u}_j = \mathbf{y} - \mathbf{X}_j\hat{\beta}_j$. If p_j is the penalty function for the j^{th} model, then the information criterion for the j^{th} model will be $I_j = L_j(\hat{\beta}_j, \hat{\sigma}_j) - p_j$. We computed I_j for all j in the set and ranked them. Evidently if $I_j > I_i$ for all i = 1, 2, ..., $(j-1)$, $(j+1)$, ..., m, then we have correct selection of the j^{th} model.

Step 5: We repeated steps 1 to 4 q times, replicated N times for a fixed number qN (we used qN = 2000 for our simulation experiments) and calculated the Monte Carlo probabilities of correct selection by using the relative frequency definition of probability.

Steps 1 to 5 are used for finding the probability of correct selection under any information criterion just by replacing p_j by the penalty function of that criterion in step 4. In the case of a penalty p_j involving unknown constants λ_1 and λ_2 we find the Monte Carlo probabilities of correct selection for all feasible values of λ_1 and

λ_2. Then we identify those values of λ_1 and λ_2 for which the mean of the expected probability of correct selection is a maximum. This was done for the second set of experiments for our proposed method.

3.3.2 CHOICE OF N AND q

In this section, we investigate different choices of N, the number of replications, and q, the number of parameter drawings, via simulation experiments, which involve selecting a model from a set of competing alternative models in linear regression settings. The aim is to find the best choice of N and q values for fixed qN. In Subsection 3.3.2.1, we outline the simulation experiments conducted for this purpose. The results of these simulation experiments are presented in Subsection 3.3.2.2.

3.3.2.1 OUTLINE OF THE SIMULATION EXPERIMENTS

We conducted the simulation experiments with two sets of data and one set of models as follows in order to find the optimal values of q and N.

Data set 1: x_{1t} is Australian retail trade quarterly data from 1959(1) to 1982(4) and x_{2t} is the same series lagged one quarter. We have used data from Australian Bureau of Statistics. Here we set $s_e^2 = 55$ and $s_b^2 = 6$.

s_c^2 and s_b^2 are arbitrary positive values held constant for the experiment. The reasons for choosing these values were discussed in Section 3.3.1.

Data set 2: x_{1t} is the real per capita GDP and x_{2t} is the investment of a country. We used the annual data from Summers and Heston (1991) revised version 5.6 and World Bank world tables. Here we set $s_c^2 = 2.5$ and $s_b^2 = 0.002$.

Model Set: We used the following four linear regression models for our evaluation.

$$M_1: \quad y_t = \beta_{10} + u_{1t}, \qquad u_{1t} \sim IN(0, \sigma_1^2); \tag{3.16}$$

$$M_2: \quad y_t = \beta_{20} + x_{1t}\beta_{21} + u_{2t}, \qquad u_{2t} \sim IN(0, \sigma_2^2); \tag{3.17}$$

$$M_3: \quad y_t = \beta_{30} + x_{2t}\beta_{31} + u_{3t}, \qquad u_{3t} \sim IN(0, \sigma_3^2); \tag{3.18}$$

$$M_5: \quad y_t = \beta_{50} + x_{1t}\beta_{51} + x_{2t}\beta_{52} + u_{5t}, \qquad u_{5t} \sim IN(0, \sigma_5^2); \tag{3.19}$$

where y_t is the t^{th} observation of the dependent variable, x_{it} is the t^{th} observation of the i^{th} regressor, β_{j0} is a constant for the j^{th} model, β_{ji} (j = 1, 2, 3 & 5; and i = 1 & 2) is a scalar regression coefficient associated with the j^{th} model and regressor x_{it}; and u_{jt} is a random disturbance term that is independently normally distributed with zero mean and variance σ_j^2.

Consequently we have two designs for our simulation experiments. Design 3.1 is the combination of data set 1 and our model set, and Design 3.2 is the combination of data set 2 and our model set.

We generated the data using the technique described in Section 3.3.1. The sample sizes used for the simulation study were $n = 20$, 50 and 96 for Design 3.1, and $n = 20$, 50 and 100 for Design 3.2. For the simulation experiments, we used the following fifteen combinations of N and q so that qN is equal to 2000.

N	2000	1000	500	250	200	100	50	40	20	10	8	5	4	2	1
q	1	2	4	8	10	20	40	50	100	200	250	400	500	1000	2000

Because our aim is to find the combination of q and N values that gives the greatest accuracy in estimating mean probability of correct selection, we repeated the whole experiment 20 times in order to calculate the standard deviation of the estimated mean probabilities of correct selection. Then we computed the average of this mean probability of correct selection over the competing models (here we have four competing models). We also computed the average of the standard deviations of mean probability of correct selection over the competing models to see the trend of this standard deviation as q changes. We used average standard deviation averaged over the range of competing models (ASD) as a measure of the efficiency and the coefficient of variation (CV) as a measure of reliability of the estimated mean probability of correct selection of the true model. That combination of q and N, where both ASD and CV are the lowest, will be judged to be the best combination to estimate the mean probability of correct selection. We also found the mathematical relationship between ASD and q, and estimated the value of q for the simulation experiment where ASD is the minimum.

3.3.2.2 RESULTS AND DISCUSSIONS

The results of these simulation experiments are presented in Tables 3.1 and 3.2. From these tables, it is evident that with an increase in q, the value of ASD decreases. For both the designs and for all sample sizes, the highest estimated mean probability of correct selection averaged over four models (AAPCS) under each criteria is obtained when the parameter is generated once and replicated 2000 times. But for this combination of q (= 1) and N (= 2000), the ASD is also the highest compared to all other combinations of q and N, for all cited criteria and sample sizes considered for the experiment. The CVs obtained for this combination of q and N are also the highest (minimum 10% and maximum 33%) compared to all other combinations of q and N, for all cited criteria, sample sizes and for both the designs. This indicates that the estimated mean probability of correct selection obtained from a single drawing and replicated 2000 times is not efficient and reliable. In almost all cases, the lowest ASD is obtained when the parameter is generated 2000 times and replicated once with the estimated AAPCS being very close to those obtained from the combination $q = 1$ and $N = 2000$. For this combination of q and N, the CVs are also the lowest (minimum 1% and maximum 2.5%) for all cited criteria, sample sizes and for both the designs. This means that the estimated average probability of correct selection obtained from the maximum number of drawings of parameters and replicated once, is more efficient, less variable and more reliable than that of a single drawing of the parameters replicated the maximum number of times.

To find the mathematical relationship between ASD and q, we first plotted the data and got the impression that the relationship between q and ASD may be represented by the following mathematical model:

$$\text{ASD} = A^{*} q^{B} v_t^{*}, \qquad (3.20)$$

where A^{*} and B are parameters and v_t^{*} is a random disturbance term. Model (3.20) may be written in the following log linear form:

$$\ln(\text{ASD}) = A + B\ln(q) + v_t, \qquad (3.21)$$

where $A = \ln A^{*}$, $v_t = \ln v_t^{*}$, and v_t is $\text{IN}(0, \sigma_0^2)$.

We estimated this model for both the designs and for all sample sizes under all cited criteria. We let $\hat{A}$ and $\hat{B}$ be the estimated values of A and B, respectively. It is observed that $\hat{A}$ and $\hat{B}$ are highly significant and the adjusted coefficients of determination ($\bar{R}^2$) are also high and highly significant (Table 3.3) for both the designs and all sample sizes under all cited criteria. But for the sample sizes 20 and 50 under all cited criteria and for both the designs, there is strong evidence of significant positive autocorrelation in the residuals, as the calculated values of the Durbin-Watson (DW) test statistic are always less than their respective tabulated lower bounds of the critical values, suggesting the possibility of functional misspecification. So to establish a more accurate relationship between ASD and q for all the sample sizes and designs, we tested several alternative models. From these test results, we came to the conclusion that the following model explains the relationship satisfactorily for all sample sizes, under all cited criteria and for both the designs:

$$\ln(\text{ASD}) = a + b\ln(q) + c(\ln(q))^2 + u_t, \tag{3.22}$$

where u_t is $\text{IN}(0, \sigma^2)$.

We estimated the values of a, b and c for all cited criteria and sample sizes for both designs. We let $\hat{a}$, $\hat{b}$ and $\hat{c}$ be the estimated values of a, b and c, respectively. By differentiating the right hand side of (3.22) with respect to $\ln(q)$, and equating it to zero, and using the estimated values of b and c, we can find $\hat{q}$, the estimated value of q, for which the value of ln(ASD) and hence ASD is the minimum. This value turns out to be

$$\hat{q} = e^{-\frac{\hat{b}}{2\hat{c}}}. \tag{3.23}$$

The estimated values of a, b and c with their respective significant levels and $\hat{q}$ are given in Tables 3.4 and 3.5 for Designs 3.1 and 3.2, respectively.

From Tables 3.4 and 3.5, it is observed that the adjusted coefficient of determination ($\bar{R}^2$) is always highly significant ($\bar{R}^2 > 0.97$ for Design 3.1 and $\bar{R}^2 > 0.95$ for Design 3.2), and the estimated coefficients are highly significant, but none of the DW statistics are significant for all combinations of data sets and sample sizes. This indicates that the relationship between ASD and q is satisfactorily represented by the mathematical model (3.22), and the estimated coefficients are also efficient. For all sample sizes and for both the data sets except for $n = 20$ of Design 3.2, $\hat{q}$, the estimated value of q, produces the minimum ASD and $\hat{q}$ is always higher than that of the value of qN (2000). This provides further proof that generating the parameter vector the maximum number of times (here 2000 times) and replicating only once ($N = 1$), is the best way to obtain an efficient estimate of mean probability of correct selection for a fixed number qN.

3.3.3 PERFORMANCE OF THE PROPOSED CRITERIA

This section describes the models and designs used to examine the performance of the proposed method over the listed existing IC. We defined five penalty functions in (3.9), (3.10), (3.12), (3.13) and (3.11), and their corresponding IC, named NIC1_j, NIC2_j, NIC3_j, NIC4_j and NIC5_j, for the j^{th} model are,

$$\text{NIC1}_j = L(\hat{\theta}_j) - \lambda_1 k_j, \tag{3.24}$$

$$\text{NIC2}_j = L(\hat{\theta}_j) - \lambda_1 \ln(n - k_j), \tag{3.25}$$

$$\text{NIC3}_j = L(\hat{\theta}_j) - \lambda_1 k_j \ln(n - k_j), \tag{3.26}$$

$$\text{NIC4}_j = L(\hat{\theta}_j) - \lambda_1 k_j^{\lambda_2}, \tag{3.27}$$

$$\text{NIC5}_j = L(\hat{\theta}_j) - \left[\lambda_1 k_j + \lambda_2 \ln(n - k_j)\right]. \tag{3.28}$$

We compare the performance of these newly proposed IC with the listed existing IC in a classical linear regression setting.

3.3.3.1 MODELS FOR THE MONTE CARLO STUDIES

The following four linear regression models, along with the four models described in Subsection 3.3.2.1, were used in the study to examine the performance of the proposed IC compared to the listed existing IC.

$$M_4: \quad y_t = \beta_{40} + x_{3t}\beta_{41} + u_{4t}, \qquad u_{4t} \sim \text{IN}(0, \sigma_4^2); \tag{3.29}$$

$$M_6: \quad y_t = \beta_{60} + x_{1t}\beta_{61} + x_{3t}\beta_{62} + u_{6t}, \qquad u_{6t} \sim \text{IN}(0, \sigma_6^2); \tag{3.30}$$

$$M_7: \quad y_t = \beta_{70} + x_{2t}\beta_{71} + x_{3t}\beta_{72} + u_{7t}, \qquad u_{7t} \sim \text{IN}(0, \sigma_7^2); \tag{3.31}$$

$$M_8: \quad y_t = \beta_{80} + x_{1t}\beta_{81} + x_{2t}\beta_{82} + x_{3t}\beta_{83} + u_{8t}, \qquad u_{8t} \sim \text{IN}(0, \sigma_8^2); \tag{3.32}$$

where y_t is the t^{th} observation on the dependent variable, x_{it} is the t^{th} observation on the i^{th} regressor, β_{j0} is a constant for j^{th} model, β_{ji} (j = 4, 6, 7 & 8; and i = 1, 2, & 3) is a scalar regression coefficient associated with j^{th} model and regressor x_{it}; and u_{jt} is a random disturbance term following the normal distribution with mean zero and variance σ_j^2.

3.3.3.2 THE DESIGNS FOR THE SIMULATION EXPERIMENTS

We have three sets of data and two sets of competing models, one with four competing models and the other with eight competing models. Altogether we have six designs, Design 3.2 to Design 3.7, for the simulation experiments. A description of Design 3.2 was given in the Subsection 3.3.2.1. A brief description of the remaining five designs is given below.

Design 3.3: x_{1t} and x_{2t} are randomly and independently generated values from the N(0,1) distribution. Here we consider four non-nested models M_1, M_2, M_3 and M_5 given by (3.16), (3.17), (3.18) and (3.19) with $s_\epsilon^2 = s_b^2 = 1$ for samples of sizes 20, 50 and 100.

Design 3.4: x_{1t}, x_{2t}, and x_{3t} are randomly and independently generated values from the N(0,1) distribution. Here we consider eight non-nested models M_1, M_2, M_3, M_4, M_5, M_6, M_7 and M_8 given by (3.16), (3.17), (3.18), (3.29), (3.19), (3.30), (3.31) and (3.32), respectively, with $s_\epsilon^2 = s_b^2 = 0.1$ for samples of sizes 20, 50 and 100.

Design 3.5: This is an extension of Design 3.2 to more models using one extra variable x_{3t} as the price level consumption of the t^{th} country. Here we consider eight non-nested models M_1, M_2, M_3, M_4, M_5, M_6, M_7 and M_8 given by (3.16), (3.17), (3.18), (3.29), (3.19), (3.30), (3.31) and (3.32), respectively, with $s_e^2 = 2.5$ and $s_b^2 = 0.02$ for samples of sizes 20, 50 and 100.

Design 3.6: Australian cross-section data for 1961 and 1976 was classified according to eight categories of sex/marital status and eight categories of age. x_{1t} is the household population and x_{2t} is the number of households whose head belongs to the given population category. We have used the data from Williams and Sams (1981). Here we consider four non-nested models M_1, M_2, M_3 and M_5 given by (3.16), (3.17), (3.18) and (3.19), respectively with $s_e^2 = 0.1$ and $s_b^2 = 0.02$ for samples of sizes 20, 50 and 100.

Design 3.7: This is an extension of Design 3.6 to more models using one extra variable x_{3t} as the household headship ratio which is the proportion of people in any given population category. Here we consider eight non-nested models M_1, M_2, M_3, M_4, M_5, M_6, M_7 and M_8 given by (3.16), (3.17), (3.18), (3.29), (3.19), (3.30), (3.31) and (3.32), respectively, with $s_e^2 = 0.1$ and $s_b^2 = 0.15$ for samples of sizes 20, 50 and 100.

We estimated the probabilities of correct selection for AIC, BIC, HQ, GCV, RBAR, HOC, NIC1, NIC2, NIC3, NIC4 and NIC5 from 2000 drawing of parameters for each design separately.

3.3.3.3 RESULTS AND DISCUSSION

Tables 3.6, 3.7, 3.8, 3.9, 3.10 and 3.11 contain the probabilities of correct selection of the true model for each of the Designs 3.2, 3.3, 3.4, 3.5, 3.6 and 3.7, respectively, under different criteria, namely AIC, BIC, HQ, GCV, RBAR, HOC, NIC1, NIC2, NIC3, NIC4 and NIC5. The mean average probability of correct selection (MAPCS) and the standard deviation among the average probabilities of correct selection (APCS) for selecting various models within a particular design under each criteria are also provided. A criterion with maximum mean average probability of correct selection and minimum variation among the probabilities of correctly selecting the true model for a fixed sample size is, always the most desirable.

Several interesting phenomena are apparent in the tables. The performance of the selected existing criteria varies from data set to data set. Even for a particular data set, the performance of the selected criteria varies from sample size to sample size. The probabilities of correct selection gradually increase as the sample size n increases for all designs under consideration, which is desirable. The variation among the probabilities of correct selection for selecting various models within a particular design under any criteria decreases as the sample size increases for all designs. Among the six existing criteria that we have considered, for nine out of 18 (six designs and three sample sizes) experiments, the MAPCS are the highest for AIC followed by BIC, for which eight experiments produce the highest MAPCS. But in most of the experiments, the variation among the probabilities of correct selection is the lowest for RBAR.

In terms of correctly choosing the true model, the new proposed criteria NIC1, NIC2, NIC3, NIC4 and NIC5 perform better than all the listed existing criteria for all designs and for all sample sizes. Among the new proposed criteria, the performances of NIC4 and NIC5 are marginally better than those of the remaining three proposed criteria, but the performances of NIC4 and NIC5 are very similar for all designs and sample sizes. The penalties that maximise the MAPCS obtained from the new criteria are different from those of the existing criteria. The variation among the APCS obtained using proposed criteria is higher for some designs and lower for other designs. Because of the existing criteria, AIC and BIC are the most widely used, and thus we will compare the lowest MAPCS (LMAPCS), obtained from the proposed criteria with AIC's and BIC's MAPCS. We choose the lowest MAPCS, so that our proposed IC are not favoured in the comparison.

For Design 3.2, with $n = 50$ and 100, the highest MAPCS is obtained from NIC4, and NIC5 produces the highest MAPCS for sample size 20. Among the new criteria, for $n = 20$ and 100, NIC2, and, for $n = 50$, NIC1 produces the lowest MAPCS, which are higher than the corresponding highest MAPCS obtained from the existing criteria. BIC produces the highest MAPCS among the existing criteria for $n = 20$ and 50, while for $n = 100$, HQ produces the highest MAPCS among the existing criteria.

There are 3.9, 2.9 and 3.6 percent increases in LMAPCS over that of AIC observed for sample sizes 20, 50 and 100, respectively. These increases are 1, 0.6 and 1.4 percent over the MAPCS of BIC for n = 20, 50 and 100, respectively. The variation among the APCS obtained from the proposed criteria is higher than that of all the existing criteria for n = 20, and lower than that of BIC for n = 50 and 100.

For Design 3.3, the highest MAPCS is obtained from NIC5 for n = 50 and 100, and both NIC4 and NIC5 produces the highest MAPCS for the sample size 20. Among the new criteria, the lowest MAPCS is obtained from NIC3 and NIC2 for n = 20 and 50, respectively, while both NIC3 and NIC2 produce the lowest MAPCS for the sample size 100. The lowest MAPCS obtained from the new criteria for different sample sizes are higher than the corresponding highest MAPCS obtained from the existing criteria. Among the existing criteria, AIC, BIC and GCV choose the true model with the highest MAPCS for n = 20, 50 and 100, respectively. There are 1.1, 0.3 and 0.3 percent improvements of LMAPCS over that of AIC observed for n = 20, 50 and 100, respectively. These improvements over BIC are 8.8, 7.1 and 6.9 percent, respectively. For n = 20 and 50, the variations among the APCS for the proposed criteria are lower than those of all cited existing criteria except RBAR. But for sample size 100, the variations among the APCS obtained from the proposed criteria are lower than those of BIC and HQ and higher than those of the other criteria.

For Design 3.4, the highest MAPCS is obtained from NIC4 for sample sizes 20 and 50, while both NIC4 and NIC5 produce the highest MAPCS for sample size 100. Among the new criteria, NIC3, NIC2 and NIC1 produces the lowest MAPCS for n = 20, 50 and 100, respectively, which are higher than the corresponding highest MAPCS obtained from the existing criteria. For all sample sizes, among the existing criteria the highest MAPCS is obtained from AIC. There is 0.8 percent increase in LMAPCS over that of AIC for n = 20, and for n = 50 and 100, LMAPCS are very similar to those of AIC. But these increases over the MAPCS of BIC are 9, 9.5 and 9.4 percent for sample sizes 20, 50 and 100, respectively. The variation among the APCS under the new criteria is always lower than that of all existing criteria except RBAR for n = 20 and 50. For n = 100, the proposed criteria produced lower variations compared to BIC and HQ, and higher than those of RABR, and in other cases the results are very similar to those of existing criteria.

For Design 3.5, NIC4 produces the highest MAPCS for n = 50 and 100, while for sample size 20, the highest MAPCS is obtained from NIC5. Among the new criteria for all sample sizes, the lowest MAPCS is obtained from NIC2, which are higher than the corresponding largest MAPCS obtained from the existing criteria. For all sample sizes among the existing criteria, BIC produces the highest MAPCS, which also has the highest variation among the APCS. The increase in MAPCS that results from using the worst of the new IC over AIC are 11.7, 14.2 and 17.1 percent for n = 20, 50 and 100, respectively. The increase of LMAPCS over that of BIC are 2.4, 1.1

and 0.4 percent for n = 20, 50 and 100, respectively. For all cases, the variations among the APCS obtained from the proposed technique are higher than those of existing criteria except for n = 100, where RBAR has the highest variation among the APCS.

For Design 3.6, and sample sizes 20 and 50, the highest MAPCS is obtained from NIC5, while NIC4 produces the highest MAPCS for n = 100. For all sample sizes among the new criteria, the lowest MAPCS is obtained from NIC3. Among the existing criteria, the highest MAPCS for sample sizes 20 and 50 is obtained from RBAR with the lowest variation among the APCS, while AIC produces the highest MAPCS for n = 100 with the second highest variation among the APCS. There are 2.3, 1.6 and 0.7 percent increases of the LMAPCS over those of AIC and 9.1, 14.8 and 11.9 percent increases over those of BIC observed for n = 20, 50 and 100, respectively. For all sample sizes, the variations among the APCS obtained from the new criteria are always lower than those of the existing criteria.

For Design 3.7 and all sample sizes, the highest MAPCS is obtained from NIC4. Among the new criteria, the lowest MAPCS is obtained from NIC2 and NIC1 for sample sizes 20 and 100, respectively, while both NIC1 and NIC2 produce the lowest MAPCS among the new criteria for sample size 50. For all sample sizes, among the existing criteria, the highest MAPCS is obtained from AIC. There are 1.1, 0.4, and 0.5 percent increases of the LMAPCS observed for the new technique over those of

AIC for n = 20, 50 and 100, respectively. These increases over the MAPCS of BIC are 6.5, 13, and 13.8 percent for n = 20, 50 and 100, respectively. The variations among the APCS obtained from new method are lower than those of all listed existing criteria except RBAR.

3.4 CONCLUDING REMARKS

The main purpose of this chapter was to introduce a new method for selecting the true model from a set of competing alternative models, that performs better on average compared to existing IC. The general form of the penalty functions of AIC, BIC and HQ for the j^{th} model is $p_j = \lambda_1 k_j$ and that for RBAR, GCV and HOC is $p_j = \lambda_2 \ln(n - k_j)$. The penalty functions of all six of these IC can be genaralised to a single penalty function $p_j = \lambda_1 k_j + \lambda_2 \ln(n - k_j)$. For the listed existing criteria, λ_1 and λ_2 are determined by the sample size n. For example if $\lambda_1 = 1$ and $\lambda_2 = 0$, the penalty is the AIC penalty, if $\lambda_1 = 0$ and $\lambda_2 = -\frac{n}{2}$, it is the RBAR penalty. In our proposed method, we did not restrict λ_1 and λ_2 to these six sets of values but allowed λ_1 and λ_2 to take any values that maximise the MAPCS. Another two penalty functions are also defined as $p_j = \lambda_1 k_j \ln(n - k_j)$ and $p_j = \lambda_1 k_j^{\lambda_2}$. We investigated the performance of the new method with these proposed penalties over the listed existing criteria in a linear regression setting.

We conducted simulation experiments to investigate how the efficiency of the estimated AAPCS increases with increases in the number of parameter drawings (q) for a simulation experiment with a fixed number of overall simulations, namely qN, where N is the number of simulations conducted for each drawing of the parameters. The average standard deviation averaged over the number of competing models (ASD) was used as a measure of efficiency and CV as a measure of reliability of the estimated AAPCS. From the simulation experiments, it was apparent that the number of parameter drawings for a simulation experiment and ASD is negatively correlated. For all cited criteria and for all sample sizes, the ASD and CV are the highest when $q = 1$ and $N = 2000$, which indicates that when $q = 1$, the AAPCS has the greatest variability and unreliability. With an increase in q, the value of ASD and CV decreases and in almost all cases the lowest ASD and CV is obtained when $q = 2000$, the maximum value of q, thus indicating that the AAPCS obtained using maximum q is more efficient and reliable.

Our regression results indicate that the relationship between ASD and q is well represented by the model $\ln(\text{ASD}) = a + b\ln(q) + c(\ln(q))^2 + u_t$. The estimated value of q, where ASD is the minimum, is always higher than the maximum value of qN (here 2000) used for our simulation experiment except for $n = 20$ in Design 3.2. This means that maximum number of parameter drawings with single replication is the best way to obtain an efficient estimate of mean probability of correct selection for a fixed qN. At this combination of q and N, the CV is also the lowest in almost all cases, indicating that the estimated mean probability of correct selection of the true model obtained using the maximum number of parameter drawings is more reliable.

We conducted a second set of simulation experiments to investigate the performance of our proposed new criteria over the existing listed criteria. The simulation results demonstrate that the performance of the proposed criteria always dominates the existing criteria in terms of MAPCS of the true model. Also, in general, the variations among the APCS obtained from new IC are smaller than those of all the listed existing IC except RBAR. From the simulation results, it is revealed that the performances of the six listed existing model selection criteria, vary from situation to situation and from data set to data set. Even for a particular data set, the performance of the selected criteria varies from sample size to sample size. In all designs under study, the MAPCS increases as the sample sizes increases. Within a particular design under any criteria, the variation among the APCS decreases as the sample size increases. Among the listed existing criteria in most cases, RBAR produces the lowest variation among the APCS. Although none of the existing criteria performs well in all situations in terms of MAPCS, the performance of all proposed new criteria are always better than the best of the existing listed criteria.

The performances of all new proposed IC are very similar, though NIC4 and NIC5 perform better than the others. The estimation of two parameters is required for NIC4 and NIC5, and this can be time consuming. The improvements of MAPCS obtained by NIC4 and NIC5 over NIC1, NIC2 and NIC3 are not significant. So considering the computational time and improvement in MAPCS, we recommend the use of any one of the criteria NIC1, NIC2 and NIC3.

Table 3.1 Average over 4 models of mean over 20 iterations of estimated mean probabilities of correct selection and their standard deviations for these 20 iterations averaged over 4 models for Design 3.1.

Number of beta drawings	AIC		BIC		GCV		HOC		HQ		RBAR		MCP		JIC	
	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD
Sample size = 20																
1	0.4564	0.1527	0.4553	0.1647	0.4581	0.1565	0.4586	0.1577	0.4580	0.1563	0.4337	0.1232	0.4567	0.1532	0.4572	0.1537
2	0.4334	0.1072	0.4264	0.1133	0.4338	0.1095	0.4335	0.1103	0.4336	0.1097	0.4187	0.0898	0.4336	0.1074	0.4335	0.1080
4	0.4229	0.0904	0.4167	0.0947	0.4228	0.0920	0.4226	0.0926	0.4229	0.0920	0.4079	0.0779	0.4230	0.0905	0.4230	0.0908
8	0.4294	0.0601	0.4223	0.0638	0.4303	0.0617	0.4302	0.0620	0.4303	0.0616	0.4123	0.0494	0.4296	0.0604	0.4299	0.0609
10	0.4295	0.0453	0.4225	0.0467	0.4299	0.0463	0.4294	0.0464	0.4296	0.0459	0.4131	0.0397	0.4296	0.0456	0.4297	0.0456
20	0.4295	0.0343	0.4224	0.0373	0.4291	0.0350	0.4287	0.0353	0.4289	0.0351	0.4129	0.0289	0.4295	0.0343	0.4296	0.0344
40	0.4316	0.0287	0.4251	0.0292	0.4316	0.0294	0.4315	0.0298	0.4313	0.0293	0.4158	0.0259	0.4316	0.0286	0.4316	0.0285
50	0.4307	0.0288	0.4247	0.0333	0.4308	0.0300	0.4305	0.0300	0.4307	0.0299	0.4122	0.0253	0.4309	0.0290	0.4311	0.0292
100	0.4307	0.0196	0.4238	0.0214	0.4313	0.0203	0.4312	0.0206	0.4313	0.0205	0.4156	0.0169	0.4306	0.0200	0.4308	0.0199
200	0.4318	0.0140	0.4265	0.0150	0.4321	0.0142	0.4322	0.0139	0.4322	0.0142	0.4137	0.0137	0.4319	0.0139	0.4322	0.0141
250	0.4307	0.0154	0.4254	0.0142	0.4313	0.0157	0.4311	0.0148	0.4312	0.0150	0.4136	0.0150	0.4307	0.0153	0.4308	0.0152
400	0.4350	0.0137	0.4292	0.0126	0.4350	0.0141	0.4350	0.0140	0.4349	0.0142	0.4158	0.0137	0.4350	0.0139	0.4348	0.0138
500	0.4328	0.0139	0.4274	0.0138	0.4333	0.0139	0.4330	0.0138	0.4332	0.0139	0.4152	0.0107	0.4332	0.0139	0.4330	0.0136
1000	0.4314	0.0111	0.4259	0.0107	0.4312	0.0113	0.4312	0.0111	0.4312	0.0113	0.4134	0.0111	0.4313	0.0112	0.4315	0.0110
2000	0.4329	0.0111	0.4263	0.0099	0.4330	0.0107	0.4329	0.0105	0.4330	0.0105	0.4171	0.0127	0.4331	0.0108	0.4330	0.0109
Sample size = 50																
1	0.6769	0.2022	0.7040	0.2571	0.6803	0.2060	0.6811	0.2068	0.6979	0.2303	0.6004	0.1457	0.6770	0.2023	0.6932	0.2215
2	0.6649	0.1462	0.6952	0.1744	0.6685	0.1482	0.6693	0.1486	0.6883	0.1609	0.5890	0.1112	0.6651	0.1463	0.6822	0.1565
4	0.6648	0.1060	0.7025	0.1311	0.6686	0.1077	0.6698	0.1083	0.6905	0.1182	0.5853	0.0785	0.6649	0.1060	0.6840	0.1144
8	0.6814	0.0776	0.7213	0.0941	0.6853	0.0790	0.6863	0.0793	0.7080	0.0870	0.5998	0.0570	0.6815	0.0777	0.7010	0.0842
10	0.6806	0.0685	0.7164	0.0817	0.6845	0.0695	0.6857	0.0697	0.7059	0.0750	0.5999	0.0532	0.6808	0.0685	0.6994	0.0728
20	0.6758	0.0409	0.7126	0.0495	0.6800	0.0416	0.6810	0.0417	0.7008	0.0447	0.5947	0.0314	0.6760	0.0410	0.6945	0.0431
40	0.6628	0.0297	0.6962	0.0354	0.6665	0.0299	0.6674	0.0299	0.6867	0.0323	0.5857	0.0235	0.6630	0.0297	0.6805	0.0319
50	0.6730	0.0258	0.7085	0.0296	0.6768	0.0261	0.6775	0.0263	0.6976	0.0277	0.5935	0.0228	0.6731	0.0259	0.6910	0.0277
100	0.6741	0.0260	0.7105	0.0292	0.6783	0.0258	0.6795	0.0256	0.6997	0.0271	0.5959	0.0221	0.6743	0.0261	0.6935	0.0262
200	0.6724	0.0157	0.7086	0.0184	0.6766	0.0161	0.6776	0.0162	0.6976	0.0174	0.5931	0.0150	0.6725	0.0157	0.6914	0.0164
250	0.6706	0.0147	0.7045	0.0177	0.6745	0.0145	0.6756	0.0144	0.6948	0.0162	0.5919	0.0128	0.6708	0.0148	0.6883	0.0160
400	0.6711	0.0154	0.7066	0.0150	0.6748	0.0154	0.6759	0.0153	0.6965	0.0154	0.5928	0.0146	0.6712	0.0153	0.6896	0.0160
500	0.6750	0.0143	0.7125	0.0138	0.6791	0.0143	0.6801	0.0140	0.7010	0.0140	0.5943	0.0126	0.6752	0.0143	0.6945	0.0137
1000	0.6726	0.0129	0.7094	0.0118	0.6767	0.0130	0.6777	0.0130	0.6983	0.0127	0.5929	0.0127	0.6729	0.0130	0.6920	0.0122
2000	0.6704	0.0109	0.7074	0.0099	0.6745	0.0109	0.6755	0.0109	0.6955	0.0104	0.5905	0.0109	0.6705	0.0109	0.6886	0.0103

Table 3.1 Average over 4 models of mean over 20 iterations of estimated mean probabilities of correct selection and their standard deviations for these 20 iterations averaged over 4 models for Design 3.1 (continued).

Number of beta drawings	AIC		BIC		GCV		HOC		HQ		RBAR		MCP		JIC	
	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD
Sample size = 96																
1	0.8340	0.0801	0.9359	0.1002	0.8376	0.0809	0.8386	0.0811	0.8973	0.0922	0.7058	0.0608	0.8341	0.0800	0.8865	0.0900
2	0.7996	0.0943	0.8830	0.1226	0.8029	0.0953	0.8038	0.0953	0.8549	0.1097	0.6810	0.0707	0.7997	0.0943	0.8459	0.1064
4	0.8222	0.0544	0.9129	0.0719	0.8258	0.0551	0.8267	0.0552	0.8807	0.0635	0.6988	0.0388	0.8222	0.0545	0.8707	0.0612
8	0.8113	0.0430	0.9015	0.0569	0.8146	0.0433	0.8156	0.0433	0.8686	0.0504	0.6919	0.0332	0.8114	0.0430	0.8588	0.0491
10	0.8168	0.0323	0.9095	0.0398	0.8202	0.0325	0.8211	0.0326	0.8761	0.0362	0.6941	0.0262	0.8169	0.0323	0.8658	0.0352
20	0.8171	0.0236	0.9079	0.0286	0.8207	0.0240	0.8215	0.0240	0.8741	0.0259	0.6940	0.0203	0.8172	0.0236	0.8647	0.0256
40	0.8149	0.0221	0.9027	0.0271	0.8187	0.0221	0.8197	0.0221	0.8716	0.0247	0.6939	0.0176	0.8150	0.0221	0.8620	0.0241
50	0.8105	0.0173	0.9002	0.0208	0.8142	0.0172	0.8152	0.0171	0.8679	0.0186	0.6868	0.0151	0.8106	0.0172	0.8580	0.0181
100	0.8146	0.0140	0.9060	0.0155	0.8183	0.0141	0.8192	0.0141	0.8728	0.0148	0.6909	0.0137	0.8147	0.0140	0.8627	0.0147
200	0.8120	0.0131	0.9021	0.0138	0.8156	0.0133	0.8165	0.0132	0.8694	0.0138	0.6900	0.0119	0.8121	0.0131	0.8600	0.0135
250	0.8144	0.0120	0.9056	0.0122	0.8179	0.0119	0.8187	0.0120	0.8728	0.0125	0.6929	0.0111	0.8145	0.0120	0.8631	0.0126
400	0.8124	0.0103	0.9037	0.0094	0.8159	0.0103	0.8167	0.0104	0.8707	0.0099	0.6927	0.0103	0.8124	0.0102	0.8609	0.0097
500	0.8115	0.0109	0.9029	0.0087	0.8153	0.0112	0.8161	0.0112	0.8693	0.0095	0.6894	0.0112	0.8116	0.0109	0.8596	0.0099
1000	0.8120	0.0094	0.9019	0.0071	0.8156	0.0092	0.8165	0.0092	0.8697	0.0087	0.6901	0.0106	0.8120	0.0094	0.8599	0.0093
2000	0.8116	0.0091	0.9021	0.0067	0.8151	0.0089	0.8160	0.0089	0.8693	0.0077	0.6896	0.0098	0.8117	0.0090	0.8598	0.0078

Table 3.2 Average over 4 models of mean over 20 iterations of estimated mean probabilities of correct selection and their standard deviations for these 20 iterations averaged over 4 models for Design 3.2.

Number of beta drawings	AIC		BIC		GCV		HOC		HQ		RBAR		MCP		JIC	
	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD
Sample size = 20																
1	0.5006	0.0633	0.5180	0.0618	0.5061	0.0634	0.5075	0.0632	0.5061	0.0634	0.4605	0.0564	0.5011	0.0632	0.5019	0.0632
2	0.4859	0.0630	0.5007	0.0655	0.4906	0.0640	0.4915	0.0643	0.4909	0.0639	0.4493	0.0509	0.4863	0.0631	0.4869	0.0633
4	0.4904	0.0306	0.5073	0.0282	0.4950	0.0298	0.4961	0.0293	0.4949	0.0299	0.4514	0.0303	0.4910	0.0306	0.4919	0.0305
8	0.4926	0.0223	0.5093	0.0218	0.4973	0.0228	0.4987	0.0228	0.4971	0.0227	0.4550	0.0220	0.4932	0.0224	0.4938	0.0224
10	0.4938	0.0225	0.5103	0.0203	0.4989	0.0218	0.5000	0.0218	0.4983	0.0216	0.4544	0.0216	0.4944	0.0225	0.4951	0.0224
20	0.4918	0.0165	0.5071	0.0164	0.4960	0.0162	0.4969	0.0163	0.4962	0.0161	0.4525	0.0153	0.4924	0.0166	0.4934	0.0164
40	0.4931	0.0160	0.5096	0.0148	0.4974	0.0167	0.4992	0.0165	0.4975	0.0168	0.4543	0.0146	0.4935	0.0162	0.4942	0.0164
50	0.4880	0.0148	0.5049	0.0144	0.4924	0.0149	0.4940	0.0149	0.4926	0.0148	0.4493	0.0154	0.4885	0.0147	0.4890	0.0149
100	0.4922	0.0136	0.5085	0.0107	0.4971	0.0131	0.4982	0.0129	0.4970	0.0132	0.4541	0.0130	0.4928	0.0136	0.4937	0.0134
200	0.4917	0.0105	0.5080	0.0099	0.4969	0.0103	0.4982	0.0099	0.4969	0.0104	0.4518	0.0114	0.4923	0.0102	0.4933	0.0103
250	0.4931	0.0116	0.5104	0.0101	0.4980	0.0117	0.4994	0.0111	0.4982	0.0115	0.4559	0.0105	0.4936	0.0119	0.4943	0.0120
400	0.4916	0.0114	0.5077	0.0089	0.4958	0.0108	0.4969	0.0105	0.4957	0.0107	0.4524	0.0111	0.4920	0.0113	0.4931	0.0111
500	0.4916	0.0118	0.5087	0.0092	0.4964	0.0115	0.4978	0.0114	0.4964	0.0115	0.4537	0.0113	0.4920	0.0116	0.4926	0.0117
1000	0.4943	0.0108	0.5089	0.0090	0.4981	0.0105	0.4995	0.0104	0.4984	0.0104	0.4563	0.0107	0.4949	0.0107	0.4953	0.0106
2000	0.4909	0.0102	0.5088	0.0083	0.4959	0.0092	0.4973	0.0088	0.4960	0.0093	0.4517	0.0094	0.4914	0.0101	0.4920	0.0099
Sample size = 50																
1	0.5448	0.0952	0.5568	0.0924	0.5463	0.0956	0.5465	0.0956	0.5566	0.0974	0.5004	0.0809	0.5449	0.0952	0.5536	0.0976
2	0.5316	0.0744	0.5432	0.0749	0.5331	0.0748	0.5333	0.0750	0.5420	0.0770	0.4862	0.0615	0.5317	0.0744	0.5397	0.0765
4	0.5451	0.0676	0.5584	0.0670	0.5471	0.0680	0.5477	0.0679	0.5570	0.0699	0.4972	0.0544	0.5452	0.0676	0.5545	0.0689
8	0.5537	0.0368	0.5670	0.0339	0.5555	0.0369	0.5558	0.0369	0.5654	0.0378	0.5038	0.0324	0.5537	0.0368	0.5632	0.0370
10	0.5421	0.0357	0.5549	0.0307	0.5440	0.0360	0.5445	0.0358	0.5535	0.0351	0.4970	0.0319	0.5421	0.0356	0.5513	0.0358
20	0.5505	0.0265	0.5637	0.0247	0.5522	0.0269	0.5529	0.0268	0.5629	0.0258	0.5026	0.0233	0.5504	0.0266	0.5599	0.0264
40	0.5492	0.0190	0.5608	0.0163	0.5511	0.0191	0.5516	0.0191	0.5603	0.0180	0.5017	0.0182	0.5493	0.0191	0.5578	0.0183
50	0.5497	0.0198	0.5626	0.0185	0.5519	0.0196	0.5526	0.0196	0.5620	0.0193	0.5000	0.0186	0.5499	0.0199	0.5593	0.0200
100	0.5471	0.0164	0.5602	0.0140	0.5488	0.0166	0.5492	0.0166	0.5594	0.0163	0.4994	0.0146	0.5472	0.0165	0.5562	0.0165
200	0.5520	0.0118	0.5666	0.0102	0.5542	0.0120	0.5544	0.0121	0.5645	0.0107	0.5039	0.0113	0.5520	0.0119	0.5621	0.0115
250	0.5480	0.0132	0.5616	0.0103	0.5494	0.0132	0.5499	0.0135	0.5607	0.0133	0.5005	0.0128	0.5481	0.0132	0.5575	0.0136
400	0.5460	0.0103	0.5593	0.0084	0.5479	0.0103	0.5485	0.0103	0.5581	0.0099	0.4984	0.0114	0.5461	0.0103	0.5557	0.0098
500	0.5491	0.0108	0.5613	0.0088	0.5510	0.0108	0.5512	0.0105	0.5607	0.0100	0.5015	0.0105	0.5492	0.0108	0.5579	0.0101
1000	0.5470	0.0113	0.5629	0.0088	0.5491	0.0112	0.5496	0.0111	0.5600	0.0115	0.4993	0.0125	0.5471	0.0113	0.5571	0.0110
2000	0.5489	0.0099	0.5620	0.0067	0.5507	0.0097	0.5512	0.0096	0.5609	0.0082	0.4992	0.0108	0.5491	0.0099	0.5585	0.0082

Table 3.2 Average over 4 models of mean over 20 iterations of estimated mean probabilities of correct selection and their standard deviations for these 20 iterations averaged over 4 models for Design 3.2 (continued).

Number of beta drawings	AIC		BIC		GCV		HOC		HQ		RBAR		MCP		JIC	
	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD	Average of average	Average SD
Sample size = 100																
1	0.6066	0.1313	0.6212	0.1449	0.6079	0.1323	0.6083	0.1324	0.6255	0.1442	0.5453	0.1036	0.6066	0.1314	0.6235	0.1423
2	0.6126	0.0938	0.6195	0.1082	0.6141	0.0944	0.6143	0.0944	0.6284	0.1034	0.5516	0.0734	0.6126	0.0938	0.6271	0.1020
4	0.6028	0.0599	0.6129	0.0626	0.6040	0.0601	0.6045	0.0601	0.6194	0.0634	0.5421	0.0493	0.6028	0.0599	0.6182	0.0630
8	0.6039	0.0530	0.6143	0.0578	0.6053	0.0532	0.6056	0.0534	0.6205	0.0569	0.5443	0.0410	0.6039	0.0530	0.6193	0.0566
10	0.6170	0.0484	0.6319	0.0493	0.6183	0.0486	0.6186	0.0487	0.6358	0.0506	0.5519	0.0408	0.6169	0.0485	0.6334	0.0504
20	0.6224	0.0412	0.6341	0.0428	0.6235	0.0413	0.6238	0.0413	0.6408	0.0439	0.5565	0.0335	0.6224	0.0412	0.6393	0.0434
40	0.6162	0.0274	0.6308	0.0261	0.6175	0.0272	0.6178	0.0272	0.6356	0.0276	0.5513	0.0220	0.6162	0.0274	0.6337	0.0272
50	0.6207	0.0260	0.6329	0.0254	0.6219	0.0264	0.6223	0.0264	0.6395	0.0267	0.5558	0.0221	0.6208	0.0260	0.6379	0.0271
100	0.6190	0.0216	0.6324	0.0195	0.6201	0.0216	0.6204	0.0216	0.6373	0.0207	0.5544	0.0171	0.6190	0.0217	0.6359	0.0210
200	0.6171	0.0148	0.6320	0.0131	0.6184	0.0148	0.6187	0.0148	0.6371	0.0137	0.5507	0.0134	0.6171	0.0148	0.6349	0.0144
250	0.6173	0.0161	0.6313	0.0131	0.6185	0.0159	0.6190	0.0159	0.6361	0.0156	0.5530	0.0143	0.6173	0.0161	0.6345	0.0158
400	0.6171	0.0129	0.6323	0.0107	0.6184	0.0128	0.6188	0.0128	0.6377	0.0124	0.5514	0.0117	0.6171	0.0129	0.6353	0.0128
500	0.6190	0.0125	0.6321	0.0104	0.6204	0.0125	0.6208	0.0125	0.6381	0.0115	0.5536	0.0121	0.6190	0.0126	0.6369	0.0119
1000	0.6192	0.0097	0.6330	0.0081	0.6204	0.0099	0.6207	0.0099	0.6381	0.0095	0.5548	0.0112	0.6192	0.0097	0.6361	0.0097
2000	0.6185	0.0111	0.6341	0.0071	0.6198	0.0113	0.6202	0.0113	0.6382	0.0093	0.5550	0.0120	0.6185	0.0111	0.6361	0.0097

Table 3.3 The estimated models$^{\otimes}$ of the relationship between ASD and q under different criteria

Sample size	Criteria	Design 3.1				Design 3.2			
		$\hat{A}$	$\hat{B}$	$\bar{R}^2$	DW	$\hat{A}$	$\hat{B}$	$\bar{R}^2$	DW
20	AIC	-2.081	-0.365	0.958	0.552	-3.086	-0.239	0.832***	0.595
	BIC	-1.979***	-0.385***	0.970***	0.768**	-3.070***	-0.270***	0.868***	0.745**
	GCV	-2.049***	-0.369***	0.963***	0.571**	-3.069***	-0.-47***	0.850***	0.688**
	HOC	-2.035***	-0.373***	0.964***	0.575**	-3.065***	-0.252***	0.855***	0.712**
	HQ	-20.47***	-0.370***	0.964***	0.549**	-3.072***	-0.246***	0.848***	0.676**
	RBAR	-2.308***	-0.335***	0.932***	0.713**	-3.178***	-0.228***	0.863***	0.456**
	MCP	-2.076***	-0.366***	0.960***	0.540**	-3.081***	-0.240***	0.836***	0.628**
	JIC	-2.068***	-0.368***	0.961***	0.558**	-3.078***	-0.242***	0.839***	0.644**
50	AIC	-1.788	-0.403	0.957	0.581	-2.524	-0.321	0.934	0.674
	BIC	-1.513***	-0.444***	0.979***	0.706**	-2.508***	-0.362***	0.953***	0.814**
	GCV	-1.768***	-0.406***	0.958***	0.552**	-2.514***	-0.323***	0.938***	0.689**
	HOC	-1.763***	-0.408***	0.958***	0.546**	-2.513***	-0.324***	0.940***	0.731**
	HQ	-1.646***	-0.424***	0.969***	0.560**	-2.485***	-0.338***	0.935***	0.911**
	RBAR	-2.141***	-0.358***	0.942***	0.603**	-2.744***	-0.282***	0.908***	0.578**
	MCP	-1.787***	-0.403***	0.957***	0.580**	-2.524***	-0.321***	0.935***	0.689**
	JIC	-1.681***	-0.421***	0.971***	0.607**	-2.479***	-0.337***	0.946***	0.967**
96 for Design 3.1 and 100 for Design 3.2	AIC	-2.537	-0.328	0.939	0.916	-2.214	-0.346	0.974	1.291^{ns}
	BIC	-2.515***	-0.407***	0.974***	1.727^{ns}	-2.043***	-0.407***	0.988***	1.697^{ns}
	GCV	-2.524***	-0.331***	0.940***	0.932*	-2.211***	-0.345***	0.972***	1.231nd
	HOC	-2.523***	-0.331***	0.941***	0.927*	-2.210***	-0.345***	0.972***	1.205nd
	HQ	-2.329***	-0.369***	0.959***	1.254^{ns}	-2.098***	-0.375***	0.980***	1.363^{ns}
	RBAR	-2.888***	-0.271***	0.904***	0.727**	-2.518***	-0.303***	0.944***	0.624**
	MCP	-2.537***	-0.239***	0.939***	0.912*	-2.214***	-0.346***	0.974***	1.299^{ns}
	JIC	-2.371***	-0.360***	0.955***	1.174^{ns}	-2.118***	-0.368***	0.980***	1.346^{ns}

*** Significant at 0.1 percent level
** Significant at 1 percent level
* Significant at 5 percent level
ns Not significant
nd No decision
$\otimes$ $\ln(\text{ASD}) = \hat{A} + \hat{B}\ln q$
q must be greater than or equal to 1

Table 3.4 The estimated models$^{\otimes}$ of the relationship between ASD and q under different criteria for Design 3.1.

Sample size	Criteria	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\bar{R}^2$	DW	$\hat{q}^{\#}$
20	AIC	-1.800	-0.594	0.0301	0.988	1.709^{ns}	19286
	BIC	-1.754***	-0.569***	0.0242**	0.987***	1.720^{ns}	127542
	GCV	-1.788***	-0.582***	0.0280***	0.989***	1.691^{ns}	32626
	HOC	-1.778***	-0.583***	0.0276***	0.988***	1.571^{ns}	38623
	HQ	-1.788***	-0.582***	0.0278***	0.989***	1.593^{ns}	35159
	RBAR	-1.979***	-0.603***	0.0353***	0.979***	1.682^{ns}	5121
	MCP	-1.801***	-0.590***	0.0295***	0.988***	1.665^{ns}	22026
	JIC	-1.796***	-0.591***	0.0293***	0.989***	1.665^{ns}	23988
50	AIC	-1.478	-0.565	0.0333	0.987	1.692^{ns}	4834
	BIC	-1.280***	-0.633***	0.0249***	0.993***	1.817^{ns}	331321
	GCV	-1.459***	-0.659***	0.0333***	0.987***	1.583^{ns}	19829
	HOC	-1.451***	-0.661***	0.0335***	0.987***	1.542^{ns}	19258
	HQ	-1.369***	-0.649***	0.0297***	0.991***	1.622^{ns}	55599
	RBAR	-1.807***	-0.630***	0.0359***	0.989***	1.898^{ns}	6466
	MCP	-1.478***	-0.655***	0.0332***	0.987***	1.695^{ns}	19234
	JIC	-1.406***	-0.644***	0.0295***	0.991***	1.740^{ns}	55009
96	AIC	-2.264	-0.551	0.0293	0.972	2.243^{ns}	12121
	BIC	-2.029***	-0.507***	0.0132^{ns}	0.977***	2.159^{ns}	218991934
	GCV	-2.254***	-0.551***	0.0290*	0.972***	2.251^{ns}	13360
	HOC	-2.254***	-0.550***	0.0288*	0.973***	2.249^{ns}	14025
	HQ	-2.120***	-0.539***	0.0225*	0.974***	2.211^{ns}	159178
	RBAR	-2.573***	-0.528***	0.0338***	0.969***	2.495^{ns}	2467
	MCP	-2.263***	-0.552***	0.0294*	0.972***	2.237^{ns}	11941
	JIC	-2.148***	-0.542***	0.0239*	0.973***	2.257^{ns}	84029

*** Significant at 0.1 percent level
** Significant at 1 percent level
* Significant at 5 percent level
ns Not significant
\# Estimated value of q where ASD is the minimum
$\otimes$ $\ln(\mathrm{ASD}) = \hat{a} + \hat{b}\ln q + \hat{c}(\ln q)^2$
q must be greater than or equal to 1

Table 3.5 The estimated models$^{\otimes}$ of the relationship between ASD and q under different criteria for Design 3.2.

Sample size	Criteria	$\hat{a}$	$\hat{b}$	$\hat{c}$	$\bar{R}^2$	DW	$\hat{q}^{\#}$
20	AIC	-2.681	-0.568	0.0434	0.962	2.335^{ns}	695
	BIC	-2.682^{***}	-0.587^{***}	0.0417^{***}	0.964^{***}	2.553^{nd}	1140
	GCV	-2.693^{***}	-0.554^{***}	0.0404^{***}	0.955^{***}	2.300^{ns}	950
	HOC	-2.692^{***}	-0.556^{***}	0.0400^{***}	0.955^{***}	2.328^{ns}	1043
	HQ	-2.692^{***}	-0.556^{***}	0.0408^{***}	0.957^{***}	2.302^{ns}	910
	RBAR	-2.829^{***}	-0.512^{***}	0.0375^{***}	0.973^{***}	1.843^{ns}	922
	MCP	-2.682^{***}	-0.566^{***}	0.0429^{***}	0.961^{***}	2.418^{ns}	733
	JIC	-2.685^{***}	-0.561^{***}	0.0421^{***}	0.956^{***}	2.364^{ns}	783
50	AIC	-2.212	-0.575	0.0334	0.981	2.172^{ns}	5474
	BIC	-2.236^{***}	-0.583^{***}	0.0292^{**}	0.981^{***}	2.513^{ns}	21653
	GCV	-2.211^{***}	-0.570^{***}	0.0325^{***}	0.982^{***}	2.205^{ns}	6433
	HOC	-2.215^{***}	-0.566^{***}	0.0319^{***}	0.982^{***}	2.277^{ns}	7126
	HQ	-2.186^{***}	-0.582^{***}	0.0321^{**}	0.974^{***}	2.417^{ns}	8651
	RBAR	-2.400^{***}	-0.562^{***}	0.0369^{***}	0.981^{***}	2.219^{ns}	2029
	MCP	-2.215^{***}	-0.573^{***}	0.0332^{***}	0.981^{***}	2.196^{ns}	5594
	JIC	-2.205^{***}	-0.560^{***}	0.0294^{**}	0.978^{***}	2.528^{ns}	13682
100	AIC	-2.035	-0.491	0.0192	0.987	2.235^{ns}	357345
	BIC	-1.926^{***}	-0.502^{***}	0.0125^{*}	0.992^{***}	2.537^{ns}	525573182
	GCV	-2.023^{***}	-0.499^{***}	0.0202^{**}	0.986^{***}	2.123^{ns}	231303
	HOC	-2.022^{***}	-0.499^{***}	0.0202^{*}	0.987^{***}	2.129^{nd}	231303
	HQ	-1.938^{***}	-0.505^{***}	0.0171^{*}	0.989^{***}	2.285^{ns}	2587187
	RBAR	-2.243^{***}	-0.528^{***}	0.0295^{***}	0.986^{***}	1.449^{ns}	7701
	MCP	-2.035^{***}	-0.491^{***}	0.0192^{*}	0.987^{***}	2.244^{ns}	357345
	JIC	-1.955^{***}	-0.500^{***}	0.0174^{*}	0.990^{***}	2.363^{ns}	1737254

*** Significant at 0.1 percent level
** Significant at 1 percent level
* Significant at 5 percent level
ns Not significant
nd No decision
\# Estimated value of q where ASD is the minimum
$\otimes$ $\ln(\text{ASD}) = \hat{a} + \hat{b}\ln q + \hat{c}(\ln q)^2$
q must be greater than or equal to 1

Table 3.6 Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for samples of different sizes and under different criteria for Design 3.2.

Sample size	Model	Average probabilities of correct selection with penalties (given in the parenthesis for new criteria only) under different criteria										
		AIC	BIC	HQ	GCV	RBAR	HOC	NIC1	NIC2	NIC3	NIC4	NIC5
20	M_1	0.6715	0.8115	0.7120	0.7095	0.4705	0.7205	0.8745 (1.8600)	0.8605 (-95.3998)	0.8765 (1.9728)	0.8750 (1.8900)	0.8795 (-28.16439)
	M_2	0.8100	0.8850	0.8260	0.8315	0.6785	0.8365	0.9160 (3.7200)	0.9165 (-93.6480)	0.9140 (3.8731)	0.9160 (3.7539)	0.9165 (-26.34372)
	M_3	0.2170	0.1595	0.2050	0.2085	0.2805	0.2070	0.1335 (3.7200)	0.1425 (-93.6480)	0.1310 (3.8731)	0.1330 (3.7539)	0.1350 (-26.34372)
	M_5	0.2945	0.1940	0.2720	0.2635	0.4385	0.2565	0.1505 (5.5800)	0.1510 (-91.7961)	0.1530 (5.6948)	0.1510 (5.6081)	0.1510 (-24.49213)
	Mean	**0.4983**	**0.5125**	**0.5038**	**0.5033**	**0.4670**	**0.5051**	**0.5186**	**0.5176**	**0.5186**	**0.5188**	**0.5205**
	SD	0.2874	0.3891	0.3110	0.3134	0.1636	0.3198	0.4353	0.4289	0.4353	0.4354	0.4362
50	M_1	0.7080	0.9055	0.8110	0.7195	0.4780	0.7225	0.8570 (1.6000)	0.8570 (-299.6702)	0.8705 (1.6735)	0.8570 (1.5900)	0.8555 (-153.9669)
	M_2	0.8255	0.9435	0.8910	0.8365	0.6690	0.8370	0.9160 (3.1900)	0.9225 (-298.0825)	0.9235 (3.3292)	0.9190 (3.1800)	0.9190 (-151.3924)
	M_3	0.3090	0.2060	0.2645	0.3045	0.3620	0.3040	0.2445 (3.1900)	0.2445 (-298.0825)	0.2315 (3.3292)	0.2454 (3.1800)	0.2460 (-151.3924)
	M_5	0.3580	0.1945	0.2735	0.3430	0.5000	0.3405	0.2445 (4.7600)	0.2400 (-296.4614)	0.2375 (4.9667)	0.2445 (4.7700)	0.2445 (-149.8008)
	Mean	**0.5501**	**0.5624**	**0.5600**	**0.5509**	**0.5023**	**0.5510**	**0.5655**	**0.5660**	**0.5658**	**0.5665**	**0.5663**
	SD	0.2555	0.4185	0.3376	0.2670	0.1266	0.2687	0.3714	0.3748	0.3831	0.3721	0.3716
100	M_1	0.7175	0.9370	0.8455	0.7255	0.4770	0.7265	0.8610 (1.6200)	0.8905 (-817.9313)	0.8820 (1.7462)	0.8645 (1.6800)	0.8780 (-339.0689)
	M_2	0.8470	0.9670	0.9190	0.8535	0.6800	0.8545	0.9255 (3.2400)	0.9430 (-816.1242)	0.9365 (3.4846)	0.9255 (3.3137)	0.9365 (-337.3476)
	M_3	0.4145	0.3325	0.3900	0.4140	0.4185	0.4130	0.3880 (3.2400)	0.3720 (-816.1242)	0.3770 (3.4846)	0.3870 (3.3137)	0.3815 (-337.3476)
	M_5	0.4810	0.2765	0.3875	0.4775	0.6120	0.4770	0.3745 (4.8600)	0.3425 (-814.2986)	0.3545 (5.2152)	0.3745 (4.9305)	0.3550 (-335.6186)
	Mean	**0.6150**	**0.6283**	**0.6355**	**0.6176**	**0.5469**	**0.6178**	**0.6373**	**0.6370**	**0.6375**	**0.6379**	**0.6378**
	SD	0.2021	0.3747	0.2865	0.2069	0.1202	0.2079	0.2968	0.3240	0.3147	0.2980	0.3123

Penalties for NIC1, NIC2, NIC3, NIC4 and NIC5 are $\lambda_1 k_j$, $\lambda_2 \ln(n - k_j)$, $\lambda_1 k_j \ln(n - k_j)$, $\lambda_1 k_j^{\lambda_2}$ and $\lambda_1 k_j + \lambda_2 \ln(n - k_j)$, respectively.

Table 3.7 Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for samples of different sizes and under different criteria for Design 3.3.

Sample size	Model	Average probabilities of correct selection with penalties (given in the parenthesis for new criteria only) under different criteria										
		AIC	BIC	HQ	GCV	RBAR	HOC	NIC1	NIC2	NIC3	NIC4	NIC5
20	M_1	0.9080	0.9685	0.9200	0.9180	0.7205	0.9245	0.8455 (0.7890)	0.8305 (-40.6333)	0.8575 (0.8539)	0.8000 (0.4200)	0.8460 (-27.9966)
	M_2	0.6425	0.5955	0.6345	0.6420	0.6645	0.6365	0.6615 (1.5780)	0.6695 (-39.8871)	0.6555 (1.6764)	0.6870 (1.0931)	0.6660 (-27.2075)
	M_4	0.5935	0.5395	0.5885	0.5900	0.6340	0.5925	0.6155 (1.5780)	0.6230 (-39.8871)	0.6065 (1.6764)	0.6415 (1.0931)	0.6180 (-27.2075)
	M_5	0.4290	0.3200	0.4005	0.3920	0.5495	0.3835	0.4830 (2.3670)	0.4830 (-39.0983)	0.4820 (1.4649)	0.4790 (1.9128)	0.4775 (-26.3889)
	Mean	**0.6433**	**0.6059**	**0.6359**	**0.6355**	**0.6421**	**0.6343**	**0.6514**	**0.6515**	**0.6504**	**0.6519**	**0.6519**
	SD	0.1987	0.2694	0.2148	0.2170	0.0714	0.2228	0.1499	0.1433	0.1562	0.1331	0.1521
50	M_1	0.9265	0.9950	0.9715	0.9310	0.7420	0.9330	0.9115 (0.9100)	0.9075 (-169.294)	0.9140 (0.9340)	0.9170 (0.7700)	0.9110 (-98.8614)
	M_2	0.7600	0.6975	0.7425	0.7595	0.7375	0.7595	0.7590 (1.8200)	0.7600 (-168.397)	0.7590 (1.8582)	0.7690 (1.7087)	0.7605 (-97.9556)
	M_4	0.7145	0.6460	0.6855	0.7125	0.7220	0.7135	0.7245 (1.8200)	0.7275 (-168.397)	0.7235 (1.8582)	0.7270 (1.7087)	0.7265 (-97.9556)
	M_5	0.5750	0.4500	0.5150	0.5685	0.6895	0.5650	0.5915 (2.7300)	0.5915 (-167.481)	0.5915 (2.7721)	0.5735 (2.7238)	0.5910 (-97.0388)
	Mean	**0.7440**	**0.6971**	**0.7286**	**0.7429**	**0.7228**	**0.7428**	**0.7466**	**0.7466**	**0.7470**	**0.7466**	**0.7473**
	SD	0.1449	0.2254	0.1686	0.1494	0.0238	0.1516	0.1315	0.1297	0.1326	0.1413	0.1314
100	M_1	0.9330	0.9985	0.9805	0.9390	0.7545	0.9400	0.9460 (1.0600)	0.9545 (-510.058)	0.9575 (1.1488)	0.9435 (0.9900)	0.9460 (-229.206)
	M_2	0.7965	0.7455	0.7885	0.7970	0.7750	0.7970	0.7985 (2.1200)	0.8000 (-508.931)	0.7975 (2.2925)	0.8010 (2.0357)	0.7990 (-228.148)
	M_4	0.7880	0.7135	0.7595	0.7885	0.7470	0.7875	0.7875 (2.1200)	0.7850 (-508.931)	0.7845 (2.2925)	0.7885 (2.0357)	0.7875 (-228.14[illegible])
	M_5	0.7000	0.5620	0.6425	0.6980	0.7790	0.6975	0.6950 (3.1800)	0.6870 (-507.793)	0.6870 (3.4310)	0.6940 (3.1034)	0.6950 (-227.086)
	Mean	**0.8044**	**0.7549**	**0.7928**	**0.8056**	**0.7638**	**0.8055**	**0.8068**	**0.8066**	**0.8066**	**0.8068**	**0.8069**
	SD	0.0962	0.1811	0.1402	0.0996	0.0156	0.1003	0.1038	0.1106	0.1120	0.1029	0.1038

Penalties for NIC1, NIC2, NIC3, NIC4 and NIC5 are $\lambda_1 k_j$, $\lambda_2 \ln(n - k_j)$, $\lambda_1 k_j \ln(n - k_j)$, $\lambda_1 k_j^{\lambda_2}$ and $\lambda_1 k_j + \lambda_2 \ln(n - k_j)$, respectively.

Table 3.8 Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for samples of different sizes and under different criteria for Design 3.4.

Sample size	Model	Average probabilities of correct selection with penalties (given in the parenthesis for new criteria only) under different criteria										
		AIC	BIC	HQ	GCV	RBAR	HOC	NIC1	NIC2	NIC3	NIC4	NIC5
20	M_1	0.8630	0.9545	0.8915	0.8875	0.6185	0.8945	0.7770 (0.8000)	0.7590 (-41.2221)	0.8070 (0.8833)	0.6875 (0.3500)	0.7710 (-31.6099)
	M_2	0.5590	0.5310	0.5560	0.5670	0.5330	0.5650	0.5525 (1.6000)	0.5625 (-40.4652)	0.5465 (1.7342)	0.5885 (0.9763)	0.5595 (-30.8360)
	M_3	0.5780	0.5415	0.5745	0.5830	0.5400	0.5840	0.5780 (1.6000)	0.5840 (-40.4652)	0.5735 (1.7342)	0.6110 (0.9763)	0.5830 (-30.8360)
	M_4	0.6205	0.6055	0.6230	0.6290	0.5680	0.6265	0.6120 (1.6000)	0.6175 (-40.4652)	0.6050 (1.7342)	0.6415 (0.9763)	0.6175 (-30.8360)
	M_5	0.4005	0.3265	0.3845	0.3845	0.4635	0.3780	0.4290 (3.4000)	0.4335 (-39.6650)	0.4235 (3.5499)	0.4390 (1.7791)	0.4315 (-30.0287)
	M_6	0.3850	0.3060	0.3710	0.3650	0.4520	0.3595	0.4210 (3.4000)	0.4245 (-39.6650)	0.4145 (3.5499)	0.4275 (1.7791)	0.4210 (-30.0287)
	M_7	0.4290	0.3505	0.4100	0.4025	0.4825	0.3980	0.4580 (3.4000)	0.4635 (-39.6650)	0.4525 (3.5499)	0.4720 (1.7791)	0.4620 (-30.0287)
	M_8	0.3290	0.2385	0.3135	0.2930	0.4280	0.2845	0.3730 (3.2000)	0.3545 (-38.8162)	0.3760 (3.3271)	0.3395 (3.7234)	0.3560 (-29.1840)
	Mean	**0.5205**	**0.4818**	**0.5155**	**0.5139**	**0.5107**	**0.5113**	**0.5251**	**0.5249**	**0.5248**	**0.5258**	**0.5252**
	SD	0.1732	0.2318	0.1882	0.1929	0.0650	0.1976	0.1322	0.1307	0.1408	0.1228	0.1340
50	M_1	0.8900	0.9920	0.9570	0.8965	0.6565	0.8980	0.8770 (0.9500)	0.8715 (-175.1319)	0.8805 (0.9730)	0.8800 (1.0900)	0.8730 (-136.0037)
	M_2	0.7055	0.6650	0.6990	0.7110	0.6200	0.7100	0.7055 (1.9000)	0.7065 (-174.2041)	0.7035 (1.9356)	0.6940 (2.0482)	0.7065 (-135.0720)
	M_3	0.7450	0.7195	0.7495	0.7485	0.6320	0.7490	0.7380 (1.9000)	0.7405 (-174.2041)	0.7380 (1.9356)	0.7285 (2.0482)	0.7395 (-135.0720)
	M_4	0.7200	0.6905	0.7285	0.7250	0.6300	0.7250	0.7175 (1.9000)	0.7180 (-174.2041)	0.7170 (1.9356)	0.7120 (2.0482)	0.7180 (-135.0720)
	M_5	0.5745	0.4580	0.5280	0.57[illegible]5	0.6180	0.5710	0.5795 (3.8500)	0.5810 (-173.2566)	0.5795 (3.8876)	0.5830 (3.9622)	0.5815 (-134.1252)
	M_6	0.5785	0.4615	0.5345	0.5710	0.6145	0.5690	0.5870 (3.8500)	0.5885 (-173.2566)	0.5860 (3.8876)	0.5905 (3.9622)	0.5890 (-134.1252)
	M_7	0.5805	0.4690	0.5330	0.5730	0.6090	0.5725	0.5815 (3.8500)	0.5830 (-173.2566)	0.5800 (3.8876)	0.5865 (3.9622)	0.5830 (-134.1252)
	M_8	0.5335	0.4125	0.4825	0.5200	0.6310	0.5175	0.5440 (3.8000)	0.5405 (-173.2889)	0.5455 (3.8286)	0.5590 (3.8486)	0.5420 (-133.1625)
	Mean	**0.6659**	**0.6085**	**0.6515**	**0.6646**	**0.6264**	**0.6640**	**0.6663**	**0.6662**	**0.6663**	**0.6667**	**0.6666**
	SD	0.1206	0.1969	0.1615	0.1272	0.0148	0.1284	0.1132	0.1121	0.1140	0.1089	0.1121
100	M_1	0.9040	0.9960	0.9750	0.9080	0.6575	0.9120	0.9040 (1.[illegible]08)	0.9030 (-450.3218)	0.9050 (1.0109)	0.8950 (0.8500)	0.9050 (-123.3382)
	M_2	0.7740	0.7365	0.7745	0.7775	0.6645	0.7765	0.7690 (2.0164)	0.7760 (-449.3268)	0.7740 (2.0174)	0.7785 (1.8094)	0.7760 (-123.3341)
	M_3	0.7840	0.7540	0.7820	0.7845	0.6655	0.7845	0.7860 (2.0164)	0.7855 (-449.3268)	0.7830 (2.0174)	0.7895 (1.8094)	0.7845 (-123.3341)
	M_4	0.7855	0.7500	0.7925	0.7880	0.6590	0.7880	0.7858 (2.0164)	0.7865 (-449.3268)	0.7855 (2.0174)	0.7895 (1.8094)	0.7865 (-123.3341)
	M_5	0.6605	0.5350	0.6095	0.6590	0.6620	0.6590	0.6619 (3.0195)	0.6610 (-448.3217)	0.6605 (3.0193)	0.6620 (3.8150)	0.6610 (-121.3272)
	M_6	0.6660	0.5575	0.6205	0.6630	0.6710	0.6630	0.6676 (3.0195)	0.6660 (-448.3217)	0.6660 (3.0193)	0.6675 (3.8150)	0.6660 (-121.3272)
	M_7	0.6420	0.5260	0.5990	0.6400	0.6485	0.6395	0.6422 (3.0195)	0.6420 (-448.3217)	0.6420 (3.0193)	0.6435 (3.8150)	0.6420 (-121.3272)
	M_8	0.6550	0.5105	0.5900	0.6495	0.7360	0.6475	0.6550 (4.0167)	0.6530 (-447.3061)	0.6555 (4.0166)	0.6495 (3.8518)	0.6540 (-120.3174)
	Mean	**0.7339**	**0.6707**	**0.7179**	**0.7337**	**0.6705**	**0.7338**	**0.7339**	**0.7341**	**0.7339**	**0.7344**	**0.7344**
	SD	0.0929	0.1695	0.1366	0.0958	0.0273	0.0971	0.0924	0.0931	0.0930	0.0918	0.0934

Penalties for NIC1, NIC2, NIC3, NIC4 and NIC5 are $\lambda_1 k_j$, $\lambda_2 \ln(n - k_j)$, $\lambda_1 k_j \ln(n - k_j)$, $\lambda_1 k_j^{\lambda_2}$ and $\lambda_1 k_j + \lambda_2 \ln(n - k_j)$, respectively.

Table 3.9 Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for samples of different sizes and under different criteria for Design 3.5.

Sample size	Model	Average probabilities of correct selection with penalties (given in the parenthesis for new criteria only) under different criteria										
		AIC	BIC	HQ	GCV	RBAR	HOC	NIC1	NIC2	NIC3	NIC4	NIC5
20	M_1	0.5520	0.7200	0.5945	0.5885	0.3115	0.6015	0.8780 (3.3200)	0.8585 (-119.5442)	0.8830 (3.4733)	0.8785 (3.2900)	0.8780 (-9.6778)
	M_2	0.6440	0.7830	0.6820	0.6970	0.4595	0.7095	0.9005 (4.6400)	0.9005 (-117.3491)	0.8965 (4.8558)	0.9015 (4.6119)	0.9010 (-7.3615)
	M_3	0.3800	0.4095	0.3930	0.4030	0.3090	0.4060	0.3815 (4.6400)	0.3930 (-117.3491)	0.3770 (4.8558)	0.3820 (4.6119)	0.3820 (-7.3615)
	M_4	0.5685	0.6315	0.5985	0.6150	0.4250	0.6235	0.7505 (4.6400)	0.7550 (-117.3491)	0.7455 (4.8558)	0.7510 (4.6119)	0.7515 (-7.3615)
	M_5	0.4360	0.4180	0.4380	0.4435	0.4330	0.4425	0.3530 (6.9600)	0.3545 (-115.0285)	0.3540 (7.1397)	0.3520 (6.9459)	0.3525 (-5.0329)
	M_6	0.6990	0.7475	0.7135	0.7215	0.6200	0.7350	0.7665 (6.9600)	0.7745 (-115.0285)	0.7630 (7.1397)	0.7685 (6.9459)	0.7685 (-5.0329)
	M_7	0.4195	0.3870	0.4135	0.41[illegible]0	0.4205	0.4125	0.3290 (6.9600)	0.3315 (-115.0285)	0.3295 (7.1397)	0.3295 (6.9459)	0.3300 (-5.0329)
	M_8	0.5180	0.4520	0.5040	0.4855	0.6035	0.4805	0.3605 (9.2800)	0.3415 (-113.5671)	0.3725 (9.3159)	0.3570 (9.2879)	0.3570 (-3.6904)
	Mean	**0.5271**	**0.5748**	**0.5421**	**0.5469**	**0.4478**	**0.5514**	**0.5899**	**0.5886**	**0.5901**	**0.5900**	**0.5901**
	SD	0.1116	0.1723	0.1230	0.1291	0.1155	0.1330	0.2554	0.2543	0.2536	0.2564	0.2561
50	M_1	0.5780	0.8540	0.7320	0.5890	0.3165	0.5975	0.9310 (3.4900)	0.9240 (-461.1807)	0.9365 (3.6075)	0.9320 (3.4400)	0.9320 (-126.5901)
	M_2	0.7030	0.8885	0.8115	0.7215	0.4645	0.7240	0.9370 (4.9800)	0.9375 (-458.7373)	0.9425 (5.1874)	0.9405 (4.9481)	0.9405 (-124.0696)
	M_3	0.4920	0.5485	0.5320	0.5035	0.3660	0.5040	0.5410 (4.9800)	0.5460 (-458.7373)	0.5415 (5.1874)	0.5435 (4.9481)	0.5430 (-124.0696)
	M_4	0.6625	0.8360	0.7550	0.6810	0.4370	0.6830	0.8725 (4.9800)	0.8730 (-458.7373)	0.8745 (5.1874)	0.8750 (4.9481)	0.8750 (-124.0696)
	M_5	0.5675	0.5620	0.5860	0.5740	0.5210	0.5745	0.5330 (7.4700)	0.5325 (-456.2425)	0.5290 (7.7388)	0.5295 (7.4826)	0.5295 (-121.5349)
	M_6	0.7705	0.8685	0.8240	0.7840	0.6520	0.7845	0.8805 (7.4700)	0.8835 (-456.2425)	0.8805 (7.7388)	0.8835 (7.4826)	0.8835 (-121.5349)
	M_7	0.5755	0.5570	0.5830	0.5785	0.5495	0.5785	0.5290 (7.4700)	0.5300 (-456.2425)	0.5235 (7.7388)	0.5265 (7.4826)	0.5265 (-121.5349)
	M_8	0.7080	0.5945	0.6595	0.7005	0.7850	0.6975	0.5510 (9.9600)	0.5470 (-453.6940)	0.5495 (10.261)	0.5470 (10.034)	0.5470 (-118.9852)
	Mean	**0.6321**	**0.7136**	**0.6854**	**0.6415**	**0.5114**	**0.6429**	**0.7219**	**0.7217**	**0.7222**	**0.7222**	**0.7221**
	SD	0.0932	0.1596	0.1113	0.0942	0.1526	0.0937	0.1974	0.1966	0.2007	0.1997	0.1997
100	M_1	0.6140	0.9025	0.7795	0.6220	0.3390	0.6235	0.9175 (3.5000)	0.9170 (-1121.209)	0.9190 (3.5273)	0.9145 (3.2000)	0.9245 (-103.3678)
	M_2	0.6975	0.9460	0.8410	0.7075	0.4575	0.7115	0.9570 (5.0000)	0.9570 (-1118.732)	0.9570 (5.0435)	0.9570 (4.6188)	0.9585 (-100.8143)
	M_3	0.5710	0.6845	0.6510	0.5775	0.4065	0.5775	0.6860 (5.0000)	0.6865 (-1118.732)	0.6845 (5.0435)	0.6905 (4.6188)	0.6855 (-100.8143)
	M_4	0.6945	0.9010	0.8175	0.7025	0.4890	0.7045	0.9105 (5.0000)	0.9105 (-1118.732)	0.9105 (5.0435)	0.9110 (4.6188)	0.9120 (-100.8143)
	M_5	0.6710	0.6730	0.6860	0.6740	0.5935	0.6740	0.6695 (7.5000)	0.6695 (-1116.229)	0.6695 (7.5483)	0.6705 (7.1276)	0.6645 (-98.25840)
	M_6	0.7930	0.9125	0.8570	0.8000	0.6505	0.8005	0.9210 (7.5000)	0.9215 (-1116.229)	0.9210 (7.5483)	0.9235 (7.1276)	0.9235 (-98.25840)
	M_7	0.6730	0.7055	0.7040	0.6755	0.5735	0.6750	0.7000 (7.5000)	0.7005 (-1116.229)	0.6995 (7.5483)	0.7010 (7.1276)	0.6980 (-98.25840)
	M_8	0.7800	0.6825	0.7315	0.7745	0.8370	0.7740	0.6715 (10.000)	0.6690 (-1113.701)	0.6715 (10.042)	0.6670 (9.6968)	0.6670 (-95.70000)
	Mean	**0.6868**	**0.8009**	**0.7584**	**0.6917**	**0.5433**	**0.6926**	**0.5041**	**0.8039**	**0.8041**	**0.8044**	**0.8042**
	SD	0.0749	0.1236	0.0765	0.0731	0.1567	0.0731	0.1319	0.1321	0.1323	0.1317	0.1351

Penalties for NIC1, NIC2, NIC3, NIC4 and NIC5 are $\lambda_1 k_j$, $\lambda_2 \ln(n - k_j)$, $\lambda_1 k_j \ln(n - k_j)$, $\lambda_1 k_j^{\lambda_2}$ and $\lambda_1 k_j + \lambda_2 \ln(n - k_j)$, respectively.

Table 3.10 Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for samples of different sizes and under different criteria for Design 3.6.

Sample size	Model	Average probabilities of correct selection with penalties (given in the parenthesis for new criteria only) under different criteria										
		AIC	BIC	HQ	GCV	RBAR	HOC	NIC1	NIC2	NIC3	NIC4	NIC5
20	M_1	0.9285	0.9795	0.9400	0.9395	0.7685	0.9440	0.8950 (0.8200)	0.8830 (-43.9888)	0.8945 (0.8539)	0.8950 ([illegible].8200)	0.8915 (-13.3855)
	M_2	0.4815	0.4300	0.4760	0.4805	0.5265	0.4745	0.5020 (1.6400)	0.5070 (-43.1994)	0.4980 (1.6764)	0.5020 ([illegible])	0.5045 (-11.5776)
	M_3	0.4670	0.3995	0.4520	0.4545	0.5120	0.4515	0.4865 (1.6400)	0.4910 (-43.1994)	0.4825 (1.6764)	0.4865 ([illegible])	0.4875 (-11.5776)
	M_5	0.1185	0.0625	0.1040	0.0955	0.2330	0.0915	0.1630 (3.4600)	0.1615 (-41.3649)	0.1670 (3.4649)	0.1630 (3.4600)	0.1635 (-10.7561)
	Mean	**0.4989**	**0.4679**	**0.4930**	**0.4925**	**0.5100**	**0.4904**	**0.5116**	**0.5106**	**0.5105**	**0.5116**	**0.5118**
	SD	0.3320	0.3796	0.3431	0.3459	0.2190	0.3496	0.2996	0.2949	0.2980	0.2996	0.2978
50	M_1	0.9430	0.9950	0.9765	0.9465	0.7805	0.9480	0.8625 (0.[illegible]900)	0.8580 (-128.4301)	0.8625 (0.7005)	0.8615 (0.7900)	0.8625 (-40.3941)
	M_2	0.6765	0.5885	0.6475	0.6745	0.6680	0.6735	0.6805 (1.3800)	0.6830 (-127.7496)	0.6790 (1.3936)	0.6760 (1.4742)	0.6815 (-39.7076)
	M_3	0.5525	0.4420	0.5075	0.5505	0.5910	0.5495	0.5855 (1.3800)	0.5890 (-127.7496)	0.5835 (1.3936)	0.5805 (1.4742)	0.5880 (-39.7076)
	M_5	0.3120	0.1730	0.2385	0.2985	0.4525	0.2975	0.3975 (2.0700)	0.3965 (-127.0549)	0.3980 (2.0791)	0.4085 (3.1234)	0.3975 (-39.0166)
	Mean	**0.6210**	**0.5496**	**0.5925**	**0.6175**	**0.6230**	**0.6171**	**0.6315**	**0.6316**	**0.6308**	**0.6316**	**0.6324**
	SD	0.2626	0.3432	0.3071	0.2694	0.1378	0.2704	0.1938	0.1923	0.1936	0.1890	0.1937
100	M_1	0.9355	0.9985	0.9840	0.9355	0.7860	0.9355	0.9040 (0.8200)	0.9040 (-373.2047)	0.9040 (0.8271)	0.9170 (1.0900)	0.9040 (-243.2614)
	M_2	0.7660	0.7000	0.7425	0.7690	0.7345	0.7690	0.7690 (1.6400)	0.7695 (-371.3824)	0.7685 (1.6506)	0.7650 (1.9784)	0.7700 (-243.4433)
	M_3	0.6855	0.5940	0.6495	0.6840	0.6765	0.6835	0.6930 (1.6400)	0.6940 (-371.3824)	0.6925 (1.6506)	0.6865 (1.9784)	0.6940 (-243.4433)
	M_5	0.4845	0.2905	0.3870	0.4795	0.6050	0.4785	0.5255 (3.4600)	0.5240 (-370.5516)	0.5255 (3.4703)	0.5250 (3.8038)	0.5250 (-241.8197)
	Mean	**0.7179**	**0.6458**	**0.6908**	**0.7170**	**0.7005**	**0.7166**	**0.7229**	**0.7229**	**0.7226**	**0.7234**	**0.7233**
	SD	0.1873	0.2923	0.2467	0.1897	0.0778	0.1901	0.1579	0.1585	0.1579	0.1632	0.1581

Penalties for NIC1, NIC2, NIC3, NIC4 and NIC5 are $\lambda_1 k_j$, $\lambda_2 \ln(n - k_j)$, $\lambda_1 k_j \ln(n - k_j)$, $\lambda_1 k_j^{\lambda_2}$ and $\lambda_1 k_j + \lambda_2 \ln(n - k_j)$, respectively.

Table 3.11 Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for samples of different sizes and under different criteria for Design 3.7.

Sample size	Model	Average probabilities of correct selection with penalties (given in the parenthesis for new criteria only) under different criteria										
		AIC	BIC	HQ	GCV	RBAR	HOC	NIC1	NIC2	NIC3	NIC4	NIC5
20	M_1	0.8935	0.9690	0.9110	0.9110	0.6945	0.9165	0.8500 (0.8400)	0.8100 (-40.63326)	0.8520 (0.8833)	0.8620 (1.0800)	0.8300 (-9.98998)
	M_2	0.7300	0.7495	0.7410	0.7460	0.6510	0.7485	0.7105 (1.6800)	0.7105 (-39.88713)	0.7060 (1.7342)	0.7050 (1.9602)	0.7090 (-9.18534)
	M_3	0.6945	0.7095	0.7035	0.7095	0.5860	0.7105	0.6710 (1.6800)	0.6580 (-39.88713)	0.6655 (1.7342)	0.6660 (1.9602)	0.6680 (-9.18534)
	M_4	0.2395	0.1855	0.2300	0.2325	0.2915	0.2290	0.2595 (1.6800)	0.2785 (-39.88713)	0.2560 (1.7342)	0.2530 (1.9602)	0.2685 (-9.18534)
	M_5	0.4425	0.3720	0.4265	0.4225	0.5025	0.4180	0.4670 (3.5200)	0.4735 (-39.09834)	0.4670 (3.5499)	0.4660 (3.7781)	0.4715 (-8.36957)
	M_6	0.1620	0.1025	0.1480	0.1425	0.2395	0.1405	0.1885 (3.5200)	0.1995 (-39.09834)	0.1915 (3.5499)	0.1910 (3.7781)	0.1920 (-8.36957)
	M_7	0.1225	0.0700	0.1065	0.1025	0.2025	0.0985	0.1545 (3.5200)	0.1655 (-39.09834)	0.1595 (3.5499)	0.1585 (3.7781)	0.1615 (-8.36957)
	M_8	0.0810	0.0350	0.0705	0.0600	0.1555	0.0570	0.1040 (3.3600)	0.1055 (-38.26172)	0.1200 (3.3271)	0.1185 (3.5579)	0.1075 (-7.54132)
	Mean	**0.4207**	**0.3991**	**0.4171**	**0.4158**	**0.4154**	**0.4148**	**0.4256**	**0.4251**	**0.4272**	**0.4275**	**0.4260**
	SD	0.3159	0.3624	0.3284	0.3324	0.2169	0.3356	0.2888	0.2747	0.2847	0.2874	0.2816
50	M_1	0.9045	0.9870	0.9550	0.9105	0.6785	0.9125	0.8190 (0.7500)	0.8055 (-134.2678)	0.8255 (0.7784)	0.8540 (1.4400)	0.8180 (-13.9514)
	M_2	0.8410	0.8570	0.8610	0.8510	0.7215	0.8505	0.7935 (1.5000)	0.7910 (-133.5564)	0.7955 (1.5485)	0.7800 (3.2912)	0.7925 (-13.2092)
	M_3	0.7610	0.7575	0.7720	0.7665	0.6665	0.7685	0.7305 (1.5000)	0.7265 (-133.5564)	0.7310 (1.5485)	0.7175 (3.2912)	0.7295 (-13.2092)
	M_4	0.3655	0.2455	0.3115	0.3620	0.4060	0.3590	0.3960 (1.5000)	0.4015 (-133.5564)	0.3915 (1.5485)	0.3735 (3.2912)	0.3975 (-13.2092)
	M_5	0.6455	0.5665	0.6275	0.6450	0.6430	0.6430	0.6530 (3.2500)	0.6550 (-133.8301)	0.6520 (3.3101)	0.6390 (3.0063)	0.6530 (-11.4655)
	M_6	0.3225	0.1920	0.2585	0.3135	0.4045	0.3120	0.3640 (3.2500)	0.3690 (-133.8301)	0.3615 (3.3101)	0.3650 (3.0063)	0.3655 (-11.4655)
	M_7	0.2330	0.1110	0.1765	0.2260	0.3265	0.2260	0.2750 (3.2500)	0.2810 (-133.8301)	0.2725 (3.3101)	0.2765 (3.0063)	0.2785 (-11.4655)
	M_8	0.1905	0.0715	0.1335	0.1805	0.3350	0.1775	0.2500 (3.0000)	0.2515 (-132.0881)	0.2495 (3.0629)	0.2870 (3.6454)	0.2510 (-10.7202)
	Mean	**0.5329**	**0.4735**	**0.5119**	**0.5319**	**0.5227**	**0.5311**	**0.5351**	**0.5351**	**0.5349**	**0.5366**	**0.5357**
	SD	0.2872	0.3634	0.3293	0.2944	0.1691	0.2957	0.2382	0.2330	0.2407	0.2358	0.2367
100	M_1	0.9175	0.9955	0.9770	0.9200	0.7190	0.9215	0.8790 (0.8200)	0.8755 (-367.6096)	0.8790 (0.8271)	0.9120 (1.2800)	0.8770 (-146.5538)
	M_2	0.8695	0.8865	0.8950	0.8740	0.7210	0.8745	0.8415 (1.6400)	0.8415 (-366.7974)	0.8410 (1.6506)	0.8495 (3.2597)	0.8415 (-145.7389)
	M_3	0.8235	0.8430	0.8495	0.8260	0.6980	0.8265	0.8015 (1.6400)	0.8015 (-366.7974)	0.8015 (1.6506)	0.8105 (3.2597)	0.8005 (-145.7389)
	M_4	0.5155	0.3685	0.4520	0.5160	0.5185	0.5140	0.5240 (1.6400)	0.5265 (-366.7974)	0.5235 (1.6506)	0.5080 (3.2597)	0.5265 (-145.7389)
	M_5	0.7345	0.6620	0.7150	0.7335	0.7040	0.7335	0.7360 (3.4600)	0.7380 (-365.9769)	0.7355 (3.4703)	0.7335 (3.1510)	0.7360 (-144.9208)
	M_6	0.4565	0.3055	0.3910	0.4535	0.5200	0.4525	0.4795 (3.4600)	0.4795 (-365.9769)	0.4795 (3.4703)	0.4715 (3.[illegible]510)	0.4805 (-144.9208)
	M_7	0.3775	0.2195	0.3005	0.3740	0.4640	0.3735	0.4105 (3.4600)	0.4110 (-365.9769)	0.4105 (3.4703)	0.3960 ([illegible]510)	0.4130 (-144.9208)
	M_8	0.3135	0.1410	0.2340	0.3060	0.4610	0.3055	0.3610 (3.2800)	0.3600 (-365.1479)	0.3615 (3.2863)	0.3585 ([illegible]93)	0.3610 (-144.0991)
	Mean	**0.6260**	**0.5527**	**0.6018**	**0.6254**	**0.6007**	**0.6252**	**0.6291**	**0.6292**	**0.6290**	**0.6299**	**0.6295**
	SD	0.2377	0.3337	0.2913	0.2412	0.1196	0.2420	0.2076	0.2071	0.2075	0.2202	0.2065

Penalties for NIC1, NIC2, NIC3, NIC4 and NIC5 are $\lambda_1 k_j$, $\lambda_2 \ln(n - k_j)$, $\lambda_1 k_j \ln(n - k_j)$, $\lambda_1 k_j^{\lambda_2}$ and $\lambda_1 k_j + \lambda_2 \ln(n - k_j)$, respectively.

CHAPTER 4

MAXIMISATION OF MEAN AVERAGE PROBABILITY OF CORRECT SELECTION USING ADDITIVE AND MULTIPLICATIVE PENALTIES[1]

4.1 INTRODUCTION

In the previous chapter, we discussed some of the widely used information criteria to select the best model from a set of competing alternative models. But one of the main problems of existing criteria is that their performance varies from data set to data set and none of the existing criteria performs well in all situations. So from a user's point of view, which criteria one should use to select the best model for a particular data set is a question that is unresolved. Also, there is no guarantee that an existing criterion will select the true model with the highest average probability in all situations.

In this chapter we investigate the problem of maximisation of mean average probability of correct selection (MAPCS) of the model using additive and

[1] A paper based on some of the findings reported in this chapter and Chapter 3 has been accepted for publication in *Computer Aided Econometrics*, edited by D.E.A Giles, see King and Bose (2002).

multiplicative penalties. The theory of the use of APCS and the technique used to estimate APCS were discussed in Section 3.2. We used the Simulated Annealing Optimisation (SAO) technique to find penalties that maximise the MAPCS. Here we are not imposing a particular functional form of the penalty; instead we give upper and lower boundaries of the penalty, a set of starting values, and a temperature reduction value. Then, we use the SAO technique (discussed in Section 2.4 of Chapter 2) to find the penalties for the competing models, which maximise the MAPCS of the true model. This chapter is divided into five sections. In Section 4.2, we describe the technique of maximisation of APCS using additive penalties with maximised log-likelihood functions. The maximisation of MAPCS using multiplicative penalties applied to each model's mean squared error is described in Section 4.3. The designs of Monte Carlo experiments are given in Section 4.4. Section 4.5 contains results and discussion, and some concluding remarks are presented in the last section.

4.2 ADDITIVE PENALTY

Suppose we are interested in selecting a model from m alternative models, M_1, M_2, . . ., M_m, for a given data set. Let the model M_j, $j = 1, 2, \ldots, m$, be represented by

$$\mathbf{y} = \mathbf{X}_j\beta_j + \mathbf{u}_j, \tag{4.1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations on the dependent variable, $\mathbf{X}_j$ is an $n \times k_j^*$ matrix, $k_j^* = (k_j - 1)$, β_j is a vector of k_j^* parameters, and $\mathbf{u}_j$ is an $n \times 1$

vector of random disturbances distributed as $N(0, \sigma_j^2 \mathbf{I})$. $\mathbf{X}_j$ contains a column vector of ones in its first column and observations on $(k_j^* - 1)$ non-stochastic variables in the remaining $(k_j^* - 1)$ columns. Let the log-likelihood function for model M_j be $L_j(\beta_j, \sigma_j^2)$ and the maximized value of $L_j(\beta_j, \sigma_j^2)$ be $L_j(\hat{\beta}_j, \hat{\sigma}_j^2)$, where $\hat{\beta}_j$ and $\hat{\sigma}_j^2$ are the maximum likelihood estimates of β_j and σ_j^2, respectively. Then

$$L_j(\beta_j, \sigma_j^2) = -\frac{n}{2}\left[\ln \sigma_j^2 + \ln(2\pi) + \frac{(\mathbf{y} - \mathbf{X}_j\beta_j)'(\mathbf{y} - \mathbf{X}_j\beta_j)}{n\sigma_j^2}\right]. \tag{4.2}$$

The maximum of the log-likelihood, $L_j(\hat{\beta}_j, \hat{\sigma}_j^2)$, can be written as

$$L_j(\hat{\beta}_j, \hat{\sigma}_j^2) = -\frac{n}{2}\left[\ln \hat{\sigma}_j^2 + \ln(2\pi) + 1\right], \tag{4.3}$$

where $\hat{\sigma}_j^2 = \dfrac{(\mathbf{y} - \mathbf{X}_j\hat{\beta}_j)'(\mathbf{y} - \mathbf{X}_j\hat{\beta}_j)}{n}$ is the maximum likelihood estimator of σ_j^2 and $\hat{\beta}_j = (\mathbf{X}_j'\mathbf{X}_j)^{-1}\mathbf{X}_j'\mathbf{y}$ is the maximum likelihood estimator of β_j.

Let p_j denote the penalty for model M_j. In almost all IC based model selection procedures, the model with the largest I_j is selected, where I_j is given by

$$I_j = L_j(\hat{\beta}_j, \hat{\sigma}_j^2) - p_j. \tag{4.4}$$

This I_j is called the penalized maximised log-likelihood. Then the j^{th} model will be selected if

$$I_j > I_i, \quad \forall i = 1,2,\ldots,(j-1),(j+1),\ldots,m. \tag{4.5}$$

For all existing criteria, p_j are a function of n, the sample size, and k_j, the number of parameters in the j^{th} model. The functional form of p_j for different existing criteria was presented in Chapter 3. One of the main disadvantages of using these penalty functions is; that they are independent of the data set, i.e. for a particular model, the penalty remains the same for different data sets. To overcome this problem, we suggest the use of the SAO technique to find the optimum penalty, which maximises the MAPCS for the data set in hand. This approach does not involve a functional form of penalty, instead it requires an arbitrary starting value of the penalty set with appropriate upper and lower bounds of the penalties, and a temperature reduction value. A detailed description and discussion of the advantages of this optimisation technique was given in Section 2.4 of Chapter 2. The theory and the computational techniques for calculating MAPCS were discussed in Section 3.2 of Chapter 3.

4.3 MULTIPLICATIVE PENALTY

Equation (4.3) can be written as a function of the residual sum of squares as follows:

$$L_j(\hat{\beta}_j, \hat{\sigma}_j^2) = -\frac{n}{2}\left[\ln\frac{S_j^2}{n} + \ln(2\pi) + 1\right], \tag{4.6}$$

where $S_j^2 = (\mathbf{y} - \mathbf{X}_j\hat{\beta}_j)'(\mathbf{y} - \mathbf{X}_j\hat{\beta}_j)$ is the residual sum of squares for the j^{th} model.

From this equation it is obvious that the selection of a model using the largest value of (4.6) is equivalent to the selection of a model with the smallest value of S_j^2. But one of the main problems in using the smallest S_j^2 is that the residual sum of squares always decreases or remains constant (i.e., never increases) with the inclusion of additional regressors. So there is a tendency towards selecting models with a needlessly large number of regressors. Hence in order to use residual sum of squares for model selection purposes, we need some adjustment so that models with an unnecessarily large number of parameters will not be favoured. This is achieved by using a penalty function. Rahman (1998) defined the penalised error sum of squares as follows:

$$J_j = S_j^2 q_j, \tag{4.7}$$

where q_j is called the multiplicative penalty and J_j is the penalised error sum of squares. The model with the smallest J_j (multiplicative information criterion for the j^{th} model) will be selected if

$$J_j < J_i, \quad \forall i = 1, 2, \ldots, (j-1), (j+1), \ldots, m. \tag{4.8}$$

The functional form of a possible q_j is also given by Rahman (1998) as

$$q_j = (n - k_j)^{a_1} (a_2)^{\frac{k_j}{n}}, \tag{4.9}$$

where a_1 and a_2 are arbitrary constants.

He mentioned that all existing criteria can be expressed as a special case of this new criterion in linear regression settings and he analytically showed how the widely used criteria AIC, BIC, HQ, GCV, HOC, Mallows' C_p and RBAR, were a special case of this criterion. He also mentioned that by choosing the appropriate values of a_1 and a_2, it is possible to develop an infinite number of new criteria, which will perform well in a range of situations. But a problem with this penalty function is the need to find the values of a_1 and a_2 for a particular data set. Also the penalty function is a function of n, the sample size, and k_j, the number of free parameters, but is independent of data values, i.e. for the same set of competing models, a change of data sets does not have any impact on the penalty function. Another problem with this penalty function is that, for a particular data set, like the existing additive IC, the same dimensional models have the same penalty. To overcome these problems, we redefine the multiplicative information criteria J_j for the j^{th} model as J_j^*, where

$$J_j^* = E_j^2 \, q_j^*, \tag{4.10}$$

in which $E_j^2 = \dfrac{S_j^2}{n-k_j^*}$ is the mean squared error and $q_j^* = (n-k_j)^{a_1-1}(a_2)^{\frac{k_j}{n}}$ is the multiplicative penalty.

Then, instead of searching for the values of a_1 and a_2, we use the SAO technique to find the value of q_j^*, which maximises the MAPCS by minimising J_j^* in equation

(4.10). The advantage of using the mean squared error over the error sum of squares is that we are taking account of the residual degrees of freedom.

4.4 THE DESIGNS OF THE MONTE CARLO STUDIES

In order to evaluate the performance of the additive and multiplicative penalties discussed in Sections 4.2 and 4.3 and compare them to the performance of the selected existing criteria, we conducted some simulation experiments. In these experiments the assumed models are the same as used in Chapter 3. We use the same technique as described in Section 3.3 of the previous chapter for generating the data and computing the APCS. The designs used for the experiments are as follows.

Design 4.1: x_{1t} is the real per capita GDP of the t^{th} country and x_{2t} is GDP as a percent of USA GDP. We used the annual data from Summers and Heston (1991) revised version 5.6 and World Bank world tables. Here we consider four non-nested models M_1, M_2, M_3 and M_5 given by (3.16), (3.17), (3.18) and (3.19) with $s_e^2 = 0.1$ and $s_b^2 = 0.5$.

Design 4.2: x_{1t} is Australian retail trade quarterly data commencing the first quarter of 1959 and x_{2t} is the same series lagged one quarter. We have used the data from the Australian Bureau of Statistics. Here we consider four non-nested models M_1, M_2, M_3 and M_5 given by (3.16), (3.17), (3.18) and (3.19) with $s_e^2 = 55$ and $s_b^2 = 6$.

Design 4.3: x_{1t} and x_{2t} are randomly and independently generated values from the N(0,1) distribution. Here we consider four non-nested models M_1, M_2, M_3 and M_5 given by (3.16), (3.17), (3.18) and (3.19) with $s_e^2 = 0.12$ and $s_b^2 = 0.08$.

Design 4.4: This is an extension of Design 4.1 to more models using one extra variable x_{3t} which is the price level of consumption of the t^{th} country. Here we consider eight non-nested models M_1, M_2, M_3, M_4, M_5, M_6, M_7 and M_8 given by (3.16), (3.17), (3.18), (3.29), (3.19), (3.30), (3.31) and (3.32), respectively, with $s_e^2 = 0.35$ and $s_b^2 = 0.02$.

Design 4.5: This is an extension of Design 4.2. Here we add one more variable x_{3t}, which is x_{1t} (in Design 4.2) lagged two quarters. Eight non-nested models, i.e. M_1, M_2, M_3, M_4, M_5, M_6, M_7 and M_8, given by (3.16), (3.17), (3.18), (3.29), (3.19), (3.30), (3.31) and (3.32), respectively, with $s_e^2 = 0.1$ and $s_b^2 = 0.15$, are considered.

Design 4.6: This is an extension of Design 4.3 with x_{3t} randomly and independently generated from the N(0,1) distribution. Here we consider eight non-nested models M_1, M_2, M_3, M_4, M_5, M_6, M_7 and M_8, given by (3.16), (3.17), (3.18), (3.29), (3.19), (3.30), (3.31) and (3.32), respectively, with $s_e^2 = 0.12$ and $s_b^2 = 0.08$.

The sample sizes used for the Designs 4.1, 4.3, 4.4 and 4.6 are 20, 50 and 100; and for Designs 4.2 and 4.5, sample sizes are 20, 50 and 96.

To apply the SAO technique with maximised log-likelihood functions and additive penalties, we experimented with the use of different sets of starting values, upper and lower boundaries, and temperature reduction values, which are requirements of the SAO technique. It is well known that among the existing criteria, AIC and BIC have wider use in econometrics and physical sciences for model selection. So we used the relative penalties[2] of AIC and BIC as starting values of the penalties for the SAO technique. In addition, we also used zero penalties as the starting values. One of the conditions of the SAO technique is that the starting value of the parameter (here penalties) cannot be outside the boundary values of the parameters. So when selecting the boundaries of the penalties, we had to keep this condition in mind. We considered the maximum number of free parameter (k) among the competing models, the number of competing model (m), km and an arbitrary value of 10 as upper boundaries with starting penalties relative AIC and zeros. For these starting values, we used two types of lower boundaries, namely zeros and the negative of upper boundaries. We used $k\ln(n)/2$, $m\ln(n)/2$, an arbitrary value of 10 and km as upper boundaries with starting penalties being relative BIC. In this case we also have two

[2] Suppose we have four competing alternative models M_1, M_2, M_3 and M_4 with number of free parameters k_1, k_2, k_3 and k_4, respectively. Then the relative penalties of AIC for the models M_1, M_2, M_3 and M_4 are 0, $k_2 - k_1$, $k_3 - k_1$ and $k_4 - k_1$, respectively.

types of lower boundaries, namely zeros and the negative of k, m, 10 and km. When maximising the MAPCS using maximised log-likelihoods with additive penalties, we used the following initial values for Designs 4.1, 4.2 and 4.3.

Three sets of starting values of the penalties were $\{0, 0, 0, 0\}$,
$\{(k_1-1), (k_2-1), (k_3-1), (k_4-1)\}$, and
$\left\{\frac{k_1-1}{2}\ln(n), \frac{k_2-1}{2}\ln(n), \frac{k_3-1}{2}\ln(n), \frac{k_4-1}{2}\ln(n)\right\}$, where k_j, j = 1, 2, 3, 4, is the number of free parameters in the j^{th} model.

We used the eight boundary sets (0, 3), (0, 4) (0, 10), (0, 12), (− 3, 3), (− 4, 4), (− 10, 10), and (− 12, 12) with the first and second starting penalty sets, and the eight boundary sets (0, 3ln(n)/2), (0, 2ln(n)), (0, 10), (0, 12), (− 3, 3ln(n)/2), (− 4, 2ln(n)), (− 10, 10), (− 12, 12) with the third penalty. In all cases, we used four temperature reduction values namely 0.1, 0.01, 0.001, and 0.0001; so, we have 96 combinations of initial sets of values for the SAO technique to estimate APCS.

The starting values and boundary values for Designs 4.4, 4.5 and 4.6 (eight competing models) were also chosen considering the same reasons given for Designs 4.1, 4.2 and 4.3.

The three sets of starting values for additive penalties for Designs 4.4, 4.5 and 4.6 are, {0, 0, 0, 0, 0, 0, 0, 0}, $\{(k_j - 1)\}$ and $\left\{\frac{k_j - 1}{2}\ln(n)\right\}$, where k_j, j = 1, 2, 3, 4, 5, 6, 7, 8, is the number of free parameters in the j^{th} model.

We used the eight boundary sets (0, 4), (0, 8) (0, 10), (0, 32), (− 4, 4), (− 8, 8), (− 10, 10), and (− 32, 32), with the first and second starting penalty sets, and the eight boundary sets (0, 10), (0, 32), (0, 2ln(n)), (0, 4ln(n)/2), (− 10, 10), (− 32, 32), (− 4, 2ln(n)) and (− 8, 4ln(n)) with the third penalty set. In all cases, we used four temperature reduction values, namely 0.1, 0.01, 0.001, and 0.0001; so, we have 96 combinations of initial sets of values for the SAO technique to estimate the APCS.

In the case of multiplicative penalties, we have to multiply the mean squared error by the penalties selected in such a way that the MAPCS is maximised with the minimisation of mean squared error and the penalised value of mean squared error being non-negative. Considering these constraints, we have to select the starting and boundary values of the penalties. Because the penalised mean squared error cannot be negative and the mean squared error is positive, the value of the penalties must be positive. As mentioned earlier, AIC and BIC are widely used IC, so we used AIC and BIC as the starting values for the SAO technique. In addition, we used no penalty as the starting value. Here we use zero as the lower boundary. The upper boundaries are

k, m and some arbitrary value; and a combination of these. These boundaries were selected so that the conditions required by the SAO technique were satisfied.

The three sets of starting values for multiplicative penalties for Designs 4.1, 4.2 and 4.3 are, $\{1, 1, 1, 1\}$, $\{k_1, k_2, k_3, k_4\}$, and $\left\{\frac{k_1}{2}\ln(n), \frac{k_2}{2}\ln(n), \frac{k_3}{2}\ln(n), \frac{k_4}{2}\ln(n)\right\}$, where k_j, $j = 1, 2, 3, 4$, is the number of free parameters in the j^{th} model.

We used the six boundary sets, (0, 1), (0, 3), (0, 10), (0,12), (0, 20) and (0, 36), with the first starting penalty set; the four boundary sets, (0, 10), (0, 12), (0, 27) and (0, 48), with the second starting penalty set; and the four boundary sets (0, 3ln(n)/2), (0, 3ln(n)), (0, 4ln(n)) and (0, 5ln(n)), with the third starting penalty set. For each combination of boundary and penalty sets, we used the same four temperature reduction values used for the additive penalty. Thus we have 56 combinations of initial sets of values.

The three sets of starting values of multiplicative penalties for Designs 4.4, 4.5 and 4.6 are, $\{1, 1, 1, 1, 1, 1, 1, 1\}$, $\{k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8\}$, and

$\left\{\frac{k_1}{2}\ln(n), \frac{k_2}{2}\ln(n), \frac{k_3}{2}\ln(n), \frac{k_4}{2}\ln(n) \frac{k_5}{2}\ln(n), \frac{k_6}{2}\ln(n), \frac{k_7}{2}\ln(n), \frac{k_8}{2}\ln(n)\right\}$, where k_j, j = 1, 2, 3, 4, 5, 6, 7, 8, is the number of free parameters in the j^{th} model.

We used the five boundary sets (0, 1), (0, 4), (0, 8), (0, 20) and (0, 32) with the first penalty set; the five boundary sets (0, 2ln(n)), (0, 4), (0, 8), (0, 20) and (0, 32) with the second penalty set; and the five boundary sets (0, 2ln(n)), (0, 4ln(n)), (0, 8ln(n)), (0, 20) and (0, 32) with the third penalty set. For each combination of boundary and penalty sets, we used the same four temperature reduction values used for the additive penalty. This adds up to 60 combinations of initial sets of values.

4.5 MONTE CARLO RESULTS

The results of the simulation experiments for the SAO technique with additive and multiplicative penalties are presented in Tables 4.1a – c to Tables 4.6a – c. In each table there are three types of penalties: Type 1 – additive, which are penalties of the existing criteria; Type 2 – additive, which use maximised log-likelihood with the SAO technique; and Type 3 – multiplicative, which use mean squared error with the SAO technique. There are eight Type 1 criteria, namely AIC, BIC, GCV, HOC, HQ RBAR, MCP and JIC. The largest MAPCS, smallest variation among the APCS, modal MAPCS and median MAPCS obtained from the SAO technique are used as criteria under Type 2 and Type 3 penalties. A comparative study between Type 1 and Type 2 penalties for all designs under consideration is given in Section 4.5.1. The comparison between Type 1 and Type 3 penalties is presented in the Section 4.5.2 and in Section 4.5.3 we give a comparative study of Type 2 and Type 3 penalties.

4.5.1 COMPARISON OF TYPE 1 AND TYPE 2 PENALTIES

The average probabilities of choosing the true model along with MAPCS and the penalties for Designs 4.1, 4.2, 4.3, 4.4, 4.5 and 4.6 are presented in Tables 4.1a – c, 4.2a – c, 4.3a – c, 4.4a – c, 4.5a – c and 4.6a – c, respectively.

The Monte Carlo experiments indicate that the application of the SAO technique has a great effect on the performance of model selection in terms of APCS. For Designs 4.1, 4.2 and 4.3 and for all sample sizes under consideration, the MAPCS obtained using Type 2 penalties is always greater than that of the largest MAPCS (here that of BIC) among the listed existing IC (Table 4.1a to Table 4.3c). Also the variation among the APCS under Type 2 penalties is smaller than that of cited IC for all cases of Design 4.1 and some cases of Design 4.2 and 4.3. For all designs and sample sizes, in comparison to BIC, there is an increase in APCS for the model with the largest number of regressors (M_5) and a decrease in APCS for the model with the smallest number of regressors (M_1) for the largest MAPCS obtained from the SAO technique; but the picture is reversed in the case of the largest MAPCS obtained from the existing IC (here that of BIC). The mode[3] and median MAPCS obtained from the SAO technique are very close to the largest MAPCS obtained from the SAO technique and in some cases these are identical to the largest MAPCS. For Design

[3]We have 96 MAPCS for 96 combinations of initial parameter values for the SAO technique applied to Type 2 penalties. We computed the mode and median of these 96 MAPCS.

4.1, the MAPCS corresponding to the smallest standard deviation of APCS among the competing models obtained from the SAO technique is very close to the largest MAPCS for all sample sizes. For Designs 4.2 and 4.3, the gap between the largest MAPCS and the APCS corresponding to the smallest standard deviation among the MAPCS obtained from the SAO technique increases as the sample size increases, but the value of standard deviation decreases with an increase in sample size. In general it is observed that among the existing criteria, RBAR selects the true model with the lowest MAPCS and the variation among its APCS is also the lowest. But this lowest variation is generally higher than the corresponding smallest variation among the APCS obtained from the SAO technique with higher MAPCS obtained from this technique compared to RBAR.

The results obtained for Designs 4.4, 4.5 and 4.6 are very similar to those of Designs 4.1, 4.2 and 4.3 even though the number of competing models are double. For all designs and sample sizes, the largest MAPCS obtained using Type 2 penalties are always higher than that of BIC, the largest MAPCS among the existing listed IC with lower variation among the APCS compared to BIC (Table 4.4a to Table 4.6c). A decrease in the APCS for the model with the lowest number of regressors (M_1) and an increase in the APCS for the model with the highest number of regressors (M_8) is observed compared to the corresponding APCS obtained from BIC, for all sample sizes of Design 4.4, n = 20 of Design 4.5, and n = 50 and 100 of Design 4.6. For the remaining sample sizes of Designs 4.5 and 4.6, the APCS for the model with the lowest number of regressors and the model with the highest number of regressors are increasing compared to BIC. For all sample sizes and designs, the mode and median MAPCS obtained from the SAO technique are very close to the largest MAPCS obtained from the SAO technique and some cases these are identical to the largest MAPCS obtained from the SAO technique. An increase in the difference between the largest MAPCS and the MAPCS corresponding to the smallest standard deviation among the APCS is observed with an increase in sample size, but the numerical value of the standard deviation decreases as the sample size increases. Among the existing criteria, RBAR chooses the true model with the lowest MAPCS and in some cases the variation among the APCS is also the lowest. But this lowest variation among the PCS is always higher than the smallest variation among the APCS obtained from the SAO technique with larger MAPCS compared to RBAR.

The number of competing models in Designs 4.1, 4.2 and 4.3 are the same; and we used three different sets of data to test the performance of the application of the SAO technique with additive penalties. The relative penalty for a particular model under the existing criteria remains the same for Designs 4.1, 4.2 and 4.3 as the existing penalties are a function of n, the sample size, and k, the number of free parameters. But for the same set of starting values, boundary values and temperature reduction factors, the relative penalty obtained using the SAO technique for the Designs 4.1, 4.2 and 4.3 are different. For example, for sample size 20, starting value (0, 0, 0, 0), boundary (0,10), and temperature reduction factor 0.1, the relative penalties for models M_1, M_2, M_3 and M_5 that maximise the MAPCS are (0, 7.6894, 0.5024, 8.0203), (0, 1.6146, 1.3136, 3.7060) and (0, 2.1953, 1.8668, 3.5430) for Designs 4.1, 4.2 and 4.3, respectively. It indicates that the penalties do not depend only on n and k, but also the data generating process. It is also notable that although the models M_2 and M_3 have the same number of parameters and the penalties of these models under the existing criteria are the same, while the penalties are different under the SAO technique. Another notable finding from the simulation experiments using Type 2 penalties is that for a particular data set and a set of competing models, exactly the same MAPCS and the same variation among the APCS is obtained from different sets of relative penalties. For example, for Design 4.1 and sample size 50, the largest APCS and SD among the APCS are 0.6290 and 0.1687, respectively. These MAPCS and SD values are obtained from 21 different relative penalty sets, e.g. (0, 8.9402, 1.2482, 9.2081), (0, 11.4229, 1.2439, 11.6918), (0, 17.5325, 1.2462, 17.8005). This result implies that there is no unique set of penalties for a particular data set to maximise the MAPCS. Further it is probable that there may be no unique functional form for the penalties for a particular data set. A similar picture is also observed for Designs 4.4, 4.5 and 4.6.

4.5.2 COMPARISON OF TYPE 1 AND TYPE 3 PENALTIES

Here we compare the MAPCS and variation among the APCS obtained using the Type 1 penalties, which are existing penalties, and the Type 3 penalties (multiplicative), which are those obtained using the SAO technique applied to penalised mean squared error. The results of the experiments are presented in Tables 4.1a – c, 4.2a – c, 4.3a – c, 4.4a – c, 4.5a – c and 4.6a – c.

The simulation results demonstrate that in 100 percent of the cases under consideration for Designs 4.1, 4.2, and 4.3 and for all sample sizes, the MAPCS obtained using the SAO technique with multiplicative penalties (Type 3 penalties) is greater than that of the largest MAPCS among the cited criteria (here BIC). It is also evident from the simulation results that in 100% of the cases for Design 4.1, the variation among the APCS is also lower than the variation for the largest MAPCS among the cited criteria. But in the case of Designs 4.2 and 4.3, the variation among the APCS is comparatively higher than that of the existing criteria. For Design 4.1 with $n = 20$, all sample sizes of Design 4.2 and sample sizes 20 and 100 of Design 4.3, the mode[4] and the median MAPCS coincide with the largest MAPCS obtained from Type 3 penalties. In other cases, the mode and the median MAPCS are very similar to the largest MAPCS obtained using Type 3 penalties. The MAPCS corresponding to the smallest standard deviation among the APCS is very close to the largest MAPCS. The variation among the APCS increases as the sample size increases for Designs 4.2 and 4.3, but the reverse picture is observed for Design 4.1.

[4]We have 56 MAPCS for 56 combinations of initial parameter values for the SAO technique applied to Type 3 penalties. We computed the mode and median of these 56 MAPCS.

The results obtained from Designs 4.4, 4.5 and 4.6 are very similar to the results obtained from Designs 4.1, 4.2 and 4.3 although the number of competing models is double in Designs 4.4, 4.5 and 4.6. In 100 percent of the cases, the MAPCS obtained using Type 3 penalties is greater than the corresponding largest MAPCS obtained from the existing IC, for n = 20 and 50 of Design 4.4, for all sample sizes of Design 4.5, and n = 20 of Design 4.6. For the remaining designs and sample sizes, in more than 95 percent of the cases, the MAPCS obtained from Type 3 penalties is higher than the corresponding largest MAPCS obtained from the existing IC. For all sample sizes of Design 4.4, and n = 20 of Design 4.5, an increase of APCS in the model with the largest number of regressors (M_8) and a decrease of APCS in the model with the smallest number of regressors (M_1) is observed, compared to the largest MAPCS from the existing IC. The reverse picture is observed for the remaining sample sizes of Design 4.5 and all sample sizes of Design 4.6. For all designs and sample sizes, the mode[5] and median MAPCS are very close to the largest MAPCS obtained from Type 3 penalties and in some of the cases these are identical. The MAPCS corresponding to the smallest variation among the APCS is very close to the largest MAPCS obtained using Type 3 penalties.

As for additive penalties, in the case of multiplicative penalties we find, for the same set of starting values, boundary and temperature reduction factors, the relative

[5]We have 60 MAPCS for 60 combinations of initial parameter values for the SAO technique applied to Type 3 penalties. We computed the mode and median of these 60 MAPCS.

penalties obtained using the SAO technique for Designs 4.1, 4.2 and 4.3 are different. For example, for n = 20, starting penalties (1, 1, 1, 1), boundary (0, 3), and temperature reduction factor 0.1, the relative penalties for models M_1, M_2, M_3 and M_5 that maximise the MAPCS are (1, 3.1681, 0.9959, 0.30906), (1, 1.1127, 1.0804, 1.2957) and (1, 1.1806, 1.1412, 1.2887) for Designs 4.1, 4.2 and 4.3, respectively. Like Type 2 penalties, for the same dimensional model, the penalties are different under the SAO technique although the penalties of these models under the existing criteria are the same. Another remarkable finding from the simulation experiments is that for a particular data set and set of competing models, exactly the same MAPCS and same variation among the APCS is obtained from different sets of relative penalties for the SAO technique. For example, for Design 4.1 and sample size 100, the APCS 0.7331 and SD 0.2021 is obtained from 12 combinations of initial parameter values for the SAO technique. But the relative penalties that maximise this MAPCS, are different e.g. (1.000, 3.5512, 1.0257, 3.5727), and (1.000, 1.6630, 1.0257, 1.6731). It confirms the finding concerning Type 2 penalties, that there is no unique set of penalties that maximise the MAPCS. A similar picture is also observed for Designs 4.4, 4.5 and 4.6.

4.5.3 COMPARISON OF TYPE 2 AND TYPE 3 PENALTIES

The Monte Carlo experiments indicate that the results obtained from additive penalties used with maximised log-likelihood functions (Type 2) and from multiplicative penalties used with mean squared error (Type 3), are very similar for all designs and sample sizes under consideration. It implies that there is no significant effect of the form of penalties on the MAPCS of the model. So, from the user's point of view, one can use either of the penalties with the SAO technique to select the true model from a set of competing alternative models in linear regression settings.

4.6 CONCLUDING REMARKS

In this chapter, we introduced the application of the SAO technique to model selection. We used two types of penalty with the SAO technique to maximise the MAPCS, namely additive penalties applied to maximised log-likelihood functions and multiplicative penalties applied to mean squared error. The employment of this technique has a great effect on the performance of model selection procedures with respect to the selection of the true model in terms of MAPCS. Simulation results show that the MAPCS obtained using the SAO technique is always higher than that of the existing criteria for all designs and sample sizes. Also, the variation among APCS obtained from the SAO technique decreases or remains very close to that of existing criteria. In some experimental designs, an exceptional improvement in the APCS for the model with the highest number of regressors is observed. In these cases, a drastic reduction of the variation among the APCS is also observed.

The simulation results demonstrate that the optimal penalty for a particular model does not depend only on the sample size and the number of free parameters, but also on the data generating process. The results for the SAO technique indicate that exactly the same MAPCS and variation among the APCS for a particular data set and a set of competing models can be obtained with different sets of penalties. It means that there is no unique set of penalties which maximise the MAPCS. We found from our simulation results that for the same dimensional model, the penalties obtained using the SAO technique vary from data set to data set although they are the same in the case of existing criteria. Billah and King (1998) conducted simulation experiments for linear regression models with white noise, AR(1), AR(2) and MA(1) disturbances to estimate the penalties which maximise the MAPCS of the model. It is also apparent from their simulation results that although the AR(1) and MA(1) disturbance models have the same number of parameters, the penalties that maximise the MAPCS are different. Thus, in the context of selecting the best model from a set of models with equal number of parameters, the idea of penalising their corresponding maximised log-likelihood or mean squared error functions is inevitable in order to maximise the MAPCS. In the next chapter, we will discuss this issue of model selection.

Our simulation results indicate that the application of the multiplicative penalty to mean squared error and the additive penalty to the maximised log-likelihood function have very similar effects on the selection of the true model among a set of competing alternative models. So from the user's point of view, one can use either of these techniques for model selection purposes.

In this chapter, it has been asserted that the application of the SAO technique to additive penalties applied to maximised log-likelihoods and multiplicative penalties applied to mean squared error, can be used to select the true model with higher APCS compared to existing criteria. It would seem that the numerical evidence from our simulation experiments is a good reason to recommend the use of the SAO technique for model selection purposes with either additive penalties applied to maximised log-likelihoods or multiplicative penalties applied to mean squared error.

Table 4.1a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.1 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.6800	0.8060	0.1580	0.2220	0.4665	0.3244	1.0000	1.0000	2.0000							
	BIC	0.8225	0.8815	0.1150	0.1345	0.4884	0.4206	1.4979	1.4979	2.9957							
	GCV	0.7125	0.8305	0.1495	0.1865	0.4698	0.3521	1.0813	1.0813	2.2245							
	HOC	0.7195	0.8395	0.1450	0.1820	0.4715	0.3593	1.1123	1.1123	2.2901							
	HQ	0.7135	0.8205	0.1455	0.1960	0.4689	0.3476	1.0972	1.0972	2.1944							
	RBAR	0.4710	0.6790	0.2195	0.3695	0.4347	0.1928	0.5407	0.5407	1.1123							
	MCP	0.6810	0.8080	0.1565	0.2155	0.4655	0.3272	1.0059	1.0059	2.0220							
	JIC	0.6850	0.8105	0.1550	0.2125	0.4657	0.3305	1.0193	1.0193	2.0540							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.6245	0.5435	0.4550	0.5080	0.5328	0.0712	13.0759	0.4639	13.4040	0	1.4979	1.4979	2.9957	-12	12.0000	0.0100
	Smallest SD among APCS	0.5535	0.5440	0.4845	0.5080	0.5225	0.0320	2.6281	0.3685	2.9563	0	1.0000	1.0000	2.0000	0	3.0000	0.1000
	Modal MAPCS	0.6255	0.5435	0.4535	0.5080	0.5326	0.0721	8.2473	0.4660	8.5757	0	1.4979	1.4979	2.9957	-10	10.0000	0.0001
	Median MAPCS	0.6490	0.5460	0.4265	0.5055	0.5318	0.0926	5.1629	0.5032	5.4936	0	1.4979	1.4979	2.9957	0	5.9915	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.6485	0.5400	0.4325	0.5120	0.5333	0.0893	3.1681	0.9959	3.0906	1	1.0000	1.0000	1.0000	0	3.0000	0.1000
	Smallest SD among APCS	0.5845	0.5460	0.4940	0.5055	0.5325	0.0412	2.2897	0.9852	2.2352	1	1.0000	1.0000	1.0000	0	1.0000	0.1000
	Modal MAPCS	0.6485	0.5400	0.4325	0.5120	0.5333	0.0893	3.1681	0.9959	3.0906	1	1.0000	1.0000	1.0000	0	3.0000	0.1000
	Median MAPCS	0.6480	0.5400	0.4320	0.5120	0.5330	0.0893	1.7426	0.9959	1.6999	1	2.0000	2.0000	3.0000	0	12.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.1b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.1 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.7560	0.8230	0.3165	0.2080	0.5259	0.3088	1.0000	1.0000	2.0000							
	BIC	0.9260	0.9450	0.2840	0.0800	0.5587	0.4430	1.9560	1.9560	3.9120							
	GCV	0.7630	0.8320	0.3155	0.2010	0.5279	0.3161	1.0310	1.0310	2.0836							
	HOC	0.7670	0.8340	0.3150	0.1990	0.5287	0.3185	1.0418	1.0418	2.1058							
	HQ	0.8450	0.8865	0.3085	0.1430	0.5457	0.3760	1.3641	1.3641	2.7281							
	RBAR	0.5145	0.6660	0.3070	0.3660	0.4634	0.1508	0.5155	0.5155	1.0418							
	MCP	0.7560	0.8240	0.3165	0.2075	0.5260	0.3093	1.0009	1.0009	2.0035							
	JIC	0.8175	0.8650	0.3150	0.1630	0.5401	0.3537	1.2357	1.2357	2.4769							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8770	0.5170	0.5935	0.5285	0.6290	0.1687	8.9402	1.2482	9.2081	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
	Smallest SD among APCS	0.7765	0.5180	0.6275	0.5275	0.6124	0.1201	2.6381	0.8357	2.9068	0.0000	0.0000	0.0000	0.0000	0	3.0000	0.0100
	Modal MAPCS	0.8770	0.5170	0.5935	0.5285	0.6290	0.1687	8.9402	1.2482	9.2081	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
	Median MAPCS	0.8745	0.5170	0.5920	0.5285	0.6280	0.1676	5.4455	1.2475	5.7132	0.0000	0.0000	0.0000	0.0000	0	10.0000	0.0100
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8770	0.5290	0.5930	0.5145	0.6284	0.1692	1.6936	1.0293	1.6774	1.0000	1.0000	1.0000	1.0000	0	1.0000	0.1000
	Smallest SD among APCS	0.8730	0.5290	0.5910	0.5145	0.6269	0.1674	1.1923	1.0294	1.1809	1.9560	2.9120	3.9120	5.8680	0	19.5601	0.0010
	Modal MAPCS	0.8770	0.5290	0.5930	0.5145	0.6284	0.1692	1.6936	1.0293	1.6774	1.0000	1.0000	1.0000	1.0000	0	1.0000	0.1000
	Median MAPCS	0.8770	0.5285	0.5930	0.5140	0.6281	0.1694	2.9855	1.0293	2.9570	1.0000	1.0000	1.0000	1.0000	0	36.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.1c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 4.1 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.7615	0.8450	0.5215	0.4040	0.6330	0.2052	1.0000	1.0000	2.0000							
	BIC	0.9495	0.9580	0.5190	0.2095	0.6590	0.3631	2.3026	2.3026	4.6052							
	GCV	0.7645	0.8520	0.5225	0.3970	0.6340	0.2107	1.0152	1.0152	2.0409							
	HOC	0.7665	0.8520	0.5220	0.3970	0.6344	0.2112	1.0204	1.0204	2.0514							
	HQ	0.8745	0.9170	0.5305	0.2990	0.6552	0.2939	1.5272	1.5272	3.0544							
	RBAR	0.5380	0.6935	0.4825	0.5435	0.5644	0.0904	0.5076	0.5076	1.0204							
	MCP	0.7615	0.8450	0.5215	0.4035	0.6329	0.2054	1.0002	1.0002	2.0009							
	JIC	0.8585	0.9065	0.5270	0.3210	0.6533	0.2784	1.4051	1.4051	2.8128							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9340	0.7990	0.7380	0.4635	0.7336	0.1978	15.2409	1.7396	16.0419	0.0000	0.0000	0.0000	0.0000	-12	12.0000	0.1000
	Smallest SD among APCS	0.8450	0.7530	0.7405	0.5040	0.7106	0.1454	2.3308	1.1201	2.9778	0.0000	0.0000	0.0000	0.0000	0	3.0000	0.0010
	Modal MAPCS	0.9340	0.7990	0.7375	0.4635	0.7335	0.1978	7.4871	1.7373	8.2880	0.0000	0.0000	0.0000	0.0000	0	10.0000	0.1000
	Median MAPCS	0.9340	0.7990	0.7355	0.4635	0.7330	0.1978	6.1524	1.7391	6.9534	0.0000	2.3026	2.3026	4.6052	-3	6.9078	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9365	0.8040	0.7355	0.4570	0.7332	0.2022	1.2291	1.0257	1.2366	2.3026	4.6052	4.6052	6.9078	0	23.0259	0.1000
	Smallest SD among APCS	0.9365	0.7850	0.7355	0.4745	0.7329	0.1923	3.7588	1.0257	3.7764	2.3026	4.6052	4.6052	6.9078	0	13.8155	0.0010
	Modal MAPCS	0.9365	0.8035	0.7355	0.4570	0.7331	0.2021	1.6630	1.0257	1.6731	2.3026	4.6052	4.6052	6.9078	0	6.9078	0.0001
	Median MAPCS	0.9365	0.7855	0.7355	0.4745	0.7330	0.1924	1.7747	1.0257	1.7830	2.3026	4.6052	4.6052	6.9078	0	6.9078	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.2a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.2 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.6690	0.6545	0.3850	0.2605	0.4923	0.2023	1.0000	1 0000	2.0000							
	BIC	0.7970	0.6850	0.3485	0.1700	0.5001	0 2911	1.4979	1 4979	2.9957							
	GCV	0.6985	0.6710	0.3900	0.2295	0.4973	0.2265	1.0813	1.0813	2.2245							
	HOC	0.7095	0.6745	0.3885	0.2245	0.4992	0.2329	1.1123	1.1123	2.2901							
	HQ	0.7035	0.6635	0.3835	0.2370	0.4969	0.2242	1.0972	1.0972	2.1944							
	RBAR	0.4505	0.5775	0.4190	0.3785	0.4564	0.0860	0.5407	0.5407	[illegible]							
	MCP	0.6715	0.6565	0.3850	0.2545	0 4919	0.2059	1.0059	1 0059	[illegible]							
	JIC	0.6765	0.6580	0.3870	0.2490	0.4926	0 2095	1 0193	1.0193	2.0540							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8015	0.7255	0 3975	0 1095	0.5085	0.3186	1.6079	1.3140	3.6832	0.0000	0.0000	0.0000	0.0000	-12	12.0000	0.1000
	Smallest SD among APCS	0.7940	0.6075	0 3635	0.2465	0 5029	0 2455	1 5974	1 3141	2 5162	0.0000	1.0000	1.0000	2.0000	0.	3.0000	0.0001
	Modal MAPCS	0.8015	0.7255	0.3975	0.1095	0.5085	0.3186	1.6079	1.3140	3.6832	0.0000	0.0000	0 0000	0 0000	-12	12 0000	0.1000
	Median MAPCS	0.8015	0.7255	0.3975	0 1095	0.5085	0.3186	1.6079	1 3140	3 6832	0.0000	0.0000	0 0000	0.0000	-12	12.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8015	0.7255	0.3975	0.1095	0.5085	0.3186	1.1125	1.0804	1.2930	1.4979	2.9957	2 9957	4 4936	0	14.9787	0 1000
	Smallest SD among APCS	0.8015	0.7255	0.3975	0.1095	0 5085	0.3186	1.1125	1.0804	1.2930	1.4979	2 9957	2.9957	4 4936	0	14 9787	0 1000
	Modal MAPCS	0.8015	0.7255	0.3975	0.1095	0.5085	0.3186	1.1125	1.0804	1.2930	1.4979	2.9957	2.9957	4.4936	0	14.9787	0.1000
	Median MAPCS	0.8015	0.7255	0.3975	0 1095	0.5085	0.3186	1.1125	1.0804	1 2930	1.4979	2.9957	2.9957	4 4936	0	14.9787	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.2b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.2 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.7190	0.7780	0.6655	0.5925	0.6887	0.0789	1.0000	1.0000	2.0000							
	BIC	0.9095	0.8830	0.7100	0.4530	0.7389	0.2101	1.9560	1.9560	3.9120							
	GCV	0.7290	0.7920	0.6715	0.5840	0.6941	0.0884	1.0310	1.0310	2.0836							
	HOC	0.7325	0.7940	0.6720	0.5815	0.6950	0.0906	1.0418	1.0418	2.1058							
	HQ	0.8125	0.8335	0.6925	0.5315	0.7175	0.1387	1.3641	1.3641	2.7281							
	RBAR	0.4900	0.6390	0.5800	0.6915	0.6001	0.0864	0.5155	0.5155	1.0418							
	MCP	0.7190	0.7785	0.6660	0.5915	0.6887	0.0795	1.0009	1.0009	2.0035							
	JIC	0.7830	0.8230	0.6870	0.5525	0.7114	0.1203	1.2357	1.2357	2.4769							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	3.3296	2.8476	5.2849	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
	Smallest SD among APCS	0.8690	0.8165	0.6840	0.5460	0.7289	0.1446	1.7732	1.6949	2.9996	0.0000	0.0000	0.0000	0.0000	0	3.0000	0.0010
	Modal MAPCS	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	3.3296	2.8476	5.2849	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
	Median MAPCS	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	3.3296	2.8476	5.2849	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	1.1194	1.0978	1.1852	1.0000	1.0000	1.0000	1.0000	0	3.0000	0.1000
	Smallest SD among APCS	0.9635	0.8740	0.6785	0.4600	0.7440	0.2236	1.0922	1.0988	1.1551	1.0000	1.0000	1.0000	1.0000	0	12.0000	0.0001
	Modal MAPCS	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	1.1194	1.0978	1.1852	1.0000	1.0000	1.0000	1.0000	0	3.0000	0.1000
	Median MAPCS	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	1.1194	1.0978	1.1852	1.0000	1.0000	1.0000	1.0000	0	3.0000	0.1000

* Additive penalty for model $M1$ is zero and multiplicative penalty for model $M1$ is one

** TRF Temperature reduction factor

Table 4.2c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 96 for Design 4.2 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.7415	0.8155	0.8035	0.8830	0.8109	0.0580	1.0000	1.0000	2.0000							
	BIC	0.9390	0.9445	0.9160	0.8255	0.9063	0.0552	2.2822	2.2822	4.5643							
	GCV	0.7455	0.8195	0.8080	0.8815	0.8136	0.0557	1.0159	1.0159	2.0426							
	HOC	0.7480	0.8205	0.8085	0.8815	0.8146	0.0547	1.0213	1.0213	2.0536							
	HQ	0.8665	0.8905	0.8740	0.8595	0.8726	0.0133	1.5183	1.5183	3.0366							
	RBAR	0.5290	0.6600	0.6610	0.9155	0.6914	0.1618	0.5079	0.5079	1.0213							
	MCP	0.7415	0.8155	0.8035	0.8830	0.8109	0.0580	1.0003	1.0003	2.0009							
	JIC	0.8375	0.8745	0.8595	0.8625	0.8585	0.0154	1.3951	1.3951	2.7928							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	4.7801	4.1070	7.6732	0.0000	0.0000	0.0000	0.0000	-4	4.0000	0.1000
	Smallest SD among APCS	0.8650	0.8805	0.8820	0.8600	0.8719	0.0110	1.5598	1.4537	2.9997	0.0000	0.0000	0.0000	0.0000	0	3.0000	0.1000
	Modal MAPCS	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	4.7801	4.1070	7.6732	0.0000	0.0000	0.0000	0.0000	-4	4.0000	0.1000
	Median MAPCS	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	4.7801	4.1070	7.6732	0.0000	0.0000	0.0000	0.0000	-4	4.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	1.0931	1.0778	1.1486	1.0000	2.0000	2.0000	3.0000	0	48.0000	0.0100
	Smallest SD among APCS	0.9970	0.9395	0.9480	0.8070	0.9229	0.0813	1.1118	1.0960	1.1569	1.0000	2.0000	2.0000	3.0000	0	48.0000	0.0001
	Modal MAPCS	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	1.0931	1.0778	1.1486	1.0000	2.0000	2.0000	3.0000	0	48.0000	0.0100
	Median MAPCS	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	1.0931	1.0778	1.1486	1.0000	2.0000	2.0000	3.0000	0	48.0000	0.0100

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.3a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.3 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.6355	0.5995	0.5725	0.5085	0.5790	0.0536	1.0000	1.0000	2.0000							
	BIC	0.7815	0.6210	0.5750	0.4205	0.5995	0.1486	1.4979	1.4979	2.9957							
	GCV	0.6650	0.6155	0.5815	0.4810	0.5857	0.0778	1.0813	1.0813	2.2245							
	HOC	6755	0.6165	0.5825	0.4740	0.5871	0.0846	1.1123	1.1123	2.2901							
	HQ	0.6695	0.6090	0.5735	0.4900	0.5855	0.0750	1.0972	1.0972	2.1944							
	RBAR	0.4305	0.5365	0.5320	0.6100	0.5272	0.0737	0.5407	0.5407	1.1123							
	MCP	0.6400	0.6025	0.5750	0.5075	0.5813	0.0559	1.0059	1.0059	2.0220							
	JIC	0.6450	0.6065	0.5750	0.5050	0.5829	0.0593	1.0193	1.0193	2.0540							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	2.1992	1.8608	3.6479	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
	Smallest SD among APCS	0.8000	0.6100	0.5330	0.4590	0.6005	0.1466	1.5052	1.6540	2.8529	0.0000	0.0000	0.0000	0.0000	0	3.0000	0.0100
	Modal MAPCS	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	2.1992	1.8608	3.6479	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
	Median MAPCS	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	2.1992	1.8608	3.6479	0.0000	0.0000	0.0000	0.0000	0	12.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	1.1806	1.1412	1.2887	1.0000	1.0000	1.0000	1.0000	0	10.0000	0.1000
	Smallest SD among APCS	0.8315	0.5915	0.5590	0.4330	0.6038	0.1665	1.1389	1.1221	1.2321	1.0000	2.0000	2.0000	3.0000	0	48.0000	0.0010
	Modal MAPCS	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	1.1806	1.1412	1.2887	1.0000	1.0000	1.0000	1.0000	0	10.0000	0.1000
	Median MAPCS	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	1.1806	1.1412	1.2887	1.0000	1.0000	1.0000	1.0000	0	10.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.3b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.3 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.7050	0.7140	0.6610	0.6610	0.6853	0.0282	1.0000	1.0000	2.0000							
	BIC	0.8985	0.7660	0.6980	0.5605	0.7308	0.1408	1.9560	1.9560	3.9120							
	GCV	0.7140	0.7235	0.6720	0.6565	0.6915	0.0323	1.0310	1.0310	2.0836							
	HOC	0.7165	0.7265	0.6740	0.6545	0.6929	0.0342	1.0418	1.0418	2.1058							
	HQ	0.7945	0.7460	0.6875	0.6140	0.7105	0.0778	1.3641	1.3641	2.7281							
	RBAR	0.4735	0.6145	0.5735	0.7455	0.6018	0.1127	0.5155	0.5155	1.0418							
	MCP	0.7050	0.7145	0.6615	0.6605	0.6854	0.0284	1.0009	1.0009	2.0035							
	JIC	0.7670	0.7395	0.6880	0.6290	0.7059	0.0608	1.2357	1.2357	2.4769							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9060	0.7665	0.7035	0.5560	0.7330	0.1452	2.0229	1.9573	4.0287	0.0000	0.0000	0.0000	0.0000	0	10.0000	0.1000
	Smallest SD among APCS	0.8005	0.7780	0.7010	0.5930	0.7181	0.0937	1.3051	1.4416	2.9445	0.0000	0.0000	0.0000	0.0000	0	3.0000	0.1000
	Modal MAPCS	0.9060	0.7665	0.7035	0.5560	0.7330	0.1452	2.0229	1.9573	4.0287	0.0000	0.0000	0.0000	0.0000	0	10.0000	0.1000
	Median MAPCS	0.9060	0.7665	0.7035	0.5560	0.7330	0.1452	2.0229	1.9573	4.0287	0.0000	0.0000	0.0000	0.0000	0	10.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9065	0.7665	0.7035	0.5560	0.7331	0.1454	1.0624	1.0593	1.1271	1.0000	2.0000	2.0000	3.0000	0	10.0000	0.1000
	Smallest SD among APCS	0.9015	0.7630	0.7075	0.5590	0.7328	0.1417	1.0624	1.0568	1.1244	1.0000	1.0000	1.0000	1.0000	0	36.0000	0.0001
	Modal MAPCS	0.9070	0.7655	0.7035	0.5560	0.7330	0.1455	1.0626	1.0593	1.1272	1.0000	2.0000	2.0000	3.0000	0	12.0000	0.0010
	Median MAPCS	0.9070	0.7655	0.7035	0.5560	0.7330	0.1455	1.0626	1.0593	1.1272	1.0000	2.0000	2.0000	3.0000	0	12.0000	0.0010

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.3c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 4.3 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Additive: Existing criteria (Type 1)	AIC	0.7025	0.7665	0.7125	0.7515	0.7332	0.0306	1.0000	1.0000	2.0000							
	BIC	0.9315	0.8325	0.7580	0.6375	0.7899	0.1240	2.3026	2.3026	4.6052							
	GCV	0.7090	0.7710	0.7170	0.7475	0.7361	0.0286	1.0152	1.0152	2.0409							
	HOC	0.7090	0.7710	0.7175	0.7470	0.7361	0.0284	1.0204	1.0204	2.0514							
	HQ	0.8450	0.8095	0.7605	0.6980	0.7782	0.0637	1.5272	1.5272	3.0544							
	RBAR	0.4575	0.6440	0.6055	0.8275	0.6336	0.1522	0.5076	0.5076	1.0204							
	MCP	0.7025	0.7665	0.7125	0.7515	0.7332	0.0306	1.0002	1.0002	2.0009							
	JIC	0.8215	0.8010	0.7550	0.7090	0.7716	0.0502	1.4051	1.4051	2.8128							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8905	0.8330	0.8165	0.6305	0.7926	0.1126	2.1758	1.5387	4.2676	0.0000	0.0000	0.0000	0.0000	0	12.0000	.01000
	Smallest SD among APCS	0.8485	0.7945	0.7700	0.7065	0.7799	0.0589	1.6293	1.4306	2.9826	0.0000	1.0000	1.0000	2.0000	0	3.0000	0.0010
	Modal MAPCS	0.8905	0.8330	0.8160	0.6305	0.7925	0.1126	2.1811	1.5461	4.2745	0.0000	1.0000	1.0000	2.0000	-3	3.0000	0.1000
	Median MAPCS	0.8905	0.8330	0.8160	0.6305	0.7925	0.1126	2.1811	1.5461	4.2745	0.0000	1.0000	1.0000	2.0000	-3	3.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8950	0.8320	0.8175	0.6270	0.7929	0.1156	1.0350	1.0213	1.0692	1.0000	1.0000	1.0000	1.0000	0	10.0000	0.1000
	Smallest SD among APCS	0.8895	0.8160	0.8235	0.6400	0.7923	0.1067	1.0366	1.0193	1.0655	1.0000	1.0000	1.0000	1.0000	0	12.0000	0.0001
	Modal MAPCS	0.8950	0.8320	0.8175	0.6270	0.7929	0.1156	1.0350	1.0213	1.0692	1.0000	1.0000	1.0000	1.0000	0	10.0000	0.1000
	Median MAPCS	0.8950	0.8320	0.3175	0.6270	0.7929	0.1156	1.0350	1.0213	1.0692	1.0000	1.0000	1.0000	1.0000	0	10.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 4.4a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.4 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.5445	0.6280	0.2395	0.5015	0.2985	0.6275	0.2335	0.3330	0.4257	0.1681	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.7240	0.7735	0.2325	0.5660	0.2300	0.6555	0.1855	0.2420	0.4511	0.2519	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936
	GCV	0.5865	0.6825	0.2485	0.5325	0.2765	0.6440	0.2250	0.2915	0.4359	0.1935	1.0813	1.0813	1.0813	2.2245	2.2245	2.2245	3.4370
	HOC	0.6000	0.6910	0.2515	0.5390	0.2780	0.6485	0.2210	0.2860	0.4394	0.1983	1.1123	1.1123	1.1123	2.2901	2.2901	2.2901	3.5417
	HQ	0.5885	0.6630	0.2400	0.5205	0.2825	0.6295	0.2265	0.3140	0.4331	0.1852	1.0972	1.0972	1.0972	2.1944	2.1944	2.1944	3.2916
	RBAR	0.3175	0.4295	0.2200	0.3775	0.3450	0.5635	0.2735	0.4380	0.3706	0.1075	0.5407	0.5407	0.5407	1.1123	1.1123	1.1123	1.7185
	MCP	0.5465	0.6340	0.2415	0.5055	0.2955	0.6290	0.2320	0.3280	0.4265	0.1706	1.0059	1.0059	1.0059	2.0220	2.0220	2.0220	3.0538
	JIC	0.5550	0.6410	0.2425	0.5080	0.2930	0.6305	0.2310	0.3230	0.4280	0.1738	1.0193	1.0193	1.0193	2.0540	2.0540	2.0540	3.1061
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.6815	0.7115	0.4390	0.5095	0.3670	0.5375	0.3600	0.3455	0.4939	0.1433	4.7017	0.7324	1.5757	5.5869	6.5864	2.2801	7.4196
	Smallest SD among APCS	0.5190	0.6290	0.4905	0.5150	0.4040	0.5270	0.3115	0.3915	0.4734	0.0992	1.8258	0.4346	1.2583	2.5177	3.3279	2.0440	3.9924
	Modal MAPCS	0.6830	0.7120	0.4415	0.5045	0.3675	0.5375	0.3550	0.3455	0.4933	0.1441	4.1549	0.7339	1.6192	5.0393	6.0374	2.3214	6.8702
	Median MAPCS	0.6770	0.7015	0.4350	0.5070	0.3625	0.5490	0.3570	0.3490	0.4922	0.1416	3.8050	0.7321	1.5763	4.6895	5.5458	2.2803	6.3889
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.6855	0.7110	0.4445	0.5230	0.3670	0.5375	0.3375	0.3455	0.4939	0.1472	1.5420	1.0195	1.1138	1.5910	1.7583	1.1374	1.7988
	Smallest SD among APCS	0.6835	0.6650	0.4420	0.5050	0.4235	0.5310	0.3555	0.3395	0.4931	0.1295	1.4543	1.0194	1.1142	1.4755	1.6746	1.1288	1.7131
	Modal MAPCS	0.6865	0.7575	0.4475	0.5230	0.3660	0.5285	0.3370	0.3515	0.4934	0.1455	1.6716	1.0193	1.1151	1.7249	1.9060	1.1389	1.9438
	Median MAPCS	0.7825	0.6580	0.3725	0.4885	0.4405	0.5095	0.3485	0.3415	0.4927	0.1570	1.7006	1.0510	1.1533	1.7209	2.0067	1.1686	2.0464

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.4a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.4 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	8.0000	0.1000
	Smallest SD among APCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0010
	Modal MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	8.0000	0.1000
	Median MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.0001
	Smallest SD among APCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.0010
	Modal MAPCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	20.0000	0.1000
	Median MAPCS	1.4979	2.9957	2.9957	2.9957	4.4936	4.4936	4.4936	5.9915	0.0000	32.0000	0.1000

** TRF Temperature reduction factor

Table 4.4b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.4 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.5775	0.6915	0.3595	0.6210	0.3875	0.7475	0.3975	0.4700	0.5315	0.1485	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.8535	0.8860	0.3410	0.7820	0.3020	0.8140	0.3325	0.3125	0.5779	0.2755	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680
	GCV	0.5910	0.7100	0.3660	0.6340	0.3815	0.7605	0.3955	0.4595	0.5373	0.1566	1.0310	1.0310	1.0310	2.0836	2.0836	2.0836	3.1589
	HOC	0.5965	0.7165	0.3675	0.6365	0.3790	0.7600	0.3950	0.4565	0.5384	0.1584	1.0418	1.0418	1.0418	2.1058	2.1058	2.1058	3.1929
	HQ	0.7120	0.7950	0.3710	0.7015	0.3505	0.7900	0.3725	0.4045	0.5621	0.2036	1.3641	1.3641	1.3641	2.7281	2.7281	2.7281	4.0922
	RBAR	0.3215	0.4665	0.2890	0.4230	0.3965	0.6115	0.4105	0.5760	0.4368	0.1124	0.5155	0.5155	0.5155	1.0418	1.0418	1.0418	1.5795
	MCP	0.5795	0.6925	0.3595	0.6210	0.3875	0.7485	0.3970	0.4690	0.5318	0.1491	1.0009	1.0009	1.0009	2.0035	2.0035	2.0035	3.0084
	JIC	0.6730	0.7605	0.3725	0.6805	0.3620	0.7795	0.3840	0.4260	0.5548	0.1847	1.2357	1.2357	1.2357	2.4769	2.4769	2.4769	3.7238
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.7875	0.7685	0.5530	0.7260	0.5260	0.7060	0.4620	0.4450	0.6218	0.1402	4.9667	0.9045	2.4313	5.7810	7.7208	3.5824	8.7929
	Smallest SD among APCS	0.5950	0.6810	0.5380	0.6675	0.5035	0.7055	0.4875	0.4905	0.5836	0.0910	1.7772	0.6178	1.2365	2.5019	3.1335	2.0651	3.9976
	Modal MAPCS	0.7955	0.7585	0.5560	0.7090	0.5200	0.7165	0.4650	0.4485	0.6211	0.1387	5.4151	0.9033	2.7849	6.2300	7.7770	3.8951	8.8491
	Median MAPCS	0.7845	0.7660	0.5505	0.7245	0.5220	0.7095	0.4565	0.4465	0.6200	0.1407	3.9549	0.9037	2.4215	4.7714	6.5256	3.5734	7.5972
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.7900	0.7660	0.5560	0.7310	0.5255	0.7060	0.4555	0.4440	0.6218	0.1419	1.5637	1.0156	1.0808	1.5819	1.7095	1.1099	1.7465
	Smallest SD among APCS	0.7355	0.7560	0.6030	0.7180	0.5195	0.7175	0.4675	0.4485	0.6207	0.1276	1.2183	1.0081	1.0804	1.2324	1.3074	1.1060	1.3357
	Modal MAPCS	0.7895	0.7325	0.5560	0.7300	0.5555	0.7065	0.4560	0.4450	0.6214	0.1346	1.2955	1.0156	1.0810	1.3048	1.4109	1.1098	1.4414
	Median MAPCS	0.8410	0.7540	0.5205	0.7095	0.5190	0.7145	0.4500	0.4480	0.6196	0.1523	1.9159	1.0226	1.1129	1.9381	2.0563	1.1425	2.1007

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.4b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.4 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	-8.0000	8.0000	0.0010
	Smallest SD among APCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0001
	Modal MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	-10.0000	10.0000	0.1000
	Median MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	20.0000	0.0001
	Smallest SD among APCS	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	32.0000	0.0001
	Modal MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.1000
	Median MAPCS	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	20.0000	0.0010

** TRF Temperature reduction factor

Table 4.4c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 4.4 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.5980	0.7005	0.4750	0.6650	0.5105	0.7875	0.5170	0.6320	0.6107	0.1070	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.9100	0.9230	0.5005	0.8325	0.4215	0.8740	0.4630	0.4610	0.6732	0.2288	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078
	GCV	0.6030	0.7100	0.4810	0.6735	0.5125	0.7950	0.5170	0.6250	0.6146	0.1090	1.0152	1.0152	1.0152	2.0409	2.0409	2.0409	3.0772
	HOC	0.6050	0.7110	0.4820	0.6750	0.5130	0.7970	0.5175	0.6245	0.6156	0.1094	1.0204	1.0204	1.0204	2.0514	2.0514	2.0514	3.0931
	HQ	0.7755	0.8345	0.5230	0.7650	0.4920	0.8420	0.5030	0.5560	0.6614	0.1560	1.5272	1.5272	1.5272	3.0544	3.0544	3.0544	4.5815
	RBAR	0.3435	0.4670	0.3640	0.4550	0.4860	0.6520	0.4860	0.7355	0.4986	0.1337	0.5076	0.5076	0.5076	1.0204	1.0204	1.0204	1.5386
	MCP	0.5980	0.7010	0.4750	0.6650	0.5110	0.7875	0.5170	0.6310	0.6107	0.1070	1.0002	1.0002	1.0002	2.0009	2.0009	2.0009	3.0021
	JIC	0.7450	0.8160	0.5215	0.7460	0.5030	0.8285	0.5090	0.5700	0.6549	0.1424	1.4051	1.4051	1.4051	2.8128	2.8128	2.8128	4.2232
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8955	0.8545	0.6525	0.7545	0.5735	0.7690	0.5975	0.6070	0.7130	0.1230	7.5854	1.4027	3.3902	8.8119	10.4567	4.5752	11.5769
	Smallest SD among APCS	0.6770	0.7735	0.6115	0.7580	0.5480	0.7165	0.5820	0.6775	0.6680	0.0817	1.6096	0.9039	1.2082	2.6533	3.2269	2.3071	3.9943
	Modal MAPCS	0.8925	0.8705	0.6560	0.7540	0.5690	0.7580	0.5980	0.6040	0.7127	0.1253	7.1941	1.3768	3.3912	8.4791	10.7632	4.5705	11.8838
	Median MAPCS	0.8755	0.8635	0.6440	0.7560	0.5650	0.7760	0.6210	0.5910	0.7115	0.1224	25.9405	1.3755	2.5081	27.2264	28.9419	3.6079	30.1568
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8925	0.8640	0.6560	0.7540	0.5635	0.7810	0.5985	0.5930	0.7128	0.1279	1.2124	1.0175	1.0594	1.2312	1.2709	1.0736	1.2888
	Smallest SD among APCS	0.8955	0.8590	0.6525	0.7540	0.5810	0.7235	0.5990	0.6060	0.7088	0.1206	2.7717	1.0180	1.0595	2.8115	3.0264	1.0737	3.0584
	Modal MAPCS	0.8920	0.8650	0.6560	0.7540	0.5665	0.7585	0.5990	0.6035	0.7118	0.1245	1.7619	1.0175	1.0592	1.7894	1.8618	1.0734	1.8844
	Median MAPCS	0.8955	0.8645	0.6525	0.7540	0.5830	0.7215	0.5990	0.5935	0.7079	0.1228	2.1047	1.0181	1.0596	2.1350	2.3482	1.0739	2.3767

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.4c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 4.4 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078	0.0000	18.4207	0.1000
	Smallest SD among APCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0100
	Modal MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	32.0000	0.1000
	Median MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-32.0000	32.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.0001
	Smallest SD among APCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	9.2103	0.0010
	Modal MAPCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Median MAPCS	2.3026	4.6052	4.6052	4.6052	6.9078	6.9078	6.9078	9.2103	0.0000	20.0000	0.0100

** TRF Temperature reduction factor

Table 4.5a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.5 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.5320	0.3800	0.3635	0.3470	0.2760	0.2625	0.2560	0.2585	0.3344	0.0946	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.7105	0.4060	0.3755	0.3655	0.2280	0.2185	0.2160	0.1905	0.3388	0.1726	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936
	GCV	0.5730	0.3955	0.3745	0.3625	0.2670	0.2590	0.2475	0.2210	0.3375	0.1156	1.0813	1.0813	1.0813	2.2245	2.2245	2.2245	3.4370
	HOC	0.5880	0.4000	0.3760	0.3635	0.2655	0.2570	0.2435	0.2160	0.3387	0.1219	1.1123	1.1123	1.1123	2.2901	2.2901	2.2901	3.5417
	HQ	0.5760	0.3865	0.3630	0.3555	0.2675	0.2585	0.2470	0.2380	0.3365	0.1131	1.0972	1.0972	1.0972	2.1944	2.1944	2.1944	3.2916
	RBAR	0.2990	0.2990	0.3220	0.2910	0.3155	0.3010	0.2905	0.3590	0.3096	0.0228	0.5407	0.5407	0.5407	1.1123	1.1123	1.1123	1.7185
	MCP	0.5325	0.3820	0.3640	0.3460	0.2740	0.2615	0.2525	0.2505	0.3329	0.0965	1.0059	1.0059	1.0059	2.0220	2.0220	2.0220	3.0538
	JIC	0.5385	0.3840	0.3655	0.3470	0.2715	0.2610	0.2515	0.2465	0.3332	0.0994	1.0193	1.0193	1.0193	2.0540	2.0540	2.0540	3.1061
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.6320	0.4095	0.4280	0.3425	0.2650	0.2550	0.1870	0.2305	0.3437	0.1444	1.2311	1.1016	1.4259	2.4117	2.5093	2.8040	3.7523
	Smallest SD among APCS	0.5695	0.3850	0.4620	0.3745	0.2410	0.2810	0.2320	0.1955	0.3426	0.1290	1.2359	0.8975	1.1031	2.5051	2.2600	2.3944	3.8528
	Modal MAPCS	0.6255	0.4100	0.4115	0.3345	0.2685	0.2555	0.2465	0.1950	0.3434	0.1381	1.2383	1.1034	1.3851	2.4199	2.5278	2.4925	3.9504
	Median MAPCS	0.6245	0.4170	0.4170	0.3395	0.2650	0.2555	0.2345	0.1930	0.3433	0.1405	1.2046	1.1027	1.3873	2.4320	2.5371	2.5755	4.0098
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.3490	0.3570	0.3645	0.3470	0.3625	0.3600	0.3430	0.3990	0.3603	0.0175	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Smallest SD among APCS	0.3490	0.3570	0.3645	0.3470	0.3625	0.3600	0.3430	0.3990	0.3603	0.0175	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Modal MAPCS	0.3490	0.3570	0.3645	0.3470	0.3625	0.3600	0.3430	0.3990	0.3603	0.0175	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	Median MAPCS	0.5935	0.4250	0.4380	0.4085	0.2140	0.2470	0.2230	0.1915	0.3426	0.1444	1.0715	1.0559	1.0543	1.1881	1.1501	1.1555	1.2719

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.5a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.5 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	**Criteria**	**Input values for simulated annealing**										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	-8.0000	8.0000	0.0100
	Smallest SD among APCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	8.0000	0.0100
	Modal MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0001
	Median MAPCS	0.0000	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936	0.0000	11.9829	0.0001
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Smallest SD among APCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Modal MAPCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Median MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.0010

** TRF Temperature reduction factor

Table 4.5b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.5 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.6240	0.5700	0.5905	0.5945	0.5540	0.5290	0.5635	0.5725	0.5747	0.0286	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.8770	0.7125	0.7325	0.7300	0.5280	0.5115	0.5400	0.4565	0.6360	0.1467	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680
	GCV	0.6450	0.5870	0.6040	0.6085	0.5530	0.5305	0.5675	0.5580	0.5817	0.0368	1.0310	1.0310	1.0310	2.0836	2.0836	2.0836	3.1589
	HOC	0.6505	0.5900	0.6050	0.6130	0.5530	0.5315	0.5685	0.5575	0.5836	0.0385	1.0418	1.0418	1.0418	2.1058	2.1058	2.1058	3.1929
	HQ	0.7530	0.6470	0.6640	0.6765	0.5535	0.5300	0.5690	0.5200	0.6141	0.0831	1.3641	1.3641	1.3641	2.7281	2.7281	2.7281	4.0922
	RBAR	0.3620	0.4010	0.4220	0.4090	0.5010	0.4765	0.5230	0.6650	0.4699	0.0958	0.5155	0.5155	0.5155	1.0418	1.0418	1.0418	1.5795
	MCP	0.6255	0.5710	0.5910	0.5955	0.5540	0.5295	0.5635	0.5720	0.5753	0.0290	1.0009	1.0009	1.0009	2.0035	2.0035	2.0035	3.0084
	JIC	0.7145	0.6290	0.6460	0.6505	0.5535	0.5300	0.5675	0.5365	0.6034	0.0662	1.2357	1.2357	1.2357	2.4769	2.4769	2.4769	3.7238
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9255	0.6885	0.7245	0.7180	0.5625	0.5330	0.5300	0.4530	0.6419	0.1519	2.5078	2.3976	2.4758	4.1494	4.2778	4.4728	6.2599
	Smallest SD among APCS	0.7720	0.6845	0.6485	0.6845	0.5685	0.4940	0.5335	0.5475	0.6166	0.0952	1.2583	1.4758	1.4544	2.6422	2.9266	2.9284	3.9298
	Modal MAPCS	0.9135	0.6940	0.7250	0.7275	0.5665	0.5320	0.5160	0.4580	0.6416	0.1503	2.2996	2.2837	2.2696	3.9580	4.1045	4.4485	6.2200
	Median MAPCS	0.9135	0.6940	0.7250	0.7275	0.5665	0.5320	0.5160	0.4580	0.6416	0.1503	2.2996	2.2887	2.2696	3.9580	4.1045	4.4485	6.2200
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9260	0.6895	0.7230	0.7160	0.5610	0.5335	0.5320	0.4540	0.6419	0.1515	1.0831	1.0786	1.0818	1.1330	1.1384	1.1465	1.2118
	Smallest SD among APCS	0.9260	0.6895	0.7230	0.7160	0.5610	0.5335	0.5320	0.4540	0.6419	0.1515	1.0831	1.0786	1.0818	1.1330	1.1384	1.1465	1.2118
	Modal MAPCS	0.9275	0.6910	0.7225	0.7210	0.5640	0.5345	0.5130	0.4595	0.6416	0.1532	1.0831	1.0820	1.0821	1.1331	1.1387	1.1557	1.2130
	Median MAPCS	0.9275	0.6910	0.7225	0.7210	0.5640	0.5345	0.5130	0.4595	0.6416	0.1532	1.0831	1.0820	1.0821	1.1331	1.1387	1.1557	1.2130

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.5b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.5 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680	-10.0000	10.0000	0.0001
	Smallest SD among APCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0010
	Modal MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	8.0000	0.0010
	Median MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	8.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.1000
	Smallest SD among APCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.1000
	Modal MAPCS	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	20.0000	0.0100
	Median MAPCS	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	20.0000	0.0100

** TRF Temperature reduction factor

Table 4.5c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 96 for Design 4.5 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.6895	0.6565	0.6860	0.6630	0.7480	0.7490	0.7210	0.8735	0.7233	0.0702	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.9525	0.8850	0.8905	0.8840	0.8340	0.8030	0.8055	0.8135	0.8585	0.0531	2.2822	2.2822	2.2822	4.5643	4.5643	4.5643	6.8465
	GCV	0.6990	0.6660	0.6935	0.6715	0.7535	0.7535	0.7235	0.8720	0.7291	0.0667	1.0159	1.0159	1.0159	2.0426	2.0426	2.0426	3.0805
	HOC	0.7025	0.6685	0.6945	0.6750	0.7555	0.7545	0.7250	0.8720	0.7309	0.0657	1.0213	1.0213	1.0213	2.0536	2.0536	2.0536	3.0971
	HQ	0.8525	0.7895	0.8130	0.7945	0.8065	0.7845	0.7780	0.8460	0.8081	0.0278	1.5183	1.5183	1.5183	3.0366	3.0366	3.0366	4.5548
	RBAR	0.3975	0.4445	0.4600	0.4485	0.6270	0.6345	0.6070	0.9040	0.5654	0.1658	0.5079	0.5079	0.5079	1.0213	1.0213	1.0213	1.5402
	MCP	0.6895	0.6575	0.6860	0.6635	0.7480	0.7490	0.7210	0.8735	0.7235	0.0700	1.0003	1.0003	1.0003	2.0009	2.0009	2.0009	3.0023
	JIC	0.8245	0.7635	0.7915	0.7705	0.8000	0.7780	0.7700	0.8520	0.7937	0.0308	1.3951	1.3951	1.3951	2.7928	2.7928	2.7928	4.1934
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9970	0.9230	0.9395	0.9300	0.8380	0.8150	0.7870	0.7690	0.8748	0.0831	4.4921	4.4030	4.5470	7.7022	7.5448	8.2879	11.4742
	Smallest SD among APCS	0.7145	0.8480	0.8470	0.8285	0.7920	0.7465	0.7720	0.8475	0.7995	0.0515	0.8333	0.8190	0.8562	2.6041	2.7918	2.5156	3.9997
	Modal MAPCS	0.9970	0.9230	0.9395	0.9300	0.8380	0.8150	0.7870	0.7690	0.8748	0.0831	4.4921	4.4030	4.5470	7.7022	7.5448	8.2879	11.4742
	Median MAPCS	0.9970	0.9240	0.9350	0.9310	0.8390	0.8140	0.7825	0.7745	0.8746	0.0825	4.4892	4.5413	4.5984	7.6989	7.6223	8.4281	11.3691
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9970	0.9225	0.9395	0.9305	0.8375	0.8150	0.7870	0.7695	0.8748	0.0830	1.0884	1.0856	1.0888	1.1513	1.1469	1.1645	1.2313
	Smallest SD among APCS	0.9970	0.9235	0.9350	0.9310	0.8395	0.8140	0.7825	0.7745	0.8746	0.0825	1.0833	1.0845	1.0860	1.1455	1.1438	1.1636	1.2245
	Modal MAPCS	0.9970	0.9225	0.9395	0.9305	0.8375	0.8150	0.7870	0.7695	0.8748	0.0830	1.0884	1.0856	1.0888	1.1513	1.1469	1.1645	1.2313
	Median MAPCS	0.9970	0.9255	0.9345	0.9305	0.8400	0.8135	0.7825	0.7745	0.8747	0.0826	1.0825	1.0854	1.0866	1.1457	1.1445	1.1642	1.2247

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.5c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 96 for Design 4.5 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-32.0000	32.0000	0.1000
	Smallest SD among APCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0010
	Modal MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-32.0000	32.0000	0.1000
	Median MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-10.0000	10.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.1000
	Smallest SD among APCS	2.2822	4.5643	4.5643	4.5643	6.8465	6.8465	6.8465	9.1287	0.0000	9.1287	0.0100
	Modal MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.1000
	Median MAPCS	2.2822	4.5643	4.5643	4.5643	6.8465	6.8465	6.8465	9.1287	0.0000	36.5148	0.0010

** TRF Temperature reduction factor

Table 4.6a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.6 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.5310	0.3275	0.5425	0.5705	0.3325	0.3515	0.5650	0.3740	0.4493	0.1116	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.7065	0.3285	0.6260	0.6865	0.2885	0.3190	0.5770	0.3025	0.4793	0.1858	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936
	GCV	0.5655	0.3445	0.5730	0.6125	0.3255	0.3530	0.5775	0.3385	0.4613	0.1302	1.0813	1.0813	1.0813	2.2245	2.2245	2.2245	3.4370
	HOC	0.5775	0.3460	0.5835	0.6195	0.3220	0.3525	0.5775	0.3330	0.4639	0.1352	1.1123	1.1123	1.1123	2.2901	2.2901	2.2901	3.5417
	HQ	0.5690	0.3345	0.5610	0.6010	0.3280	0.3460	0.5660	0.3525	0.4572	0.1258	1.0972	1.0972	1.0972	2.1944	2.1944	2.1944	3.2916
	RBAR	0.3070	0.2810	0.4020	0.3940	0.3530	0.3840	0.5090	0.4890	0.3899	0.0795	0.5407	0.5407	0.5407	1.1123	1.1123	1.1123	1.7185
	MCP	0.5325	0.3300	0.5455	0.5750	0.3335	0.3550	0.5680	0.3675	0.4509	0.1129	1.0059	1.0059	1.0059	2.0220	2.0220	2.0220	3.0538
	JIC	0.5385	0.3295	0.5500	0.5815	0.3325	0.3555	0.5705	0.3630	0.4526	0.1161	1.0193	1.0193	1.0193	2.0540	2.0540	2.0540	3.1061
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.7615	0.4425	0.6540	0.6940	0.3435	0.3455	0.4360	0.3675	0.5056	0.1702	1.1824	1.8443	2.5753	3.1273	3.9067	5.0477	5.9209
	Smallest SD among APCS	0.5170	0.5160	0.6735	0.6520	0.3215	0.3730	0.3760	0.4835	0.4891	0.1290	0.5890	1.0991	1.5113	2.3820	2.5229	3.5102	3.9726
	Modal MAPCS	0.7520	0.4450	0.6540	0.6855	0.3450	0.3595	0.4355	0.3665	0.5054	0.1648	1.1542	1.8081	2.5115	3.0893	3.7540	4.9743	5.8478
	Median MAPCS	0.7580	0.4515	0.6545	0.6625	0.3440	0.3720	0.4350	0.3605	0.5047	0.1619	1.1451	1.8474	2.7002	3.1285	3.8322	5.1673	6.0580
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.7575	0.4395	0.6525	0.6865	0.3450	0.3595	0.4345	0.3680	0.5054	0.1662	1.0663	1.1375	1.2163	1.2211	1.3005	1.4708	1.5086
	Smallest SD among APCS	0.7320	0.4735	0.6540	0.6485	0.3410	0.3875	0.4390	0.3635	0.5049	0.1514	1.0490	1.1374	1.2344	1.2211	1.2906	1.4793	1.5195
	Modal MAPCS	0.7575	0.4485	0.6545	0.6505	0.3420	0.3845	0.4390	0.3645	0.5051	0.1584	1.0621	1.1392	1.2367	1.2232	1.2958	1.4826	1.5228
	Median MAPCS	0.7620	0.4550	0.6440	0.6790	0.3405	0.3570	0.4355	0.3655	0.5048	0.1654	1.0621	1.1527	1.2386	1.2374	1.3245	1.4965	1.5369

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.6a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 4.6 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.0010
	Smallest *SD* among APCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0100
	Modal MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-32.0000	32.0000	0.1000
	Median MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.4979	2.9957	2.9957	2.9957	4.4936	4.4936	4.4936	5.9915	0.0000	23.9659	0.1000
	Smallest SD among APCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.0001
	Modal MAPCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	32.0000	0.1000
	Median MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	20.0000	0.0010

** TRF Temperature reduction factor

Table 4.6b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.6 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.5815	0.4510	0.6105	0.5990	0.4730	0.4900	0.6810	0.5505	0.5546	0.0787	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.8675	0.4735	0.7660	0.7895	0.4165	0.4375	0.7155	0.4190	0.6106	0.1913	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680
	GCV	0.5940	0.4575	0.6270	0.6190	0.4785	0.4945	0.6905	0.5400	0.5626	0.0828	1.0310	1.0310	1.0310	2.0836	2.0836	2.0836	3.1589
	HOC	0.6010	0.4600	0.6295	0.6225	0.4790	0.4925	0.6895	0.5370	0.5639	0.0834	1.0418	1.0418	1.0418	2.1058	2.1058	2.1058	3.1929
	HQ	0.7165	0.4770	0.6960	0.7055	0.4605	0.4795	0.6985	0.4955	0.5911	0.1213	1.3641	1.3641	1.3641	2.7281	2.7281	2.7281	4.0922
	RBAR	0.3105	0.3460	0.4265	0.4105	0.4730	0.4705	0.5695	0.6570	0.4579	0.1134	0.5155	0.5155	0.5155	1.0418	1.0418	1.0418	1.5795
	MCP	0.5815	0.4510	0.6125	0.5995	0.4730	0.4905	0.6810	0.5500	0.5549	0.0789	1.0009	1.0009	1.0009	2.0035	2.0035	2.0035	3.0084
	JIC	0.6750	0.4665	0.6695	0.6735	0.4655	0.4910	0.6950	0.5135	0.5812	0.1051	1.2357	1.2357	1.2357	2.4769	2.4769	2.4769	3.7238
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8560	0.5595	0.7680	0.7210	0.4675	0.5220	0.6500	0.4485	0.6241	0.1486	1.4094	2.0569	3.0438	3.6339	4.3033	5.5823	7.1213
	Smallest SD among APCS	0.6155	0.6325	0.7045	0.7010	0.4880	0.5030	0.5965	0.5635	0.6006	0.0808	0.6797	1.3275	1.4290	2.4737	2.6022	3.1670	3.9984
	Modal MAPCS	0.8225	0.6085	0.7420	0.7125	0.4655	0.5320	0.6575	0.4480	0.6236	0.1347	1.1256	2.3000	2.8629	3.7902	4.0360	5.3913	6.9278
	Median MAPCS	0.8390	0.5735	0.7660	0.7255	0.4640	0.5195	0.6510	0.4490	0.6234	0.1450	1.3257	2.0568	2.8626	3.6388	4.1367	5.4097	6.9457
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8670	0.5625	0.7565	0.7215	0.4605	0.5230	0.6620	0.4390	0.6240	0.1524	1.0369	1.0740	1.1055	1.1201	1.1384	1.1982	1.2514
	Smallest SD among APCS	0.8025	0.5890	0.7080	0.7205	0.5020	0.5130	0.6520	0.4825	0.6212	0.1180	1.0247	1.0638	1.0746	1.0915	1.1066	1.1479	1.1845
	Modal MAPCS	0.8560	0.5600	0.7680	0.7200	0.4660	0.5215	0.6510	0.4490	0.6239	0.1487	1.0364	1.0638	1.1064	1.1096	1.1393	1.1985	1.2476
	Median MAPCS	0.8190	0.6075	0.7405	0.7235	0.4690	0.5225	0.6660	0.4390	0.6234	0.1374	1.0247	1.0722	1.0954	1.1134	1.1280	1.1867	1.2395

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.6b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 4.6 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680	-4.0000	7.8240	0.1000
	Smallest SD among APCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0001
	Modal MAPCS	0.0000	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680	-10.0000	10.0000	0.0001
	Median MAPCS	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	32.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	32.0000	0.1000
	Smallest SD among APCS	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	32.0000	0.0100
	Modal MAPCS	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.1000
	Median MAPCS	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	32.0000	0.1000

** TRF Temperature reduction factor

Table 4.6c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 4.6 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Additive: Existing criteria (Type 1)	AIC	0.5860	0.5310	0.6470	0.6660	0.5245	0.5740	0.7245	0.6465	0.6124	0.0700	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
	BIC	0.8865	0.5875	0.8220	0.8540	0.4905	0.5390	0.7685	0.5065	0.6818	0.1670	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078
	GCV	0.5905	0.5385	0.6515	0.6755	0.5265	0.5770	0.7300	0.6390	0.6161	0.0702	1.0152	1.0152	1.0152	2.0409	2.0409	2.0409	3.0772
	HOC	0.5925	0.5390	0.6550	0.6755	0.5265	0.5765	0.7315	0.6385	0.6169	0.0707	1.0204	1.0204	1.0204	2.0514	2.0514	2.0514	3.0931
	HQ	0.7540	0.5855	0.7570	0.7785	0.5295	0.5825	0.7590	0.5840	0.6663	0.1043	1.5272	1.5272	1.5272	3.0544	3.0544	3.0544	4.5815
	RBAR	0.3105	0.3870	0.4345	0.4430	0.4880	0.5110	0.6160	0.7415	0.4914	0.1351	0.5076	0.5076	0.5076	1.0204	1.0204	1.0204	1.5386
	MCP	0.5860	0.5310	0.6470	0.6670	0.5245	0.5740	0.7245	0.6435	0.6126	0.0701	1.0002	1.0002	1.0002	2.0009	2.0009	2.0009	3.0021
	JIC	0.7285	0.5805	0.7360	0.7600	0.5320	0.5855	0.7530	0.5940	0.6587	0.0939	1.4051	1.4051	1.4051	2.8128	2.8128	2.8128	4.2232
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8700	0.6725	0.8025	0.7955	0.5630	0.6235	0.7065	0.5500	0.6979	0.1174	1.5503	2.3922	3.3009	4.0272	4.7063	6.0731	7.6431
	Smallest SD among APCS	0.6715	0.6365	0.7400	0.7920	0.5590	0.5820	0.6535	0.6600	0.6618	0.0764	1.0279	1.3596	1.3104	2.4956	2.6699	3.1693	3.9985
	Modal MAPCS	0.8490	0.6855	0.8055	0.8010	0.5630	0.6295	0.6955	0.5520	0.6976	0.1130	1.4065	2.2542	3.0763	3.8907	4.4596	6.0289	7.5572
	Median MAPCS	0.8705	0.6725	0.7970	0.7785	0.5585	0.6155	0.7265	0.5605	0.6974	0.1151	1.5488	2.3967	3.3606	4.0343	4.7452	5.5759	7.1319
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8670	0.6725	0.8025	0.8120	0.5630	0.6175	0.7000	0.5490	0.6979	0.1195	1.0211	1.0385	1.0537	1.0621	1.0742	1.1066	1.1301
	Smallest SD among APCS	0.8665	0.6710	0.7995	0.8120	0.5560	0.6105	0.6825	0.5815	0.6974	0.1159	1.0210	1.0385	1.0508	1.0621	1.0717	1.0993	1.1159
	Modal MAPCS	0.8665	0.6710	0.7995	0.8120	0.5560	0.6105	0.6825	0.5815	0.6974	0.1159	1.0210	1.0385	1.0508	1.0621	1.0717	1.0993	1.1159
	Median MAPCS	0.8975	0.6740	0.7985	0.7780	0.5660	0.6080	0.6650	0.5395	0.6908	0.1243	1.0233	1.0441	1.1133	1.0675	1.1334	1.1755	1.2002

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Table 4.6c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 4.6 together with relative penalty values and input values of SAO technique (continued).

Type of penalty	Criteria	Input values for simulated annealing										
		Starting values of penalties								Boundaries		TRF**
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.0000	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078	0.0000	32.0000	0.1000
	Smallest SD among APCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0100
	Modal MAPCS	0.0000	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078	0.0000	9.2103	0.0100
	Median MAPCS	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	-8.0000	8.0000	0.0001
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	2.3026	4.6052	4.6052	4.6052	6.9078	6.9078	6.9078	9.2103	0.0000	20.0000	0.1000
	Smallest SD among APCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.1000
	Modal MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.1000
	Median MAPCS	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	4.0000	0.0001

** TRF Temperature reduction factor

CHAPTER 5

MAXIMISATION OF MEAN AVERAGE PROBABILITY OF CORRECT SELECTION FOR EQUI-DIMENSIONAL COMPETING MODELS

5.1 INTRODUCTION

In Chapter 4, we applied the Simulated Annealing Optimisation (SAO) technique to estimate the optimal penalties to select the best model from a set of competing alternative models with an unequal number of regressors, with the objective of maximising the mean average probability of correct selection (MAPCS). In this chapter, we investigate the issue of model selection when the competing models have an equal number of parameters. For the existing IC based model selection procedures when the competing models have the same number of parameters, there is no need to use a penalty function because they result in the same penalty. In this situation, the problem reduces to choosing the model with the largest maximised log-likelihood.

For model selection when the competing models are white noise, first-order autoregressive (AR(1)), first-order moving average (MA(1)) and second-order autoregressive (AR(2)) disturbances in the linear regression model, Billah and King

(1998) applied a grid search method for finding the optimal penalty. The results obtained from their simulation experiments show that, though the AR(1) and MA(1) disturbance models have the same number of parameters, the penalty values that maximise the APCS are different. For selecting between AR(1) and MA(1) disturbances in the linear regression model, Grose and King (1994) observed that the MA(1) model is favoured because of the functional form of its log-likelihood. They argued that for model selection in small samples, a penalty for differences in the functional form of the log-likelihoods is needed to improve the probability of correct selection (PCS) in addition to the penalty for the number of parameters in the model. In other words, the penalty functions need to be able to take into account the form of the log-likelihood functions.

The simulation results we presented in the previous chapter indicate that for the models with the same number of parameters, the penalties that maximise the MAPCS of different data sets are different. Also for a particular data set, penalties for models with same number of parameters are different. From the literature on model selection and the simulation results we presented in the previous chapter, it is apparent that penalty functions should depend not only on n, the sample size, and k, the number of free parameters, but also on the form of the log-likelihood function. Unfortunately all the existing penalty functions are functions of n and k but not of the form of the log-likelihood function. This is the reason, when competing models are equi-dimensional, that the penalty function has no effect. In our view, a good penalty function should be a function of the log-likelihood in addition to n and k. To overcome the problem of the independence of penalty function from the data. here we apply the SAO technique to find the penalties. which maximise the MAPCS of the model. when the competing models have the same number of parameters to be estimated. Here. instead of using a particular functional form of penalty. we use appropriate upper and lower limits of penalties. and starting values of penalties. Then we apply the SAO technique to find the penalties. which maximise the MAPCS.

In this chapter we use both the additive and multiplicative penalties discussed in the previous chapter to select the correct model. when the competing models have an equal number of parameters. We use the SAO technique to maximise the MAPCS of the models. A detailed description of this method of optimisation technique was given in Section 2.4 of Chapter 2. The theory and computational technique of MAPCS are discussed in Section 3.2 of Chapter 3.

The plan of the chapter is as follows. In Section 5.2, we outline the Monte Carlo experiments. A variety of computer simulation results are presented in Section 5.3, and Section 5.4 contains some concluding remarks.

5.2 THE DESIGNS OF THE MONTE CARLO STUDIES

The main objective of the simulation experiments of this chapter is to investigate the performance of SAO technique to select the correct model from a set of equi-dimensional competing alternative models. Another objective is to see how the penalties obtained using SAO technique differ from zero (in all existing IC the penalties are zero when the competing models are equi-dimensional) when the competing models have same number of parameters.

The following three sets of models and two sets of data were used to conduct Monte Carlo experiments to examine the performance of the SAO technique with additive and multiplicative penalties when competing models are equi-dimensional.

Data set 1: x_{1t} is the real per capita GDP, x_{2t} is the investment of a country, x_{3t} is the price level consumption and x_{4t} is 1960 GDP as a percent of USA GDP of 1960. We used the annual data from Summers and Heston (1991) revised version 5.6 and World Bank tables.

Data set 2: We generated data from the normal distribution with standard deviation 4, 9, 11 and 17. Then, x_{1t}~IN(0, 4^2); x_{2t} ~IN(0, 9^2); x_{3t} ~IN(0, 11^2); and x_{4t}~IN(0, 17^2).

We used three model sets for simulation experiments with the above two data sets. Model set 1 is the single regressor model. Model set 2 and Model set 3 have two and three independent variables, respectively.

Model set 1.

$$M_1: \quad y_t = \beta_{10} + x_{1t}\beta_{11} + u_{1t}, \qquad u_{1t} = \mathrm{IN}(0, \sigma_1^2); \tag{5.1}$$

$$M_2: \quad y_t = \beta_{20} + x_{2t}\beta_{21} + u_{2t}, \qquad u_{2t} = \mathrm{IN}(0, \sigma_2^2); \tag{5.2}$$

$$M_3: \quad y_t = \beta_{30} + x_{3t}\beta_{31} + u_{3t}, \qquad u_{3t} = \mathrm{IN}(0, \sigma_3^2); \tag{5.3}$$

$$M_4: \quad y_t = \beta_{40} + x_{4t}\beta_{41} + u_{4t}, \qquad u_{4t} = \mathrm{IN}(0, \sigma_4^2). \tag{5.4}$$

Model set 2.

$$M_5: \quad y_t = \beta_{50} + x_{1t}\beta_{51} + x_{2t}\beta_{52} + u_{5t}, \qquad u_{5t} = \mathrm{IN}(0, \sigma_5^2); \tag{5.5}$$

$$M_6: \quad y_t = \beta_{60} + x_{2t}\beta_{62} + x_{3t}\beta_{63} + u_{6t}, \qquad u_{6t} = \mathrm{IN}(0, \sigma_6^2); \tag{5.6}$$

$$M_7: \quad y_t = \beta_{70} + x_{3t}\beta_{73} + x_{4t}\beta_{74} + u_{7t}, \qquad u_{7t} = \mathrm{IN}(0, \sigma_7^2); \tag{5.7}$$

$$M_8: \quad y_t = \beta_{80} + x_{4t}\beta_{84} + x_{1t}\beta_{81} + u_{8t}, \qquad u_{8t} = \mathrm{IN}(0, \sigma_8^2). \tag{5.8}$$

Model set 3.

$$M_9: \quad y_t = \beta_{90} + x_{1t}\beta_{91} + x_{2t}\beta_{92} + x_{3t}\beta_{93} + u_{9t}, \qquad u_{9t} = \mathrm{IN}(0, \sigma_9^2); \tag{5.9}$$

$$M_{10}: \quad y_t = \beta_{100} + x_{2t}\beta_{102} + x_{3t}\beta_{103} + x_{4t}\beta_{104} + u_{10t}, \qquad u_{10t} = \mathrm{IN}(0, \sigma_{10}^2); \tag{5.10}$$

$$M_{11}: \quad y_t = \beta_{110} + x_{3t}\beta_{113} + x_{4t}\beta_{114} + x_{1t}\beta_{111} + u_{11t}, \qquad u_{11t} = \mathrm{IN}(0, \sigma_{11}^2); \tag{5.11}$$

$$M_{12}: \quad y_t = \beta_{120} + x_{4t}\beta_{124} + x_{1t}\beta_{121} + x_{2t}\beta_{122} + u_{12t}, \qquad u_{12t} = \mathrm{IN}(0, \sigma_{12}^2). \tag{5.12}$$

Here y_t is the t^{th} observation on the dependent variable, x_{it} is the t^{th} observation on the i^{th} regressor, β_{j0} is a constant, β_{ji} is a scalar regression coefficient associated with the j^{th} model, j = 1, 2, ..., 12 and x_{ti} is the t^{th} value of the i^{th} regressor, i = 1, 2, 3, 4; and u_{jt} is a random disturbance term following a normal distribution with mean zero and variance σ_j^2.

We used the various combinations of data sets and model sets as designs. This gave $2\times3 = 6$ designs, as follows:

Design	Data set	Model set
5.1	1	1
5.2	2	1
5.3	1	2
5.4	2	2
5.5	1	3
5.6	2	3

We used the same data generating technique as described in Section 3.3.1 of Chapter 3 for these experiments using s_e^2 = 50.0 and 15, and s_b^2 = 0.02, 0.05 and 0.1. The sample sizes 20, 50 and 100 were used for the simulation experiments for all designs.

5.3 RESULTS OF THE SIMULATION EXPERIMENTS

The results of the Monte Carlo simulation experiments are presented in Tables 5.1a – c to 5.6a – c. There are three types of penalties in each table as mentioned in the previous chapter. The relative penalties for all existing criteria (Type 1 penalties) are zero and the APCS obtained form all the existing criteria are the same as the competing models are equi-dimensional. We present a comparative study of Type 1 and Type 2 penalties (additive penalty with maximised log-likelihood and the SAO technique) in Section 5.3.1. Section 5.3.2 contains the comparison of the simulation results obtained from Type 1 and Type 3 penalties (multiplicative penalty with mean squared error and the SAO technique). We compare the simulation results obtained from the newly proposed Type 2 and Type 3 penalties in Section 5.3.3.

5.3.1 COMPARISON OF TYPE 1 AND TYPE 2 PENALTIES

There are several interesting and notable results from the simulation experiments. For all designs and sample sizes under study, in 100 percent of the 96 combinations of the initial parameter values for the SAO technique, the MAPCS obtained from the simulation experiments using the SAO technique and Type 2 penalty for equi-dimensional models are higher than those obtained using the existing criteria. The variations among the APCS are always lower than those of the existing criteria for single regressor models. For other models, the variations among the APCS of the true models are generally higher than those of the existing criteria with an exception of Design 5.5 for sample size 100. The relative penalties that maximise the MAPCS using the SAO technique are not the same and are different from zero. An exceptional improvement in MAPCS was observed for all sample sizes for Design 5.5. In general we see that the model having the largest APCS from existing criteria, has the smallest APCS from the SAO technique and vice versa. It is evident from the simulation results that the MAPCS obtained from the different combinations of the initial parameter values for a particular sample size under any particular design are very similar. It implies that the APCS is in general insensitive to the initial values of the parameters of the SAO technique when the competing models are equi-dimensional.

Tables 5.1a – c and 5.2a – c contain the simulation results for Designs 5.1 and 5.2, where the competing models have a single non-constant regressor. For these designs, the gap between the largest[1] MAPCS and the smallest MAPCS obtained from the SAO technique is very small. The smallest MAPCS obtained from the SAO technique for Design 5.1 are 11, 14 and 17 percent higher than those of the MAPCS obtained from the existing criteria for n = 20, 50 and 100, respectively. For Design 5.2, the smallest MAPCS obtained from the SAO technique are 5, 4 and 4 percent higher than those of the existing criteria for n = 20, 50 and 100, respectively. For both designs, the variation among the APCS obtained from Type 2 penalties is lower than the variation obtained from the existing criteria. The decrease of variation among the APCS obtained from the SAO technique over the existing criteria for sample sizes 20, 50 and 100 is 27, 35 and 45 percent for Design 5.1 and 71, 56 and 73 percent for Design 5.2, respectively. The mode and the median MAPCS obtained

[1] There are 96 MAPCS for 96 combinations of initial parameter values for Type 2 penalties obtained using the SAO technique. We computed the largest, smallest, mode and median of these 96 MAPCS.

from the SAO technique are identical to the largest MAPCS obtained from the SAO technique for all sample sizes of Design 5.1 and sample size 20 of Design 5.2. For the other sample sizes of Design 5.2, the mode and median MAPCS are very similar to the largest MAPCS obtained from the SAO technique. For both designs and for all sample sizes, the MAPCS corresponding to the smallest variation among the APCS obtained from the SAO technique is very similar to the corresponding largest MAPCS obtained from the SAO technique.

The simulation results of the competing models with two non-constant regressors of Designs 5.3 and 5.4 are presented in Tables 5.3a – c to 5.4a – c. Like for the single regressor models, for both designs, the differences between the largest MAPCS and the smallest MAPCS is very small. The smallest MAPCS obtained from the SAO technique are 16, 9 and 9 percent higher than those of the existing criteria for sample sizes 20, 50 and 100, respectively for Design 5.3. The increases of MAPCS over the existing criteria are 8, 10, and 10 percent for sample sizes 20, 50 and 100, respectively, for Design 5.4. For both designs, the variations among the APCS are comparatively larger than those obtained from existing criteria. The mode and median MAPCS are very similar to the largest MAPCS and in some situations these are identical to the largest MAPCS. For both designs, the MAPCS corresponding to the smallest variation among the APCS are very similar to the largest MAPCS. This implies that the performance of Type 3 and Type 4 penalties are similar and from a user's point of view, any of these may be used for selecting the best model when the competing models are equi-dimensional.

The results of the simulation experiments for the three non-constant regressor models of Designs 5.5 and 5.6 are presented in Tables 5.5a – c to Table 5.6a – c. The results indicate that for both designs, for 100 percent of the combinations of the initial parameter values for the SAO technique, the MAPCS is higher than that of the existing criteria. An exceptional improvement of the MAPCS obtained from the SAO technique over the existing criteria is observed for Design 5.5. For both designs, the largest MAPCS and the smallest MAPCS obtained from the SAO technique are very similar. The smallest MAPCS obtained from the SAO technique for the sample sizes 20, 50 and 100 are 56, 57 and 56 percent larger than that of the existing criteria for Design 5.5. These increases in MAPCS for Design 5.6 are 14, 12 and 9 percent for the sample sizes 20, 50 and 100, respectively. For all designs and sample sizes, the variation among the APCS is larger than that of the existing criteria with the exception for Design 5.5 and n = 100, where the variation among the APCS is smaller than that of the existing criteria. The mode and the median MAPCS are very similar to the largest MAPCS for both designs and all sample sizes and in some cases these are identical to the largest MAPCS. The MAPCS corresponding to the smallest variation among the APCS are very close to the largest MAPCS for both designs.

For all designs, exactly the same MAPCS and SD among the APCS is obtained from different sets of relative penalties. For example, for Design 5.1 with n = 20, the largest MAPCS is 0.6470 with a SD among the APCS of 0.2434. This MAPCS and SD is obtained from 26 different relative penalty sets, two examples being (0, -1.5532, -1.1620, -1.8220), and (0, -1.5495, -1.1575, -1.1879). This result confirms the comment we made in the previous chapter, that there is no unique set of penalties for a particular data set that maximises the MAPCS. It is also observed that for all designs exactly the same MAPCS and SD are obtained from different sets of initial values for the SAO technique. For example, as mentioned earlier, for Design 5.1 the largest MAPCS with SD 0.2434 is obtained from 26 different penalties. These penalties are obtained from 26 different initial sets of values for the SAO technique. The smallest MAPCS obtained from the SAO technique for Design 5.1 and n = 20, is 0.6464 with a SD among the APCS of 0.2501, which is very close to the corresponding largest MAPCS and SD. This means that when the competing models have the same number of parameters, the SAO technique is largely insensitive to the initial values of the SAO technique.

5.3.2 COMPARISON OF TYPE 1 AND TYPE 3 PENALTIES

In this section, we compare the MAPCS obtained from the application of the SAO technique with mean squared error and multiplicative penalties, with those from the existing criteria. The simulation results obtained from the designs considered are presented in the second part of Tables 5.1a – c to 5.6a – c.

It is apparent from the simulation results that in 100 percent of the combinations of the initial parameter values for the SAO technique with mean squared error for all designs and sample sizes, the MAPCS obtained using the SAO technique are higher than those of the MAPCS obtained from the existing criteria. For all designs and sample sizes, the MAPCS obtained using the SAO technique with maximised log-likelihood and additive penalties are very similar to the MAPCS obtained from the SAO technique with mean squared error and multiplicative penalties. Sometimes the largest, mode, median MAPCS and the MAPCS corresponding to the smallest variation among the APCS, are all identical for Type 2 and Type 3 penalties. Thus for all designs and sample sizes, the comparisons between Type 1 and Type 3 penalties are very similar to the comparisons between Type 1 and Type 2 penalties discussed in the previous section.

5.3.3 COMPARISON OF TYPE 2 AND TYPE 3 PENALTIES

The results obtained from the simulation experiments for Type 2 and Type 3 penalties indicate that MAPCS obtained from both types of penalties are very similar. The largest MAPCS from the additive and multiplicative penalties are identical in 16 out of 18 experiments (six designs each with three sample sizes) and in the remaining two experiments, the numerical values of the largest MAPCS are very close. The modal MAPCS are identical in 12 out of 18 experiments and in the other experiments the values are very close. A similar picture is observed for median MAPCS and the MAPCS corresponding to the smallest variation among the APCS. The smallest

MAPCS obtained from these two types of penalties are also very close. It implies that to maximise the MAPCS of the true model from a set of competing alternative equi-dimensional models, the effect of the form of the penalty is insignificant. Therefore, either of the methods with the SAO technique can be used to maximise the MAPCS of the true model from a set of competing equi-dimensional models.

5.4 CONCLUDING REMARKS

In this chapter, we investigated the performance of the SAO technique to select the true model from a set of equi-dimensional competing alternative models. To compare the performance of the technique with the existing criteria, we conducted simulation experiments with two sets of data in three different sets for equi-dimensional models. In 100 percent of the combinations of the initial parameter values for the SAO technique, the MAPCS is higher than those of the existing criteria for all designs and the sample sizes considered. The variation among the APCS of the true model is smaller compared to that of the existing criteria when the competing models have a single regressor. In all other experiments, this variation is generally larger than those of the existing criteria with the exception of Design 5.5 with sample size 100. An exceptional increase (56%) of the MAPCS over the existing criteria is observed for all sample sizes of Design 5.5.

The relative penalties are zero for the existing criteria, but in the new method the relative penalties that maximise the MAPCS are different from zero. Also, for the same set of models, these penalties are different for different data sets. This finding confirms our previous finding that the penalties should not only be functions of the sample size and the number of free parameters, but also the form of the log-likelihood function. Exactly equal MAPCS were obtained from different relative penalties, which implies that there is no unique set of penalties that maximise the MAPCS.

From the simulation results, it is observed that the model having the largest MAPCS for the existing criteria, has the smallest MAPCS from the SAO technique and vice versa. It is apparent from the simulation results that for a particular sample size under any particular design, the MAPCS obtained from different combinations of the initial parameters are very similar. The difference between the largest and the smallest MAPCS obtained from the SAO technique is very small for all designs and for all sample sizes considered. This implies, for equi-dimensional competing alternative models, that the APCS of the true model is generally insensitive to the starting parameter values of the SAO technique. We may conclude from our simulation results that for selecting the best model from a set of equi-dimensional competing alternative models, the MAPCS obtained from the SAO technique will always be higher than that of existing criteria. We also recommend the use of the SAO technique to compute penalties for selecting the best model, when the competing models are equi-dimensional.

The results obtained from the simulation experiments indicate that the APCS obtained using the SAO technique with maximised log-likelihood functions and additive penalties, and mean squared error with multiplicative penalties, are very similar. So from the user's point of view, one can use either technique to select the true model from a set of competing alternative equi-dimensional models.

The performance of Type 2 and Type 3 penalties relative to the existing IC for selecting the true model from a set of equi-dimensional competing models, is uniformly better in all the experiments we conducted. From the results of the simulation experiments, it may be concluded that for the equi-dimensional competing models, the application of the SAO technique with additive or multiplicative penalties always guarantees the selection of the true model with higher MAPCS compared to the existing criteria in all situations for linear regression settings. We presented numerical evidence in favour of using the SAO technique over the existing IC to select the true model from a set of equi-dimensional competing alternative models. Such evidence is the best grounds for using this technique to improve APCS of the true model, instead of not using a penalty as the existing criteria effectively do in this situation.

Table 5.1a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 5.1 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.9845	0.3905	0.7200	0.2350	0.5825	0.3357	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	-1.5532	-1.1620	-1.8220	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.9685	0.3855	0.6395	0.5925	0.6465	0.2414	-2.1418	-1.7502	-2.4104	1.4979	1.4979	1.4979	1.4979	0.0000	8.0000	0.0001
	Modal MAPCS	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	-1.5532	-1.1620	-1.8220	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	-1.5532	-1.1620	-1.8220	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	0.8561	0.8903	0.8334	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.9730	0.3900	0.6110	0.6115	0.6464	0.2414	0.8563	0.9070	0.8335	1.0000	2.0000	2.0000	3.0000	0.0000	32.0000	0.0001
	Modal MAPCS	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	0.8561	0.8903	0.8334	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	0.8561	0.8903	0.8334	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 5.1b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 5.1 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.9920	0.5000	0.8880	0.2065	0.6466	0.3618	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	-0.3541	0.8881	-0.8217	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.1000
	Smallest SD of APCS	0.9890	0.4595	0.7860	0.7130	0.7394	0.2145	-0.5173	0.8542	-0.8258	0.0000	0.0000	0.0000	0.0000	0.0000	2.0000	0.0010
	Modal MAPCS	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	-0.3541	0.8881	-0.8217	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.1000
	Median MAPCS	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	-0.3541	0.8881	-0.8217	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	0.9860	1.0362	0.9677	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
	Smallest SD of APCS	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	0.9860	1.0362	0.9677	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
	Modal MAPCS	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	0.9860	1.0362	0.9677	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
	Median MAPCS	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	0.9860	1.0362	0.9677	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 5.1c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 5.1 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.9970	0.6180	0.9315	0.2055	0.6880	0.3617	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	-0.8164	0.6215	-1.4880	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.1000
	Smallest SD of APCS	0.9940	0.5285	0.8585	0.8405	0.8054	0.1969	-0.8678	0.5861	-1.5303	0.0000	1.0000	1.0000	2.0000	0.0000	10.0000	0.0100
	Modal MAPCS	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	-0.8164	0.6215	-1.4880	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.1000
	Median MAPCS	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	-0.8164	0.6215	-1.4880	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	0.9799	1.0085	0.9669	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
	Smallest SD of APCS	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	0.9799	1.0085	0.9669	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
	Modal MAPCS	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	0.9799	1.0085	0.9669	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
	Median MAPCS	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	0.9799	1.0085	0.9669	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 5.2a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 5.2 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.2690	0.4950	0.5085	0.6305	0.4758	0.1507	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.5570	0.4770	0.4560	0.5110	0.5002	0.0441	0.3764	0.5043	0.9540	0.0000	0.0000	0.0000	0.0000	0.0000	8.0000	0.1000
	Smallest SD of APCS	0.5535	0.4760	0.4560	0.5150	0.5001	0.0432	0.3767	0.5043	0.9144	0.0000	1.0000	1.0000	2.0000	0.0000	8.0000	0.1000
	Modal MAPCS	0.5570	0.4770	0.4560	0.5110	0.5002	0.0441	0.3764	0.5043	0.9540	0.0000	0.0000	0.0000	0.0000	0.0000	8.0000	0.1000
	Median MAPCS	0.5570	0.4770	0.4560	0.5110	0.5002	0.0441	0.3764	0.5043	0.9540	0.0000	0.0000	0.0000	0.0000	0.0000	8.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.5570	0.4770	0.4560	0.5110	0.5002	0.0441	1.0384	1.0517	1.1001	1.0000	1.0000	1.0000	1.0000	0.0000	2.0000	0.1000
	Smallest SD of APCS	0.5535	0.4760	0.4560	0.5150	0.5001	0.0432	1.0384	1.0518	1.0957	1.0000	1.0000	1.0000	1.00[illegible]	0.0000	20.0000	0.1000
	Modal MAPCS	0.5570	0.4770	0.4560	0.5110	0.5002	0.0441	1.0384	1.0517	1.1001	1.0000	1.0000	1.0000	1.0000	0.0000	2.0000	0.1000
	Median MAPCS	0.5535	0.4760	0.4560	0.5150	0.5001	0.0432	1.0384	1.0518	1.0957	1.0000	1.0000	1.0000	1.0000	0.0000	20.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 5.2b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 5.2 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.3360	0.5725	0.6275	0.7180	0.5635	0.1631	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.6720	0.5275	0.5245	0.6165	0.5851	0.0719	0.5147	0.8774	1.0103	0.0000	0.0000	0.0000	0.0000	-10.0000	10.0000	0.1000
	Smallest SD of APCS	0.6465	0.5235	0.5420	0.6265	0.5846	0.0609	0.5123	0.7462	0.8995	0.0000	1.0000	1.0000	2.0000	-2.0000	2.0000	0.0010
	Modal MAPCS	0.6720	0.5275	0.5245	0.6165	0.5851	0.0719	0.5147	0.8774	1.0103	0.0000	0.0000	0.0000	0.0000	-10.0000	10.0000	0.1000
	Median MAPCS	0.6685	0.5265	0.5285	0.6165	0.5850	0.0697	0.5143	0.8476	1.0102	1.9560	1.9560	1.9560	1.9560	0.0000	8.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.6720	0.5275	0.5245	0.6165	0.5851	0.0719	1.0208	1.0357	1.0412	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Smallest SD of APCS	0.6510	0.5235	0.5420	0.6220	0.5846	0.0615	1.0207	1.0302	1.0386	3.9120	3.9120	3.9120	3.9120	0.0000	3.9120	0.0010
	Modal MAPCS	0.6720	0.5275	0.5245	0.6165	0.5851	0.0719	1.0208	1.0357	1.0412	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Median MAPCS	0.6720	0.5275	0.5245	0.6165	0.5851	0.0719	1.0208	1.0357	1.0412	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 5.2c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 5.2 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.4290	0.6575	0.7125	0.8180	0.6542	0.1643	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	0.4189	0.6262	1.1782	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	0.4189	0.6262	1.1782	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Modal MAPCS	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	0.4189	0.6262	1.1782	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	0.4189	0.6262	1.1782	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	1.0084	1.0126	1.0239	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Smallest SD of APCS	0.6900	0.6320	0.6620	0.7355	0.6799	0.0440	1.0084	1.0126	1.0242	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.0010
	Modal MAPCS	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	1.0084	1.0126	1.0239	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Median MAPCS	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	1.0084	1.0126	1.0239	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Table 5.3a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 5.3 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.5270	0.4030	0.3745	0.4785	0.4457	0.0697	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.2440	0.3505	0.7285	0.7470	0.5175	0.2581	-3.4352	-3.7546	-0.5835	2.9957	2.9957	2.9957	2.9957	-4.0000	5.9915	0.0001
	Smallest SD of APCS	0.3310	0.3215	0.7410	0.6680	0.5154	0.2204	-1.8554	-2.2314	-0.3124	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.0001
	Modal MAPCS	0.2440	0.3505	0.7285	0.7470	0.5175	0.2581	-3.4352	-3.7546	-0.5835	2.9957	2.9957	2.9957	2.9957	-4.0000	5.9915	0.0001
	Median MAPCS	0.2800	0.3510	0.7285	0.7095	0.5172	0.2349	-3.4126	-3.7303	-0.4440	0.0000	1.0000	1.0000	2.0000	-12.0000	12.0000	0.0100
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.2440	0.3510	0.7280	0.7470	0.5175	0.2579	0.7101	0.6879	0.9433	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.1000
	Smallest SD of APCS	0.2800	0.3510	0.7285	0.7095	0.5172	0.2349	0.7131	0.6907	0.9566	4.4936	4.4936	4.4936	4.4936	0.0000	14.9787	0.0010
	Modal MAPCS	0.2440	0.3510	0.7280	0.7470	0.5175	0.2579	0.7101	0.6879	0.9433	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.1000
	Median MAPCS	0.2440	0.3265	0.7495	0.7485	0.5171	0.2699	0.7485	0.7209	0.9433	1.0000	2.0000	2.0000	3.0000	0.0000	27.0000	0.0010

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Table 5.3b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 5.3 together with relative penalty values and input values [illegible] SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.5820	0.5045	0.4295	0.4845	0.5001	0.0631	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.3945	0.3845	0.7400	0.7085	0.5569	0.1937	-1.4966	-1.8183	-0.2936	3.9120	3.9120	3.9120	3.9120	0.0000	12.0000	0.1000
	Smallest SD of APCS	0.4210	0.3885	0.7345	0.6815	0.5564	0.1769	-1.4294	-1.7417	-0.2272	3.9120	3.9120	3.9120	3.9120	0.0000	7.8240	0.1000
	Modal MAPCS	0.4190	0.3850	0.7450	0.6750	0.5560	0.1806	-2.0712	-2.3929	-0.2262	3.9120	3.9120	3.9120	3.9120	-12.0000	12.0000	0.1000
	Median MAPCS	0.4190	0.3850	0.7450	0.6750	0.5560	0.1806	-2.0712	-2.3929	-0.2262	3.9120	3.9120	3.9120	3.9120	-12.0000	12.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.3945	0.3840	0.7405	0.7075	0.5566	0.1938	0.9403	0.9283	0.9883	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Smallest SD of APCS	0.4130	0.3865	0.7485	0.6370	0.5463	0.1755	0.6784	0.6697	0.9918	1.0000	1.0000	1.0000	1.0000	0.0000	12.0000	0.0001
	Modal MAPCS	0.4195	0.3850	0.7445	0.6755	0.5561	0.1804	0.9238	0.9120	0.9910	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.0001
	Median MAPCS	0.4150	0.3900	0.7425	0.6670	0.5536	0.1775	0.8733	0.8624	0.9910	5.8680	5.8680	5.8680	5.8680	0.0000	5.8680	0.0001

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Table 5.3c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 5.3 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.6230	0.6085	0.4785	0.5005	0.5526	0.0737	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.4480	0.4490	0.8080	0.7330	0.6095	0.1884	-1.9048	-2.3262	-0.3119	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.0100
	Smallest SD of APCS	0.4515	0.4490	0.8075	0.7290	0.6093	0.1864	-1.8037	-2.2247	-0.3028	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.0010
	Modal MAPCS	0.4460	0.4500	0.8105	0.7300	0.6091	0.1889	-3.3369	-3.7578	-0.3118	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Median MAPCS	0.4460	0.4500	0.8105	0.7300	0.6091	0.1889	-3.3369	-3.7578	-0.3118	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.4480	0.4490	0.8080	0.7330	0.6095	0.1884	0.9625	0.9544	0.9938	1.0000	2.0000	2.0000	3.0000	0.0000	12.0000	0.1000
	Smallest SD of APCS	0.4580	0.4405	0.7970	0.7250	0.6051	0.1825	0.9825	0.9743	0.9942	6.9078	6.9078	6.9078	6.9078	0.0000	6.9078	0.0001
	Modal MAPCS	0.4400	0.4500	0.8095	0.7370	0.6091	0.1919	0.9527	0.9447	0.9935	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.0100
	Median MAPCS	0.4575	0.4445	0.8025	0.7235	0.6070	0.1831	0.9757	0.9675	0.9942	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.0001

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Table 5.4a **Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 5.4 together with relative penalty values and input values of SAO technique.**

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.3790	0.4805	0.5250	0.3840	0.4421	0.0724	0.0000	0.0000	0.000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.7265	0.3540	0.3700	0.4720	0.4806	0.1720	0.7617	0.8956	0.4813	0.0000	0.0000	0.0000	0.0000	-12.0000	12.0000	0.1000
	Smallest SD of APCS	0.6180	0.5120	0.3185	0.4710	0.4799	0.1241	0.2655	0.8940	0.4780	0.0000	0.0000	0.0000	0.0000	-12.0000	12.0000	0.0100
	Modal MAPCS	0.7135	0.3755	0.3765	0.4560	0.4804	0.1599	0.6810	0.8592	0.5009	2.9957	2.9957	2.9957	2.9957	-12.0000	12.0000	0.1000
	Median MAPCS	0.7135	0.3755	0.3765	0.4560	0.4804	0.1599	0.6810	0.8592	0.5009	2.9957	2.9957	2.9957	2.9957	-12.0000	12.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.7265	0.3540	0.3700	0.4720	0.4806	0.1720	1.0790	1.0936	1.0492	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.1000
	Smallest SD of APCS	0.6805	0.4285	0.3400	0.4715	0.4801	0.1444	1.0513	1.0947	1.0502	1.0000	1.0000	1.0000	1.0000	0.0000	20.0000	0.0010
	Modal MAPCS	0.7135	0.3755	0.3775	0.4555	0.4805	0.1597	1.0702	1.0894	1.0514	1.0000	1.0000	1.0000	1.0000	0.0000	20.0000	0.1000
	Median MAPCS	0.7135	0.3755	0.3775	0.4555	0.4805	0.1597	1.0702	1.0894	1.0514	1.0000	1.0000	1.0000	1.0000	0.0000	20.0000	0.1000

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Table 5.4b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 5.4 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.4155	0.5910	0.6380	0.4610	0.5264	0.1052	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.7005	0.5455	0.4435	0.6305	0.5800	0.1109	0.5450	0.9479	0.2884	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.6945	0.5510	0.4440	0.6300	0.5799	0.1079	0.5265	0.9294	0.2819	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.1000
	Modal MAPCS	0.7025	0.5400	0.4465	0.6300	0.5797	0.1109	0.5604	0.9389	0.2902	3.9120	3.9120	3.9120	3.9120	-12.0000	12.0000	0.0010
	Median MAPCS	0.7025	0.5400	0.4465	0.6300	0.5797	0.1109	0.5604	0.9389	0.2902	3.9120	3.9120	3.9120	3.9120	-12.0000	12.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.7005	0.5450	0.4440	0.6305	0.5800	0.1107	1.0220	1.0385	1.0114	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.6965	0.5500	0.4435	0.6295	0.5799	0.1089	1.0214	1.0381	1.0115	5.8680	5.8680	5.8680	5.8680	0.0000	19.5601	0.1000
	Modal MAPCS	0.6965	0.5500	0.4435	0.6295	0.5799	0.1089	1.0214	1.0381	1.0115	5.8680	5.8680	5.8680	5.8680	0.0000	19.5601	0.1000
	Median MAPCS	0.7330	0.5715	0.4195	0.5950	0.5797	0.1284	1.0213	1.0492	1.0224	5.8680	5.8680	5.8680	5.8680	0.0000	46.9443	0.1000

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Table 5.4c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 5.4 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.5130	0.6520	0.7370	0.5780	0.6200	0.0965	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.7805	0.6545	0.5480	0.7465	0.6824	0.1042	0.5510	1.2144	0.5141	0.0000	1.0000	1.0000	2.0000	-10.0000	10.0000	0.1000
	Smallest SD of APCS	0.7675	0.6450	0.5590	0.7540	0.6814	0.0983	0.5513	1.0794	0.4080	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Modal MAPCS	0.7795	0.6545	0.5495	0.7455	0.6823	0.1030	0.5467	1.2017	0.5117	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.1000
	Median MAPCS	0.7795	0.6545	0.5495	0.7455	0.6823	0.1030	0.5467	1.2017	0.5117	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.7805	0.6545	0.5480	0.7465	0.6824	0.1042	1.0111	1.0246	1.0103	6.9078	6.9078	6.9078	6.9078	0.0000	13.8155	0.1000
	Smallest SD of APCS	0.7805	0.6450	0.5705	0.7290	0.6813	0.0926	1.0111	1.0219	1.0103	6.9078	6.9078	6.9078	6.9078	0.0000	6.9078	0.0001
	Modal MAPCS	0.7795	0.6550	0.5485	0.7460	0.6823	0.1035	1.0110	1.0245	1.0103	1.0000	2.0000	2.0000	3.0000	0.0000	12.0000	0.1000
	Median MAPCS	0.7795	0.6550	0.5485	0.7460	0.6823	0.1035	1.0110	1.0245	1.0103	1.0000	2.0000	2.0000	3.0000	0.0000	12.0000	0.1000

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Table 5.5a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 5.5 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.6410	0.3770	0.4650	0.3950	0.4695	0.1205	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.3690	0.9970	0.6530	0.9180	0.7343	0.2845	-5.2089	-0.7590	-2.2686	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.3725	0.9855	0.6625	0.9060	0.7316	0.2760	-3.9370	-0.7641	-2.0906	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.1000
	Modal MAPCS	0.3690	[illegible]	0.6530	0.9180	0.7343	0.2845	-5.2089	-0.7590	-2.2686	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.3695	0.9955	0.6530	0.9180	0.7340	0.2838	4.9027	-0.7595	-2.2689	4.4936	4.4936	4.4936	4.4936	0.0000	5.9915	0.0100
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.3690	0.9970	0.6530	0.9180	0.7343	0.2845	0.5873	0.9270	0.7970	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Smallest SD of APCS	0.3760	0.9970	0.6460	0.9165	0.7339	0.2819	0.5935	0.9310	0.8012	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
	Modal MAPCS	0.3690	0.9970	0.6530	0.9180	0.7343	0.2845	0.5873	0.9270	0.7970	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Median MAPCS	0.3690	0.9970	0.6530	0.91	0.7343	0.2845	0.5873	0.9270	0.7970	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Table 5.5b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 5.5 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.7685	0.4010	0.4955	0.4365	0.5254	0.1667	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.5175	0.9990	0.8325	0.9490	0.8245	0.2162	-4.8874	-1.2190	-2.6135	0.0000	1.0000	1.0000	2.0000	0.0000	10.0000	0.0100
	Smallest SD of APCS	0.5330	0.9935	0.8095	0.9600	0.8240	0.2099	-3.9153	-1.0541	-2.7323	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.1000
	Modal MAPCS	0.5175	0.9990	0.8325	0.9490	0.8245	0.2162	-4.8874	-1.2190	-2.6135	0.0000	1.0000	1.0000	2.0000	0.0000	10.0000	0.0100
	Median MAPCS	0.5175	0.9990	0.8325	0.9490	0.8245	0.2162	-4.8874	-1.2190	-2.6135	0.0000	1.0000	1.0000	2.0000	0.0000	10.0000	0.0100
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.5175	0.9990	0.8325	0.9495	0.8246	0.2163	0.8220	0.9523	0.9002	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Smallest SD of APCS	0.5165	0.9935	0.8330	0.9540	0.8242	0.2162	0.8488	0.9523	0.8981	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.0001
	Modal MAPCS	0.5175	0.9990	0.8325	0.9495	0.8246	0.2163	0.8220	0.9523	0.9002	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Median MAPCS	0.5175	0.9990	0.8325	0.9495	0.8246	0.2163	0.8220	0.9523	0.9002	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Table 5.5c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 5.5 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.8505	0.4350	0.5005	0.4495	0.5589	0.1964	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	-5.3640	-1.2130	-3.0750	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.6530	0.9845	0.8785	0.9685	0.8711	0.1527	-3.4805	-1.2128	-2.5651	0.0000	1.0000	1.0000	2.0000	0.0000	4.0000	0.0010
	Modal MAPCS	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	-5.3640	-1.2130	-3.0750	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	-5.3640	-1.2130	-3.0750	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	0.8978	0.9760	0.9403	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Smallest SD of APCS	0.6520	0.9985	0.8760	0.9715	0.8745	0.1574	0.8964	0.9760	0.9474	9.2103	9.2103	9.2103	9.2103	0.0000	9.2103	0.0100
	Modal MAPCS	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	0.8978	0.9760	0.9403	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
	Median MAPCS	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	0.8978	0.9760	0.9403	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Table 5.6a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 20 for Design 5.6 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.4295	0.6030	0.4520	0.4630	0.4869	0.0787	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8290	0.2640	0.5340	0.6330	0.5650	0.2351	1.6520	0.6531	0.5078	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Smallest SD of APCS	0.8275	0.3400	0.4970	0.5915	0.5640	0.2040	1.2607	0.7286	0.5748	0.0000	1.0000	1.0000	2.0000	-10.0000	10.0000	0.0100
	Modal MAPCS	0.8290	0.2640	0.5340	0.6330	0.5650	0.2351	1.6520	0.6531	0.5078	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Median MAPCS	0.8265	0.2580	0.5360	0.6390	0.5649	0.2373	1.6989	0.6420	0.4884	0.0000	1.0000	1.0000	2.0000	-10.0000	10.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8290	0.2640	0.5340	0.6330	0.5650	0.2351	1.1796	1.0675	1.0521	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.8370	0.3280	0.5140	0.5775	0.5641	0.2105	1.1434	1.0755	1.0657	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.0010
	Modal MAPCS	0.8290	0.2640	0.5340	0.6330	0.5650	0.2351	1.1796	1.0675	1.0521	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.8265	0.2580	0.5360	0.6390	0.5649	0.2373	1.1851	1.0663	1.0500	5.9915	5.9915	5.9915	5.9915	0.0000	11.9829	0.0010

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Table 5.6b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 50 for Design 5.6 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.5360	0.6960	0.5555	0.5520	0.5849	0.0746	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.8325	0.4160	0.5895	0.7950	0.6583	0.1936	1.3174	0.5771	0.1042	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Smallest SD of APCS	0.8285	0.4225	0.5925	0.7855	0.6573	0.1871	1.2521	0.5440	0.1041	5.8680	5.8680	5.8680	5.8680	-10.0000	10.0000	0.0010
	Modal MAPCS	0.8340	0.4180	0.5835	0.7965	0.6580	0.1943	1.3200	0.6113	0.1041	0.0000	0.0000	0.0000	0.0000	-16.0000	16.0000	0.0010
	Median MAPCS	0.8340	0.4180	0.5835	0.7965	0.6580	0.1943	1.3200	0.6113	0.1041	0.0000	0.0000	0.0000	0.0000	-16.0000	16.0000	0.0010
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.8325	0.4160	0.5895	0.7950	0.6583	0.1936	1.0541	1.0233	1.0042	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.7900	0.4450	0.6430	0.7395	0.6544	0.1523	1.0404	1.0113	1.0042	1.0000	1.0000	1.0000	1.0000	0.0000	16.0000	0.0100
	Modal MAPCS	0.8325	0.4160	0.5895	0.7950	0.6583	0.1936	1.0541	1.0233	1.0042	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.8325	0.4160	0.5895	0.7950	0.6583	0.1936	1.0541	1.0233	1.0042	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Table 5.6c Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models for sample size 100 for Design 5.6 together with relative penalty values and input values of SAO technique.

Type of penalty	Criteria	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Additive: Existing criteria (Type 1)	All are same	0.6290	0.7505	0.6785	0.6565	0.6786	0.0520	0.0000	0.0000	0.0000							
Additive: Using log-likelihood with SAO technique (Type 2)	Largest MAPCS	0.9270	0.5510	0.7080	0.7890	0.7438	0.1571	1.4753	0.8747	0.6798	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.8945	0.5670	0.7345	0.7760	0.7430	0.1355	1.1706	0.5862	0.4886	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0100
	Modal MAPCS	0.9270	0.5510	0.7080	0.7890	0.7438	0.1571	1.4753	0.8747	0.6798	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Median MAPCS	0.9005	0.5090	0.7440	0.8200	0.7434	0.1688	1.5228	0.5904	0.4183	0.0000	0.0000	0.0000	0.0000	0.0000	16.0000	0.1000
Multiplicative: Using mean squared error with SAO technique (Type 3)	Largest MAPCS	0.9270	0.5510	0.7080	0.7890	0.7438	0.1571	1.0300	1.0177	1.0137	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
	Smallest SD of APCS	0.8935	0.5670	0.7350	0.7765	0.7430	0.1352	1.0235	1.0116	1.0097	9.2103	9.2103	9.2103	9.2103	0.0000	18.4207	0.0100
	Modal MAPCS	0.8990	0.5085	0.7445	0.8210	0.7432	0.1687	1.0308	1.0116	1.0082	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.0001
	Median MAPCS	0.8990	0.5085	0.7445	0.8210	0.7432	0.1687	1.0308	1.0116	1.0082	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.0001

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

CHAPTER 6

EQUAL PROBABILITIES OF CORRECT SELECTION: A CONSTRAINED MINIMISATION OF VARIATION AMONG THE AVERAGE PROBABILITIES OF CORRECT SELECTION

6.1 INTRODUCTION

In the previous three chapters, we mainly concentrated on the maximisation of the mean average probability of correct selection (MAPCS), using our proposed technique used in Chapter 3 and the SAO technique used in Chapter 4 and Chapter 5. In Chapter 3, we presented the generaliased form of the penalty function of six existing IC and a technique of maximisation of the MAPCS. In Chapter 4 and Chapter 5, we presented two different types of penalties: additive penalties with maximised log-likelihood and multiplicative penalties with mean squared error; and applied the SAO technique to maximise the MAPCS. We conducted several simulation experiments with different sets of data and competing models to evaluate the performance of both types of penalties with the SAO technique. It was observed that the MAPCS obtained from these techniques were always higher than those of the existing criteria. Unfortunately like the existing criteria, the average probabilities of correct selection (APCS) for each model are uneven across different models for both types of penalties, although in some cases the variation among the APCS is much less than that of existing criteria. But a desirable property of a good model selection procedure should be that it selects the best model without favouring one model over the others, i.e. the APCS of each competing model when it is the true model should be equal. Our goal is to choose the penalty in such a way that none of the competing models is favoured over the others unknowingly. In the literature, this technique is called controlling the probabilities of correct selection. (see for example King *et al.* (1995) and Forbes *et al.* (1995)). This chapter is concerned with controlling the APCS in linear regression settings.

Forbes *et al.* (1995) proposed three techniques for controlling the probability of correct selection of one model over the others. They showed that for the variable selection problem, the relative penalty of any two competing models can be expressed as a function of the percentile of the F distribution and they called this penalty function FIC. Their second method is also for variable selection purposes and is based on quasi-maximum likelihood functions. They showed that in this case the penalty function could be expressed approximately as a function of a χ_1^2 random variable. Their third method gives the penalties that aim to control the probability of correct selection but are approximate penalties for general model selection problems.

The problem with using the above methods is that the penalty function is fixed (function of F distribution or χ_1^2 distribution) for a particular sample size and number of free parameters, and the form of log-likelihood function has no impact on its value. Also, when the models are equi-dimensional, the impact of these penalties in selecting the true model is null. King *et al.* (1995) proposed two approaches for controlling the probability of correct selection (PCS) of models and gave an algorithm for calculating the penalties, which control the PCS. Their first approach is based on the idea of a common model, but there are two problems with this approach. Firstly, there may not be a common model among the competing models. Secondly, when the competing models are nested, then there is no fixed rule for setting the selection probabilities and it is left to the user's arbitrary choice. Obviously different users may then come to different conclusions for the same data set and set of competing models. Their second method is based on a representative fixed points approach. They proposed two techniques for selecting representative fixed point models. Their first technique entirely depends on the judgement of the user, while the second technique depends on the idea of a common model, which may not exist in all situations. They argued that it is possible to control probabilities of correct selection, though their approach has some limitations as mentioned earlier. Our objective in this chapter is to develop a technique of selecting the true model by controlling the APCS in such a way that each competing model has an equal chance on average of being selected.

The organisation of this chapter is as follows. In Section 6.2 we discuss the issue of making the APCS equal and propose a new method of selecting penalties which makes the APCS equal when the number of parameters in the competing models are unequal and equal. Section 6.3 contains the descriptions of the Monte Carlo experiments and results of these experiments when competing models have an unequal number of parameters. Section 6.4 contains the description of the simulation experiments and the results of these experiments when competing models are equi-dimensional. The final section contains some concluding remarks.

6.2 PROPOSED METHOD OF MINIMISATION OF VARIATION AMONG THE APCS

In this section we propose a method of choosing the penalties in an IC based model selection procedure so that the variation among the APCS of each model is as small as possible. Theoretically, our main objective is to make the APCS of each model equal and it will happen if the variation among the APCS of models is zero. Numerically if the variation among the APCS of each model is close to zero, it may be believed that the estimated APCS of the models are nearly equal.

As noted in the previous chapter, in almost all IC based model selection procedures, the j^{th} model will be selected if

$$\left(L_j(\hat{\beta}_j,\hat{\sigma}_j^2)-p_j\right)>\left(L_i(\hat{\beta}_i,\hat{\sigma}_i^2)-p_i\right), \forall i,\ i=1,2,\ldots,(j-1),(j+1),\ldots,m, \quad (6.1)$$

or,

$$E_j^2q_j^* < E_i^2q_i^*, \forall i,\ i=1,2,\ldots,(j-1),(j+1),\ldots,m, \quad (6.2)$$

where p_j is the additive penalty, $L_j(\hat{\beta}_j, \hat{\sigma}_j^2)$ is the maximised log-likelihood function, $\hat{\beta}_j$ and $\hat{\sigma}_j^2$ are the maximum likelihood estimators of β_j and σ_j^2, respectively, q_j^* is the multiplicative penalty and E_j^2 is the mean squared error for the j^{th} model.

In equation (6.1), $L_j(\hat{\beta}_j, \hat{\sigma}_j^2), \forall j, j = 1, 2, \ldots, m$, and in (6.2), $E_j^2, \forall j, j = 1, 2, \ldots, m$, is known. Our main objective is to find the penalties p_j and $q_j^*, \forall j, j = 1, 2, \ldots, m$, in such a way that the average probabilities of correctly choosing the j^{th} model and the i^{th} models are equal, i.e.

$$\text{APCS of the } j^{th} \text{ model} = \text{APCS of the } i^{th} \text{ model}, \ \forall j \neq i, \ i, j = 1, 2, \ldots, m. \quad (6.3)$$

or approximately equal, i.e.

$$\text{APCS of the } j^{th} \text{ model} \approx \text{APCS of the } i^{th} \text{ model}, \ \forall j \neq i, \ i, j = 1, 2, \ldots, m. \quad (6.4)$$

In other words, equation (6.3) will be true if the variation among the APCS of the j^{th} model, $\forall j, j = 1, 2, \ldots, m$, is zero and (6.4) will be true if the variation among the APCS of the j^{th} model, $\forall j, j = 1, 2, \ldots, m$, is close to zero. The literature on model selection shows that for an IC based model selection procedure, none of the existing penalties can satisfy these equations. This is because some IC favour the model with the smallest number of parameters, while others favour the model with the largest number of parameters, as a result the APCS of different models are uneven in all situations. For example, BIC always favours the model with the smallest number of parameters, while RBAR favours the model with the largest number of parameters. Also from our simulation studies, it is apparent that the APCS of different models obtained from the existing criteria are far from equal. The MAPCS obtained from our proposed additive and multiplicative penalties with the SAO technique in Chapter 4 and Chapter 5 are higher than those of the existing criteria. But the APCS of models are generally unequal, though in some cases the variation among the APCS of different models is much less compared to that for the existing IC. So, in order to select the true model without favouring one model over the others, we need a set of penalties which satisfy equation (6.3) ideally or (6.4) to a reasonable extent. We used standard deviation (SD) among the APCS of the models as a measure of variation and applied the SAO technique to find penalties for a particular data set and set of competing models with the objective that the SD among the APCS of the models be the minimum.

The APCS when the the j^{th} model, M_j, is true, is defined in Chapter 3 as

$$\text{APCS}_j = \underset{\theta_j}{\text{E}}\left[\text{P}\left(\text{CS}|M_j,\theta_j\right)\right] = \int \psi\left(\theta_j\right) g\left(\theta_j\right) d\theta_j, j = 1, 2, \ldots, m,$$

where $g\left(\theta_j\right)$ is the weighting density function of the vector of parameters θ_j. Thus the mean of the average probability of correct selection will be

$$\text{MAPCS} = \frac{\sum_{j=1}^{m} \underset{\theta_j}{\text{E}}\left[\text{P}\left(\text{CS}|M_j,\theta_j\right)\right]}{m} = \frac{\sum_{j=1}^{m} \text{APCS}_j}{m}.$$

Then the standard deviation (SD) among the APCS_j, $j = 1, 2, \ldots, m$, can be written as

$$\mathrm{SD(APCS)} = E(\mathrm{APCS} - \mathrm{MAPCS})^2 = \frac{\sum_{j=1}^{m}\left(\mathrm{APCS}_j - \mathrm{MAPCS}\right)^2}{m},$$

which is our objective function to be minimised. In Chapter 3 we explained how APCS can be estimated using the Monte Carlo method with the objective that MAPCS is maximised. Here we apply the same technique but instead of maximising the MAPCS here we minimised the SD(APCS) among the APCS.

We used additive penalties with maximised log-likelihood functions and multiplicative penalties with mean squared error to compute penalties for a particular data set, which satisfy equation (6.3) or (6.4).

6.3 MINIMISATION OF VARIATION AMONG THE APCS FOR MODELS OF UNEQUAL NUMBER OF PARAMETERS

In this section, we have applied the method of minimisation of variation among the APCS, discussed in Section 6.2 with the models of an unequal number of parameters in the competing models in a linear regression setting. The purpose of this section is to evaluate the performance of the method proposed in Section 6.2 over the existing IC and the methods of maximising MAPCS discussed in the Section 4.2 and 4.3. We conducted simulation experiments to provide numerical evidence of the performance of the proposed method, and to compare this method with the existing IC and the methods discussed in Chapter 4. The designs for our simulation experiments are discussed in Section 6.3.1. In Section 6.3.2 we compare the simulation results of this proposed method with the results for the existing IC and the methods discussed in the Section 4.2 and 4.3.

6.3.1 THE DESIGNS OF THE MONTE CARLO SIMULATION EXPERIMENTS

Simulation experiments were conducted to evaluate the performance of the technique of minimisation of variation among the APCS with the models of an unequal number of parameters. The same designs used in Section 4.4 were used for the present simulation experiments. The models used for the present study are also the same models used in Chapter 3. The same data generating process discussed in Section 3.3.1 was also employed here. The initial parameter sets for the SAO technique for this chapter are the same initial parameter sets used with the additive and multiplicative penalties in Chapter 4. The main reason for using the same models, designs, data generating process and initial parameter values is to compare the results of these simulation experiments with the simulation results of the methods discussed in Chapter 4.

6.3.2 RESULTS OF THE MONTE CARLO EXPERIMENTS

The results of the simulation experiments are presented in two series of tables, Tables 6.1a to 6.6a and Tables 6.1b to 6.6b. Each table represents three methods of computing penalties. These are: method 1, existing IC; method 2, maximistion of

MAPCS; and method 3, minimisation of standard deviation (SD) among the APCS. We called the penalties of method 1, Type 1 penalties. In method 2, we have two types of penalties, the additive penalty with maximised log-likelihood, and the multiplicative penalty with mean squared error. In Chapter 4, we called these penalties Type 2 and Type 3 penalties, respectively. In method 3, we also have two types of penalties, additive penalties with maximised log-likelihood functions, and multiplicative penalties with mean squared error. Let us call these penalties Type 4 and Type 5 penalties, respectively. In Table 6.1a to 6.6a, we presented the largest MAPCS obtained using the above mentioned five types of penalties and, in Tables 6.1b to 6.6b, we presented the MAPCS corresponding to the smallest variation among the APCS obtained using Type 1 to Type 5 penalties. Comparative studies of Type 1 and Type 2 penalties, Type 1 and Type 3 penalties, and Type 2 and Type 3 penalties were given in Sections 4.5.1, 4.5.2 and 4.5.3 of Chapter 4, respectively. Simulation results show that MACPS obtained using method 2 (Type 2 and Type 3 penalties) are always large with high variation among the APCS, as compared to method 3 (Type 4 and Type 5 penalties). Therefore, here we will present a comparative study of Type 1 and Type 4 penalties, Type 1 and Type 5 penalties, and Type 4 and Type 5 penalties.

6.3.2.1 COMPARISON OF TYPE 1 AND TYPE 4 PENALTIES

The largest MAPCS obtained from Type 1 to Type 5 penalties along with APCS, standard deviation (SD) among the APCS and the penalties for Designs 4.1 to 4.6 are presented in Table 6.1a to 6.6a. Tables 6.1b to 6.6b contain the APCS and MAPCS with the smallest variation among the APCS along with SD and penalties obtained from Type 1 to Type 5 penalties for Designs 4.1 to 4.6, respectively.

The simulation results show that, in six out of 18 experiments (six designs times three sample sizes), the largest MAPCS obtained from Type 4 penalties is larger than those of BIC (the largest MAPCS among the existing criteria for all designs and sample sizes). The mean, maximum and minimum gain of the MAPCS from these six experiments over BIC are 5.7, 9.2 and 2.7 percent, respectively (Table 6.1a – 6.6b). In the remaining 12 experiments, the largest MAPCS are less than those of BIC and the mean, maximum and minimum loss from these 12 experiments over BIC are 6.0, 12.8 and 0.6 percent, respectively. It is well established that BIC always favours models with a smaller number of parameters, so the APCS of the model with the largest number of parameters is generally much less compared to the APCS of the model with the smallest number of parameters. Consequently the variation among the APCS is high for BIC and in all designs and sample sizes the variation among the APCS obtained from Type 4 penalties is smaller than that of BIC. There are eight experiments, where there is no variability among the APCS obtained from Type 4 penalties, i.e. the APCS of the models are exactly equal.

Simulation results show that in 72 percent (13 out of 18) experiments, the MAPCS corresponding to the smallest variation among the APCS obtained from using Type 4 penalties are higher than those of the MAPCS with the smallest variation among the APCS of the existing IC. In the other experiments, the numerical values of APCS are very close to those of the existing IC (Tables 6.1b – 6.6b) with the smallest variation among the APCS. Among the existing criteria, RBAR and AIC are the dominant criteria, which produce the smallest variation among the APCS. RBAR produces the lowest variation among the APCS in eight experiments, while AIC produces the lowest variation among the APCS in seven experiments. Among the existing IC, there is no experiment where the variation among the APCS is zero, i.e. in the existing criteria there are no experiments which produce equal APCS. But in 89 percent (16 out of 18) of the experiments with a Type 4 penalty, the variation among the APCS is zero, i.e. the APCS of the competing models are equal.

It is observed from the simulation results that the largest MAPCS obtained from the existing IC are generally larger than those of the MAPCS corresponding to the smallest variation among the APCS obtained from Type 4 penalties. But there are three experiments with Type 4 penalties, where the MAPCS corresponding to the smallest variation among the APCS is higher than those of the largest MAPCS obtained from the existing IC (here that of BIC). But in most of the experiments the variation among the APCS is zero or very close to zero, when we consider the MAPCS corresponding to the smallest variation among the APCS with Type 4 penalties. This means that the APCS obtained using the smallest variation among the APCS with Type 4 penalties are equal. The average, maximum and minimum gains of MAPCS corresponding to the smallest variation among the APCS with Type 4 penalties over BIC from these three experiments are 3.7, 6.9 and 1.4 percent, respectively. The mean, maximum and minimum losses of MAPCS from the remaining 15 experiments over BIC are 6.6, 12.0 and 1.2 percent, respectively. This implies that, if we use the MAPCS corresponding to the smallest variation among the APCS with a Type 4 penalty to select the true model, then on average there is a chance of losing approximately six percent in MAPCS compared to the largest MAPCS among the existing IC. But the big gain of using this technique is that none of the models is unduly favoured and the APCS of models is equal, i.e. there is no variation among the APCS of the competing models.

6.3.2.2 COMPARISON OF TYPE 1 AND TYPE 5 PENALTIES

The MAPCS obtained from Type 5 penalties are also presented in Tables 6.1a – 6.6a and Tables 6.1b – 6.6b. From the simulation experiments, it is observed that in seven out of 18 experiments the largest MAPCS obtained using Type 5 penalties are larger than those of the largest MAPCS from the existing criteria (here that of BIC). The mean, maximum and minimum gain of MAPCS obtained from using Type 5 penalties from these seven experiments over BIC are 5.2, 9.1 and 1.2 percent and the losses of MAPCS obtained from using Type 5 penalties from the remaining eleven experiments over BIC are 5.7, 11.6 and 1.0 percent, respectively. But in all experiments, the variations among the APCS obtained from Type 5 penalties are smaller than the variations among the APCS obtained from using BIC. It was observed in the previous section that all of the existing criteria produce some variation in APCS. But there are five experiments with Type 5 penalties, where the variations among the APCS are close to zero, i.e. the APCS of the competing models are almost equal.

The simulation results demonstrate that in 56 percent of the experiments (ten out of 18), the MAPCS corresponding to the smallest variation among the APCS obtained from Type 5 penalties are higher than those of the MAPCS with the smallest variation among the APCS of the existing criteria. As noted in the previous section, among the existing criteria, RBAR and AIC produce the smallest variation among the APCS in 44 and 39 percent of the experiments, respectively. In 67 percent of the experiments (12 out of 18), it is observed that there is no variation among the APCS, i.e. APCS are equal, when we used Type 5 penalties and the smallest variation among the APCS. But for the existing criteria, none of the experiments produces equal APCS.

The simulation results demonstrate that the largest MAPCS obtained from the existing IC are generally larger than those of the MAPCS obtained using Type 5 penalties. But in three experiments, the largest MAPCS obtained from Type 1 penalties are smaller than the MAPCS corresponding to the smallest variation among the APCS obtained from using Type 5 penalties. The average, maximum and minimum gains of MAPCS from these three experiments over BIC (the largest

MAPCS among the existing IC) are 3.5, 7.0 and 1.2 percent, respectively. The mean, maximum and minimum losses of MAPCS obtained from using Type 5 penalties from the remaining 15 experiments are 6.1, 12.0 and 1.2 percent, respectively. But in all experiments, the smallest variations among the APCS obtained from using Type 5 penalties are zero or close to zero, implying that the APCS are equal or close to equal.

6.3.2.3 COMPARISON OF TYPE 4 AND TYPE 5 PENALTIES

From the simulation results we presented in Section 6.3.2.2, it is apparent that the largest MAPCS obtained using Type 4 and Type 5 penalties are very similar for all designs and sample sizes. However, there are some exceptions where Type 5 penalties produced marginally higher MAPCS than those of Type 4 penalties. In fourteen out of 18 experiments (three sample sizes across six designs), the variations among the APCS are higher when using Type 5 penalties compared to using Type 4 penalties, when we consider the largest MAPCS. But in 17 out of 18 experiments, the variation among the APCS obtained from using Type 5 penalties is lower than that obtained from the largest MAPCS of the existing criteria. The differences of variation between Type 1 and Type 5 penalties are larger than those of Type 1 and Type 4 penalties as the variation among APCS obtained from using Type 5 penalties is higher than that when using Type 4 penalties.

In 67 percent of the experiments (twelve out of 18), the MAPCS corresponding to the smallest variation among the APCS of competing models obtained using Type 5 penalties are the same as those obtained from using Type 4 penalties. In other experiments, the MAPCS obtained from using Type 5 penalties are very similar to those obtained from using Type 4 penalties.

6.4 MINIMISATION OF VARIATION AMONG THE APCS FOR MODELS OF EQUAL NUMBER OF PARAMETERS

In Section 6.3, we discussed the issue of minimisation of variation among the APCS of the models, when the competing models have an unequal number of parameters. In this section, we investigate the special case of an equal number of parameters among the competing models. We consider this special case because for the existing IC, when the competing models have the same number of parameters, there is no need to use any penalty because the penalties cancel out. Thus, the problem reduces to selecting the model with the largest maximised log-likelihood because all existing penalty functions are a function of sample size and number of free parameters. In Chapter 5, we have shown that the application of the SAO technique with two different types of penalties (additive with maximised log-likelihood and multiplicative with mean squared error) has a great effect on selecting the true model with higher MAPCS, when the competing models have an equal number of parameters. It is also shown that the penalties which maximise the MAPCS are not zero although the competing models have the same number of parameters. But like the existing criteria, the APCS of models obtained using the SAO technique with these penalties are not equal and sometimes far from equal. So here we apply the idea of Section 6.2 to find the MAPCS with the objective that the variation among the APCS is as small as possible; essentially, our goal is to make the variation zero or close to zero. We apply the SAO technique to find the penalties in such a way that the SD among the APCS is the minimum. We compare these results with those obtained using the methods discussed in Sections 4.2 and 4.3, and the existing cited criteria are presented in Section 5.3 of Chapter 5.

The plan of this section is as follows. In Section 6.4.1, we discuss the design of the simulation experiments, while the simulation results are reported in Section 6.4.2.

6.4.1 THE DESIGNS OF THE MONTE CARLO SIMULATION EXPERIMENTS

The purpose of the simulation experiments of this section is to evaluate the performance of the technique discussed in Section 6.2, when the competing models have the same number of parameters to be estimated. The same designs used in Chapter 5 are used here to compute the MAPCS with the objective of minimisation of SD among the APCS. We used the same models, data generating processes and the initial parameter values for the SAO technique used in Chapter 5 for our present study. The reason for keeping everything the same for the simulation experiments, is to compare the results of these experiments with those obtained using the methods discussed in Chapter 5.

6.4.2 RESULTS OF THE MONTE CARLO EXPERIMENTS

Like Section 6.3, this section also contains two series of tables. The series 'a' contain the largest MAPCS obtained using Type 1 to Type 5 penalties, while series 'b' contain the MAPCS corresponding to the smallest variation among the APCS obtained using Type 1 to Type 5 penalties. The results of these experiments are presented in Tables 6.7a to 6.12a and Tables 6.7b to 6.12b, respectively. In each table, there are three methods and five penalty types as mentioned in Section 6.3.2. A comparative study of Type 1 and Type 2 penalties, Type 1 and Type 3 penalties, and Type 2 and Type 3 penalties was given in Sections 5.3.1, 5.3.2 and 5.3.3, respectively. Simulation results show that MACPS obtained using Type 4 and Type 5 penalties are always smaller with zero or close to zero variation among the APCS compared to Type 2 and Type 3 penalties, respectively. Therefore, in this section we will present the comparative study of Type 1 and Type 4 penalties, Type 1 and Type 5 penalties, and Type 4 and Type 5 penalties.

6.4.2.1 COMPARISON OF TYPE 1 AND TYPE 4 PENALTIES

In this section, our concern is with competing models with an equal number of parameters in linear regression settings. As mentioned earlier, in this situation the MAPCS and APCS obtained from all the existing IC are the same and the relative penalties are zero. But the MAPCS with minimum SD among the APCS obtained from using Type 4 penalties are different; and relative penalties are different from zero in all designs and sample sizes. We mentioned earlier that the MAPCS obtained using Type 2 and Type 3 penalties are always higher than those obtained from using Type 1 penalties, i.e. from the existing IC with sometimes higher variation among the APCS compared to Type 1 penalties. Here the MAPCS obtained from using Type 4 penalties are also always higher than those for Type 1 penalties for all designs and sample sizes and with smaller variation among the APCS. But the MAPCS obtained from using Type 4 penalties are always less than those of Type 2 penalties with less variation among the APCS, and in most cases the variation among the APCS are zero or close to zero, which implies that the APCS of different competing models are equal.

For Design 5.1, in respect to all sample sizes and combinations of initial parameters, the MAPCS obtained from Type 4 penalties are always higher than those of Type 1 penalties with less variation among the APCS compared to the existing criteria (Table 6.7a). The largest MAPCS obtained from using Type 4 penalties are 9.7, 13.1 and 15.1 percent higher than those of existing IC with 33.6, 44.8 and 54.5 percent less variability among the APCS for the sample sizes 20, 50 and 100, respectively. We mentioned earlier that all existing criteria are the same when selecting from the competing models that have the same number of parameters, so the largest MAPCS obtained from the existing IC, is the same as that obtained with the smallest variation among the APCS. For n = 20, 50 and 100, the MAPCS with the smallest variation among the APCS obtained using Type 4 penalties are 6.4, 9.2 and 12.4 percent higher than those of the existing criteria with 45.5, 54.5 and 59.6 percent less variability among the APCS (Table 6.7b). The gaps between the largest MAPCS and smallest MAPCS obtained from using Type 4 penalties are small and the smallest MAPCS is always higher than that of the existing IC. For n = 20, 50 and 100, the smallest MAPCS obtained using Type 4 penalties are 6.4, 9.2 and 12.4 percent higher than those of the existing IC with 45.5, 54.1 and 59.6 less variability among the APCS, respectively.

For Design 5.2, the largest MAPCS obtained from using Type 4 penalties are 3.2, 3.1 and 2.9 percent higher with 99.7, 96.0 and 87.8 percent smaller variation among the APCS than those of existing criteria for sample sizes 20, 50 and 100, respectively (Table 6.8a). For n = 20, 50 and 100, the MAPCS with the smallest variation among the APCS obtained using Type 4 penalties are 4.0, 3.0 and 2.2 percent higher than those of the existing criteria, respectively (Table 6.8b). For all sample sizes, there is no variation among the APCS, which implies that the APCS of the different models are exactly equal. The MAPCS obtained from using Type 2 penalties and Type 3 penalties are very close to those obtained from using Type 4 penalties. The smallest MAPCS obtained using Type 4 penalties are 3.1, 3.0 and 1.5 percent higher than those obtained from the existing criteria with no variation among the APCS for sample sizes 20, 50 and 100, respectively.

For Design 5.3, the largest MAPCS obtained using Type 4 penalties for the sample sizes 20, 50 and 100 are 14.0, 6.9 and 6.1 percent higher than those of the existing criteria, with 63.3, 82.1 and 77.2 percent less variability among the APCS than those of the existing criteria, respectively (Table 6.9a). It is interesting to note that for all sample sizes, the MAPCS obtained using Type 2 and Type 3 penalties are higher than those obtained from using the existing criteria and are very close to those obtained from using Type 4 penalties. But the variation among the APCS for Type 2 penalties and Type 3 penalties are higher than those obtained from Type 1 penalties, and even higher than those obtained from using Type 4 penalties. The MAPCS with the smallest variation among the APCS obtained using Type 4 penalties are 10.6, 5.8 and 4.0 percent higher than those obtained from using Type 4 penalties for n = 20, 50 and 100, respectively (Table 6.9b). The variation among the APCS obtained using Type 4 penalties is zero for n = 50 and almost zero for the other two sample sizes. This implies that the APCS of the different models are equal. The MAPCS with the smallest variation among the APCS obtained from using Type 2 and Type 3 penalties are higher than those obtained from the existing criteria and close to those obtained from using Type 4 penalties. But the variations among the APCS obtained using Type 2 penalties and Type 3 penalties are always higher than those obtained from using the existing criteria and are much higher than that obtained from using Type 4 penalties. In this design there are 13.55, 8.93 and 29.8 percent of the cases where MAPCS is less than that of the existing criteria for n = 20, 50 and 100, respectively.

For Design 5.4, the largest MAPCS obtained using Type 4 penalties are 4.6, 5.9 and 7.0 percent higher than those obtained from the existing criteria (Table 6.11a). The variability among the APCS obtained from using Type 4 penalties are almost zero for n = 20 and 50, and 74.5 percent less than that of the existing criteria for the sample size 100. The largest MAPCS with the smallest variation among the APCS obtained using Type 4 penalties are 4.5, 5.8 and 6.4 percent higher than those obtained from the existing criteria (Table 6.10b). The variations among the APCS obtained using Type 4 penalties are zero for all sample sizes, which indicates the APCS of the different models are exactly equal. The MAPCS with smallest variation obtained using Type 2 penalties and Type 3 penalties are higher than those obtained from the existing criteria and are marginally higher than those obtained from using Type 4 penalties. But the variation among the APCS is the least for Type 4 penalties compared to Type 2 penalties, Type 3 penalties and the existing criteria. The smallest MAPCS obtained using Type 4 penalties are 4.4, 5.8 and 6.0 percent higher than those obtained using the existing criteria with 99.6, 100 and 94.8 percent less variation among the APCS for sample sizes 20, 50 and 100, respectively.

For Design 5.5, the largest MAPCS and the MAPCS corresponding to the smallest variation among the APCS obtained using Type 4 penalties are the same for all sample sizes and are 20.0, 29.1 and 35.6 percent higher than those obtained from the existing criteria for n = 20, 50 and 100, respectively (Tables 6.11a and 6.11b). The variations among the APCS obtained using Type 4 penalties are equal to zero for all sample sizes, implying that the APCS of the models are equal. The MAPCS obtained using Type 2 penalties and Type 3 penalties are bigger than those obtained from using Type 4 penalties, but the variation among the APCS obtained from using Type 2 penalties and Type 3 penalties are also large compared to those of Type 4 penalties. The smallest MAPCS obtained using Type 4 penalties are 20.0, 28.8 and 35.3 percent higher than those obtained using the existing criteria with 100, 98.2 and 98.9 percent less variation among the APCS for sample sizes 20, 50 and 100, respectively.

Like for Design 5.5, for Design 5.6, the largest MAPCS and the MAPCS with the smallest variation among the APCS obtained using Type 4 penalties are the same for all sample sizes, and are 6.7, 5.1 and 3.8 percent higher than those obtained from the existing criteria for n = 20, 50 and 100, respectively (Table 6.12a, 6.12b). The variations among the APCS obtained using Type 4 penalties are equal to zero for all sample sizes, which implies that the APCS of the different models are equal. The MAPCS obtained using Type 2 and Type 3 penalties are larger than those obtained from using Type 4 penalties, but the variation among the APCS obtained from Type 2 and Type 3 penalties are also large compared to those for Type 4 penalties. The smallest MAPCS obtained using Type 4 penalties are 6.7, 5.0 and 3.7 percent higher than those obtained using the existing criteria with 99.2, 100, 98.9 percent less variation among the APCS for sample sizes 20, 50 and 100, respectively.

6.4.2.2 COMPARISON OF TYPE 4 AND TYPE 5 PENALTIES

In this section, we compare the MAPCS and the variation among the APCS of models chosen using Type 4 penalties and Type 5 penalties. Simulation results demonstrate that in 33 percent of the experiments, the largest MAPCS and the variation among the APCS of models chosen using Type 4 penalties and Type 5 penalties are exactly equal and in other experiments these values are very close (Tables 6.7a to 6.12b). In 67 percent of the experiments, the MAPCS obtained using the smallest variation among the APCS and the variation among the APCS are exactly equal and in the remaining experiments these values are very close. This means that there is no significant effect of the type of penalties on the estimated MAPCS of the model and the variation among the APCS of the competing models. So either of the penalty types can be used with the SAO technique to estimate the MAPCS to select the true model from a set of equi-dimensional competing alternative models.

6.4.2.3 COMPARISON OF TYPE 1 AND TYPE 5 PENALTIES

It is apparent from the results of Section 6.4.2.2 that the MAPCS and the variation among the APCS obtained using Type 4 and Type 5 penalties are similar for almost all designs and sample sizes. So the comparison between Type 1 and Type 5 penalties is very similar to the comparison between Type 1 and Type 4 penalties given in Section 6.4.2.1.

6.5 CONCLUDING REMARKS

The purpose of this chapter was to develop a technique to select the correct model from a set of competing alternative models of both unequal and equal numbers of parameters, in such a way that all competing models have an equal chance of being selected. Here we applied the SAO technique to find the penalties to select the correct model, with the objective that the variation among the APCS of the different models is as small as possible. We used SD among the APCS as a measure of variation and applied the SAO technique to find the minimum value of the SD, to determine the penalties to select the correct model.

The results of the simulation experiments for competing models with an unequal number of parameters show that in six out of 18 experiments, the MAPCS is on average 5.7 percent higher than the largest MAPCS obtained using the existing criteria (here BIC), while for the remaining experiments, MAPCS is on average 6.0 percent lower compared to that of the existing criteria. But in all experiments the variation among the APCS obtained from Type 4 and Type 5 penalties are less than that of BIC.

In 13 out of 18 experiments, the MAPCS corresponding to the smallest variation among the APCS are higher than those obtained from the corresponding lowest variation among the APCS in the existing IC (here RBAR and AIC). From the results of the simulation experiments, it is evident that if the MAPCS corresponding to the smallest variation among the APCS is used to select the best model, then, on average MAPCS is approximately six percent lower than the largest MAPCS in the existing IC (here, that of BIC). Generally, for the existing criteria, the APCS of the competing models are far from equal, but the APCS obtained using Type 4 and Type 5 penalties are often equal or close to equal, which is the best feature of this technique.

The simulation results demonstrate that for models with an equal number of parameters in 100 percent of the combinations of initial parameter values for the SAO technique, the MAPCS obtained from using Type 4 and Type 5 penalties are higher than those of the existing IC for all designs with an exception of Design 5.3, where these figures were on average 87 and 83 percent for Type 4 and Type 5 penalties, respectively. For all designs except for Design 5.3, and in 100 percent of the combinations of initial parameter values for the SAO technique, the variations among the APCS obtained from using Type 4 and Type 5 penalties are less than those obtained from the existing criteria. For Design 5.3, these figures are on average 92 and 88 percent for Type 4 and Type 5 penalties, respectively. There are three experiments with Type 4 penalties and four experiments with Type 5 penalties, where the APCS are equal in 100 percent of the combinations of initial parameter sets for the SAO technique.

It is apparent from the simulation results that for a particular sample size and design, the MAPCS obtained from the different combinations of initial parameter values for the SAO technique are very similar. In five out of six designs and for three different sample sizes, the MAPCS and variation among the APCS obtained using different initial parameter values are almost equal. This indicates that the MAPCS obtained from using Type 4 and Type 5 penalties are generally insensitive to the initial parameter values of the SAO technique. From our simulation results. it may be concluded that the MAPCS obtained from Type 4 and Type 5 penalties will always be higher than those of the existing criteria with no variation among the APCS or variation among the APCS close to zero. This means that for equi-dimensional competing alternative models, the application of the SAO technique with Type 4 or Type 5 penalties is the best way of selecting the true model without favouring any of the competing models.

Table 6.1a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 4.1 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (BIC)	0.8225	0.8815	0.1150	0.1345	0.4884	0.4206	1.4979	1.4979	2.9957							
Maximisation of MAPCS	Type 2	0.6245	0.5435	0.4550	0.5080	0.5328	0.0712	13.0759	0.4639	13.4040	0.0000	1.4979	1.4979	2.9957	-12.0000	12.0000	0.0100
	Type 3	0.6485	0.5400	0.4325	0.5120	0.5333	0.0893	3.1681	0.9959	3.0906	1.0000	1.0000	1.0000	1.0000	0.0000	3.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5300	0.5270	0.5300	0.5210	0.5270	0.0042	13.8227	0.3070	14.1300	0.0000	1.4979	1.4979	2.9957	-12.0000	12.0000	0.0001
	Type 5	0.5315	0.5245	0.5320	0.5230	0.5278	0.0047	3.2600	0.9768	3.1741	1.0000	2.0000	2.0000	3.0000	0.0000	12.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (BIC)	0.9260	0.9450	0.2840	0.0800	0.5587	0.4430	1.9560	1.9560	3.9120							
Maximisation of MAPCS	Type 2	0.8770	0.5170	0.5935	0.5285	0.6290	0.1687	8.9402	1.2482	9.2081	0.0000	0.0000	0.0000	0.0000	0.0000	12.0000	0.1000
	Type 3	0.8770	0.5290	0.5930	0.5145	0.6284	0.1692	1.6936	1.0293	1.6774	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.6765	0.5185	0.7185	0.5270	0.6101	0.1024	9.2764	0.5157	9.5453	0.0000	1.0000	1.0000	2.0000	0.0000	12.0000	0.0100
	Type 5	0.6795	0.5280	0.7150	0.5155	0.6095	0.1025	1.1903	1.0001	1.1788	1.0000	2.0000	2.0000	3.0000	0.0000	16.0000	0.0100
Sample size = 100																	
Existing IC	Type 1 (BIC)	0.9495	0.9580	0.5190	0.2095	0.6590	0.3631	2.3026	2.3026	4.6052							
Maximisation of MAPCS	Type 2	0.9340	0.7990	0.7380	0.4635	0.7336	0.1978	15.2409	1.7396	16.0419	0.0000	0.0000	0.0000	0.0000	-12.0000	12.0000	0.1000
	Type 3	0.9365	0.8040	0.7355	0.4570	0.7332	0.2022	1.2291	1.0257	1.2366	2.3026	4.6052	4.6052	6.9078	0.0000	23.0300	0.1000
Minimisation of SD among APCS	Type 4	0.7515	0.6320	0.8375	0.5955	0.7041	0.1111	6.5959	0.6777	6.9833	0.0000	2.3026	2.3026	4.6052	-10.0000	10.0000	0.0001
	Type 5	0.7415	0.6305	0.8415	0.5980	0.7029	0.1110	1.3866	1.0031	1.3831	1.0000	2.0000	2.0000	3.0000	0.0000	12.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.1b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 4.1 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (RBAR)	0.4710	0.6790	0.2195	0.3695	0.4347	0.1928	0.5407	0.5407	1.1123							
Maximisation of MAPCS	Type 2	0.5535	0.5440	0.4845	0.5080	0.5225	0.0320	2.6281	0.3685	2.9563	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.1000
	Type 3	0.5845	0.5460	0.4940	0.5055	0.5325	0.0412	2.2897	0.9852	2.2352	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5220	0.5220	0.5225	0.5220	0.5221	0.0002	3.7796	0.3072	4.0844	0.0000	1.0000	1.0000	2.0000	-10.0000	10.000	0.0010
	Type 5	0.5215	0.5235	0.5220	0.5240	0.5228	0.0012	1.3312	0.9768	1.2960	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (RBAR)	0.5145	0.6660	0.3070	0.3660	0.4634	0.1608	0.5155	0.5155	1.0418							
Maximisation of MAPCS	Type 2	0.7765	0.5180	0.6275	0.5275	0.6124	0.1201	2.6381	0.8357	2.9068	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.0100
	Type 3	0.8730	0.5290	0.5910	0.5145	0.6269	0.1674	1.1923	1.0294	1.1809	1.9560	3.9120	3.9120	5.8680	0.0000	19.5601	0.0010
Minimisation of SD among APCS	Type 4	0.5225	0.5225	0.5225	0.5225	0.5225	0.0000	0.8311	0.4552	1.1052	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.1000
	Type 5	0.5205	0.5205	0.5205	0.5205	0.5205	0.0000	1.0126	0.9975	1.0025	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 100																	
Existing IC	Type 1 (RBAR)	0.5380	0.6935	0.4825	0.5435	0.5644	0.0904	0.5076	0.5076	1.0204							
Maximisation of MAPCS	Type 2	0.8450	0.7530	0.7405	0.5040	0.7106	0.1454	2.3308	1.1201	2.9778	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.0010
	Type 3	0.9365	0.7850	0.7355	0.4745	0.7329	0.1923	3.7588	1.0257	3.7764	2.3026	4.6052	4.6052	6.9078	0.0000	13.8155	0.0010
Minimisation of SD among APCS	Type 4	0.6095	0.6095	0.6095	0.6095	0.6095	0.0000	0.9096	0.5489	1.2657	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.6095	0.6095	0.6095	0.6095	0.6095	0.0000	1.0081	1.0009	1.0050	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.2a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 96 for Design 4.2 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (BIC)	0.7970	0.6850	0.3485	0.1700	0.5001	0.2911	1.4979	1.4979	2.9957							
Maximisation of MAPCS	Type 2	0.8015	0.7255	0.3975	0.1095	0.5085	0.3186	1.6079	1.3140	3.6832	0.0000	0.0000	0.0000	0.0000	-12.0000	12.0000	0.1000
	Type 3	0.8015	0.7255	0.3975	0.1095	0.5085	0.3186	1.1125	1.0804	1.2930	1.4979	2.9957	2.9957	4.4936	0.0000	14.9790	0.1000
Minimisation of SD among APCS	Type 4	0.4400	0.4400	0.4400	0.4400	0.4400	0.0000	0.6965	0.4446	0.9954	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5320	0.3830	0.5615	0.3400	0.4541	0.1091	1.0963	0.9898	1.0629	1.4979	2.9957	2.9957	4.4936	0.0000	4.4940	0.0100
Sample size = 50																	
Existing IC	Type 1 (BIC)	0.9095	0.8830	0.7100	0.4530	0.7389	0.2101	1.9560	1.9560	3.9120							
Maximisation of MAPCS	Type 2	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	3.3296	2.8476	5.2849	0.0000	0.0000	0.0000	0.0000	0.0000	12.0000	0.1000
	Type 3	0.9745	0.8650	0.7185	0.4210	0.7448	0.2400	1.1194	1.0978	1.1852	1.0000	1.0000	1.0000	1.0000	0.0000	3.0000	0.1000
Minimisation of SD among APCS	Type 4	0.6530	0.6530	0.6555	0.6505	0.6530	0.0020	0.9932	0.7299	1.5666	0.0000	1.9560	1.9560	3.9120	0.0000	5.8680	0.0001
	Type 5	0.6530	0.6530	0.6555	0.6505	0.6530	0.0020	1.0193	1.0087	1.0212	1.9560	3.9120	3.9120	5.8680	0.0000	19.5600	0.0001
Sample size = 96																	
Existing IC	Type 1 (BIC)	0.9390	0.9445	0.9160	0.8255	0.9063	0.0552	2.2822	2.2822	4.5643							
Maximisation of MAPCS	Type 2	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	4.7801	4.1070	7.6732	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Type 3	0.9940	0.9530	0.9550	0.7920	0.9235	0.0897	1.0931	1.0778	1.1486	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.0100
Minimisation of SD among APCS	Type 4	0.8650	0.8650	0.8650	0.8645	0.8649	0.0002	1.6006	1.4843	2.9090	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Type 5	0.8650	0.8650	0.8650	0.8645	0.8649	0.0002	1.0230	1.0206	1.0401	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.2b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 96 for Design 4.2 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (RBAR)	0.4505	0.5775	0.4190	0.3785	0.4564	0.0860	0.5407	0.5407	1.1123							
Maximisation of MAPCS	Type 2	0.7940	0.6075	0.3635	0.2465	0.5029	0.2455	1.5974	1.3141	2.5162	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.0001
	Type 3	0.8015	0.7255	0.3975	0.1095	0.5085	0.3186	1.1125	1.0804	1.2930	1.4979	2.9957	2.9957	4.4936	0.0000	14.9790	0.1000
Minimisation of SD among APCS	Type 4	0.4400	0.4400	0.4400	0.4400	0.4400	0.0000	0.6365	0.4446	0.9954	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.4400	0.4400	0.4400	0.4400	0.4400	0.0000	1.0158	0.9905	0.9885	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (AIC)	0.7190	0.7780	0.6655	0.5925	0.6887	0.0789	1.0000	1.0000	2.0000							
Maximisation of MAPCS	Type 2	0.8690	0.8165	0.6840	0.5460	0.7289	0.1446	1.7732	1.6949	2.9996	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.0010
	Type 3	0.9635	0.8740	0.6785	0.4600	0.7440	0.2236	1.0922	1.0988	1.1551	1.0000	1.0000	1.0000	1.0000	0.0000	12.0000	0.0001
Minimisation of SD among APCS	Type 4	0.6520	0.6520	0.6520	0.6520	0.6520	0.0000	0.9815	0.7386	1.5528	0.0000	0.0000	0.0000	0.0000	0.0000	12.0000	0.1000
	Type 5	0.6520	0.6520	0.6520	0.6520	0.6520	0.0000	1.0188	1.0090	1.0206	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 96																	
Existing IC	Type 1 (JIC)	0.8375	0.8745	0.8595	0.8625	0.8585	0.0154	1.3951	1.3951	2.7928							
Maximisation of MAPCS	Type 2	0.8650	0.8805	0.8820	0.8600	0.8719	0.0110	1.5598	1.4537	2.9997	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.1000
	Type 3	0.9970	0.9395	0.9480	0.8070	0.9229	0.0813	1.1118	1.0960	1.1569	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.0001
Minimisation of SD among APCS	Type 4	0.8645	0.8645	0.8645	0.8645	0.8645	0.0000	1.5924	1.4832	2.8994	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.1000
	Type 5	0.8645	0.8645	0.8645	0.8645	0.8645	0.0000	1.0229	1.0205	1.0399	1.0000	1.0000	1.0000	1.0000	0.0000	3.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.3a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 4.3 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (BIC)	0.7815	0.6210	0.5750	0.4205	0.5995	0.1486	1.4979	1.4979	2.9957							
Maximisation of MAPCS	Type 2	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	2.1992	1.8608	3.6479	0.0000	0.0000	0.0000	0.0000	0.0000	12.0000	0.1000
	Type 3	0.8740	0.5745	0.5590	0.4110	0.6046	0.1941	1.1806	1.1412	1.2887	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5605	0.5605	0.5605	0.5605	0.5605	0.0000	0.8391	0.7869	1.5818	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5605	0.5605	0.5605	0.5610	0.5606	0.0003	1.0300	1.0249	1.0476	5.3469	2.9957	2.9957	4.4936	0.0000	8.9870	0.0001
Sample size = 50																	
Existing IC	Type 1 (BIC)	0.8985	0.7660	0.6980	0.5605	0.7308	0.1408	1.9560	1.9560	3.9120							
Maximisation of MAPCS	Type 2	0.9060	0.7665	0.7035	0.5560	0.7330	0.1452	2.0229	1.9573	4.0287	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.9065	0.7665	0.7035	0.5560	0.7331	0.1454	1.0624	1.0593	1.1271	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.6760	0.6760	0.6755	0.6760	0.6759	0.0003	1.0194	0.8441	1.8413	0.0000	1.9560	1.9560	3.9120	0.0000	5.8680	0.0001
	Type 5	0.7510	0.6510	0.8270	0.5465	0.6939	0.1219	1.1105	1.0087	1.1363	16.9521	2.0000	2.0000	3.0000	0.0000	27.0000	0.0001
Sample size = 100																	
Existing IC	Type 1 (BIC)	0.9315	0.8325	0.7580	0.6375	0.7899	0.1240	2.3026	2.3026	4.6052							
Maximisation of MAPCS	Type 2	0.8905	0.8330	0.8165	0.6305	0.7926	0.1126	2.1758	1.5387	4.2676	0.0000	0.0000	0.0000	0.0000	0.0000	12.0000	01000
	Type 3	0.8950	0.8320	0.8175	0.6270	0.7929	0.1156	1.0350	1.0213	1.0692	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.7465	0.7465	0.7465	0.7465	0.7465	0.0000	1.2720	1.0431	2.2346	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.7790	0.7330	0.8735	0.6465	0.7580	0.0946	1.0514	1.0058	1.0663	19.5147	2.0000	2.0000	3.0000	0.0000	27.0000	0.0010

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.3b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 4.3 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_5	Mean	SD	P_2	P_3	P_5	S_1	S_2	S_3	S_5	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (AIC)	0.6355	0.5995	0.5725	0.5085	0.5790	0.0536	1.0000	1.0000	2.0000							
Maximisation of MAPCS	Type 2	0.8000	0.6100	0.5330	0.4590	0.6005	0.1466	1.5052	1.6540	2.8529	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.0100
	Type 3	0.8315	0.5915	0.5590	0.4330	0.6038	0.1665	1.1389	1.1221	1.2321	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.0010
Minimisation of SD among APCS	Type 4	0.5605	0.5605	0.5605	0.5605	0.5605	0.0000	0.8391	0.7869	1.5818	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	).5605	0.5605	0.5605	0.5605	0.5605	0.0000	1.0303	1.0249	1.0481	0.9747	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (AIC)	0.7050	0.7140	0.6610	0.6610	0.6853	0.0282	1.0000	1.0000	2.0000							
Maximisation of MAPCS	Type 2	0.8005	0.7780	0.7010	0.5930	0.7181	0.0937	1.3051	1.4416	2.9445	0.0000	0.0000	0.0000	0.0000	0.0000	3.0000	0.1000
	Type 3	0.9015	0.7630	0.7075	0.5590	0.7328	0.1417	1.0624	1.0568	1.1244	1.0000	1.0000	1.0000	1.0000	0.0000	36.0000	0.0001
Minimisation of SD among APCS	Type 4	0.6755	0.6755	0.6755	0.6755	0.6755	0.0000	1.0207	0.8426	1.8425	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.6755	0.6755	0.6755	0.6755	0 6755	0.0000	1.0202	1.0132	1.0325	0.9623	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 100																	
Existing IC	Type 1 (HOC)	0.7090	0.7710	0.7175	0.7470	0.7361	0.0284	1 0204	1 0204	2 0514							
Maximisation of MAPCS	Type 2	0.8485	0.7945	0.7700	0.7065	0.7799	0.0589	1.6293	1.4306	2.9826	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.0010
	Type 3	0.8895	0.8160	0.8235	0.6400	0.7923	0.1067	1.0366	1.0193	1.0655	1.0000	1.0000	1.0000	1.0000	0.0000	12.0000	0.0001
Minimisation of SD among APCS	Type 4	0.7465	0.7465	0.7465	0.7465	0.7465	0.0000	1.2720	1.0431	2.2346	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.7465	0.7465	0.7465	0.7465	0.7465	0.0000	1.0157	1.0107	1.0247	0.9755	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.4a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 4.4 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Sample size = 20																		
Existing IC	Type 1 (BIC)	0.7240	0.7735	0.2325	0.5660	0.2300	0.6555	0.1855	0.2420	0.4511	0.2519	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936
Maximisation of MAPCS	Type 2	0.6815	0.7115	0.4390	0.5095	0.3670	0.5375	0.3600	0.3455	0.4939	0.1433	4.7017	0.7324	1.5757	5.5869	6.5864	2.2801	7.4196
	Type 3	0.6855	0.7110	0.4445	0.5230	0.3670	0.5375	0.3375	0.3455	0.4939	0.1472	1.5420	1.0195	1.1138	1.5910	1.7583	1.1374	1.7988
Minimisation of SD among APCS	Type 4	0.4695	0.5045	0.4695	0.4630	0.5150	0.4350	0.4590	0.4405	0.4695	0.0280	4.3512	0.3881	0.8700	4.7285	6.0333	1.2931	6.4884
	Type 5	0.4675	0.5065	0.4675	0.4665	0.5515	0.4080	0.4655	0.4060	0.4674	0.0477	2.1299	0.9851	1.0306	2.0844	2.5428	1.0150	2.5092
Sample size = 50																		
Existing IC	Type 1 (BIC)	0.8535	0.8860	0.3410	0.7820	0.3020	0.8140	0.3325	0.3125	0.5779	0.2755	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680
Maximisation of MAPCS	Type 2	0.7875	0.7685	0.5530	0.7260	0.5260	0.7060	0.4620	0.4450	0.6218	0.1402	4.9667	0.9045	2.4313	5.7810	7.7208	3.5824	8.7929
	Type 3	0.7900	0.7660	0.5560	0.7310	0.5255	0.7060	0.4555	0.4440	0.6218	0.1419	1.5637	1.0156	1.0808	1.5819	1.7095	1.1099	1.7465
Minimisation of SD among APCS	Type 4	0.6620	0.6215	0.6740	0.5620	0.6290	0.5520	0.5505	0.5415	0.5991	0.0538	5.9632	0.4816	3.7861	6.3914	8.6105	4.3210	9.1502
	Type 5	0.5860	0.5820	0.5875	0.5940	0.5820	0.5790	0.6070	0.5595	0.5846	0.0135	1.5821	0.9988	1.0272	1.5767	1.6408	1.0259	1.6409
Sample size = 100																		
Existing IC	Type 1 (BIC)	0.9100	0.9230	0.5005	0.8325	0.4215	0.8740	0.4630	0.4610	0.6732	0.2288	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078
Maximisation of MAPCS	Type 2	0.8955	0.8545	0.6525	0.7545	0.5735	0.7690	0.5975	0.6070	0.7130	0.1230	7.5854	1.4027	3.3902	8.8119	10.4567	4.5752	11.5769
	Type 3	0.8925	0.8640	0.6560	0.7540	0.5635	0.7810	0.5985	0.5930	0.7128	0.1279	1.2124	1.0175	1.0594	1.2312	1.2709	1.0736	1.2888
Minimisation of SD among APCS	Type 4	0.7265	0.6915	0.7325	0.6795	0.6705	0.6825	0.6665	0.6820	0.6914	0.0247	6.5278	0.6945	2.5233	7.1803	8.7208	3.2896	9.4218
	Type 5	0.7300	0.7000	0.7615	0.6390	0.7030	0.6600	0.6245	0.6400	0.6822	0.0490	1.1216	1.0027	1.1164	1.1237	1.2743	1.1221	1.2825

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.4a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 4.4 together with relative penalty values and input values of SAO technique (continued).

Method	Penalty type	Input values for simulated annealing										TRF**
		Starting values of penalties								Boundaries		
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	8.0000	0.1000
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.0001
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-10.0000	10.0000	0.0001
	Type 5	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.1000
Sample size = 50												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	-8.0000	8.0000	0.0010
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	20.0000	0.0001
Minimisation of SD among APCS	Type 4	0.0000	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680	0.0000	15.6481	0.0001
	Type 5	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	20.0000	0.1000
Sample size = 100												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078	0.0000	18.4207	0.1000
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.0001
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.0100
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.4b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 4.4 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Sample size = 20																		
Existing IC	Type 1 (RBAR)	0.3175	0.4295	0.2200	0.3775	0.3450	0.5635	0.2735	0.4380	0.3706	0.1075	0.5407	0.5407	0.5407	1.1123	1.1123	1.1123	1.7185
Maximisation of MAPCS	Type 2	0.5190	0.6290	0.4905	0.5150	0.4040	0.5270	0.3115	0.3915	0.4734	0.0992	1.8258	0.4346	1.2583	2.5177	3.3279	2.0440	3.9924
	Type 3	0.6835	0.6650	0.4420	0.5050	0.4235	0.5310	0.3555	0.3395	0.4931	0.1295	1.4543	1.0194	1.1142	1.4753	1.6746	1.1288	1.7131
Minimisation of SD among APCS	Type 4	0.4635	0.4640	0.4630	0.4630	0.4635	0.4635	0.4630	0.4635	0.4634	0.0004	4.7577	0.3912	0.8292	5.1517	5.8582	1.2480	6.3144
	Type 5	0.4625	0.4605	0.4625	0.4625	0.4605	0.4585	0.4665	0.4590	0.4616	0.0025	2.4383	0.9849	1.0280	2.3949	2.5728	1.0122	2.5341
Sample size = 50																		
Existing IC	Type 1 (RBAR)	0.3215	0.4665	0.2890	0.4230	0.3965	0.6115	0.4105	0.5760	0.4368	0.1124	0.5155	0.5155	0.5155	1.0418	1.0418	1.0418	1.5795
Maximisation of MAPCS	Type 2	0.5950	0.6810	0.5380	0.6675	0.5035	0.7055	0.4875	0.4905	0.5836	0.0910	1.7772	0.6178	1.2365	2.5019	3.1335	2.0651	3.9976
	Type 3	0.7355	0.7560	0.6030	0.7180	0.5195	0.7175	0.4675	0.4485	0.6207	0.1276	1.2183	1.0081	1.0804	1.2324	1.3074	1.1060	1.3357
Minimisation of SD among APCS	Type 4	0.5710	0.5705	0.5710	0.5720	0.5705	0.5690	0.5740	0.5695	0.5709	0.0015	2.2709	0.4778	1.2709	2.7072	3.5602	1.7593	4.0740
	Type 5	0.5860	0.5820	0.5875	0.5940	0.5820	0.5790	0.6070	0.5595	0.5846	0.0135	1.5821	0.9988	1.0272	1.5767	1.6408	1.0259	1.6409
Sample size = 100																		
Existing IC	Type 1 (AIC/MCP)	0.5980	0.7005	0.4750	0.6650	0.5105	0.7875	0.5170	0.6320	0.6107	0.1070	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
Maximisation of MAPCS	Type 2	0.6770	0.7735	0.6115	0.7580	0.5480	0.7165	0.5820	0.6775	0.6680	0.0817	1.6096	0.9039	1.2082	2.6533	3.2269	2.3071	3.9943
	Type 3	0.8955	0.8590	0.6525	0.7540	0.5810	0.7235	0.5950	0.6060	0.7088	0.1206	2.7717	1.0180	1.0595	2.8115	3.0264	1.0737	3.0584
Minimisation of SD among APCS	Type 4	0.6825	0.6855	0.6825	0.6830	0.6845	0.6795	0.6835	0.6795	0.6826	0.0021	2.9171	0.7231	1.5474	3.5315	5.4163	2.2577	6.1205
	Type 5	0.6625	0.6650	0.6635	0.6510	0.6640	0.6765	0.6510	0.6895	0.6654	0.0127	1.0260	1.0036	1.0247	1.0281	1.0546	1.0282	1.0578

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.4b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 4.4 together with relative penalty values and input values of SAO technique (continued).

Method	Penalty type	Input values for simulated annealing										TRF**
		Starting values of penalties								Boundaries		
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20												
Existing IC	Type 1 (RBAR)											
Maximisation of MAPCS	Type 2	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0010
	Type 3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.0010
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-10.0000	10.0000	0.1000
	Type 5	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	4.0000	0.1000
Sample size = 50												
Existing IC	Type 1 (RBAR)											
Maximisation of MAPCS	Type 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0001
	Type 3	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	32.0000	0.0001
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-3.0000	3.0000	0.0100
	Type 5	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	20.0000	0.1000
Sample size = 100												
Existing IC	Type 1 (AIC/MCP)											
Maximisation of MAPCS	Type 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0100
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	9.2103	0.0010
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.1000

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.5a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 96 for Design 4.5 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Sample size = 20																		
Existing IC	Type 1 (BIC)	0.7105	0.4060	0.3755	0.3655	0.2280	0.2185	0.2150	0.1905	0.3388	0.1726	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936
Maximisation of MAPCS	Type 2	0.6320	0.4095	0.4280	0.3425	0.2650	0.2550	0.1870	0.2305	0.3437	0.1444	1.2311	1.1016	1.4259	2.4117	2.5093	2.8040	3.7523
	Type 3	0.3490	0.3570	0.3645	0.3470	0.3625	0.3600	0.3430	0.3990	0.3603	0.0175	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimisation of SD among APCS	Type 4	0.3160	0.3165	0.3175	0.3165	0.3155	0.3160	0.3180	0.3160	0.3165	0.0008	0.5810	0.5883	0.5312	1.2147	1.1881	1.1350	1.9315
	Type 5	0.3490	0.3570	0.3645	0.3470	0.3625	0.3600	0.3430	0.3990	0.3603	0.0175	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Sample size = 50																		
Existing IC	Type 1 (BIC)	0.8770	0.7125	0.7325	0.7300	0.5280	0.5115	0.5400	0.4565	0.6360	0.1467	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680
Maximisation of MAPCS	Type 2	0.9255	0.6885	0.7245	0.7180	0.5625	0.5330	0.5300	0.4530	0.6419	0.1519	2.5078	2.3976	2.4758	4.1494	4.2778	4.4728	6.3899
	Type 3	0.9260	0.6895	0.7230	0.7160	0.5610	0.5335	0.5320	0.4540	0.6419	0.1515	1.0831	1.0786	1.0918	1.1330	1.1384	1.1465	1.2118
Minimisation of SD among APCS	Type 4	0.5730	0.5725	0.5715	0.5770	0.5715	0.5925	0.5595	0.5540	0.5714	0.0115	0.8052	0.9969	0.8544	1.9040	1.6271	1.9887	3.0404
	Type 5	0.5855	0.5970	0.5695	0.5960	0.5510	0.6615	0.5355	0.5155	0.5764	0.0451	1.0089	1.0391	1.0100	1.0595	1.0138	1.0659	1.0939
Sample size = 96																		
Existing IC	Type 1 (BIC)	0.9525	0.8850	0.8905	0.8840	0.8340	0.8030	0.8055	0.8135	0.8585	0.0531	2.2822	2.2822	2.2822	4.5643	4.5643	4.5643	6.8465
Maximisation of MAPCS	Type 2	0.9970	0.9230	0.9395	0.9300	0.8380	0.8150	0.7870	0.7690	0.8748	0.0831	4.4921	4.4030	4.5470	7.7022	7.5448	8.2879	11.4742
	Type 3	0.9970	0.9225	0.9395	0.9305	0.8375	0.8150	0.7870	0.7695	0.8748	0.0830	1.0884	1.0856	1.0888	1.1513	1.1469	1.1645	1.2313
Minimisation of SD among APCS	Type 4	0.8295	0.8295	0.8295	0.8285	0.8530	0.8260	0.8245	0.8160	0.8296	0.0105	1.2863	1.4084	1.2804	3.0183	2.9921	3.0599	5.3840
	Type 5	0.8495	0.8375	0.9685	0.8375	0.8295	0.8175	0.8105	0.7945	0.8431	0.0536	1.0522	1.0114	1.0528	1.0853	1.0894	1.0842	1.1404

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.5a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 96 for Design 4.5 together with relative penalty values and input values of SAO technique (continued)

Method	Penalty type	Input values for simulated annealing										TRF**
		Starting values of penalties								Boundaries		
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	-8.0000	8.0000	0.0100
	Type 3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	24.0000	0.0001
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.0100
Sample size = 50												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680	-10.0000	10.0000	0.0001
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.1000
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	24.0000	0.0100
	Type 5	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	32.0000	0.0001
Sample size = 96												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-32.0000	32.0000	0.1000
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.1000
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.0001
	Type 5	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	4.0000	0.0001

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.5b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 96 for Design 4.5 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Sample size = 20																		
Existing IC	Type 1 (RBAR)	0.2990	0.2990	0.3220	0.2910	0.3155	0.3010	0.2905	0.3590	0.3096	0.0228	0.5407	0.5407	0.5407	1.1123	1.1123	1.1123	1.7185
Maximisation of MAPCS	Type 2	0.5695	0.3850	0.4020	0.3745	0.2410	0.2810	0.2320	0.1955	0.3426	0.1290	1.2359	0.8975	1.1031	2.5051	2.2600	2.3944	3.8528
	Type 3	0.3490	0.3570	0.3645	0.3470	0.3625	0.3600	0.3430	0.3990	0.3603	0.0175	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Minimisation of SD among APCS	Type 4	0.3170	0.3165	0.3165	0.3165	0.3165	0.3160	0.3165	0.3160	0.3164	0.0003	0.5806	0.5928	0.5347	1.2156	1.1892	1.1431	1.9356
	Type 5	0.3490	0.3385	0.3335	0.3220	0.3370	0.3255	0.3415	0.3310	0.3347	0.0087	1.0016	1.0020	1.0020	1.0020	1.0020	1.0020	1.0071
Sample size = 50																		
Existing IC	Type 1 (AIC)	0.6240	0.5700	0.5905	0.5945	0.5540	0.5290	0.5635	0.5725	0.5747	0.0286	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
Maximisation of MAPCS	Type 2	0.7720	0.6845	0.6485	0.6845	0.5685	0.4940	0.5335	0.5475	0.6166	0.0952	1.2583	1.4758	1.4544	2.6422	2.9266	2.9284	3.9798
	Type 3	0.9260	0.6895	0.7230	0.7160	0.5610	0.5335	0.5320	0.4540	0.6419	0.1515	1.0831	1.0786	1.0818	1.1330	1.1384	1.1465	1.2118
Minimisation of SD among APCS	Type 4	0.5670	0.5670	0.5670	0.5670	0.5665	0.5670	0.5670	0.5670	0.5669	0.0002	0.8048	0.9248	0.8805	1.8094	1.6973	1.8562	2.8751
	Type 5	0.5665	0.5665	0.5665	0.5665	0.5665	0.5665	0.5665	0.5665	0.5665	0.0000	1.0116	1.0165	1.0146	1.0310	1.0264	1.0339	1.0530
Sample size = 96																		
Existing IC	Type 1 (HQ)	0.8525	0.7895	0.8130	0.7945	0.8065	0.7845	0.7780	0.8460	0.8081	0.0278	1.5183	1.5183	1.5183	3.0366	3.0366	3.0366	4.5548
Maximisation of MAPCS	Type 2	0.7145	0.8480	0.8470	0.8285	0.7920	0.7465	0.7720	0.8475	0.7995	0.0515	0.8333	0.8190	0.8562	2.6041	2.7918	2.5156	3.9997
	Type 3	0.9970	0.9235	0.9350	0.9310	0.8395	0.8140	0.7825	0.7745	0.8746	0.0825	1.0833	1.0845	1.0860	1.1455	1.1438	1.1636	1.2245
Minimisation of SD among APCS	Type 4	0.8245	0.8245	0.8245	0.8245	0.8245	0.8245	0.8245	0.8245	0.8245	0.0000	1.2674	1.4939	1.1999	3.2747	2.8044	2.9843	5.1177
	Type 5	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.8250	0.0000	1.0161	1.0204	1.0145	1.0480	1.0380	1.0416	1.0776

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.5b **Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the probabilities of correct selection of models under different methods for sample sizes 20, 50 and 96 for Design 4.5 together with relative penalty values and input values of SAO technique (continued).**

Method	**Penalty type**	**Input values for simulated annealing**										TRF**
		Starting values of penalties								Boundaries		
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20												
Existing IC	Type 1 (RBAR)											
Maximisation of MAPCS	Type 2	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	8.0000	0.0100
	Type 3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.0000	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936	0.0000	24.0000	0.0010
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.1000
Sample size = 50												
Existing IC	Type 1 (AIC)											
Maximisation of MAPCS	Type 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0010
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.1000
Sample size = 96												
Existing IC	Type 1 (HQ)											
Maximisation of MAPCS	Type 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0010
	Type 3	2.2822	4.5643	4.5643	4.5643	6.8465	6.8465	6.8465	9.1287	0.0000	9.1287	0.0100
Minimisation of SD among APCS	Type 4	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	24.0000	0.1000
	Type 5	2.2822	4.5643	4.5643	4.5643	6.8465	6.8465	6.8465	9.1287	0.0000	18.2574	0.1000

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.6a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 4.6 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Sample size = 20																		
Existing IC	Type 1 (BIC)	0.7065	0.3285	0.6260	0.6865	0.2885	0.3190	0.5770	0.3025	0.4793	0.1858	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936
Maximisation of MAPCS	Type 2	0.7615	0.4425	0.6540	0.6940	0.3435	0.3455	0.4360	0.3675	0.5056	0.1702	1.1824	1.8443	2.5753	3.1273	3.9067	5.0477	5.9209
	Type 3	0.7575	0.4395	0.6525	0.6865	0.3450	0.3595	0.4345	0.3680	0.5054	0.1662	1.0663	1.1375	1.2163	1.2211	1.3005	1.4708	1.5086
Minimisation of SD among APCS	Type 4	0.4995	0.5120	0.4950	0.4705	0.5000	0.4640	0.4365	0.4320	0.4762	0.0304	0.4758	1.1614	2.0909	1.6428	2.5945	3.4191	4.0634
	Type 5	0.6090	0.6605	0.4090	0.4910	0.3900	0.5030	0.3660	0.3655	0.4743	0.1127	0.9895	1.4499	1.3265	1.4448	1.3128	1.7508	1.7650
Sample size = 50																		
Existing IC	Type 1 (BIC)	0.8675	0.4735	0.7660	0.7895	0.4165	0.4375	0.7155	0.4190	0.6106	0.1913	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680
Maximisation of MAPCS	Type 2	0.8560	0.5595	0.7680	0.7210	0.4675	0.5220	0.6500	0.4485	0.6241	0.1486	1.4094	2.0569	3.0438	3.6339	4.3033	5.5823	7.1213
	Type 3	0.8670	0.5625	0.7565	0.7215	0.4605	0.5230	0.6620	0.4390	0.6240	0.1524	1.0369	1.0740	1.1055	1.1201	1.1384	1.1982	1.2514
Minimisation of SD among APCS	Type 4	0.6620	0.6610	0.6065	0.5970	0.6130	0.5830	0.5600	0.5405	0.6029	0.0433	0.6137	1.6517	2.9123	2.2565	3.6126	5.1529	5.9806
	Type 5	0.6360	0.7765	0.5450	0.5725	0.5265	0.5520	0.5235	0.5125	0.5806	0.0882	0.9960	1.1785	1.2707	1.1832	1.2815	1.3926	1.4092
Sample size = 100																		
Existing IC	Type 1 (BIC)	0.8865	0.5875	0.8220	0.8540	0.4905	0.5390	0.7685	0.5065	0.6818	0.1670	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078
Maximisation of MAPCS	Type 2	0.8700	0.6725	0.8025	0.7955	0.5630	0.6235	0.7065	0.5500	0.6979	0.1174	1.5503	2.3922	3.3009	[illegible]272	4.7063	6.0731	7.6431
	Type 3	0.8670	0.6725	0.8025	0.8120	0.5630	0.6175	0.7000	0.5490	0.6979	0.1195	1.0211	1.0385	1.0537	1.0621	1.0742	1.1066	1.1301
Minimisation of SD among APCS	Type 4	0.7425	0.7805	0.6490	0.6740	0.6185	0.6750	0.6260	0.6275	0.6741	0.0588	0.7452	3.3059	3.9003	4.0792	4.6636	6.4081	7.2420
	Type 5	0.6520	0.6525	0.6480	0.6555	0.6465	0.6555	0.6530	0.6555	0.6523	0.0035	1.0060	1.0175	1.0261	1.0211	1.0315	1.0469	1.0532

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.6a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 4.6 together with relative penalty values and input values of SAO technique (continued).

Method	Penalty type	Input values for simulated annealing										TRF**
		Starting values of penalties								Boundaries		
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.0010
	Type 3	1.4979	2.9957	2.9957	2.9957	4.4936	4.4936	4.4936	5.9915	0.0000	23.9659	0.1000
Minimisation of SD among APCS	Type 4	0.0000	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936	-3.0000	4.4936	0.0001
	Type 5	1.4979	2.9957	2.9957	2.9957	4.4936	4.4936	4.4936	5.9915	0.0000	5.9915	0.0010
Sample size = 50												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	1.9560	1.9560	1.9560	3.9120	3.9120	3.9120	5.8680	-4.0000	7.8240	0.1000
	Type 3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	32.0000	0.1000
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.0010
	Type 5	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	20.0000	0.1000
Sample size = 100												
Existing IC	Type 1 (BIC)											
Maximisation of MAPCS	Type 2	0.0000	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078	0.0000	32.0000	0.1000
	Type 3	2.3026	4.6052	4.6052	4.6052	6.9078	6.9078	6.9078	9.2103	0.0000	20.0000	0.1000
Minimisation of SD among APCS	Type 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.0001
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.0100

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.6b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 4.6 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model										Relative penalties*						
		M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	Mean	SD	P_2	P_3	P_4	P_5	P_6	P_7	P_8
Sample size = 20																		
Existing IC	Type 1 (RBAR)	0.3070	0.2810	0.4020	0.3940	0.3530	0.3840	0.5090	0.4890	0.3899	0.0795	0.5407	0.5407	0.5407	1.1123	1.1123	1.1123	1.7185
Maximisation of MAPCS	Type 2	0.5170	0.5160	0.6735	0.6520	0.3215	0.3730	0.3760	0.4835	0.4891	0.1290	0.5890	1.0991	1.5113	2.3820	2.5229	3.5102	3.9726
	Type 3	0.7320	0.4735	0.6540	0.6485	0.3410	0.3875	0.4390	0.3635	0.5049	0.1514	1.0490	1.1374	1.2344	1.2211	1.2906	1.4793	1.5195
Minimisation of SD among APCS	Type 4	0.4570	0.4570	0.4570	0.4570	0.4570	0.4570	0.4570	0.4570	0.4570	0.0000	0.4777	1.0568	1.2884	1.5328	1.7943	2.3599	2.9654
	Type 5	0.4575	0.4575	0.4585	0.4575	0.4580	0.4575	0.4570	0.4560	0.4574	0.0007	0.9937	1.0523	1.0794	1.0424	1.0723	1.1359	1.1356
Sample size = 50																		
Existing IC	Type 1 (AIC)	0.5815	0.4510	0.6105	0.5990	0.4730	0.4900	0.6810	0.5505	0.5546	0.0787	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
Maximisation of MAPCS	Type 2	0.6155	0.6325	0.7045	0.7010	0.4880	0.5030	0.5965	0.5635	0.6006	0.0808	0.6797	1.3275	1.4290	2.4737	2.6022	3.1670	3.9984
	Type 3	0.8025	0.5890	0.7080	0.7205	0.5020	0.5130	0.6520	0.4825	0.6212	0.1180	1.0247	1.0638	1.0746	1.0915	1.1066	1.1479	1.1845
Minimisation of SD among APCS	Type 4	0.5770	0.5770	0.5770	0.5770	0.5755	0.5770	0.5780	0.5785	0.5771	0.0009	1.0042	1.0302	1.0351	1.0342	1.0418	1.0707	1.0788
	Type 5	0.5775	0.5775	0.5775	0.5775	0.5770	0.5775	0.5780	0.5775	0.5775	0.0003	0.6204	1.2574	1.3911	1.8771	2.0773	2.7711	3.5012
Sample size = 100																		
Existing IC	Type 1 (AIC)	0.5860	0.5310	0.6470	0.6660	0.5245	0.5740	0.7245	0.6465	0.6124	0.0700	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000
Maximisation of MAPCS	Type 2	0.6715	0.6365	0.7400	0.7920	0.5590	0.5820	0.6535	0.6600	0.6618	0.0764	1.0279	1.3596	1.3104	2.4956	2.6699	3.1693	3.9985
	Type 3	0.8665	0.6710	0.7995	0.8120	0.5560	0.6105	0.6825	0.5815	0.6974	0.1159	1.0210	1.0385	1.0508	1.0621	1.0717	1.0993	1.1159
Minimisation of SD among APCS	Type 4	0.6535	0.6540	0.6540	0.6540	0.6540	0.6545	0.6540	0.6545	0.6541	0.0003	0.8088	1.3326	1.9484	2.0225	2.7240	3.4613	4.2872
	Type 5	0.6515	0.6490	0.6485	0.6530	0.6470	0.6535	0.6555	0.6555	0.6517	0.0032	1.0061	1.0171	1.0264	1.0208	1.0316	1.0463	1.0526

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.6b **Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 4.6 together with relative penalty values and input values of SAO technique (continued).**

Method	Penalty type	Input values for simulated annealing										TRF**
		Starting values of penalties								Boundaries		
		S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20												
Existing IC	Type 1 (RBAR)											
Maximisation of MAPCS	Type 2	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0100
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	32.0000	0.0001
Minimisation of SD among APCS	Type 4	0.0000	1.4979	1.4979	1.4979	2.9957	2.9957	2.9957	4.4936	-24.0000	24.0000	0.1000
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.1000
Sample size = 50												
Existing IC	Type 1 (AIC)											
Maximisation of MAPCS	Type 2	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0001
	Type 3	1.9560	3.9120	3.9120	3.9120	5.8680	5.8680	5.8680	7.8240	0.0000	32.0000	0.0100
Minimisation of SD among APCS	Type 4	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	24.0000	0.1000
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.0100
Sample size = 100												
Existing IC	Type 1 (AIC)											
Maximisation of MAPCS	Type 2	0.0000	1.0000	1.0000	1.0000	2.0000	2.0000	2.0000	3.0000	0.0000	4.0000	0.0100
	Type 3	1.0000	2.0000	2.0000	2.0000	3.0000	3.0000	3.0000	4.0000	0.0000	8.0000	0.1000
Minimisation of SD among APCS	Type 4	0.0000	2.3026	2.3026	2.3026	4.6052	4.6052	4.6052	6.9078	-3.0000	6.9078	0.1000
	Type 5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.7a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes *20, 50 and 100 for Design 5.1 together with relative penalty values and input values of SAO technique.*

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				*Boundaries*		*TRF***
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.9845	0.3905	0.7200	0.2350	0.5825	0.3357	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	-1.5532	-1.1620	-1.8220	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.9725	0.3850	0.6395	0.5910	0.6470	0.2434	0.8561	0.8903	0.8334	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.9705	0.4875	0.5590	0.5400	0.6392	0.2229	-1.9068	-0.8318	-1.9929	0.0000	0.0000	0.0000	0.0000	0.0000	2.0000	0.0010
	Type 5	0.9625	0.5050	0.5575	0.5240	0.6372	0.2179	0.7226	0.8070	0.7185	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.9920	0.5000	0.8880	0.2065	0.6466	0.3618	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	-0.3541	0.8881	-0.8217	0.0000	0.0000	0.0000	0.0000	-8.0000	8.0000	0.1000
	Type 3	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	0.9860	1.0362	0.9677	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
Minimisation of SD among APCS	Type 4	0.9850	0.5295	0.7865	0.6240	0.7313	0.1997	-1.2891	0.1305	-1.4314	0.0000	0.0000	0.0000	0.0000	0.0000	2.0000	0.0001
	Type 5	0.9860	0.5060	0.8100	0.6285	0.7326	0.2101	0.9620	1.0000	0.9550	3.9120	3.9120	3.9120	3.9120	0.0000	3.9120	0.0010
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.9970	0.6180	0.9315	0.2055	0.6880	0.3617	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	-0.8164	0.6215	-1.4880	0.0000	0.0000	0.0000	0.0000	-8.0000	8.[illegible]	0.1000
	Type 3	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	0.9799	1.0085	0.9369	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
Minimisation of SD among APCS	Type 4	0.9935	0.6420	0.8580	0.6750	0.7921	0.1645	-1.7682	0.0366	-1.9579	0.0000	0.0000	0.0000	0.0000	0.0000	2.0000	0.1000
	Type 5	0.9840	0.6495	0.7755	0.6925	0.7754	0.1486	0.8278	0.9202	0.8246	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.7b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 5.1 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.9845	0.3905	0.7200	0.2350	0.5825	0.3357	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.9685	0.3855	0.6395	0.5925	0.6465	0.2414	-2.1418	-1.7502	-2.4104	1.4979	1.4979	1.4979	1.4979	0.0000	8.0000	0.0001
	Type 3	0.9730	0.3900	0.6110	0.6115	0.6464	0.2414	0.8563	0.9070	0.8335	1.0000	2.0000	2.0000	3.0000	0.0000	32.0000	0.0001
Minimisation of SD among APCS	Type 4	0.8925	0.5060	0.5575	0.5235	0.6199	0.1830	-16.8792	-15.7717	-16.9344	0.0000	1.0000	1.0000	2.0000	-10.0000	10.0000	0.1000
	Type 5	0.9265	0.5050	0.5575	0.5250	0.6285	0.1998	0.3196	0.3572	0.3173	1.0000	2.0000	2.0000	3.0000	0.0000	16.0000	0.0001
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.9920	0.5000	0.8880	0.2065	0.6466	0.3618	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.9890	0.4695	0.7860	0.7130	0.7394	0.2145	-0.5173	0.8542	-0.8258	0.0000	0.0000	0.0000	0.0000	0.0000	20.0000	0.0010
	Type 3	0.9890	0.4225	0.7870	0.7610	0.7399	0.2348	0.9860	1.0362	0.9677	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
Minimisation of SD among APCS	Type 4	0.9365	0.5545	0.7000	0.5340	0.7062	0.1646	-18.2836	-14.5567	-18.3989	1.9560	1.9560	1.9560	1.9560	-10.0000	10.0000	0.1000
	Type 5	0.8125	0.5750	0.6520	0.5960	0.6589	0.1074	0.0832	0.1030	0.0830	1.0000	1.0000	1.0000	1.0000	0.0000	12.0000	0.0010
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.9970	0.6180	0.9315	0.2055	0.6880	0.3617	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.9940	0.5285	0.8585	0.8405	0.8054	0.1969	-0.8678	0.5861	-1.5303	0.0000	1.0000	1.0000	2.0000	0.0000	10.0000	0.0100
	Type 3	0.9940	0.5270	0.8585	0.8425	0.8055	0.1977	0.9799	1.0085	0.9669	1.0000	1.0000	1.0000	1.0000	0.0000	8.0000	0.1000
Minimisation of SD among APCS	Type 4	0.9785	0.6495	0.7735	0.6925	0.7735	0.1460	-17.0655	-11.6500	-17.2541	0.0000	1.0000	1.0000	2.0000	-10.0000	10.0000	0.1000
	Type 5	0.9120	0.6495	0.7515	0.6930	0.7515	0.1149	0.1114	0.1270	0.1110	1.0000	1.0000	1.0000	1.0000	0.0000	4.0000	0.0001

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.8a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 5.2 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.2690	0.4950	0.5085	0.6305	0.4758	0.1507	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.5570	0.4770	0.4560	0.5110	0.5002	0.0441	0.3764	0.5043	0.[illegible]540	0.0000	0.0000	0.0000	0.0000	0.0000	8.0000	0.1000
	Type 3	0.5570	0.4770	0.4560	0.5110	0.5002	0.0441	1.0384	1.0517	[illegible].1001	1.0000	1.0000	1.0000	1.0000	0.0000	2.0000	0.1000
Minimisation of SD among APCS	Type 4	0.4910	0.4905	0.4905	0.4915	0.4909	0.0005	0.2895	0.3017	1.0538	0.0000	1.0000	1.0[illegible]0	2.0000	0.0000	4.0000	0.0010
	Type 5	0.4885	0.4910	0.4860	0.5075	0.4932	0.0097	1.0281	1.0325	1.0960	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.0010
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.3360	0.5725	0.6275	0.7180	0.5635	0.1631	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6720	0.5275	0.5245	0.6165	0.5851	0.0719	0.5147	0.8774	1.0103	0.0000	0.0000	0.0000	0.0000	-10.0000	10.0000	0.1000
	Type 3	0.6720	0.5275	0.5245	0.6165	0.5851	0.0719	1.0208	1.0357	1.0412	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5765	0.5770	0.5805	0.5905	0.5811	0.0065	0.2412	0.4600	1.1378	0.0000	1.0000	1.0000	2.0000	-2.0000	2.0000	0.0100
	Type 5	0.5750	0.5790	0.5695	0.6030	0.5816	0.0148	1.0093	1.0204	1.0418	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.0001
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.4290	0.6575	0.7125	0.8180	0.6542	0.1643	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	0.4189	0.6262	1.1782	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	1.0084	1.0126	1.0239	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.6625	0.6680	0.6595	0.7030	0.6733	0.0201	0.2446	0.6109	1.4754	0.0000	1.0000	1.0000	2.0000	-2.0000	2.0000	0.0001
	Type 5	0.6640	0.8570	0.6575	0.7295	0.6770	0.0351	1.0057	1.0126	1.0252	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.0010

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.8b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 5.2 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_1	M_2	M_3	M_4	Mean	SD	P_2	P_3	P_4	S_1	S_2	S_3	S_4	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.2690	0.4950	0.5085	0.6305	0.4758	0.1507	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.5535	0.4760	0.4560	0.5150	0.5001	0.0432	0.3767	0.5043	0.9144	0.0000	1.0000	1.0000	2.0000	0.0000	8.0000	0.1000
	Type 3	0.5535	0.4760	0.4560	0.5150	0.5001	0.0432	1.0384	1.0518	1.0957	1.0000	1.0000	1.0000	1.0000	0.0000	20.0000	0.1000
Minimisation of SD among APCS	Type 4	0.4905	0.4905	0.4905	0.4905	0.4905	0.0000	0.2896	0.3005	1.0648	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.4905	0.4905	0.4905	0.4905	0.4905	0.0000	1.0293	1.0306	1.1125	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.3360	0.5725	0.6275	0.7180	0.5635	0.1631	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6465	0.5235	0.5420	0.6265	0.5846	0.0609	0.5123	0.7462	0.8995	0.0000	1.0000	1.0000	2.0000	-2.0000	2.0000	0.0010
	Type 3	0.6510	0.5235	0.5420	0.6220	0.5846	0.0615	1.0207	1.0302	1.0386	3.9120	3.9120	3.9120	3.9120	0.0000	3.9120	0.0010
Minimisation of SD among APCS	Type 4	0.5805	0.5805	0.5805	0.5805	0.5805	0.0000	0.2323	0.4600	1.2438	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5805	0.5805	0.5805	0.5805	0.5805	0.0000	1.0094	1.0186	1.0512	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.4290	0.6575	0.7125	0.8180	0.6542	0.1643	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6890	0.6320	0.6620	0.7370	0.6800	0.0446	0.4189	0.6262	1.1782	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.6900	0.6320	0.6620	0.7355	0.6799	0.0440	1.0084	1.0126	1.0242	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.0010
Minimisation of SD among APCS	Type 4	0.6685	0.6685	0.6685	0.6685	0.6685	0.0000	0.2444	0.5668	2.0418	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.6685	0.6685	0.6685	0.6685	0.6685	0.0000	1.0049	1.0114	1.0415	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000

* Additive penalty for model M_1 is zero and multiplicative penalty for model M_1 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.9a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under diff~~ent methods for sample sizes 20, 50 and 100 for Design 5.3 together with relative penalty values and input values of SAO tech que.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values fro simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.5270	0.4030	0.3745	0.4785	0.4457	0.0697	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.2440	0.3505	0.7285	0.7470	0.5175	0.2581	-3.4352	-3.7546	-0.5835	2.9957	2.9957	2.9957	2.9957	-4.0000	5.9915	0.0001
	Type 3	0.2440	0.3510	0.7280	0.7470	0.5175	0.2579	0.7101	0.6879	0.9433	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.1000
Minimisation of SD among APCS	Type 4	0.4780	0.5260	0.5325	0.4960	0.5081	0.0256	-3.5696	-3.5836	-0.0276	0.0000	1.0000	1.0000	2.0000	-3.0000	3.0000	0.0100
	Type 5	0.4820	0.5295	0.5200	0.4985	0.5075	0.0214	0.7650	0.7646	0.9973	1.0000	2.0000	2.0000	3.0000	0.0000	12.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.5820	0.5045	0.4295	0.4845	0.5001	0.0631	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.3945	0.3845	0.7400	0.7085	0.5569	0.1937	-1.4966	-1.8183	-0.2936	3.9120	3.9120	3.9120	3.9120	0.0000	12.0000	0.1000
	Type 3	0.3945	0.3840	0.7405	0.7075	0.5566	0.1938	0.9403	0.9283	0.9883	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5245	0.5485	0.5395	0.5265	0.5347	0.0113	-2.0280	-2.0545	-0.0412	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.0100
	Type 5	0.5320	0.5065	0.5040	0.5290	0.5179	0.0147	0.9865	0.9851	0.9985	1.0000	1.0000	1.0000	1.0000	0.0000	3.0000	0.1000
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.6230	0.6085	0.4785	0.5005	0.5526	0.0737	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.4480	0.4490	0.8080	0.7330	0.6095	0.1884	-1.9048	-2.3262	-0.3119	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.0100
	Type 3	0.4480	0.4490	0.8080	0.7330	0.6095	0.1884	0.9625	0.9544	0.9938	1.0000	2.0000	2.0000	3.0000	0.0000	12.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5645	0.6015	0.5975	0.5825	0.5865	0.0163	-3.7893	-3.8600	-0.0701	4.6052	4.6052	4.6052	4.6052	0.0000	6.9078	0.1000
	Type 5	0.5605	0.6020	0.5975	0.5795	0.5849	0.0189	0.8527	0.8515	0.9986	1.0000	2.0000	2.0000	3.0000	0.0000	27.0000	0.0010

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.9b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 5.3 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.5270	0.4030	0.3745	0.4785	0.4457	0.0697	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.3310	0.3215	0.7410	0.6680	0.5154	0.2204	-1.8[illegible]54	-2.2314	-0.3124	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.0001
	Type 3	0.2800	0.3510	0.72[illegible]5	0.7095	0.5172	0.2349	0.7131	0.6907	0.9566	4.4936	4.4936	4.4936	4.4936	0.0000	14.9787	0.0010
Minimisation of SD among APCS	Type 4	0.4955	0.4915	0.4915	0.4930	0.4929	0.0019	-0.9078	-0.9223	-0.0186	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.1000
	Type 5	0.5000	0.4630	0.4740	0.4950	0.4830	0.0175	0.9534	0.9512	0.9984	1.0000	1.0000	1.0000	1.0000	0.0000	3.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.5820	0.5045	0.4295	0.4845	0.5001	0.0631	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.4210	0.3385	0.7345	0.6815	0.5564	0.1769	-1.4294	-1.7417	-0.2272	3.9120	3.9120	3.9120	3.9120	0.0000	7.8240	0.1000
	Type 3	0.4130	0.3865	0.7485	0.6370	0.5463	0.1755	0.6784	0.6697	0.9918	1.0000	1.0000	1.0000	1.0000	0.0000	12.0000	0.0001
Minimisation of SD among APCS	Type 4	0.5290	0.5290	0.5290	0.5290	0.5290	0.0000	-0.8352	-0.8683	-0.0384	0.0000	1.0000	1.0000	2.0000	-4.0000	4.0000	0.0100
	Type 5	0.5320	0.5065	0.5040	0.5290	0.5179	0.0147	0.9865	0.9851	0.9985	1.0000	1.0000	1.0000	1.0000	0.0000	[illegible]	0.1000
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.6230	0.6085	0.4785	0.5005	0.5526	0.0737	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.4515	0.4490	0.8075	0.7290	0.6093	0.1864	1.8037	-2.2247	-0.3028	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.0010
	Type 3	0.4580	0.4405	0.7970	0.7250	0.6051	0.1625	0.9825	0.9743	0.9942	6.9078	6.9078	6.9078	6.9078	0.0000	6.9078	0.0001
Minimisation of SD among APCS	Type 4	0.5745	0.5740	0.5745	0.5750	0.5745	0.0004	-0.2432	-0.3146	-0.0603	4.6052	4.6052	4.6052	4.6052	0.0000	9.2103	0.0100
	Type 5	0.5735	0.5725	0.5735	0.5750	0.5736	0.0010	0.9956	0.9941	0.9988	6.9078	6.9078	6.9078	6.9078	0.0000	13.8155	0.1000

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.10a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 5.4 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.3790	0.4805	0.5250	0.3840	0.4421	0.0724	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.7265	0.3540	0.3700	0.4720	0.4806	0.1720	0.7617	0.8956	0.4813	0.0000	0.0000	0.0000	0.0000	-12.0000	12.0000	0.1000
	Type 3	0.7265	0.3540	0.3700	0.4720	0.4806	0.1720	1.0790	1.0936	1.0492	1.0000	2.0000	2.0000	3.0000	0.0000	48.0000	0.1000
Minimisation of SD among APCS	Type 4	0.4625	0.4625	0.4620	0.4625	0.4624	0.0003	0.1059	0.1544	0.0281	0.0000	1.0000	1.0000	2.0000	0.0000	4.0000	0.1000
	Type 5	0.4625	0.4625	0.4620	0.4625	0.4624	0.0003	1.0107	1.0156	1.0028	1.0000	2.0000	2.0000	3.0000	0.0000	16.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.4155	0.5910	0.6380	0.4610	0.5264	0.1052	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.7005	0.5455	0.4435	0.6305	0.5800	0.1109	0.5450	0.9479	0.2884	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.7005	0.5450	0.4440	0.6305	0.5800	0.1107	1.0220	1.0385	1.0114	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5560	0.5575	0.5595	0.5565	0.5574	0.0015	0.2101	0.2874	0.1043	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.0001
	Type 5	0.5560	0.5575	0.5590	0.5575	0.5575	0.0012	1.0085	1.0117	1.0042	5.8680	5.8680	5.8680	5.8680	0.0000	5.8680	0.0100
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.5130	0.6520	0.7370	0.5780	0.6200	0.0965	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.7805	0.6545	0.5480	0.7465	0.6824	0.1042	0.5510	1.2144	0.5141	0.0000	1.0000	1.0000	2.0000	-10.0000	10.0000	0.1000
	Type 3	0.7805	0.6545	0.5480	0.7465	0.6824	0.1042	1.0111	1.0246	1.0103	6.9078	6.9078	6.9078	6.9078	0.0000	13.8155	0.1000
Minimisation of SD among APCS	Type 4	0.6795	0.6885	0.6360	0.6500	0.6635	0.0246	0.2110	0.6355	0.4415	4.6052	4.6052	4.6052	4.6052	0.0000	6.9078	0.0001
	Type 5	0.6825	0.6890	0.6340	0.6475	0.6633	0.0267	1.0042	1.0134	1.0095	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.10b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 5.4 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_5	M_6	M_7	M_8	Mean	SD	P_6	P_7	P_8	S_5	S_6	S_7	S_8	LB	UB	
Sample size =20																	
Existing IC	Type 1 (All are same)	0.3790	0.4805	0.5250	0.3840	0.4421	0.0724	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6180	0.5120	0.3185	0.4710	0.4799	0.1241	0.2655	0.8940	0.4780	0.0000	0.0000	0.0000	0.0000	-12.0000	12.0000	0.0100
	Type 3	0.6805	0.4285	0.3400	0.4715	0.4801	0.1444	1.0513	1.0947	1.0502	1.0000	1.0000	1.0000	1.0000	0.0000	20.0000	0.0010
Minimisation of SD among APCS	Type 4	0.4620	0.4620	0.4620	0.4620	0.4620	0.0000	0.1057	0.1537	0.0284	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.4620	0.4620	0.4620	0.4620	0.4620	0.0000	1.0106	1.0155	1.0028	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size =50																	
Existing IC	Type 1 (All are same)	0.4155	0.5910	0.6380	0.4610	0.5264	0.1052	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6945	0.5510	0.4440	0.6300	0.5799	0.1079	0.5265	0.9294	0.2819	0.0000	1.0000	1.0000	2.0000	0.0000	3.0000	0.1000
	Type 3	0.6965	0.5500	0.4435	0.6295	0.5799	0.1089	1.0214	1.0381	1.0115	5.8680	5.8680	5.8680	5.8680	0.0000	19.5601	
Minimisation of SD among APCS	Type 4	0.5570	0.5570	0.5570	0.5570	0.5570	0.0000	0.2134	0.2993	0.1092	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5570	0.5570	0.5570	0.5570	0.5570	0.0000	1.0086	1.0120	1.0044	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size =100																	
Existing IC	Type 1 (All are same)	0.5130	0.6520	0.7370	0.5780	0.6200	0.0965	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.7675	0.6450	0.5590	0.7540	0.6814	0.0983	0.5513	1.0794	0.4080	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.7805	0.6450	0.5705	0.7290	0.6813	0.0926	1.0111	1.0219	1.0103	6.9078	6.9078	6.9078	6.9078	0.0000	6.9078	0.0001
Minimisation of SD among APCS	Type 4	0.6595	0.6595	0.6595	0.6595	0.6595	0.0000	0.2182	0.4622	0.2743	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.6595	0.6595	0.6595	0.6595	0.6595	0.0000	1.0044	1.0093	1.0055	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_5 is zero and multiplicative penalty for model M_5 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.11a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 5.5 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.6410	0.3770	0.4650	0.3950	0.4695	0.1205	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.3690	0.9970	0.6530	0.9180	0.7343	0.2845	-5.2089	-0.7590	-2.2686	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.3690	0.9970	0.6530	0.9180	0.7343	0.2845	0.5873	0.9270	0.7970	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5635	0.5635	0.5635	0.5635	0.5635	0.0000	-0.2022	-0.1329	-0.2334	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5635	0.5635	0.5635	0.5635	0.5635	0.0000	0.9799	0.9867	0.9769	0.9975	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.7685	0.4010	0.4955	0.4365	0.5254	0.1667	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.5175	0.9990	0.8325	0.9490	0.8245	0.2162	-4.8874	-1.2190	-2.6135	0.0000	1.0000	1.0000	2.0000	0.0000	10.0000	0.0100
	Type 3	0.5175	0.9990	0.8325	0.9495	0.8246	0.2163	0.8220	0.9523	0.9002	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.6785	0.6785	0.6785	0.6785	0.6785	0.0000	-0.1841	-0.2797	-0.3831	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.6785	0.6785	0.6785	0.6785	0.6785	0.0000	0.9927	0.9889	0.9848	0.9932	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.8505	0.4350	0.5005	0.4495	0.5589	0.1964	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	-5.3640	-1.2130	-3.0750	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.6480	0.9985	0.8710	0.9830	0.8751	0.1617	0.8978	0.9760	0.9403	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Minimisation of SD among APCS	Type 4	0.7580	0.7580	0.7580	0.7580	0.7580	0.0000	-0.3382	-0.3997	-0.5470	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.7580	0.7580	0.7580	0.7580	0.7580	0.0000	0.9933	0.9920	0.9891	0.9204	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.11b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 5.5 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.6410	0.3770	0.4650	0.3950	0.4695	0.1205	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.3725	0.9855	0.6625	0.9060	0.7316	0.2760	-3.9370	-0.7641	-2.0906	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.1000
	Type 3	0.3760	0.9970	0.6460	0.9165	0.7339	0.2819	0.5935	0.9310	0.8012	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5635	0.5635	0.5635	0.5635	0.5635	0.0000	-0.2022	-0.1329	-0.2334	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5635	0.5635	0.5635	0.5635	0.5635	0.0000	0.9799	0.9867	0.9769	0.9975	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.7685	0.4010	0.4955	0.4365	0.5254	0.1667	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.5330	0.9935	0.8095	0.9600	0.8240	0.2099	-3.9153	-1.0541	-2.7323	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.1000
	Type 3	0.5165	0.9935	0.8330	0.9540	0.8242	0.2162	0.8488	0.9523	0.8981	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.0001
Minimisation of SD among APCS	Type 4	0.6785	0.6785	0.6785	0.6785	0.6785	0.0000	-0.1841	-0.2797	-0.3831	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.6785	0.6785	0.6785	0.6785	0.6785	0.0000	0.9927	0.9889	0.9848	0.9932	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.8505	0.4350	0.5005	0.4495	0.5589	0.1964	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.6530	0.9845	0.8785	0.9685	0.8711	0.1527	-3.4805	-1.2128	-2.5651	0.0000	1.0000	1.0000	2.0000	0.0000	4.0000	0.0010
	Type 3	0.6520	0.9985	0.8760	0.9715	0.8745	0.1574	0.8964	0.9760	0.9474	9.2103	9.2103	9.2103	9.2103	0.0000	9.2103	0.0100
Minimisation of SD among APCS	Type 4	0.7580	0.7580	0.7580	0.7580	0.7580	0.0000	-0.3382	-0.3997	-0.5470	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.7580	0.7580	0.7580	0.7580	0.7580	0.0000	0.9933	0.9920	0.9891	0.9204	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.12a Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the largest mean average probabilities of correct selection under different methods for sample sizes 20, 50 and 100 for Design 5.6 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.4295	0.6030	0.4520	0.4630	0.4869	0.0787	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.8290	0.2640	0.5340	0.6330	0.5650	0.2351	1.6520	0.6531	0.5078	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Type 3	0.8290	0.2640	0.5340	0.6330	0.5650	0.2351	1.1796	1.0675	1.0521	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.5195	0.5195	0.5195	0.5195	0.5195	0.0000	0.1446	0.0096	0.0178	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5195	0.5195	0.5195	0.5195	0.5195	0.0000	1.0146	1.0010	1.0018	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.5360	0.6960	0.5555	0.5520	0.5849	0.0746	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.8325	0.4160	0.5895	0.7950	0.6583	0.1936	1.3174	0.5771	0.1042	0.0000	0.0000	0.0000	0.0000	-4.0000	4.0000	0.1000
	Type 3	0.8325	0.4160	0.5895	0.7950	0.6583	0.1936	1.0541	1.0233	1.0042	1.0000	2.0000	2.0000	3.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.6150	0.6145	0.6140	0.6145	0.6145	0.0004	0.1459	0.0133	0.0150	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.1000
	Type 5	0.6145	0.6140	0.6145	0.6150	0.6145	0.0004	1.0058	1.0005	1.0005	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.0100
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.6290	0.7505	0.6785	0.6565	0.6786	0.0520	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.9270	0.5510	0.7080	0.7890	0.7438	0.1571	1.4753	0.8747	0.6798	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 3	0.9270	0.5510	0.7080	0.7890	0.7438	0.1571	1.0300	1.0177	1.0137	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.1000
Minimisation of SD among APCS	Type 4	0.7045	0.7045	0.7045	0.7045	0.7045	0.0000	0.1476	0.0514	0.0328	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.7045	0.7045	0.7045	0.7045	0.7045	0.0000	1.0029	1.0010	1.0007	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

Table 6.12b Average probabilities, mean average probabilities and standard deviations of average probabilities of correct selection of models corresponding to the smallest variation among the average probabilities of correct selection of models under different methods for sample sizes 20, 50 and 100 for Design 5.6 together with relative penalty values and input values of SAO technique.

Method	Penalty type	Average probabilities of correct selection of model						Relative penalties*			Input values for simulated annealing						
											Starting values of penalties				Boundaries		TRF**
		M_9	M_{10}	M_{11}	M_{12}	Mean	SD	P_{10}	P_{11}	P_{12}	S_9	S_{10}	S_{11}	S_{12}	LB	UB	
Sample size = 20																	
Existing IC	Type 1 (All are same)	0.4295	0.6030	0.4520	0.4630	0.4869	0.0787	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.8275	0.3400	0.4970	0.5915	0.5640	0.2040	1.2607	0.7286	0.5748	0.0000	1.0000	1.0000	2.0000	-10.0000	10.0000	0.0100
	Type 3	0.8370	0.3280	0.5140	0.5775	0.5641	0.2105	1.1434	1.0755	1.0657	1.0000	1.0000	1.0000	1.0000	0.0000	10.0000	0.0010
Minimisation of SD among APCS	Type 4	0.5195	0.5195	0.5195	0.5195	0.5195	0.0000	0.1446	0.0096	0.0178	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.5195	0.5195	0.5195	0.5195	0.5195	0.0000	1.0146	1.0010	1.0018	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 50																	
Existing IC	Type 1 (All are same)	0.5360	0.6960	0.5555	0.5520	0.5849	0.0746	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.8285	0.4225	0.5925	0.7855	0.6573	0.1871	1.2521	0.5440	0.1041	5.8680	5.8680	5.8680	5.8680	-10.0000	10.0000	0.0010
	Type 3	0.7900	0.4450	0.6430	0.7395	0.6544	0.1523	1.0404	1.0113	1.0042	1.0000	1.0000	1.0000	1.0000	0.0000	16.0000	0.0100
Minimisation of SD among APCS	Type 4	0.6140	0.6140	0.6140	0.6140	0.6140	0.0000	0.1453	0.0121	0.0143	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.6140	0.6140	0.6140	0.6140	0.6140	0.0000	1.0058	1.0005	1.0006	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000
Sample size = 100																	
Existing IC	Type 1 (All are same)	0.6290	0.7505	0.6785	0.6565	0.6786	0.0520	0.0000	0.0000	0.0000							
Maximisation of MAPCS	Type 2	0.8945	0.5670	0.7345	0.7760	0.7430	0.1355	1.1706	0.5862	0.4886	0.0000	0.0000	0.0000	0.0000	0.0000	4.0000	0.0100
	Type 3	0.8935	0.5670	0.7350	0.7765	0.7430	0.1352	1.0235	1.0116	1.0097	9.2103	9.2103	9.2103	9.2103	0.0000	18.4207	0.0100
Minimisation of SD among APCS	Type 4	0.7045	0.7045	0.7045	0.7045	0.7045	0.0000	0.1476	0.0514	0.0328	0.0000	0.0000	0.0000	0.0000	0.0000	10.0000	0.1000
	Type 5	0.7045	0.7045	0.7045	0.7045	0.7045	0.0000	1.0029	1.0010	1.0007	1.0000	1.0000	1.0000	1.0000	0.0000	1.0000	0.1000

* Additive penalty for model M_9 is zero and multiplicative penalty for model M_9 is one

** TRF Temperature reduction factor

Type 2 and Type 4 are additive penalties with maximised log-likelihood function

Type 3 and Type 5 are multiplicative penalties with mean squared error

CHAPTER 7

SUMMARY AND CONCLUSIONS

One of the important decision making problems in econometrics and statistics is to choose an appropriate model to represent a particular data set from a set of alternative models. There are several ways of solving this problem and the most popular way is the use of an information criteria (IC). In an IC based model selection procedure, the model which has the largest maximised log-likelihood minus a penalty function is chosen as the best model. Several IC based model selection procedures have been proposed in the literature. The penalty function of all existing IC procedures depend on the number of free parameters in the model and, in most of the cases, the sample size. One of the main disadvantages of these IC procedures is that their performance varies from data set to data set and none perform well in all situations. Some IC procedures favour the model with the smallest number of parameters while others favour the model with the largest number of parameters. One of the unresolved questions is which criteria one should use to select the best model for a particular data set. Also, the penalty functions used in these IC are independent of the data set. Thus for the same set of models, a change of data does not have any impact on the penalty function. When the competing models have the same number of parameters, there is no need to use a penalty function and the problem reduces to choosing the model with the largest maximised log-likelihood, which is a weakness of this approach. Thus, it is our belief that better handling of the penalty function will improve the probability of selection of the correct model. The main aim of this thesis is to investigate the ways in which the information in the data can be used to calculate the penalty function, so that the mean average probability of correct selection (MAPCS) is increased or optimised.

A survey of relevant literature on model selection was presented in Chapter 2. We first reviewed the literature on model selection based on sum of squared errors and then we surveyed model selection based on IC. It was argued that none of the existing IC performs well in all situations, so a new technique of model selection, which performs well on average in all situations was needed. Finally, we introduced the SAO technique and reviewed its applications in econometrics.

The main purpose of Chapter 3 was twofold. First to introduce a new method for computing penalties for selecting the best model from a set of competing models, which assure the best average probability of correct selection. Another purpose was to find the most efficient combination of the number of parameter drawings (q) and replications (N) for estimating APCS via a simulation experiment for a fixed number of total simulations, qN. We used the average standard deviation averaged over the number of competing models (ASD) as a measure of efficiency and coefficient of variation (CV) as a measure of reliability of the estimated MAPCS.

We found that the relationship between ASD and q is well represented by the regression model $\ln(\text{ASD}) = \hat{a} + \hat{b}\ln(q) + \hat{c}(\ln(q))^2$, as the adjusted $\bar{R}^2$ are large (>0.95), the estimated coefficients of the model are also highly significant and there is no significant autocorrelation. The value of q, where the value of ASD is at a minimum, is almost in all cases greater than the maximum value of qN (here 2000). This implies that the maximum number of drawings of parameters produces the most efficient estimate of the MAPCS. For this combination of q and N, the value of CV is also the lowest.

We observed that the penalty functions of AIC, BIC, HQ, RBAR, GCV and HOC for the j^{th} model can be generalised to a single penalty function $p_j = \lambda_1 k_j + \lambda_2 \ln(n-k_j)$. For the listed existing criteria, λ_1 and λ_2 are determined by the sample size n. In our proposed method of computing penalties for IC based model selection, we allowed λ_1 and λ_2 to take any values that maximise the estimated MAPCS. We proposed five IC, and for the j^{th} model the criteria are as follows:

$$\text{NIC1}_j = L(\hat{\theta}_j) - \lambda_1 k_j,$$

$$\text{NIC2}_j = L(\hat{\theta}_j) - \lambda_1 \ln(n-k_j),$$

$$\mathrm{NIC3}_j = L(\hat{\theta}_j) - \lambda_1 k_j \ln(n - k_j),$$

$$\mathrm{NIC4}_j = L(\hat{\theta}_j) - \lambda_1 k_j^{\lambda_2} \text{ and}$$

$$\mathrm{NIC5}_j = L(\hat{\theta}_j) - \{\lambda_1 k_j + \lambda_2 \ln(n - k_j)\}.$$

We conducted simulation experiments to evaluate the performance of these proposed IC compared to the performance of existing cited IC in linear regression model settings.

The simulation results revealed that the performances of the six listed IC vary from situation to situation and from data set to data set. Even for a particular data set, the performance of the selected criteria varies from sample size to sample size. In all designs under study, the MAPCS increases as the sample sizes increases. Within a particular design under any criteria, the variation among the APCS decreases as the sample size increases. In terms of MAPCS, none of the listed criteria performs well in all situations. But for all designs and sample sizes, the MAPCS obtained using the new method and the proposed IC are always higher than that of the largest MAPCS obtained from the listed existing criteria. Also, in general, the variation among the APCS is smaller than that of all the listed criteria except RBAR.

The performances of all new proposed IC are very similar, although the performances of NIC4 and NIC5 are better than those of the others. The estimation of two parameters is required for NIC4 and NIC5 and this can be time consuming. The improvements of MAPCS obtained from NIC4 and NIC5 over NIC1, NIC2 and NIC3 are not significant. So considering the computational time and improvement of MAPCS, NIC1, NIC2 and NIC3 seem to be better than NIC4 and NIC5 for selection of linear regression models.

In Chapter 4 we introduced the application of the simulated annealing optimisation (SAO) technique for penalty function calculation in linear regression settings. Two types of penalties were used with the SAO technique to maximise the MAPCS. These are an additive penalty used with the maximised log-likelihood function and a multiplicative penalty used with the mean squared error. We conducted simulation experiments to compare the performance of the existing IC procedures with the SAO technique as a method for finding penalty function values in terms of maximising the MAPCS. Simulation results demonstrate that the MAPCS obtained using the SAO technique with additive and multiplicative penalties are always higher than those of the existing criteria. We found that the relative penalties that maximise the MAPCS for a particular model are different for different data sets. Also for competing models of the same dimension, the relative penalties that maximise the MAPCS are different, but for the existing IC they are zero. This implies that the optimal penalty does not depend only on the sample size and the number of free parameters, but also on the data generating process. Simulation results show that the APCS obtained using a multiplicative penalty with mean squared error are very similar to those obtained using an additive penalty with maximised log-likelihood function. So. one can use either of these techniques to select the best model.

In Chapter 5 we investigated the use of the SAO technique to select the best model from a set of equi-dimensional alternative models. We conducted simulation experiments to evaluate the performance of the SAO technique in selecting the best model and used three types of penalties. The penalties are existing IC (Type 1), maximisation of MAPCS using additive penalties (Type 2) and multiplicative penalties (Type 3).

From the simulation results we observed the following:

- The SAO technique with additive penalties always produced a larger MAPCS than any of the existing criteria.
- For the same set of equi-dimensional competing models the relative penalties that maximise the MAPCS are different and non-zero for different data sets. In contrast these are always equal to zero for the existing criteria.
- Exactly the same MAPCS is obtained from the different sets of relative penalties, which indicates that there is no unique set of penalties that maximises the MAPCS.
- The gap between the largest MAPCS and the smallest MAPCS obtained using the SAO technique is very small. which indicates that the maximised MAPCS is

insensitive to the initial parameter values for the SAO technique, when the competing models have the same number of parameters.

- The MAPCS obtained from additive and multiplicative penalties are very similar. Therefore, from the user's point of view, one can use either of the penalty types to select the best model.

In Chapter 6 we investigated an alternative approach of finding the penalties that makes the APCS for each model equal or nearly equal based on the SAO technique. We used the standard deviation (SD) among the APCS as a measure of variation and applied the SAO technique to find the penalties for a particular data set and set of competing models with the objective that the SD among the APCS of the different model is a minimum. For the purpose of comparison, we defined five types of penalties. These are existing IC (Type 1), maximisation of MAPCS (Type 2, additive and Type 3, multiplicative) and minimisation of variation among the APCS (Type 4, additive and Type 5, multiplicative).

From the simulation results we found that for models with an unequal number of parameters, the MAPCS obtained from additive (Type 2) and multiplicative (Type 3) penalties by minimising the variation among the APCS using SAO technique are generally lower (around six percent) than those obtained from the existing criteria (Type 1). But for most of the experiments, the MAPCS corresponding to the smallest variation among the APCS is higher (2.9%-15.1%) than the corresponding MAPCS for the existing criteria. Generally, the APCS obtained from the existing criteria are far from equal, but the APCS obtained using Type 4 and Type 5 penalties are often equal or very close to equal.

For equi-dimensional competing models, the MACPS obtained from using Type 4 and Type 5 penalties are generally higher than those of the existing criteria. The variation among the APCS obtained from Type 4 and Type 5 penalties is always lower than that of the existing criteria and in some experiments, the APCS were equal for all combinations of the initial parameter sets used for the SAO technique. From the simulation experiments it is observed that Type 4 and Type 5 penalties are insensitive to the initial parameter sets for the SAO technique for selecting the best model from a set of models of equal dimensions.

To find the best combination of the number of parameter drawings (q) and replications (N) for estimating APCS via simulation experiments for a fixed total number of simulations, qN, for selecting the best model with reliable MAPCS, we recommend the method proposed in Chapter 3. The performance of the technique we proposed to maximise the MAPCS for selecting the best model from a set of linear regression models with an unequal or an equal number of parameters is better then that of the existing criteria in all situations. The MAPCS obtained from our proposed method of minimising the variance in the average probability of correct selection is better than that of existing criteria most of the time, in particular the performance is always better when the competing models are equi-dimensional. So, for selecting the best model from a set of competing linear regression models with an unequal number of parameters, the method of maximisation of MAPCS proposed in Chapter 4 seems to be the best method. But for equi-dimensional competing linear regression models the technique of making APCS equal proposed in Chapter 6 is our recommended method.

Finally, the results from our experiments raise several interesting questions, which invite future research on the area of model selection. For example, our proposed method of model selection for linear regression models assuming ideal conditions can be extended for the models with heteroscedastic errors, serially correlated errors, non-normal error distributions and the models that violate some of the assumptions of regression analysis to test the robustness of the methods. The performance of the proposed methods can be tested for the selection of other types of models, which have not been considered in this thesis, for example non-linear models and models with lagged dependent variables. This technique may be used to select the best linear forecasting models. All the existing model selection criteria are likelihood based. For non-likelihood based methods such as selection of regression models based on generalised estimating equations (GEE), there is a lack of model selection criteria. Recently Pan (2001) proposed a modification of AIC for selection of regression models based on GEE and the use of a quasi-likelihood in place of the standard likelihood in AIC. Our method of model selection with SAO and proposed penalty types could be examined as an alternative solution with the quasi-likelihood replacing the likelihood function.

REFERENCES

Aarts, E.H., J. Korst and P.J. van Laarhoven (1997). Simulated annealing, in E.H. Aarts and J.K. Lenstra (eds.), *Local Search in Combinatorial Optimisation*, John Wiley and Sons, New York.

Akaike, H. (1969). Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21, 243-247.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, in B.N. Petrove and F. Caski (eds.), *Second International Symposium on Information Theory*, Akademiai Kiado, Budapest, 267-281.

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

Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics*, 16, 3-14.

Ali, M., A. Torn and S. Viitanen (1997). A direct search simulated annealing algorithm for optimisation involving continuous variables. *TUCS Technical Report No.* 97. Turku Centre for Computer Science, Finland.

Allen, D.M. (1971). The prediction sum of squares as a criterion for selecting predictor variables. *Technical Report No.* 23, Department of Statistics, University of Kentucky.

Allen, D.M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, 16, 125-127.

Amemiya, T. (1972). Lecture notes on econometrics, Department of Economics, Stanford University.

Amemiya, T. (1980). Selection of regressors. *International Economic Review*, 21, 331-354.

Atkinson, A.C. (1980). A note on the generalised information criterion for choice of a model. *Biometrika*, 67, 413-418.

Atkinson, A.C. (1981). Likelihood ratios, posterior odds and information criteria. *Journal of Econometrics*, 16, 15-20.

Atkinson, A.C. and V.V. Fedorov (1975). The design of experiments for discriminating between two rival models. *Biometrika*, 62, 289-304.

Azam, M.N. and M.L. King (1998). Model selection with a data-driven penalty when there is an unknown change point in the data, mimeo, Monash University, Department of Econometrics and Business Statistics, Australia.

Bai, Z.D., C.R. Rao and Y. Wu (1999). Model selection with data-oriented penalty. *Journal of Statistical Planning and Inference*, 77, 103-117.

Bhansali, R.J. and D.Y. Downham (1977). Some properties of the order of an autoregressive model selected by a generalisation of Akaike's FPE criterion. *Biometrika*, 64, 547-551.

Billah, M.B. (2001). Model selection for time series forecasting models, Unpublished Ph.D. thesis, Monash University, Australia.

Billah, M.B. and M.L. King (1998). Model selection and optimal penalty function estimation for time series models, mimeo, Department of Econometrics and Business Statistics, Monash University, Australia.

Billah, M.B. and M.L. King (2000a). Time series model selection and forecasting via optimal penalty estimation. *Pakistan Journal of Statistics*, 16, 126-145.

Billah, M.B. and M.L. King (2000b). Using simulated annealing to estimate penalty functions for time series model selection, mimeo, Department of Econometrics and Business Statistics, Monash University.

Bohachevsky, I., M.E. Johnson and M.L. Stein (1986). Generalised simulated annealing for function optimisation. *Technometrics*, 28, 209-217.

Bora-Senta, E. and S. Kounias (1986). Parameter estimation and order determination of autoregressive models, in O.D. Anderson (ed.), *Analysing Time Series*, North-Holland Publishing Company, Amsterdam, 93-108.

Breiman, L. and D. Freedman (1983). How many variables should be entered in a regression equation? *Journal of the American Statistical Association*, 78, 131-136.

Brownstone, D. (1990). How to "data mine" if you must: Bootstrapping Stein-rule model selection procedures. *Technical Report No.* MBS-08, Irvine Research Unit in Mathematical Behavioural Sciences, UC Urvine.

Bunke, H. and B. Droge (1985). A stepwise procedure for the selection of non linear models. *Mathematische Operationsforschung und Statistik, Series Statistics*, 16, 35-45.

Carnevali, P., L. Coletti and S. Patarnello (1985). Image processing by simulated annealing. *IBM Journal of Research and Development*, 29, 569-579.

Cerny, V. (1985). A thermodynamic approach to the traveling salesman problem: An efficient solution. *Journal of Optima Theory and Application*, 45, 41-51.

Chen, C.H., R.A. Davis, P.J. Brockwell and Z.D. Bai (1993). Order determination for autoregressive processes using resampling methods. *Statistica Sinica*, 3, 481-500.

Corana, A., M. Marchesi, C. Martini and S. Ridella (1987). Minimising multimodal functions of continuous variables with the simulated annealing algorithm. *ACM Transactions of Mathematical Software*, 13, 262-280.

Cramer, J.S. (1986). *Econometric Applications of Maximum Likelihood Methods*, Cambridge University Press, New York.

Crato, N. and B.K. Ray (1996). Model selection and forecasting for long-range dependent processes. *Journal of Forecasting*, 15, 107-125.

Craven, P. and G.S. Wahba (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalised cross validation. *Numerische Mathematik*, 31, 377-403.

Davidson, R. and J.G. MacKinnon (1984). Model specification tests based on artificial linear regressions. *International Economic Review*, 25, 485-501.

DeGroot, M.H. (1970). *Optimal Statistical Decisions*, McGraw-Hill, New York.

Derwent, D. (1988). A better way to control population. *Nature*, 331, 575-578.

Dhrymes, P. (1970). On the game of maximising $\overline{R}^2$. *Australian Economic Papers*, 9, 117-185.

Durbin, J. and G.S. Watson (1951). Testing for serial correlation in least squares regression. *Biometika*, 159-178.

Eastwood, A. and L.G. Godfrey (1992). The properties and constructive use of misspecification tests for multiple regression, in L.G. Godfrey (ed.), *The*

Implementation and Constructive Use of Misspecification Tests in Econometrics, Manchester University Press, Manchester, 109-175.

Finch, S.J., N.R. Mendall and H.C. Thode (1989). Probabilistic measures of accuracy of numerical search for global maximum. *Journal of the American Statistical Association*, 84, 1020-1023.

Forbes, C.S., M.L. King and A. Morgan (1995). Small sample variable selection procedures. An unpublished manuscript, Monash University.

Fox, K.J. (1995). Model selection criteria: A reference source. An unpublished manuscript, University of New South Wales.

Gaver, K.M. and M.S. Geisel (1974). Discriminating among alternative models: Bayesian and Non-Bayesian methods, in Paul Zarembka (ed.), *Frontiers of Econometrics*. Academic Press, New York, 49-77.

Geisser, S. (1974). A predictive approach to the random effect model. *Biometrika*, 61, 101-107.

Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association*, 70, 320-328.

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

Goffe, L.G., G.D. Ferrier and J. Rogers (1994). Global optimisation of statistical functions with simulated annealing. *Journal of Econometrics*, 60, 65-99.

Gorman, J.W. and R.J. Toman (1966). Selection of variables for fitting equations to data. *Technometrics*, 8, 27-51.

Granger, C.J.W., M.L. King and H. White (1995). Comments on testing economic theories and the use of model selection criteria. *Journal of Econometrics*, 67, 173-187.

Grassa, A.A. (1989). *Econometric Model Selection: A New Approach*. Kluwer Academic Press, Boston.

Gray, P., W. Hart, L. Painton, C. Philips, M. Trahan, and J. Wagner. (1997). A survey of global optimisation. Sandia National Laboratories, Albuquerque. [Online]. Available: http://www.cs.sandia.gov/opt/survey/main.html.

Grose, S.D. and M.L. King (1994). The use of information criteria for model selection between models with equal numbers of parameters. Paper presented at the 1994 Australasian Meeting of the Econometric Society, Armidale.

Hannan, E.J. (1982). Testing for autocorrelation and Akaike's criterion, in J. Gani and E.J. Hannan (eds.), *Essays in Statistical Science, Papers in Honour of P.A.P. Moran*, Applied Probability Trust, Sheffield, 403-412.

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

Harvey, A.C. (1981). *The Econometric Analysis of Time Series*. Philip Allan Publisher, The United Kingdom.

Hocking, R.R. (1976). The analysis and selection of variables in linear regression. *Biometrika*, 32, 1-49.

Holmes, J.M. and P.A. Hutton (1989). Optimal model selection when the true relationship is weak and occurs with a delay. *Economics Letters*, 30, 333-339.

Hossain, Z. (1998). Model selection problems involving interval restricted parameters in econometrics. Unpublished Ph.D. thesis, Department of Econometrics and Business Statistics, Monash University.

Hughes, A.W. (1997). Improved model selection based on AIC-type criteria. Unpublished Ph.D. thesis, Department of Econometrics and Business Statistics,, Monash University.

Hughes, A.W. and M.L. King (1994). One-sided model selection procedures. Paper presented at the 1994 Australasian Meeting of the Econometric Society, Armidale.

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

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

Hurvich, C.M. and C.-L. Tsai (1991). Bias of the corrected AIC criterion for underfitted regression and time series models. *Biometrika*, 78, 499-509.

Hurvich, C.M. and C.-L. Tsai (1993). A corrected Akaike information criterion for vector autoregressive model selection. *Journal of Time Series Analysis*, 14, 271-279.

Johnson, D., and L. McGeoch (1997). The traveling salesman problem: A case study in local optimisation, in E.H. Aarts and J.K. Lenstra (eds.), *Local Search in Combinatorial Optimisation*, John Wiley and Sons, New York.

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl and T.-C. Lee (1985). *The Theory and Practice of Econometrics*, 2nd edition, John Wiley, New York.

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl and T.-C. Lee (1988). *Introduction to the Theory and Practice of Econometrics*. 2nd edition. John Wiley, New York.

Kinal, T. and K. Lahiri (1984). A note on selection of regressors. *International Economic Review*, 25, 624-629.

King, M.L. (1987). Towards a theory of point-optimal testing. *Econometric Reviews*, 6, 169-218.

King, M.L. and G.K. Bose (2002). Finding optimum penalties for model selection in linear regression, in D.E.A. Giles (ed.), *Computer Aided Econometrics*, Marcel Dekker (forthcoming).

King, M.L., C.S. Forbes and A. Morgan (1995). Improved small sample model selection procedures. Paper presented at the 1995 World Congress of the Econometric Society, Tokyo.

Kirkpatrick, S., C.D. Gelatt and M.P. Vecchi (1983). Optimization by simulated annealing. *Science*, 220, 671-680.

Koehler, A.B. and E.S. Murphree (1988). A comparison of the Akaike and Schwarz criteria for selecting model order. *Applied Statistics*, 37, 187-195.

Kohn, R. (1983). Consistent estimation of minimal subset dimension. *Econometrica*, 51, 367-376.

Kullback, S. and R.A. Leibler (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 13, 79-86.

Kwek, K.T. (2000). Model selection for a class of conditional heteroscedastic processes, Unpublished Ph.D. thesis, Department of Econometrics and Business Statistics, Monash University.

Kwek, K.T. and M.L. King (1997a). Model selection and optimal penalty estimation for conditional heteroscedastic processes, in P. Bardsley and V.L. Martin (eds.), *Proceedings of the Econometric Society Australasian Meeting*, 2, University of Melbourne, Australia, 143-176.

Kwek, K.T. and M.L. King (1997b). Model selection of ARCH process using iterative optimal penalties. *Proceedings of the Third Annual Doctoral Research Conference*, Monash University, Clayton, Victoria, 51-62.

Kwek, K.T. and M.L. King (1998). Information criteria in conditional heteroscedastic models: A Bayesian approach to penalty function building, in C.L. Skeels (ed.), *Proceedings of the Econometric Society Australasian Meeting*, CD ROM, The Australian National University, Canberra, Australia.

Lawler, E.L. (1976). *Combinatorial Optimisation: Networks and Matroids*, Holt, Rinehart and Winston, New York.

Leamer, E.E. (1978). *Specification Searches*. Wiley, New York.

Linhart, H. and W. Zuchini (1986). *Model Selection*. Wiley, New York.

Lütkepohl, H. (1984). Forecasting contemporaneously aggregated vector ARMA processes. *Journal of Business and Economic Statistics*, 2, 201-214.

MacKinnon, J.G. (1983). Model specification tests against non-nested alternatives. (with discussion and reply), *Econometric Reviews*, 2, 85-157.

Maddala, G.S. (1992). *Introduction to Econometrics*. Macmillan Publishing Company, New York.

Malinvaud, E. (1981). Econometrics faced with the need of microeconomic policy, *Econometrica*, 49, 1363-1375.

Mallows, C.L. (1964). Choosing variables in a linear regression: A graphical aid. Paper presented at the Central Regional Meeting of the Institute of Mathematical Statistics, Manhattan, Kansas.

McAleer, M. (1987). Specification tests for separate models: A survey. in M.L. King and D.E.A. Giles (eds.), *Specification Analysis in the Linear Model: Essays in Honour of Donald Cochrane*, Routledge and Kegan Paul, New York, 146-196.

Meese, R. and J. Geweke (1984). A comparison of autoregressive univariate forecasting procedures for macroeconomic time series. *Journal of Business and Economic Statistics*, 2, 191-200.

Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller and E. Teller (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1087-1091.

Mills, J.A. and K. Prasad (1992). A comparison of model selection criteria. *Econometric Reviews*, 11, 201-233.

Nishii, R. (1986). Criteria for selection of response variables and the asymptotic properties in a multivariate calibration. *Annals of the Institute of Statistical Mathematics*, 38, 319-329.

Nishii, R. (1988). Maximum likelihood principle and model selection when the true model is unspecified. *Journal of Multivariate Analysis*, 27, 392-403.

O'Hagan, J. and B. McCabe (1975). Tests for the severity of multicollinearity in regression analysis: A comment. *Review of Economics and Statistics*, 57, 368-370.

Pan, W. (2001). Akaike's information criterion in generalised estimating equations. *Biometrics*, 57, 120-125.

Papadimitriou, C.H. and K. Steiglitz (1982). *Combinatorial Optimisation: Algorithms and Complexity*. Prentice-Hall, New York.

Pearce, D.W. (1992). *Macmillan Dictionary of Modern Economics*. Macmillan Publisher, The United Kingdom.

Pesaran, M.H. and B. Pesaran (1987). *Data-fit: An Interactive Econometric Software Package*. Oxford University Press, The United Kingdom.

Pinter, J.D. (1996). Continuous global optimisation software: A brief review [Online]. Available: http://mat.gsia.cmu.edu/pinter.fil.

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

Pronzato, L., E. Walter, A. Venot and J.F. Lebruchec (1984). A general purpose optimiser: Implementation and applications. *Mathematics and Computers in Simulations*, 24, 412-422.

Rahman, M.S. (1998). Generalised Model Selection Criterion, in C.L. Skeels (ed.), *Proceedings of the Econometric Society Australasian Meeting*, CD ROM, Australian National University, Canberra, Australia.

Rahman, M.S. and M.L. King (1997). Probability of correct identification in model selection, in P. Bardsley and V.L. Martin (eds.), *Proceedings of the Econometric Society Australasian Meeting*, 2, University of Melbourne, 613-640.

Rahman, M.S. and M.L. King (1999). Improved model selection criterion. *Communications in Statistics – Simulation and Computation*, 28, 51-71.

Ramsey, J.B. (1969). Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society*, B, 31, 350-371.

Rao, C.R. and Y. Wu (1989). A strongly consistent procedure for model selection in a regression problem. *Biometrika*, 76, 369-374.

Rissanen, J. (1978). Modelling by shortest data description. *Automatica*, 14, 465-471.

Rissanen, J. (1986). Stochastic complexity and modelling. *Annals of Statistics*, 14, 1080-1100.

Rissanen, J. (1987). Stochastic complexity and modelling. *Journal of the Royal Statistical Society*, B, 49, 223-239 and 252-265 with discussions.

Rissanen, J. (1988). Stochastic complexity and the MDL principle. *Econometric Reviews*, 85-102.

Romeo, F., V.A. Sangiovanni and C. Sechen (1984). Research on simulated annealing at Berkeley, in *Proceedings of the IEEE International Conference on Computer Design*, CCD 84, IEEE, New York, 652-657.

Rothman, D. (1968). Letter to the editor. *Technometrics*, 10, 432.

SAS Institute Inc. (1992). *Principles of Regression Analysis Course Notes*. USA.

Sawa, T. (1978). Information criteria for discriminating among alternative regression models. *Econometrica*, 46, 1273-1291.

Schmidt, P. (1971). Methods of choosing among alternative models. *Econometrics Workshop Paper No.* 7004, Michigan State University.

Schmidt, P. (1973). Calculating the power of the minimum standard error choice criter[illegible] *[illegible]ational Economic Review*, 14, 253-255.

Schmidt, [illegible] Choo[illegible]ng among alternative linear regression models. *Atlantic [illegible]*, 2, 7-13.

Schmidt, P. (1975). Choosing among alternative linear regression models: A correction and some further results. *Atlantic Economic Journal*, 3, 61-63.

Schmidt, P. and R. Tscherning (1993). Identification of fractional ARIMA models in the presence of long memory. Discussion paper. University of Munich, Munich, Germany.

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

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

Shibata, R. (1976). Selection of the order of an autoregressive model by Akaike's information criterion. *Biometrika*, 63, 117-126.

Shibata, R. (1981). An optimal selection of regression variables. *Biometrika*, 68, 45-54.

Shibata, R. (1986). Consistency of model selection and parameter estimation, in J. Gani and M.B. Priestley (eds.), *Essays in Time Series and Allied Processes, Journal of Applied Probability*, 23A, 127-141.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictors (with discussion). *Journal of the Royal Statistical Society*, B, 36, 111-147.

Stone, M. (1977). Asymptotics for and against cross-validation. *Biometrika*, 64, 29-35.

Sugiuro, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methodology*, 7, 13-26.

Summers, R. and A. Heston. (1991). The Penn World Table (Mark 5): An extended set of international comparisons, 1950-1988. *Quarterly Journal of Econometrics*, 106, 327-368.

Szu, H. and R. Hartley (1987). Fast simulated annealing. *Physics Letters*, A, 3, 157-162.

Telly, H., Th.M. Liebling and A. Mocellin (1987). Reconstruction of polycrystalline structures: A new application of combinatorial optimisation. *Computing*, 38, 1-11.

Theil, H. (1961). *Economic Forecasts and Policy*. Second edition. North-Holland, Amsterdam.

Theil, H. (1971). *Principles of Econometrics*. Wiley, New York.

Thompson, M.L. (1978). Selection of variables in multiple regression. *International Statistical Review*, 46, 1-19 and 129-146.

Vanderbilt, D. and S.G. Louie (1984). A Monte Carlo simulated annealing approach to optimisation over continuous variables. *Journal of Computational Physics*, 56, 259-271.

Vetterling, W.T., W.H. Press, S.A. Teukolsky and B.P. Flannery (1994). *Numerical Recipes in FORTRAN: The Art of Scientific Computing*. Cambridge University, New York.

Wallace, T.D. (1977). Pre-test estimation in regression: A survey. *American Journal of Agricultural Economics*, 59, 431-443.

Wang, D.F, H.W. Leong and C.-L. Liu (1988). *Simulated Annealing for VLSI Design*. Kluwer Academic Publishers, Boston, MA.

Wasserman, P.D. and T. Schwartz (1988). Neural networks, part 2: What are they and why is everybody so interested in them now? *IEE Expert*. Spring, 10-15.

White, H. (1982a). Maximum likelihood estimation of misspecified models. *Econometrica*, 50, 1-25.

White, H. (ed.). (1982b). Model specification. *Journal of Econometrics*, 20, 1-157.

White, H. (ed.). (1983). Non-nested models. *Journal of Econometrics*, 21, 1-160.

White, H. (1990). A consistent model selection procedure based on M-testing, in C.W.J. Granger (ed.), *Modelling Economic Series: Readings in Econometric Methodology*, Oxford University Press, Oxford.

White, K. (1978). A general computer program for econometric methods – SHAZAM, *Econometrica*, 46, 239-240.

White, S.R. (1984). Concepts of scale in simulated annealing, in *Proceedings of the IEEE International Conference on Computer Design*, ICCD 84, IEEE, New York, 46-651.

Williams, P. and D. Sams (1981). Household headship in Australia: Further developments to the IMPACT project's econometric model of household headship. *IMPACT Project Research Centre Working Paper*, No. BP-26, August 1981, 91-96.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*, John Wiley and Sons, New York.

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