MONASH UNIVERSITY

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

ON.................. 9 December 2003 ....

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

Sec. Research Graduate School Committee

Multiple Maxima of Likelihood Functions and their Implications for Inference in the General Linear Regression Model

A thesis submitted for the degree of
Doctor of Philosophy

by

Mahbuba Yeasmin
BSc.(Hons.), MSc.(Statistics) Jahangirnagar University
Dhaka, Bangladesh,

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

June 2003

Contents

List of Tables

List of Figures

Declaration

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

[redacted]

Mahbuba Yeasmin

Acknowledgments

I am grateful to my supervisor Professor Maxwell L. King for his guidance, encouragement that I enjoyed during the preparation of this thesis. I was fortunate to have benefited greatly from the profound insight that Professor King so generously imparted. His constructive criticisms, continuous encouragement and affection at every stage of my work are highly appreciated. Professor King was a source of inspiration for me throughout this work and I am grateful for his help and care while I was sick. I also acknowledge his editorial assistance in the writing of this thesis.

Many thanks are due to Mrs. Pauline Froggatt, Mrs. Philippa Geurens, Mrs. Julie Larsen, Mr. Myles Tooher, Ms. Inge Meldgaard, Ms. Deanna Orchard and Mrs. Mary Englefield for their administrative support in various ways that have helped me successfully complete this thesis.

I want to acknowledge with gratitude my debt to Monash University for awarding me a Monash Graduate Scholarship, Faculty Fee Paying Scholarship, and to the Department of Econometrics and Business Statistics for offering me a Departmental Scholarship. I acknowledge with gratitude to Professor King the importance of a part-time research assistant position before and throughout my PhD program, which helped me financially. I extended my thanks to Associate Professor

Paramsothy Silvapule for providing me part-time teaching position that gave me financial support and thoughtful experiences. An excellent working environment in the Department of Econometrics and Business Statistics is highly appreciated.

My never-ending thanks to my mother Mrs. Mahmuda Begum for her love, sacrifice and moral support through out my life. I am also grateful to my brothers and sisters Dr. Mizanur Rahman, Mr. Mesbahur Rahman, Mr. Mowdudur Rahman, Dr. Mostafizur Rahman, Mrs. Zannatun Nahar, Mrs. Mansura Parvin and Mrs. Mottahera Nasrin for their encouragement and love.

My main indebtedness and gratitude are to my husband, Dr. Mohammad N. Azam and my son Mim Mahbubul Azam, for all their love, sacrifice, patience, support and encouragement and for managing to survive on less than their reasonable share of attention during the entire course of my study.

Finally, I dedicate this thesis to the memory of my father and to all people of the world, who are steadfast in their work despite physical illness.

Abstract

This thesis is concerned with the issue of multiple maxima of likelihood functions and its implication for inference in the general linear regression model. Special attention is given to regression models with a first order moving average error process and to marginal likelihood methods. We also investigate whether estimating equations can help to improve the understanding of the estimating process in the context of the general linear model.

The first contribution of this thesis is to study the consequences of different strategies for maximising the likelihood function when a local maximum is distinctly possible on tests of regression coefficients. This is done in the context of maximum likelihood estimation of the general linear model with a first order moving average error process.

As a second contribution, we investigate two possible improvements to standard tests of the regression coefficients. We reassesses the importance of taking care in finding the global maximum of the likelihood function and investigate improvements through modified test statistics by using corrected degrees of freedom when the model contains a nuisance parameter and also using marginal likelihood based estimates. We investigate the effectiveness of these modifications through a Monte Carlo simulation.

The third contribution of this thesis is to look at the consequences of different strategies for maximising the likelihood function on forecasting performance of the linear regression model when maximum likelihood and marginal likelihood methods are used to estimate the nuisance parameter when there are first order moving average disturbances. In addition, we look at forecasting performance when the initial term of the forecasting error is estimated and compare it with that when the initial error is set to zero in the recursive calculations required for one-step-ahead forecasts. We compare the forecasting performance for both one-step-ahead and two-step-ahead forecasts.

The final contribution is to show the estimating equation for the least squares method applied to a transformed model gives the same expression as that of the estimating equation of the concentrated likelihood. The estimating equation can be expressed as a polynomial of fifth degree when regression design matrix is a vector in the context of the linear model with a non-stationary first order autoregressive error process and also an infinite degree polynomial of the parameter when regression design matrix is not a vector. Instead of using an infinite degree polynomial, which is analytically difficult or even impossible to solve, we suggest the use of an approximating polynomial of finite degree, which is easier to solve.

In conclusion, this thesis warns that looking for the global maximum of the likelihood function is not always best choice in the case of maximum likelihood because it may give tests with poor small sample sizes. This is not the case for the maximum marginal likelihood estimation method which we find to be best for inference in the linear regression model with a first order moving average error process. When using the marginal likelihood, the best test and forecasting results seem to come from taking considerable care to find its global maximum.

Glossary

AAMFE	Average Absolute Mean Forecast Error
AIC	Akaike Information Criterion
AMAE	Average Mean Absolute Error
AMFE	Absolute Mean Forecast Error
AMSFE	Average Mean Square Forecasted Error
AR(1)	First-order Autoregressive
BA	Boltzmann annealing
BIC	Schwarz Bayesian Information Criterion
CO	Cochrane-Orcutt
GLS	Generalised Least Squares
GS	Grid Search
HL	Hildreth-Lu
IC	Information Criteria
KL	Kullback-Leibler
LM	Lagrange multiplier
LR	Likelihood Ratio
LS	Least Squares
MA(1)	First-order Moving Average
MAFE	Mean Absolute Forecast Error
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator

MML	Maximum Marginal Likelihood
MSFE	Mean Square Forecast Error
MVU	Minimum Variance Unbiased
OLS	Ordinary Least Squares
OSAF	One-Step-Ahead Forecasts
REML	Restricted Maximum Likelihood
SA	Simulated Annealing
TSAF	Two-Step-Ahead Forecasts

CHAPTER 1

Introduction

1.1 Prologue

Econometrics in a broad sense is a set of quantitative techniques that are useful for making inferences about economic behaviour under uncertainty. It is concerned with extracting the best possible information from data in order to make inferences. Within this context, estimation is a very widely used statistical method, which depends on the scientific deduction of an unknown quantity from sample information. The aim of estimation is to approximate the values of unknown parameters using the available data in a way that provides a useable approximation and some idea of the quality of that approximation.

Applied econometricians and statisticians are usually interested in effective inferences using econometric or statistical models for the purpose of forecasting, policy analysis or to have a better understanding of the economy. For this reason, they need to study the properties of estimators with particular emphasis on the accuracy of the ensuing inference. In econometric modelling, information on the distributional properties of an estimator is important for making quality statistical inferences and being certain about the quality of parameter estimates of the model. In applied work, the probability distribution of the estimator is not always known and, in addition, exact statistical inference may be computationally difficult. In such cases, econometricians depend on the asymptotic properties of the estimator in the hope that asymptotic approximations will give reasonably accurate results. Unfortunately, their accuracy seems to be doubtful in some cases.

An example of one such case is given by King and McAleer (1987) who considered testing first-order autoregressive (AR(1)) disturbances against first-order moving average (MA(1)) disturbances in the linear regression model. They reported that the asymptotic Cox test of AR(1) errors against MA(1) errors has less than ideal small sample properties. They found that the Cox test based on asymptotic critical values at the 0.05 level can have true sizes in excess of 0.70, 0.40 and 0.20 when the sample sizes are 15, 30 and 60 respectively. Godfrey (1978), Griffiths and Surekha (1986), King (1987a), King and McAleer (1987), Honda (1988), Moulton and Randolph (1989), Chesher and Austin (1991), Kennedy and Simons (1991) and Latif and King (1993), are examples of other studies in a range of different settings in which the accuracy of asymptotic tests is questioned.

Obviously the best estimator may be regarded as the one whose distribution concentrates as closely as possible near the true value of the parameter being estimated. In statistical inference, the most popular and widely used estimation methods are those of maximum likelihood (ML) and least squares (LS). The method of LS involves finding the parameter values that minimize the sum of squared residuals. The difference between the dependent variable and its estimated expected value is known as the residual. The LS method was first named and published by Legendre in 1805 (Merriman, 1877). Gauss became familiar with his work and took an early opportunity to study the method in detail (Gauss, 1806, p.184). The method has a long history of use with a great number of variations, extensions and applications being published in the literature. Partly its popularity comes from the fact that it is a simple technique that does not require exact specification of the population distribution.

ML is a general method of estimation, which chooses the value of the unknown parameter in such a way that maximizes the likelihood function with respect to the unknown parameter. R.A. Fisher first introduced maximum likelihood in the early 1920's as a tool for estimating unknown parameters. Over the last few decades, ML has been a widely used method of estimation because, under some fairly general conditions, ML estimators are consistent, asymptotically normal, asymptotically unbiased and asymptotically efficient, see for example, Lehmann (1983) and LeCam (1990). It can be applied to a wide range of different parametric models and the likelihood ratio (LR), Wald and Lagrange multiplier (LM) test procedures have been developed based on likelihood functions.

In econometrics and statistics, model selection is another useful technique that has originated from the maximum likelihood method. Akaike (1974) proposed his widely used information criteria known as AIC and in 1973 he

(Akaike (1973)) also suggested selecting the best fitting model using this criteria to discriminate between alternative models. AIC was developed incorporating Kullback-Leibler (KL) information with the use of maximum likelihood principles and negative entropy. The form of AIC is that it uses the penalized maximized log-likelihood form given as $AIC = l(\hat{\gamma}) - p$, where $l(\hat{\gamma})$ is the maximized log-likelihood of the model, $\hat{\gamma}$ is the estimated parameter vector and p is the penalty term which is the number of free parameters included in the model under consideration. The main disadvantage of AIC is that it is an inconsistent information criterion. Schwartz (1978) modified the AIC procedure and proposed another model selection criteria based on maximum likelihood known as BIC, which overcomes the disadvantages of AIC for large samples.

In a given econometric setting, all the parameters of a model may not be equally important to the researcher. Estimating the parameters of interest without taking care of unwanted parameters, that is nuisance parameters, may give us misleading information about the model under consideration. To overcome this problem, we need to be able to estimate the nuisance parameters effectively so that we can reduce any estimation bias.

Nuisance parameters are also known as incidental parameters. They are not parameters of interest but are needed in order to make the model under consideration realistic. Traditionally we often rely on concentrating the likelihood function in order to estimate the nuisance parameters. Neyman and Scott (1948) introduced the nuisance parameter problem and later Kalbfleisch and Sprott (1970), Cooper and Thompson (1977), Corduas (1986), Tunnicliffe

Wilson (1989), McCullagh and Nelder (1989), Bellhouse (1991), King (1996), Laskar and King (1996b) and Rahman and King (1997) discussed problems of nuisance parameters in practical situations. The view that has emerged from this literature is that it is better to use the marginal likelihood rather than the concentrated likelihood when estimating some key nuisance parameters. It helps to reduce the estimation bias. For example, Ara (1995) and Rahman and King (1998) showed marginal likelihood based tests have better small sample properties in the context of tests of regression disturbances than corresponding tests based on the concentrated likelihood. For a general discussion of marginal likelihoods, see for example Kalbfleisch and Sprott (1970), Cox (1988), McCullagh and Nelder (1989), Tunnicliffe Wilson (1989), Rahman and King (1994) and Laskar and King (1998). This will be further discussed in Chapter 2.

The first derivative of the likelihood function is known as Fisher's score function. The ML estimate can be found by setting the score to zero and solving for the parameter values. The equation involving the score function set to zero is known as an estimating equation because solving it allows us to find the required estimate. Usually an estimating equation of a non-linear character makes solution for the parameters difficult. In this situation, numerical methods can help to find an appropriate solution via an iterative formula.

A number of authors have used estimating equations for choosing appropriate estimates. For example, Godambe (1960) used the score of the estimating equation to show an optimum property of regular maximum likelihood estimation. Godambe and Thompson (1974) modified the score of the estimating equation for regular maximum likelihood estimation as a method of bias correction. Mahmood (2000) showed how the estimating equation approach can be used to derive estimators which possess good small sample properties. Using an estimating equation approach, he assessed various modified likelihood methods of estimation and concluded that the marginal likelihood is the best overall likelihood for statistical inference in the general linear model.

Forecasting plays an important role in the field of econometrics, statistics and many other branches of science. One standard approach to producing a good forecast is to construct a reliable statistical or econometric model. This typically involves the estimation of unknown parameters of the model. There is a common view that a good forecast can only be obtained when the estimation of the parameters of the model can be done with minimum error. In other words, a good estimate of the parameters of the model will provide good forecasts when used in the model.

1.2 Objectives of this Thesis

In this thesis we look at the issue of global maximum versus local maxima of the likelihood function and its impact on testing regression coefficient and making forecasts in the linear model. We also investigate whether estimating equations can help to improve the understanding of the estimating process in the context of the general linear model.

The specific objectives of this thesis are to:

(i) Look at the consequences of different strategies for maximising the likelihood function when a local maximum is distinctly possible, on test sizes when estimating a nuisance parameter based on the concentrated likelihood and marginal likelihood for the general linear model.

(ii) Develop modified test statistics for testing regression coefficients by correcting degrees of freedom when the model contains nuisance parameters and to investigate the usefulness of such modifications through Monte Carlo simulations.

(iii) Investigate the consequences of different strategies for maximising the likelihood function on forecasting performance when the ML and maximum marginal likelihood (MML) methods are used to estimate the nuisance parameter in the context of the linear regression model with MA(1) errors.

(iv) Through an estimating equation approach, establish an equivalence between least squares and ML based on the concentrated likelihood function in the context of the general regression model.

(v) Explore analytical solutions of estimating equations in the context of a linear regression with non-stationary AR(1) errors.

1.3 Organization of the Thesis

In Chapter 2, we review estimation of the linear model in the presence of nuisance parameters. Here the particular problem of estimation when the errors follow an MA(1) process are discussed. We review some methods of numerical optimization of the likelihood function, namely the marginal likelihood and the concentrated likelihood, along with the estimating equation of the concentrated likelihood for estimating the parameters of the general linear regression model. We also discuss some aspects of estimating equations and multiple maxima, highlighting their role. In addition, we briefly outline the literature on the simulated annealing algorithm for finding the global maximum of an objective function.

Chapter 3 deals with the consequences on the size of the standard test of regression coefficients of accepting local maxima instead of the global maximum of a concentrated likelihood function involving MA(1) regression disturbances. We compare the estimated sizes of the test when estimation is done through different strategies for different values of the MA(1) error parameter and different sample sizes using Monte Carlo simulation. We conclude that, in general, our calculated test sizes are unacceptably high.

In Chapter 4, we investigate two possible improvements to the test of the regression coefficients aimed at improving size and reassess the importance of taking care in finding the global maximum of the likelihood. One possible improvement is to use the method of MML. We also construct test statistics using various degrees of freedom based on estimation of all parameters (except the variance term) and calculate the estimated sizes of tests using different strategies for estimating the moving average parameter with different degrees of freedom for different values of the moving average parameter.

In Chapter 5, first we look at the effect of different strategies for maximising the likelihood function on forecasting performance when estimation of the moving average parameter is based on the ML and MML methods. We then see whether taking extra care in finding the global maxima of likelihood functions improves the accuracy of forecasting performance. We also look at whether there is any difference between the forecasting performance of the model when the first term of the forecasting error is replaced by its estimated value or by zero as suggested by King and McAleer (1987) in the context of the linear regression model with MA(1) errors. We use the concentrated likelihood and the marginal likelihood to estimate the moving average parameter through different strategies to assess the forecasting performance. We also compare the forecasting performance of the model for different values of the parameter in the case of one-step-ahead forecasts and two-step-ahead forecasts.

In Chapter 6, we consider the concentrated likelihood and marginal likelihood for the general linear model and give the score of the concentrated likelihood and the marginal likelihood. We explore whether there is equivalence between the LS method and the maximum likelihood method and examine estimating equations for the LS method based on a specially transformed model with that from the concentrated likelihood function. We discuss estimating the covariance matrix parameter in the linear regression model with a non-stationary AR(1) error process. We show that the estimating equation is a polynomial of infinite degree in the case of non-stationary AR(1) errors. To find the roots of this polynomial, we use different mathematical expressions to approximate the polynomial up to degree four. We also show a special case where the estimating equation turns into a fifth degree polynomial when the design matrix is a single vector for this nonstationary AR(1) process. We make a suggestion for finding the roots of the polynomial and choosing an appropriate root as our desired estimate. We also derive an iterative formula to find roots of the higher degree polynomial.

We conclude this thesis with Chapter 7. It includes a summary of the results, conclusions and contributions of this thesis, and gives some suggestions for further research.

CHAPTER 2

Review of the Literature

2.1 Introduction

This thesis covers four main subject areas. These are finding the global maximum of likelihood functions, testing regression coefficients, estimation of nuisance parameters in the general linear regression model with an MA(1) error process and the consequences of using global and local maxima of the likelihood function on forecasting performance. The aim of this chapter is to survey literature on these and relevant areas with particular emphasis on estimation in the case of the general linear regression model. We discuss the problem of numerical optimisation of likelihood functions, and in particular the use of marginal likelihood and concentrated likelihood functions. We investigate estimating equations for parameter estimation of the general linear regression model and review some aspects of multiple maxima of likelihoods, highlighting their role in statistical inference.

Following the above introduction, we begin the review of the literature in Section 2.2 with a brief discussion on likelihood based inference in econometrics and statistics. In this section we discuss maximum likelihood estimation, the score vector, the information matrix and the Newton-Raphson method. In Section 2.3 we review the simulated annealing algorithm used for optimization. In Section 2.4 we discuss the concentrated likelihood function of the general linear model and survey the problem of multiple maxima of likelihood functions. The marginal likelihood approach is outlined in Section 2.5. In Section 2.6 we review the use of estimating equations and exploit the properties of this approach to address theoretical and empirical issues of estimation problems based on likelihood functions. Finally, Section 2.7 contains some concluding remarks.

2.2 Likelihood Based Inference

In econometrics and statistics, the likelihood function plays an important role. Fisher (1922) first introduced maximum likelihood as a tool for estimating unknown parameters. After Fisher's contribution, the likelihood became a central concept of parametric statistical methods both in the Bayesian and in the classical approaches. Since then, likelihood methods have become increasingly popular. In frequency based inference, the main reason for their widespread use is that sampling distributions of statistics such as the maximum likelihood estimator and the likelihood ratio test statistic have simple and well understood (first-order) asymptotic approximations in many relevant models. Much of the work on likelihood inference over the past few decades has been aimed at refining and improving upon first-order approximations by moving to higher-order asymptotics.

Maximum likelihood is probably the most popular estimation procedure in context of multi-parameter models. Its main aim is to find the parameter value(s) that makes the observed data most likely. This is because the likelihood of the parameters given the data is defined to be equal to the probability of the data given the parameters. The purpose of this section is to provide a brief introduction to the ideas behind maximum likelihood estimation.

Let $y_1,\cdots,y_n$ be n independent random variables with probability density functions $f_i(y_i;\gamma)$ depending on a vector-valued parameter γ. The joint density of n independent observations $y=(y_1,\cdots,y_n)'$ is

$$f(y;\gamma)=\prod_{i=1}^{n}f_i(y_i;\gamma)=L(\gamma;y). \tag{2.1}$$

This expression, viewed as a function of the unknown parameter γ given the data y, is called the likelihood function. Often we work with the natural logarithm of the likelihood function, the so-called log-likelihood function:

$$log\,L(\gamma;y)=\sum_{i=1}^{n}log\,f_i(y_i;\gamma). \tag{2.2}$$

In linear regression or analysis of variance, we typically turn to the principle of least squares for estimation of the unknown parameters of the model. The idea of least squares is to choose parameter estimates that minimize the average squared difference between observed and predicted values of the data being modelled. That is, to maximize the fit of the model to the data by choosing the model that is closest, in a least squares sense, to the data.

For many other models such as logistic, Poisson, and proportional hazard regression, least squares usually cannot be used as an estimation method because of the non-linear nature of these models. Instead, we turn to the method of maximum likelihood estimation. Maximum likelihood estimation involves searching over all possible sets of parameter values for a specified model and finding the set of parameter values which, given the observed sample, is most likely. This is done by maximizing the likelihood (or equivalently the log-likelihood) function. Formally, the maximum-likelihood estimator (MLE) is the value $\hat{\gamma}$ such that $\log L(\hat{\gamma};y) \geq \log L(\gamma;y)$ for all γ.

2.2.1 The score function

The first derivative of the log-likelihood function is called Fisher's score function, and is denoted by

$$u(\gamma) = \frac{\partial \log L(\gamma;y)}{\partial \gamma}. \tag{2.3}$$

Note that the score is a vector of first partial derivatives, one for each element of γ. If the log-likelihood is concave, one can find the maximum likelihood estimator by setting the score to zero, i.e. by solving the system of equations: $u(\hat{\gamma}) = 0$. Rao (1947) discussed the use of the score function for estimation of several parameters. Bartlett (1953a, 1953b) discussed the use of the score function while approximating confidence intervals.

2.2.2 The information matrix

The score is a random vector with some interesting statistical properties. In particular, the score evaluated at the true parameter value γ has variance-covariance matrix given by the information matrix:

$$var[u(\gamma)] = E[u(\gamma)u'(\gamma)] = I(\gamma). \tag{2.4}$$

The information matrix can be obtained as minus the expected value of the matrix of second derivatives of the log-likelihood:

$$I(\gamma) = -E\left[\frac{\partial^2 \log L(\gamma; y)}{\partial\gamma\partial\gamma'}\right]. \tag{2.5}$$

The matrix of negative observed second derivatives is sometimes called the observed information matrix. Sometimes it is difficult and even impossible to find the expected value of (2.5). If it is tedious to take the second derivative, then information matrix can be estimated as

$$\hat{I}(\gamma) = \left(-\frac{\partial \log L(\gamma; y)}{\partial\gamma}\right)\left(\frac{\partial \log L(\gamma; y)}{\partial\gamma}\right)'. \tag{2.6}$$

which converges stochastically to $I(\gamma)$ in an open neighbourhood of the true value of γ, which we denote by γ^t. The expected value of the score vector evaluated at the true parameter value is

$$E\left[\frac{\partial \log L(\gamma^t; y)}{\partial\gamma}\right] = 0 \tag{2.7}$$

and the variance-covariance matrix (2.4) of the score vector evaluated at the true parameter value can be written as

$$Var\left[\frac{\partial \log L(\gamma^t; y)}{\partial\gamma}\right] = E\left[\frac{\partial^2 \log L(\gamma^t; y)}{\partial\gamma\partial\gamma'}\right] = I(\gamma^t). \tag{2.8}$$

For the asymptotic distribution of the estimator $\hat{\gamma}$, we need to establish the relationship between the estimator and $\frac{\partial \log L(\gamma; y)}{\partial \gamma}$ under mild regularity conditions. Theil (1971, Chapter 8), Cox and Hinkley (1974, Chapter 9), Holly (1982), Amemiya (1985, Chapter 4), Godfrey (1988, Chapter 1) and Gourieroux and Monfort (1995, Chapter 7) among others discussed these mild regularity conditions. The following regularity conditions were compiled by Godfrey (1988, page 6):

1. "The true parameter value γ^t is interior to the parameter space which is finite dimensional, closed and bounded.
2. The probability distributions defined by two different values of γ are distinct.
3. The first- and second-order partial derivatives of $\log L(\gamma; y)$ with respect to the elements of γ are continuous throughout some neighbourhood of the true parameter value. Moreover, the third-order partial derivatives are such that the quantities $n^{-1}\frac{\partial^3 \log L(\gamma; y)}{\partial \gamma_i \partial \gamma_j \partial \gamma_k}$ $(i, j, k = 1, \cdots, m)$ exist and are bounded by integrable functions in such a neighbourhood.
4. If the score vector $\frac{\partial \log L(\gamma; y)}{\partial \gamma}$ is denoted by $u(\gamma)$ and the Hessian $\frac{\partial^2 \log L(\gamma; y)}{\partial \gamma \partial \gamma'}$ by $D_n(\gamma)$, then when the argument γ is the true parameter value γ^t,

$$E\left[u\left(\gamma^t\right)\right] = 0 \tag{2.9}$$

and

$$V\left[u\left(\gamma^t\right)\right] = E\left[D_n\left(\gamma^t\right)\right] = I\left(\gamma^t\right), \tag{2.10}$$

where $E[.]$ and $V[.]$ denote expectation and variance-covariance matrices, respectively, and $I(\gamma^t)$ is the information matrix. Equations (2.9) and (2.10) require that the relevant integrals converge and that the effective range of integration be independent of γ.

5. The matrix $I(\gamma)$ is positive definite at $\gamma = \gamma^t$ and in an open neighbourhood of the true value which implies that γ^t is locally identifiable (see Rothenberg, 1971).

6. The eigenvalues of $I(\gamma^t)$ all tends to infinity as $n \to \infty$ so that information accrues steadily as the sample size increases without limit.

7. Under the γ^t distribution, the probability limit of $-\left[I(\gamma^t)\right]^{-1} D_n(\gamma^t)$ is I_m, the $m \times m$ identity matrix. Also, the probability limit of $[D_n(\gamma)]^{-1} D_n(\gamma^t)$ is I_m for γ in a neighbourhood of γ^t.

8. The vector $\left[I(\gamma^t)\right]^{-1/2} u(\gamma^t)$ is asymptotically normally distributed with zero mean vector and covariance matrix I_m."

Using the Central Limit Theorem,

$$n^{-\frac{1}{2}} \frac{\partial \log L(\gamma^t; y)}{\partial \gamma} \xrightarrow{d} N\left(0, \; lim_{n\to\infty} n^{-1} I(\gamma^t)\right) \tag{2.11}$$

where "$\xrightarrow{d}$" represents convergence in distribution and

$$\sqrt{n}\left(\hat{\gamma} - \gamma^t\right) \xrightarrow{d} N\left(0, \left[lim_{n\to\infty} n^{-1} \hat{I}(\gamma^t)\right]^{-1}\right). \tag{2.12}$$

There are many results in the literature concerning the asymptotic properties of the ML estimator. Rao (1973, p. 364) presented the consistency of the local MLE, which was originally proved by Cramer (1946). Wald (1949) proved the consistency of the global MLE without assuming compactness of the parameter space, but his conditions are difficult to verify in practice. Many other references concerning the asymptotic properties of the MLE can be found in survey articles by Norden (1972, 1973).

The asymptotic distributional properties of ML estimates of $\hat{\gamma}$ play an important role in the derivation of the most important and widely used tests that are based on the likelihood principle. These are the likelihood ratio (LR), Wald and LM tests. Neyman (1935) introduced the LR test, Wald (1943) proposed the Wald test and Aitchinson and Silvey (1958) outlined the LM test. Rao (1948) introduced the score test which turned out to be the same as the LM test.

2.2.3 Newton-Raphson technique

In this subsection we will discuss a method for maximizing the likelihood function called the Newton-Raphson method based on an iterative procedure. Usually the calculation of the MLE requires an iterative procedure. Consider expanding the score function evaluated at the MLE $\hat{\gamma}$ around a trial value γ_0 using a first-order Taylor series, so that

$$u(\hat{\gamma}) \approx u(\gamma_0) + \frac{\partial u(\gamma)}{\partial \gamma}(\hat{\gamma} - \gamma_0). \tag{2.13}$$

Setting the left-hand-side to zero and solving for $\hat{\gamma}$ gives the first-order approximation

$$\hat{\gamma} \approx \gamma_0 - H^{-1}(\gamma_0)u(\gamma_0). \tag{2.14}$$

This result provides the basis for an iterative approach for computing the MLE known as the Newton-Raphson technique. Given an initial value of γ_0 for $\hat{\gamma}$, we use (2.14) to obtain an improved estimate and repeat the process until the difference between successive estimates is sufficiently close to zero. This procedure tends to converge quickly if the log-likelihood is well behaved (close to quadratic) in a neighbourhood of the maximum and if the starting value is reasonably close to the maximum likelihood estimate. An improved estimate can be found by using

$$\hat{\gamma} = \gamma_0 - I^{-1}(\gamma_0)u(\gamma_0) \tag{2.15}$$

which is known as Fisher scoring. In almost every econometric software package, an optimization algorithm is available. The GAUSS (See Aptech, 1995) constrained optimization module uses the Newton-Raphson technique as one of its options for optimization. We used it in simulations reported in subsequent chapters. In the next subsection we will review the simulated annealing algorithm which is another optimization technique used in this thesis.

2.3 Review of the Simulated Annealing (SA) Algorithm

In econometrics and statistics, many methods of estimation depend upon optimization techniques to estimate the parameters in the model. However, almost all conventional algorithms sometimes fail to find the optimum value of the objective function. Conventional algorithms, such as Newton-Raphson, attempt to move up hill in an iterative manner. More specifically, starting from a point, these algorithms determine the best direction and step length to head up hill. Popular statistical and econometric packages use these algorithms to solve optimization problems. Reviews of these packages can be found in Judge et al. (1985) and Press et al. (1986).

Generally, conventional optimization algorithms assume the nature of the function to be optimized is approximately quadratic. Unfortunately, some functions frequently do not follow this assumption. A common problem of the traditional algorithms is that although these algorithms converge; they may converge to a local maximum instead of the global maximum. In this situation, researchers generally try to solve these problems by using different approaches, for example, trying different starting values (see Cramer (1986) and Finch et al. (1989)). Fortunately, the simulated annealing algorithm assumes very little about the function and can tackle the optimization problem very efficiently (see Corana et al. (1987) and Goffe et al. (1994)). The advantage of this algorithm is that it is explicitly designed for functions with multiple maxima and it also works well for complex functions. The simulated annealing algorithm discovers the function's complete surface and while moving both up hill and down hill, tries to optimize the function.

In this section, we review the simulated annealing global optimization algorithm, which is a well-established stochastic neighbourhood search approach that has the capacity to jump out of the neighbourhood of a local optima in order to reach the global optimum. The name of the algorithm is drawn from an analogy between solving an optimization problem and simulating the annealing of a solid. The efficiency of the SA algorithm in solving combinatorial optimization problems is well known, see Corana et al. (1987).

The SA algorithm was initially proposed by Metropolis et al. (1953) as a means of finding the equilibrium configuration of a collection of atoms at a given temperature. Rosenbrock (1960) introduced an automatic method similar to the SA algorithm for finding the maximum value of a function and mentioned that the SA algorithm performs very well. Nelder and Mead (1965) derived a new version of the the SA algorithm called the simplex method for maximization and found that the SA algorithm always gives the global maximum or at least close to the global maximum.

Kirkpatrick et al. (1983) were the first who used the SA algorithm based on the analogy with thermodynamics, especially with the way metals, and some liquids, cool and crystallize. Geman et al. (1984) first gave a necessary and sufficient condition for the convergence of the annealing method to the global maximum. Their method is usually called either Boltzmann annealing (BA) or classical simulated annealing. Pronzato et al. (1984) developed a general form of the SA algorithm. Carnevali et al. (1985) successfully used the SA algorithm in an image-processing problem. Press et al. (1986) gave a detailed description of how the SA algorithm works.

Corana et al. (1987) mentioned that the SA algorithm assumes very little about the behaviour of the objective function and can tackle the optimization problem very efficiently. Szu et al. (1987) proposed the fast annealing method, which is a semi-local search and consists of occasional long jumps. They introduced minimization of continuous multi-modal functions using the SA algorithm and found that it is more reliable than classical optimization algorithms, being nearly always able to find the optimum, or at least a point very close to it. However, this algorithm appears to be the best with respect to the combination of ease of use and robustness. They also mentioned that unlike the case for classical optimization algorithms, it is not a necessary condition for the objective function to be approximately quadratic and differentiable when the SA algorithm is implemented.

Derwent (1988) showed an improved and better way to control pollution with the application of simulated annealing. Hajek (1988) developed the theory of cooling schedules for maximization using simulated annealing. Wasserman and Schwartz (1988) showed how simulated annealing is applicable to neural networks. Wong et al. (1988) applied simulated annealing in the context of computer and circuit design.

Ingber (1989) presented the very fast-simulated re-annealing method. They argued that their algorithm permits a fast exponential cooling schedule, while fast annealing has only an inverse-cooling schedule, and BA has only an inverse logarithmic cooling schedule. Jeng and Woods (1990) extended the use of the simulated annealing algorithm to the case of continuous functions. Johnson et al. (1990, 1991, 1992) discussed the performance of the SA algorithm on four problems:

the travelling salesman problem, graph partitioning problem, graph colouring problem and number partitioning problem. In general, the performance of the SA algorithm was mixed, in some problems it outperformed the best known heuristics for these problems, and in other cases, specialized heuristics performed better.

Goffe et al. (1994) noted that the advantage of this algorithm is that it is explicitly designed for functions with multiple maxima and also works well for complex functions. Therefore, the SA algorithm is much more user friendly than traditional algorithms found in the econometric literature. In the context of estimation of econometric models, they compared four algorithms introduced by Corana et al. (1987) with that of the SA algorithm.

They gave detailed modifications to the SA algorithm for econometric applications and mentioned that the SA algorithm can handle a very complex function and it provides valuable information about the function through the step length vector. The most important advantage of the SA algorithm is that it can properly optimize functions that are very complex and nearly impossible to optimize. They also mentioned the drawback of the SA algorithm, which requires a very high-powered computer and in present times, this is not a major problem. As a consequence, the SA algorithm is an attractive optimization algorithm for difficult functions. White (1994) introduced the perception of scale in simulated annealing and mentioned that in the presence of a large number of local maxima, it works very well.

Trouve (1996) noted that the behaviour of simulated annealing algorithms is related to the decomposition of cycles and showed how to perform the decomposition and achieve exact computation of the algorithm. Fang et al. (1997), studied simulated annealing using large deviation theory. They also specified the closed form of the critical constant in terms of the potential function of the mean model. Olivier (1998) defined a general methodology of simulated annealing with a constraint that deals with minimization of a loss function. He compared the Metropolis algorithm, simulated annealing, and the iterated energy transformation method and looked at the asymptotic theory of the algorithm. He showed that the SA algorithm works better than other algorithms.

Cohn and Fielding (1999) studied the theoretical aspects of simulated annealing that deal with the convergence to the global optimum of the objective function with probability tending to one. They mentioned that in practice, the convergent algorithms are too slow and performed a detailed analysis of various temperature schedules with application to travelling salesman problems of various sizes. Moral and Miclo (1999) discussed the convergence of generalized simulated annealing with time-inhomogeneous communication cost functions. They proposed two general genetic algorithms with a simple proof of the convergence toward the global optimum of the fitness function.

Fielding (2000) mentioned that for moderate sample sizes with a fixed temperature, the SA algorithm based on a heuristic formula for determining the optimal temperature is superior to other algorithms and gave some practical examples including the travelling salesman, quadratic assignment and graph partitioning problems. He commented in favour of the simulated annealing algorithm.

Baykasoglu et al. (2001) described in detailed how to apply the SA algorithm to solve dynamic layout problems. They found that the SA algorithm works very well at finding the global optimum. King et al. (2000) used the SA algorithm to estimate penalty functions for time series information criteria (IC) model selection. They suggested the use of the SA algorithm in conjunction with MML when faced with choosing between different regression disturbance models. Azam (2002) used the SA algorithm for finding optimal penalties for different changepoint regression models and compared it with different IC procedures and the grid search (GS) method. He concluded that the GS procedure out performed all IC procedures at a high computational cost. He found that the performance of the SA algorithm is similar to that of the GS method while its computational cost is lower. Tucci (2002) compared the performance of optimizing algorithms such as the SA, EZGARD, and genetic algorithm or tabu search in econometric problems. These algorithms are global optimization algorithms and the SA algorithm works very well compared to other algorithms. In the next subsection we will discuss briefly how the simulated annealing algorithm works.

2.3.1 Description of how the SA algorithm works

In the previous section we gave a short overview of the literature on the simulated annealing algorithm. Now we will discuss working steps of the algorithm. As we mentioned, the SA algorithm is an optimization method that finds the global maximum even in the presence of a large number of local maxima. It starts from an initial value of the parameter to be estimated; the algorithm takes a step and evaluates the objective function, which in our case is the likelihood function. The algorithm takes both up hill and down hill moves, transitions out of a local maximum and stops at the global maximum. Compared to classical methods, the SA algorithm requires less rigorous assumptions regarding the objective function and it can handle ridges and plateaux with more ease. In the remainder of this subsection, we discuss how the SA algorithm can be implemented to find the global maximum.

Let γ be the $N \times 1$ parameter vector and let $f(\gamma)$ denote our objective function to be optimized. We also assume that the optimal value of γ is such that $l < \gamma < u$, where l and u are the lower and upper bounds of the parameter vector determined by the user. The algorithm starts with an initial value γ_0, it chooses a new value say, $\tilde{\gamma}$ in the neighbourhood of γ_0. The value of $f(\tilde{\gamma})$, the objective function, is calculated at this trial point and is compared to its value at the initial point. In our case we accept up hill moves because we are maximizing our objective function. If the change $f(\tilde{\gamma}) - f(\gamma_0)$ represents an increase in the value of objective function, then the new value is accepted and the algorithm continues from this trial point. Note that the step is always centered at the trial point. If the change in the objective function represents a reduction then the down hill move may be accepted with probability $exp((f(\tilde{\gamma}) - f(\gamma_0))/T)$, where T represent a parameter called temperature. If the trial value is rejected and another value is chosen for a trial evaluation.

Let V_i be the maximum step length for the ith component of γ. This is adjusted periodically so that approximately half of all points are accepted. The algorithm requires specification of a cooling schedule. Let T be the temperature and T_0 be its initial value. In general, the initial temperature T_0 should be relatively high so that most trials are accepted and there is little chance of the algorithm converging to a local maximum in the early stages. The algorithm requires a temperature reduction factor, which is used to reduce the temperature, and a termination criterion, which used to terminate the algorithm. The following are different steps for implementation of simulated annealing.

2.3.1.1 Step 1 (Initialization of different factors)

The first and foremost task of using SA algorithm is to decide on the initial values for different factors. The initial factors are: γ_0 the initial value of the $N \times 1$ parameter vector, V_0 the starting $N \times 1$ step length vector, T_0 the initial temperature, r_t the temperature reduction factor, N_s the number of cycles after $N_s N$ function evaluations of the function after which the step length vector adjusted, N_T the number of iterations to each temperature reduction, ε the error tolerance for termination, N_ε the number of final function values used to decide upon termination, l the lower bound for the parameter vector, u the upper bound for the parameter vector and c the factor that controls the step length adjustment for the i th component of γ , $i = 1,\cdots,N$.

2.3.1.2 Step 2 (Calculation of objective function)

Using an initial value of γ say γ_0 in the function, calculate $f(\gamma_0)$ and store γ_0 as γ and $f(\gamma_0)$ as f.

2.3.1.3 Step 3 (Searching for a new parameter value)

Generate a new parameter value say, $\tilde{\gamma}$ by changing element i of γ as $\tilde{\gamma}_i = \gamma_i + uV_i$ where u is a random variable uniformly distributed from -1 to 1 and V_i is the ith element of the step vector V, $i = 1, \cdots, N$. If the parameter value $\tilde{\gamma}$ is outside the bounds then repeat calculations until another new $\tilde{\gamma}$ is found within these bounds.

2.3.1.4 Step 4 (Metropolis criterion)

Compute the value of the function at $\tilde{\gamma}$, $\tilde{f} = f(\tilde{\gamma})$ and compare with the previous function value f. If $\tilde{f} > f$, then accept the new function value; i.e., store $\tilde{\gamma}$ as γ and $\tilde{f}$ as f. On the other hand, if $\tilde{f} \le f$, the Metropolis criterion decides on acceptance or rejection of the value with probability $P_r = \exp\left(\left(\tilde{f} - f\right)/T\right)$. This is done by generating P_u, a uniformly distributed random number from $[0, 1]$. If $P_r > P_u$, the new value is accepted; otherwise it is rejected.

2.3.1.5 Step 5 (Adjustment of V)

In order to accept 50 percent of all moves, V_i, where $i = 1,2,\cdots,N$ is adjusted after N_s steps. If more than 60 percent of points are accepted for V_i, then V_i is enlarged by the factor $1+2.5c_i(m_i / N_s - 0.6)$, where m_i is the number of points accepted and c_i is the factor that controls step length adjustment for the ith parameter. If less than 40 percent of points are accepted, then V_i is decreased by the factor $1+2.5c_i(0.4 - m_i / N_s)$. Otherwise, V_i remains unchanged.

2.3.1.6 Step 6 (Temperature reduction)

After repeating steps 3 to 5; N_T times, i.e., after $N_T N_s N$ function evaluations, the temperature is reduced by the temperature reduction factor r_T: $T' = r_T T$ and return to step 3.

2.3.1.7 Step 7 (Termination criterion)

Let f_k be the most recent function value from the kth temperature reduction, $f_{k-k'}$, $k' = 1,\cdots,N_\varepsilon$, be the last N_ε values at the temperature reduction step. Then if $|f_k - f_{k-k'}| \leq \varepsilon \qquad k' = 1,\cdots,N_\varepsilon$, stop the search.

In conclusion, we can use the SA algorithm to estimate parameters when there are multiple optima. Overall, the main advantages of the SA algorithm are ease of use, adaptability to the function under consideration and suitability for computations. We will use the SA algorithm as one of the strategies to estimate parameters of the model considered in this thesis.

2.4 Multiple Maxima

Multiple maxima is a problem of numerical optimisation. The major problem of numerical optimization of a likelihood function is that sometimes the optimization procedure stops at a local optimized value rather than global maximum. The aim of this section is to provide a review of different likelihood based maximization issues in the context of the general linear model. The literature on this area is very vast and we will highlight that which reflects on our interest in the linear regression model with MA(1) errors and autocorrelated errors in general.

2.4.1 The model and related approach

Consider the linear regression model with non-spherical disturbances

$$y = X\beta + u \tag{2.16}$$

where y is an $n \times 1$ vector, X is an $n \times k$ matrix of known nonstochastic values and of full column rank, and β is a k dimensional vector of unknown parameters. The elements of u are assumed to follow the MA(1) process:

$$u_t = \varepsilon_t + \gamma\varepsilon_{t-1},\ -1 \le \gamma \le 1, \text{ where } \varepsilon_t \sim IIN\left(0, \sigma^2\right) \tag{2.17}$$

which implies that $u \sim N\left(0, \sigma^2 \Sigma(\gamma)\right)$ where $\Sigma(\gamma)$ is the tri-diagonal symmetric matrix of the form

$$\Sigma(\gamma) = \begin{bmatrix} 1+\gamma^2 & \gamma & 0 & \cdots & 0 \\ \gamma & 1+\gamma^2 & \gamma & \cdots & 0 \\ 0 & \gamma & 1+\gamma^2 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & & 1+\gamma^2 \end{bmatrix}. \tag{2.18}$$

The likelihood and the log likelihood for this model are respectively

$$L(y;\gamma,\sigma^2,\beta)$$
$$= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} |\Sigma(\gamma)|^{-\frac{1}{2}} exp\left\{-\frac{1}{2\sigma^2}(y-X\beta)'\,\Sigma(\gamma)^{-1}(y-X\beta)\right\} \tag{2.19}$$

and

$$l(y;\gamma\,\sigma^2,\beta)$$
$$= -\frac{n}{2}log\,2\pi - \frac{n}{2}log\,\sigma^2 - \frac{1}{2}log|\Sigma(\gamma)| - \frac{1}{2\sigma^2}(y-X\beta)'\,\Sigma(\gamma)^{-1}(y-X\beta). \tag{2.20}$$

We have chosen to work with the MA(1) error model which is widely used in the literature. There are a number of instances in which MA(1) processes arise from theoretical considerations. Rational lag estimation and some kinds of expectation modelling may result in an MA error process, see Trivedi (1970). These methods typically incorporate lagged effects arising from habit persistence, technological constraints or expectations effects which link anticipation with experience. Other examples may be found in papers authored by Zellner and Montmarquette (1971), Rowley and Wilton (1973) and Kenward (1975). For further discussion see King (1983) and Silvapulle and King (1991).

Nicholls, Pagan and Terrell (1975) reviewed the estimation and use of the linear model with MA disturbances and mentioned the computational difficulty involved in estimating parameters of the model. Fortunately, due to the availability of high-powered computers, this problem is not a big issue these days. There is a considerable literature on testing for MA(1) disturbances in the linear model. Breusch (1978) and Godfrey (1978) independently derived tests for MA disturbances using the LM test approach of Silvey (1959) and Aitchison and Silvey (1960). Other extensive studies include Tanaka (1990), Saikkonen and Luukkonen (1993), Davis et al. (1995) for details. See King (1987b) for a survey of the literature on testing various forms of autocorrelation in the linear regression model included MA(1) disturbances. There is also the literature on testing AR(1) disturbances against MA(1) disturbances and vice versa. For example, see Walker (1967), King (1983a, 1983b, 1987b), King and McAleer (1987), Godfrey and Tremayne (1988), McAleer et al. (1988), Hall and McAleer (1989), Burke et al. (1990), McAleer et al. (1990), Smith and Tremayne (1990), Frances (1992), Godfrey and Tremayne (1992), Silvapulle and King (1993), Baltagi and Li (1995) and McKenzie et al. (1999).

2.4.2 Concentrated likelihood

The maximum likelihood estimators of β, σ^2 and γ are obtained by jointly maximizing equation (2.20) with respect to β, σ^2 and γ. The likelihood estimating equations are

$$\begin{aligned}\frac{\partial l(y;\gamma\,\sigma^2,\beta)}{\partial\beta} &= -\frac{2}{2\sigma^2}\frac{\partial(y-X\beta)'}{\partial\beta}\Sigma(\gamma)^{-1}(y-X\beta)\\ &= \frac{1}{\sigma^2}X'\Sigma(\gamma)^{-1}(y-X\beta)=0\end{aligned} \tag{2.21}$$

$$\frac{\partial l(y;\gamma\,\sigma^2,\beta)}{\partial\sigma^2} = -\frac{n}{2\sigma^2}+\frac{2}{2\sigma^4}(y-X\beta)'\,\Sigma(\gamma)^{-1}(y-X\beta)=0 \tag{2.22}$$

$$\frac{\partial l(y;\gamma\,\sigma^2,\beta)}{\partial\gamma} = -\frac{1}{2}\frac{\partial\, log\left|\Sigma(\gamma)\right|}{\partial\gamma}-\frac{2}{2\sigma^2}\frac{\partial(y-X\beta)'\,\Sigma(\gamma)^{-1}(y-X\beta)}{\partial\gamma}=0\,. \tag{2.23}$$

In order to evaluate the latter score vector, note that

$$\frac{\partial \log|\Sigma(\gamma)|}{\partial \gamma} = tr\left[\Sigma^{-1}(\gamma)\frac{\partial \Sigma(\gamma)}{\partial \gamma}\right] \tag{2.24}$$

where $\frac{\partial \Sigma(\gamma)}{\partial \gamma}$ is the $n \times n$ matrix of derivatives of the elements of $\Sigma(\gamma)$ with respect to γ. Also

$$\frac{\partial (y - X\beta)' \Sigma^{-1}(\gamma)(y - X\beta)}{\partial \gamma} = (y - X\beta)' \frac{\partial \Sigma^{-1}(\gamma)}{\partial \gamma}(y - X\beta). \tag{2.25}$$

Therefore equation (2.23) can be written as

$$\frac{1}{2} tr\left[\Sigma^{-1}(\gamma)\frac{\partial \Sigma(\gamma)}{\partial \gamma}\right] + \frac{1}{\sigma^2}\frac{\partial (y - X\beta)' \Sigma(\gamma)^{-1}(y - X\beta)}{\partial \gamma} = 0. \tag{2.26}$$

Equations (2.21), (2.22) and (2.26) are almost impossible to solve simultaneously. We have $k+2$ equations and some of which involve the traces of products of inverses and matrices of derivatives.

For any value of γ, equation (2.21) can be solved as

$$\hat{\beta}(\gamma) = \left(X'\Sigma^{-1}(\gamma)X\right)^{-1} X'\Sigma^{-1}(\gamma)y. \tag{2.27}$$

So for any value of γ, the maximum likelihood estimate of β is $\hat{\beta}(\gamma)$. In a similar way, we can see that $\frac{\partial \log l(y;\gamma,\sigma^2,\beta)}{\partial \sigma^2} = 0$ implies

$$\sigma^2(\beta,\gamma) = \frac{1}{n}(y - X\beta)' \Sigma^{-1}(\gamma)(y - X\beta). \tag{2.28}$$

If we replace β in this equation by the estimated value conditional on γ given by (2.27), then we have

$$\hat{\sigma}^2(\gamma) = \frac{1}{n}\left(y - X\hat{\beta}(\gamma)\right)' \Sigma^{-1}(\gamma)\left(y - X\hat{\beta}(\gamma)\right). \tag{2.29}$$

Now the remaining problem is to solve for γ to get the estimate of γ and consequently the final estimates $\hat{\beta}(\hat{\gamma})$ and $\hat{\sigma}^2(\hat{\gamma})$. We can substitute the maximum likelihood estimates of β and σ^2 conditional on γ into the log likelihood function to get

$$l\left(y;\gamma, \sigma^2(\gamma), \hat{\beta}(\gamma)\right) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\hat{\sigma}^2(\gamma) - \frac{1}{2}\log\left|\Sigma(\gamma)\right| - \frac{n}{2} = l_c(y,\gamma). \tag{2.30}$$

This equation is only a function of γ and is called the log concentrated likelihood. One can estimate the parameter γ by maximizing the concentrated likelihood numerically by using a suitable computer algorithm. Once the value for γ is found, which we will denote by $\hat{\gamma}$, then $\hat{\beta}(\hat{\gamma})$ and $\hat{\sigma}^2(\hat{\gamma})$ provide the maximum likelihood estimates of β and σ^2, respectively. In the next section, we will discuss some related literature on the question of whether the optimized value of the likelihood function is really a local or global maximum.

2.4.3 Local and global maximum

In the literature, researchers have highlighted the importance of finding the global maximum of the likelihood function and, at the same time, have in some cases ignored the existence of local maxima while estimating a model of interest. We will discuss some related issues in this regard.

Hildreth and Lu (1960) discussed multiple maxima and provided an account of the existence of double maxima of the sum of squared errors of the linear regression model with autoregressive disturbances. The sum of squared errors were computed as a function of the autoregressive parameter ρ and the chosen estimate of intercept, slope and $\hat{\rho}$ is that value of ρ corresponding to the minimum sum of squared errors. They found that dual minima of the sum of squared errors at $\hat{\rho}$ = -0.90 and $\hat{\rho} = 0.30$ for their data set. Sargan (1964) noticed theoretical possibilities of occurrence of multiple maxima of the likelihood function in a lagged dependent variable regression model

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t, \tag{2.31}$$

where $u_t = \gamma u_{t-1} + e_t$ and $e_t \sim N(0, \sigma^2)$. He acknowledged the theoretical possibility of multiple maxima but noted that in 53 applications, no such occurrence was found. Theil (1971) and Hendry and Trivedi (1972) mentioned that multiple maxima can occur in the likelihood function of such a lagged dependent variable regression model.

Hendry and Srba (1977) demonstrated that multiple maxima occur in the likelihood function of a lagged dependent variable regression model (2.31) in small samples and can be eliminated as the sample size increases. In other words, multiple maxima in this model tend to disappear asymptotically. They also showed that the possibilities of dual maxima in small samples and the presence of lagged dependent variables together create small sample bias. They observed that the occurrence of multiple maxima needs special attention in order to estimate the coefficient of the lagged dependent variable in small samples.

Cooper and Thompson (1977) mentioned that the sampling distribution of the exact maximum likelihood estimator $\hat{\gamma}$ of γ in the first-order MA(1) time series model gives values of $\hat{\gamma}$ at or very near the boundaries ± 1. Ansley and Newbold (1980) investigated the small sample properties of various estimators using simulation in the context of autoregressive-moving average time-series models. They obtained the sampling distribution of exact maximum likelihood estimators of γ in the first-order moving average model. In their simulation study, they claimed that considerable concentration occurred for values of $\hat{\gamma}$ close to the invertibility boundaries $\pm$ 1, other than that, the sampling distribution seems to be unimodal.

Jonathan and Ledolter (1981) investigated the finite sample properties of the maximum likelihood estimator $\hat{\gamma}$ of γ in the MA(1) time series model. When the estimated value of the parameter is on the boundary, they gave a theoretical explanation for it. They also mentioned that the global maximum can occur at ± 1 when the true value of the parameter is near ± 1.

Oxley and Roberts (1982) focused on the possibility of multiple minima of the sum of squared residuals and the consequences of locating a local rather than global solution in a lagged dependent variable regression model. According to them, instead of using the Cochrane-Orcutt (CO) iterative procedure to identify all solutions, global and local, the Hildreth-Lu (HL) grid search technique should be used to solve this problem. Pesaran (1983) made an observation on Sargan and

Bhargava's (1983) results concerning the probability of observing a local maxima of the likelihood function of the regression model with MA(1) errors. He demonstrated that the ML estimator of γ is consistent even when the true value of γ is on the boundary. Similar results can also be derived in the case of autoregressive moving average parameters.

Oxley and Roberts (1986) investigated the occurrence of multiple roots and explained the consequences of them in the context of the lagged dependent variable regression model using the HL iterative procedure. They found that the iterative procedure can converge to the local maxima instead of the global maximum. Their results demonstrated that the problem of multiple roots is not severe in the case of small samples. As the sample size increases, the occurrence of multiple roots also increases except for the combination of higher values of the intercept and the autoregressive parameter.

Hoeschele (1988) investigated the convergence of local maxima when an iterative procedure is used to estimate the parameter with unbalanced data and mixed linear models with two or more variance components using maximum likelihood (ML) and restricted maximum likelihood (REML). Analytically he showed that for unbalanced data, ML and the Bayesian method can have two local maxima for small random effects within the permissible parameter range, whereas REML and another Bayesian method always have a single maximum. He suggested that replacing the ML method by the REML method for variance component analysis typically gives better results.

Mardia and Watkins (1989) experienced some difficulties with ML estimation. They found a problem with ML was the lack of a second derivative over the parameter space for the spherical correlation scheme of a spatial linear model. As a result they found the likelihood function might have multiple roots and they cast doubt about the accuracy of the contour plot of the concentrated likelihood.

Davis and Dunsmuir (1996) investigated the maximum likelihood estimator for an MA(1) process when the moving average parameter is on or near the unit circle. They derived the asymptotic distribution of the maximum likelihood estimator and showed that a local maximum is close to $\gamma = \pm 1$. Asymptotically these two estimators are different. Since the asymptotic distribution of the maximum likelihood of γ under H_0: $\gamma = 1$ is not known, the development of the likelihood ratio test and Wald tests are not possible. Asymptotic theory was used to construct the generalised likelihood ratio test for testing the null hypothesis of the moving average parameter being on the unit circle.

In this section, problems of multiple maxima have been discussed. There is considerable evidence in the literature of this being a practical problem for various time-series models. It is an issue we will explore further in this thesis. In the next section, we review the literature related to marginal likelihood methods.

2.5 Marginal Likelihood Method

With respect to estimation in a econometric model of the form (2.27), one of the important issues is dealing with unwanted parameters known as nuisance parameters or incidental parameters. Various methods have been suggested to improve the quality of estimation by eliminating these nuisance parameters.

Fraser (1967) first introduced the concept of the marginal likelihood in the context of statistical inference. Later Kalbfleisch and Sprott (1970) derived the marginal likelihood method as an important device to eliminate nuisance parameters. Construction of the marginal likelihood involves modifying the likelihood in such a way that it splits the likelihood function into two distinct parts, with one part containing the information about parameter of interest and the other part containing no information about the parameter of interest. The former part of the likelihood, which contains information about the parameter, works as the marginal likelihood.

The marginal likelihood approach is useful because typically it helps reduce the estimation bias, see for example, Kalbfleisch and Sprott (1970), Tunnicliffe Wilson (1989), Laskar and King (1998), Ara and King (1993) and Rahman and King (1997). This method is popular for estimating the parameters of the variance-covariance matrix in the linear regression model. Levenbach (1972) used the marginal likelihood approach for estimating the first-order autoregressive model. Levenbach (1973) also applied the method to estimation of heteroscedasticity in disturbances in the linear model. Kalbfleisch and Prentice (1973) obtained the marginal likelihood for regression parameters of the general linear model. Cooper and Thomson (1977) used the marginal likelihood method to estimate the parameter of the MA(1) model and found that the application of the marginal likelihood method reduced the estimation bias.

Bellhouse (1978) demonstrated that the application of the marginal likelihood to ARMA models and lagged dependent variable regression models is worthwhile and discussed handling the nuisance parameters for such models. Corduas (1986) showed that the marginal likelihood method helps to remove the estimation bias in the presence of trend-like regressors and AR(1) errors. Tunnicliffe Wilson (1989) derived the marginal likelihood for γ in (2.16) as

$$l_m(y;\gamma) = |\Sigma(\gamma)|^{-1/2} \left| X'\Sigma(\gamma)^{-1} X \right|^{-1/2} \hat{\sigma}^2(\gamma)^{-m/2} \tag{2.32}$$

where $m = n - k$. He used this method in trend estimation, seasonal forecasting and mixed spectrum analysis to reduce the bias in the estimation of the parameter of interest. He also used this technique for comparing different time series models.

Grose (1992) used the marginal likelihood for estimating the coefficient of the lagged dependent variable in the dynamic regression model. She reported that the estimator based on the marginal likelihood is less biased compared to the OLS estimator.

Ara and King (1993) derived general formulae for the likelihood ratio, Lagrange multiplier, Wald and asymptotically locally most powerful tests for linear regression disturbances using the marginal likelihood and investigated the small sample properties of these tests for testing the parameters of the fourth-order autoregressive disturbance process and the presence of Hildreth-Houck random coefficients. They pointed out that the problem of testing different γ values is invariant under transformations of the model. They also demonstrated that that above mentioned test can be constructed by treating the maximal invariant statistic as the observed data. Further, they demonstrated that the marginal likelihood and the likelihood of the maximal invariant are the same. In addition, Ara and King (1995) investigated the small sample sizes and power of the likelihood ratio, Lagrange multiplier, Wald and asymptotically locally most powerful tests for a subvector of the parameter vector based on the marginal likelihood. They reported better improvements in small sample sizes and powers of the marginal likelihood based tests compared to those of Ara and King (1993). This significant improvement of small sample sizes and powers of the tests occurred due to better handling of nuisance parameters.

Latif and King (1993) introduced the marginal likelihood approach for time-series forecasting based on the linear regression model in the presence of AR(1) disturbances. They suggested a weighted average of predictions, assuming different values of the AR(1) parameter with weights proportional to the marginal likelihood of that parameter. Their simulation results show that their approach produces better forecasts compared to existing procedures, which is a consequence of the application of marginal likelihood.

Shephard (1993) applied Bellhouse's (1991) method to estimation of the regression model with stochastic trend components and demonstrated that the probability of estimating the trend to be deterministic is very sensitive to the type of likelihood in the context of inference.

Rahman and King (1998) developed the marginal likelihood based Lagrange multiplier and asymptotically locally most powerful tests for situations in which the parameter vector of the error structure is partitioned into two parts, with one being the parameter of interest and another being nuisance parameters. They observed that for their problem, all nuisance parameters could not be eliminated using the likelihood of the maximal invariant or the marginal likelihood. Instead, they constructed tests, in which their maximum marginal likelihood estimators replaced those nuisance parameters, which could not be eliminated.

Laskar and King (1998) demonstrated that the use of the marginal likelihood can help to reduce the estimation bias at the same time produces more reliable estimates than of those from the concentrated likelihood in the context of AR(1) and MA(1) regression disturbances.

Laskar (1999) investigated the small sample properties of estimators and tests based on different likelihood functions. He mentioned that marginal likelihood based LM tests for both MA(1) and AR(1) processes and the LR test for AR(1) processes are quite impressive and most accurate in the sense that they have best sizes.

Sartori et al. (2003) proposed several different adjustments of the concentrated likelihood when the conditional and marginal likelihoods are available, to take proper account of the effects of estimating nuisance parameters. Their approach involves the use of the full likelihood of the model calculated with a known value of an orthogonal nuisance parameter. They proposed an orthogonal parameterization and claimed that their method gives an appropriate likelihood called the directed adjusted concentrated likelihood, which is independent of parameterisation of the nuisance parameter.

In this section we briefly surveyed the literature on the marginal likelihood method which is a well regarded estimator of the parameters of the variance-covariance matrix in the linear regression model for a wide range of applications. The consensus from the literature is that the method of marginal likelihood reduces estimation bias and provides better small sample testing procedures. The traditional approach to dealing with nuisance parameters is to reduce the problem to one in which the nuisance parameters have been eliminated, ideally with as little sacrifice of information as possible about the parameter of interest. In a limited, but often-important set of problems, reduction to a relevant conditional likelihood or a marginal likelihood that is free of the nuisance parameter is possible. The work of Ara and King (1993), King and Rahman (1997), Laskar and King (1997a, 1997b) are especially relevant, since some ideas used in this thesis trace back to their work.

In the next subsectio· we will review the literature on estimating equations.

2.6 Estimating Equation Approach

Godambe and Kale (1991) defined estimating functions as any real function f of both the observations y and the parameter γ. The estimated value of γ is $\tilde{\gamma}$ which can be found by solving the equation

$$f = f(y, \tilde{\gamma}) = 0 \tag{2.32}$$

which is called the estimating equation. This provides a unifying structure connecting some of the classical estimation methods such as the method of maximum likelihood and least squares. The following formulations illustrate these estimators as the roots of their respective estimating equations.

In a single parameter model, the maximum likelihood estimator $\hat{\gamma}$ can be found as the solution to the score equation $\dfrac{\partial l(\gamma;y)}{\partial\gamma}=0$, where $l(\gamma;y)$ is the log-likelihood. The least squares estimator is defined as the value of γ which minimizes the sum of squares of errors. After differentiation, the normal equations are obtained and the solution of which is the least squares estimator of γ. Godambe and Kale (1991) mentioned that estimation is usually approached by stating a set of optimality criteria such as sufficiency, unbiasedness, minimum variance to name but a few. They also pointed out that most estimation methods are ad hoc in the sense that the selection of a method does not follow from the optimality criteria. Given a new problem, there is no theory to say which estimation procedure should be used.

An advantage of approaching problems through estimating equations is that optimality properties can be defined in terms of the estimating equations themselves. Another justification is that major parts in the proofs of the consistency and asymptotic normality of the MLE involve facts about the score function. In particular, in an independently identically distributed sample set up, the score function is an independently identically distributed sum, so that the law of large numbers, and the central limit theorem hold quite generally. These important properties of the score function are worthy of study in their own right, see, Small and McLeish (1994).

Recently, attention has turned to the problem of multiple roots in the context of general estimating equations. Small and Yang (2000) mentioned that estimating equations can have more than one root. They also pointed out that in many cases, the theory supports an estimating equation having a unique consistent root. But they cast doubt that in reality, there may be considerable suspicion about the uniqueness of the root. Therefore the obvious question is which root should be used as an appropriate estimate of the parameters.

Huzurbazar (1948) mentioned that proofs of the consistency and asymptotic efficiency of the maximum likelihood estimator were more properly proofs of the existence of a consistent and asymptotically efficient root of the likelihood equations, and such a root could also be a local maximum. He also showed that a consistent root of the likelihood equation is asymptotically unique and corresponds to a local maximum of the likelihood function. Kraft and LeCam (1956) discussed multiple roots of the maximum likelihood estimating equation. They mentioned that the global maximum of the likelihood could give an inconsistent root of the score function.

Godambe (1960) and Durbin (1960) introduced the concept of the theory of estimating functions and Godambe (1960) has given optimality properties for the score function in a parametric framework. He derived a certain optimality property of the MLE using the estimating equation and demonstrated that the likelihood equation is an optimal estimating equation.

Barnett (1966) demonstrated a method essential for finding all the roots of the likelihood equations. His method systematically locates all the roots of the likelihood and chooses the one, which corresponds to the absolute maximum of the likelihood function. He also used five different iterative procedures for finding the roots of the likelihood equation and conducted a simulation study of the likelihood equation for the Cauchy location model. These methods were the Newton-Raphson iteration method, Fisher's scoring of parameters, the fixed derivative Newton method, the method of false position and a method used by Cohen (1957) for estimating the parameter of a truncated normal distribution. He found that the performance of the method of false position is best for locating a root of a single parameter model or in a model with multiple roots in which one of the parameters can be solved easily.

Chaubey and Gabor (1981) discussed the solution to the problem of multiple roots and mentioned that the concentrated likelihood function might have more than one maximum. Ferguson (1982) discussed inconsistent estimates from maximum likelihood. He mentioned that the global maximum of the likelihood function could give an inconsistent root of the score function but some roots of the score function might be a consistent estimator of the parameter of interest.

Lehmann (1983) gave a detailed account of the theory of efficient likelihood estimation when there are multiple local optimum in the likelihood function. He mentioned that many asymptotically efficient estimators can be constructed for the regular model but how to choose the best one remains an unsolved problem. Basford and McLachlan (1985) discussed the usefulness of multiple roots in a mixture model. The existence of multiple roots in a mixture model works as a diagnostic tool to prescribe different interpretations of the data.

Crowder (1986) discussed in detail the consistency of estimating equations. He also noted that under mild regularity conditions, the estimating equation possesses consistent roots. Godambe and Thomson (1986) discussed how to use an estimating equation to estimate the parameters of the super-population and survey population.

Liang and Zeger (1986) introduced a class of estimating equations that gives consistent estimates of regression parameters and of their asymptotic variances in the class of generalized linear models for cluster correlated data. They mentioned that when the independent variables or covariates in such models are subject to measurement errors, the parameter estimates obtained from these estimating equations are no longer consistent. They constructed an estimator with smaller asymptotic bias assuming that the measurement error variance is either known or estimable. They gave the asymptotic distribution of the bias-corrected estimator, a consistent estimator of its asymptotic variance and studied a binary logistic regression model in detail.

Stefanski and Carrol (1987) illustrated optimal and conditional scores for the general linear regression model with measurement error, which act as nuisance parameters. They eliminated the effect of nuisance parameters of the model by using the conditional joint density function on a complete sufficient statistic in such a way that the estimating function does not depend on the error.

Finch et al. (1989) introduced a numerical search procedure for finding the global maximum of the likelihood functions. Their method is useful for finding the roots by iterative search from a random starting value. They evaluated their search procedure to estimate parameters of a mixture of two normal distributions by the method of maximum likelihood. They mentioned that their search procedure is useful considering the computational time involved but difficult to get the global maximum with a high degree of reliability. In an example, they used 150 starting points in their computation. The first 50 gave some sign of trouble and second 100 refined their estimate and gave the optimal value with certainty.

Ghosh (1991) extended the idea of the use of estimating equations to estimate the super and survey populations of Godambe et al. (1986) and used it in survey sampling methods to estimate model parameters. Godambe et al. (1991) asserted that the theory of estimating equations unifies three major estimation methods in econometric and statistical literature, such as the method of LS, the method of ML and the method of minimum variance unbiased (MVU) estimates developed by Gauss (1809, 1823). They further demonstrated that estimating equations give a much wider class of estimates of the parameter of interest when there are multiple roots in the likelihood function.

Godambe and Kunte (1993) presented optimal estimation of the multiplicative treatment effect under biased allocation of treatments using the estimating equation of Godambe and Thomson (1989).

Zidek et al. (1998) developed methods for nonlinear regression analysis that are applicable for the analysis of clustered data. It has dual applications to cluster analysis and error in the measurement of the explanatory variables. They calculated second-order moments for measurement error of the explanatory variable that enable a generalized estimating equation approach for fitting and testing nonlinear models linking the response to the explanatory variables and random effects. They also used Taylor expansion methods to overcome difficulties, if any, and gave an application of this methodology that concerns the degree of association of hospital admissions for acute respiratory health problems and air pollution.

Markatou et al. (1998) introduced a modified likelihood called a weighted likelihood function where the roots of the equation were found by bootstrap searching. They mentioned that the weighted likelihood function might have multiple roots. They pooled bivariate data and estimated the weighted likelihood estimate for location. Their results showed that the weighting method successfully gives two roots by separating the data. They also incorporated a wide variety of estimating equations. The roots of the estimating equations divide into reasonable and unreasonable roots. The local maximum are unreasonable roots which are supported by a subset of data. They suggested that replication with 100 bootstrap samples is enough to detect all reasonable roots.

Heyde and Morton (1998) discussed multiple roots in estimating equations and suggested how to pick the correct root when there are more than one root of the equation. They used different methods, such as, asymptotics, analogues of empirical information and goodness-of-fit type formula. They mentioned these formulae are very useful if the estimating equation is a polynomial of degree less than five when the sample size is small. They claimed that their methods are straightforward to select roots that give sensible estimates of the parameter. They also mentioned that when a root of the estimating equation can be determined by an analytical formula and shown to be asymptotically consistent, then it would be an excellent choice for estimation of the parameter. Tzavelas (1998) proved the uniqueness of the likelihood estimator and hinted how to determine the consistent root of the estimating equation.

Small and Yang (1999) have noted that estimating equations can have multiple roots and the problem is to choose a root that estimates the parameter. They mentioned that in most of the cases there exists a unique consistent root and most of the roots cluster around consistent estimators of the parameters. They illustrated this with the example of the use of root intensity functions of first and second order of the score function for the Cauchy location model.

Small et al. (2000) provided a comprehensive review of the literature on eliminating multiple root problems in estimation and gave a considerable number of examples such as: (1) Estimation of the correlation coefficient of the bivariate normal distribution where there is as many as three real roots in the interval $(-1, 1)$. It is a small sample issue, which disappears for sufficiently large sample sizes. (2) Cauchy location models whose likelihood equation is a polynomial equation of degree $2n-1$; in this example, the problem of multiple solutions for the likelihood equation does not disappear asymptotically. (3) Inconsistent global maximum of the likelihood function. The usual conditions that are imposed for the asymptotic efficiency of the ML estimate only ensure that the consistent root of the likelihood equations is efficient; there is no guarantee that the global maxima of likelihood corresponds to a consistent root. (4) Estimating the normal mean in stratified sampling. The geometric structure of roots of this case is essentially the same as for the problem in the Cauchy location model but the probability of multiple roots arising in the two models are different. (5) Regression with measurement error. In the generalized linear models where the covariates cannot be observed directly, but can only be measured with a certain amount of measurement error. (6) Weighted likelihood equations of estimating functions with multiple roots. They concluded that when a general estimating equation has multiple roots, in general the task of detecting roots is more problematic than it would be for a likelihood-based estimating equation.

In this section, we have reviewed the literature on estimating equations in the context of linear models. Different likelihood based methods can be linked with the framework of estimating equation approaches. From the discussion of this section, we conclude that estimating equations are of considerable interest and consequently in general are worthy of study.

2.7 Conclusion

This chapter reviewed the literature related to likelihood based-inference namely, maximum likelihood, concentrated likelihood and marginal likelihood for estimating the parameters of interest of a general linear regression model. It also reviewed different methods and issues such as the simulated annealing algorithm, global maximum, local maxima, estimating equations and multiple roots. From the review, it is evident that the most popular and widely used methods in econometrics are those based on maximizing the likelihood, concentrated likelihood and marginal likelihood because of their simplicity and ease of use.

Before starting the work reported in this thesis, we conducted a survey of recent econometric texts to see what they had to say on the issue of global and local maximum of likelihood functions. Our survey suggested that the consequences of accepting a local maximum instead of the global maximum are not well articulated. This is also borne out in the review reported above. While the consequences for the estimation of parameters of interest might seem obvious, less obvious is what effect using a nuisance parameter estimate from a local maximum could have on the small sample properties of standard testing and forecasting procedures. There are some examples in the literature, which suggest it may not always appropriate to use the global maximum to obtain the MLE. Asymptotically this problem seems to disappear and intuitively it does seem sensible to seek the global maximum of the likelihood function. These issues will be taken up in Chapters 3 to 5 of this thesis.

Estimating equations seem to be a powerful method for improving our understanding of estimators. Using an estimating equation approach, we will seek to establish equivalence between least squares and ML based on the concentrated likelihood function in the context of the general regression model in Chapter 6. We will also explore analytical solution of the estimating equations expressed as a polynomial in the case of the regression model with non-stationary AR(1) errors.

CHAPTER 3

Test Sizes and the Issue of Finding the Global Maximum of the Likelihood Function[1]

3.1 Introduction

As discussed in Chapter 2, when numerical methods are used to maximise the likelihood function, we can sometimes end up with a local maximum rather than the global maximum. While the consequences for the estimation of parameters of interest might seem obvious, less obvious is what effect using a nuisance parameter estimate from a local maxima could have on the small sample properties of standard testing procedures.

The aim of this chapter is to show that not carefully looking for the global maximum when estimating nuisance parameters can affect a test's size. This is done in the context of testing a coefficient in the linear regression model with first-order moving average (MA(1)) errors. The moving average parameter, γ, is a nuisance parameter in this setting. In this chapter we investigate the effect on the size of tests of regression coefficients of estimating this nuisance parameter using

---

[1]. A paper based on material in this chapter and Chapter 4 has been accepted for publication in the *Journal of Statistical Computation and Simulation*, see Yeasmin and King (2003).

four different strategies. The first involves accepting the estimate that comes from maximizing the concentrated likelihood using constrained optimisation from a least squares based starting point. The second involves taking the best result from an additional three fixed starting points but only when the initially estimated value of γ is a boundary point. The third involves the same approach but taking even greater care to find the global maximum by using 43 different starting values. The fourth involves the use of simulated annealing (SA) with the aim of finding the global maximum.

The remainder of this chapter is organised as follows. In Section 3.2, we discuss the model and construct our test statistic based on maximization of the likelihood function for the unknown moving average parameter. In Section 3.3, we report calculations of estimated sizes of the test statistics for different sample sizes, and different values of γ using Monte Carlo simulation. Section 3.4 contains a discussion of the results of the Monte Carlo study. Section 3.5 presents some concluding remarks.

3.2 The Model and the Test Statistic

3.2.1 The model

Consider the linear regression model with non-spherical disturbances

$$y = X\beta + u \tag{3.1}$$

where y is an $n \times 1$ vector, X is an $n \times k$ matrix of known nonstochastic values and of full column rank, and β is a k-dimensional vector of unknown parameters. The elements of u are assumed to follow the MA(1) process

$$u_t = \varepsilon_t + \gamma\varepsilon_{t-1},\ -1 \le \gamma \le 1,\ \text{where } \varepsilon_t \sim IIN\left(0, \sigma^2\right) \tag{3.2}$$

which implies that $u \sim N\left(0, \sigma^2\Sigma(\gamma)\right)$ where $\Sigma(\gamma)$ is an tri-diagonal symmetric matrix of the form

$$\Sigma(\gamma) = \begin{bmatrix} 1+\gamma^2 & \gamma & 0 & \cdots & 0 \\ \gamma & 1+\gamma^2 & \gamma & \cdots & 0 \\ 0 & \gamma & 1+\gamma^2 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & & 1+\gamma^2 \end{bmatrix}. \tag{3.3}$$

One way to tackle the estimation problem for this model is to transform using the Cholesky decomposition. Let $\Sigma(\gamma) = L(\gamma)L(\gamma)'$ where $L(\gamma)$ is the upper triangular matrix of $\Sigma(\gamma)$, defined as

$$L(\gamma) = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{12} & l_{22} & 0 & \cdots & 0 \\ 0 & l_{23} & l_{33} & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \ldots & & l_{nn} \end{bmatrix}.$$

The non-zero elements of $L(\gamma)$ can be obtained recursively by using

$$l_{11} = \left(1+\gamma^2\right)^{1/2}$$

$$l_{i,i-1} = \frac{\gamma}{l_{i-1,i-1}}$$

and

$$l_{ii} = \left(1+\gamma^2 - l_{i,i-1}^2\right)^{\frac{1}{2}}$$

where $i = 2, 3, \cdots, n$.

The inverse of the matrix $L(\gamma)$ is used to transform model (3.1) in order to find the generalised least squares (GLS) estimator of β. This implies $L(\gamma)y^* = y$ and $L(\gamma)X^* = X$ where y^* and X^* are the transformed vector and matrix respectively. Using these expressions it is easy to obtain the formula for y^* which can be written as

$$y_1^* = y_1 / (1+\gamma^2)^{1/2}$$

$$y_i^* = (y_i - l_{ii-1}y_{i-1}^*) / l_{ii} \quad \text{where } i = 2, \cdots, n.$$

This same transformation is also applied to each column of X in turn in order to obtain X^*. The resultant transformed model can be written as

$$y^* = X^*\beta + u^*$$

and the transformed disturbance vector $u*$ has the properties $\mathrm{E}(u^*) = 0$ and $\mathrm{E}\left(u^* u^{*\prime}\right) = \sigma^2 I$.

The GLS estimator can be found by applying the usual ordinary least squares estimator to the above transformed model. Unless the value of γ is known, the above technique cannot be applied. For a known value of γ, the GLS estimator $\hat{\beta}(\gamma)$ defined in equation (2.22) in Chapter 2 is the best linear unbiased estimator of β. Unfortunately, the value of the parameter γ is typically unknown and therefore it needs to be estimated. For example, we can obtain the estimated value of γ by maximising the concentrated likelihood of γ. The log concentrated likelihood is

$$l_c(y;\gamma) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\left(s^2(\gamma)/n\right) - \frac{1}{2}\log\left|\Sigma(\gamma)\right| - \frac{n}{2} \tag{3.4}$$

where σ^2 and β have been replaced by their estimated values with

$$s^2(\gamma) = \left(y - X\hat{\beta}(\gamma)\right)' \Sigma^{-1}(\gamma)\left(y - X\hat{\beta}(\gamma)\right) \tag{3.5}$$

and

$$\hat{\beta}(\gamma) = \left(X'\Sigma^{-1}(\gamma)X\right)^{-1} X'\Sigma^{-1}(\gamma)y. \tag{3.6}$$

To obtain a final estimate of β, we need to estimate the value of γ by maximizing the right hand side of (3.4) and replacing γ by its estimated value in equation (3.5), which can be written as

$$\hat{\beta}(\hat{\gamma}) = \left(X'\Sigma^{-1}(\hat{\gamma})X\right)^{-1} X'\Sigma^{-1}(\hat{\gamma})y \tag{3.7}$$

where $\hat{\gamma}$ denotes the maximum likelihood estimator (MLE) of γ from maximizing (3.4).

3.2.2 The test statistics

Our interest is in investigating the effect of estimates based on local versus global maximum in the context of testing the jth regression coefficient in the model given by (3.1) and (3.2), namely testing $H_0: \beta_j = \beta_{j0}$ against $H_1: \beta_j \neq \beta_{j0}$ where $1 \le j \le k$ and β_{j0} is a known constant. The test statistic we are interested in can be written as

$$t = \frac{\hat{\beta}_j(\hat{\gamma}) - \beta_{j0}}{Se\left(\hat{\beta}_j(\hat{\gamma})\right)}, \tag{3.8}$$

where $\hat{\beta}_j(\hat{\gamma})$ is the jth element of (3.7) and

$$Se\left(\hat{\beta}_j(\hat{\gamma})\right) = \left[\left(s^2(\hat{\gamma})/n\right)\left(X'\Sigma^{-1}(\hat{\gamma})X\right)^{-1}_{jj}\right]^{1/2} \tag{3.9}$$

in which $\left(X'\Sigma^{-1}(\hat{\gamma})X\right)^{-1}_{jj}$ denotes the jth diagonal element of $\left(X'\Sigma^{-1}(\hat{\gamma})X\right)^{-1}$.

The form of log likelihood (3.4) reveals that we are unable to maximize the likelihood analytically in order to estimate γ. In this situation, the best way to solve the estimation problem is by optimising the log likelihood using a suitable numerical optimization algorithm. The model given by (3.1) and (3.2) has a minor identification problem in that $\sigma^2\Sigma(\gamma)$ and $\sigma_*^2\Sigma(\gamma_*)$ take exactly the same value whenever $\sigma_*^2 = \sigma^2\gamma^2$ and $\gamma_* = \dfrac{1}{\gamma}$. In order to confirm this, we can easily show that

$$\begin{aligned}\hat{\beta}(1/\gamma) &= \left(X'\Sigma^{-1}(1/\gamma)X\right)^{-1}X'\Sigma^{-1}(1/\gamma)y \\ &= \left(X'\gamma^2\Sigma^{-1}(\gamma)X\right)^{-1}X'\gamma^2\Sigma^{-1}(\gamma)y \\ &= \left(X'\Sigma^{-1}(\gamma)X\right)^{-1}X'\Sigma^{-1}(\gamma)y \\ &= \hat{\beta}(\gamma),\end{aligned}$$

$$\begin{aligned}s^2(1/\gamma) &= \left(y - X\hat{\beta}(1/\gamma)\right)'\Sigma^{-1}(1/\gamma)\left(y - X\hat{\beta}(1/\gamma)\right) \\ &= \left(y - X\hat{\beta}(\gamma)\right)'\gamma^2\Sigma^{-1}(\gamma)\left(y - X\hat{\beta}(\gamma)\right) \\ &= \gamma^2\left(y - X\hat{\beta}(\gamma)\right)'\Sigma^{-1}(\gamma)\left(y - X\hat{\beta}(\gamma)\right) \\ &= \gamma^2 s^2(\gamma),\end{aligned}$$

$$
\begin{aligned}
l_c(y;1/\gamma) &= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\left(s^2(1/\gamma)/n\right) - \frac{1}{2}\log\left|\Sigma(1/\gamma)\right| - \frac{n}{2} \\
&= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\left(\gamma^2 s^2(\gamma)/n\right) - \frac{1}{2}\log\left|1/\gamma^2\Sigma(\gamma)\right| - \frac{n}{2} \\
&= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\gamma^2 - \frac{n}{2}\log\left(s^2(\gamma)/n\right) - \frac{1}{2}\log\left(1/\gamma^2\right)^n \\
&\quad - \frac{1}{2}\log\left|\Sigma(\gamma)\right| - \frac{n}{2} \\
&= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\left(s^2(\gamma)/n\right) - \frac{1}{2}\log\left|\Sigma(\gamma)\right| - \frac{n}{2} \\
&= l_c(y;\gamma).
\end{aligned}
$$

As noted in Section 2.3.3 of Chapter 2, our likelihood has a turning point (which may be a maximum or a minimum) at $|\gamma| = 1$. This can result in the global maximum being at $|\gamma| = 1$, a point that is well recognised in the literature (see Kang (1975), Dunsmuir (1981) and Cryer and Ledolter (1981)). It can also result in a local maximum being at $|\gamma| = 1$. Consequently, when as can often happen, maximising the likelihood results in $|\hat{\gamma}| = 1$, one might be suspicious that one has a local maximum rather than a global maximum. Cryder and Ledolter (1981) discussed the theoretical explanation for the tendency of the maximum likelihood to have this pile up effect at the boundaries. Davis and Dunsmuir (1996) demonstrated that when $\gamma = 1$, the statistical problem is a non-regular one and the usual asymptotic theory does not apply to the maximum likelihood estimator. They also explained that at the boundary point, the standard theory of a maximum likelihood gives a very poor approximation to the actual distribution of the estimator. The consequences of getting appropriate maxima of the likelihood function will be investigated by using the different strategies described in the following subsection.

3.2.3 Strategies for estimation

Often econometricians will estimate a parameter and consider its estimated value. If it looks reasonable, they might accept the parameter particularly if they feel they have been able to use a good starting value for the optimization. If it looks unreasonable or suspicious, they might try one or more different starting values in order to assure themselves that they have the global maximum. Alternatively, they could use one of the optimisation methods specifically designed not to converge to a local maximum, such as simulated annealing.

With these thoughts in mind, we considered four different strategies for our particular problem. For each we start with a reasonable starting value that comes from a rough estimate of the degree of correlation in the OLS residuals. The first strategy involves putting this into a standard optimisation routine (using the GAUSS Newton-Raphson procedure) and accepting the estimate that results. The second involves this same process but an examination of the final estimate. If it is on the boundary, it is regarded as suspicious and the optimisation is repeated 3 further times with a spread of 3 different starting values. The final estimate is that which gives the largest maximised likelihood. The third strategy is just the second with three different starting values being replaced by 43 different starting values. It represents a higher degree of care being taken when the initial estimate is on the boundary and therefore regarded as suspicious. Finally there is the attempt to find the global maximum without any value judgements being made using SA.

3.3 Monte Carlo Experiment

A Monte Carlo experiment was conducted to investigate the effect of estimated sizes of the t test of regression coefficients based on estimation through the four different strategies for finding the global maximum of the concentrated likelihood function for the MA(1) parameter as outlined above.

3.3.1 Design of simulation experiments

In order to explore the consequences of accepting local maxima on test sizes, we consider the four strategies outlined above, namely (i) using one starting value in a standard optimization technique and accepting the outcome as the global maximum. (ii) using three further starting values of $\gamma = -0.5, 0, 0.5$ if the estimated value under (i) is on the boundary and then choosing the estimate with the largest maximised likelihood. (iii) as for (ii) but using 43 starting values, namely $\gamma = \pm 0.98, \pm 0.96, \cdots, \pm 0.78, \pm 0.75, \pm 0.70, \cdots, \pm 0.50, \pm 0.40, \pm 0.30, \cdots, 0$ if the estimated value from (i) is on the boundary and (iv) using simulated annealing. In Figure 3.1 the first three of the above mentioned strategies are presented in a concise form. In all four cases, the initial starting value was set to

$$\hat{\gamma}_0 = \begin{cases} -0.9999, & \text{if } \hat{\rho} < -0.4999, \\ 2\hat{\rho}, & \\ 0.9999, & \text{if } \hat{\rho} > 0.4999, \end{cases}$$

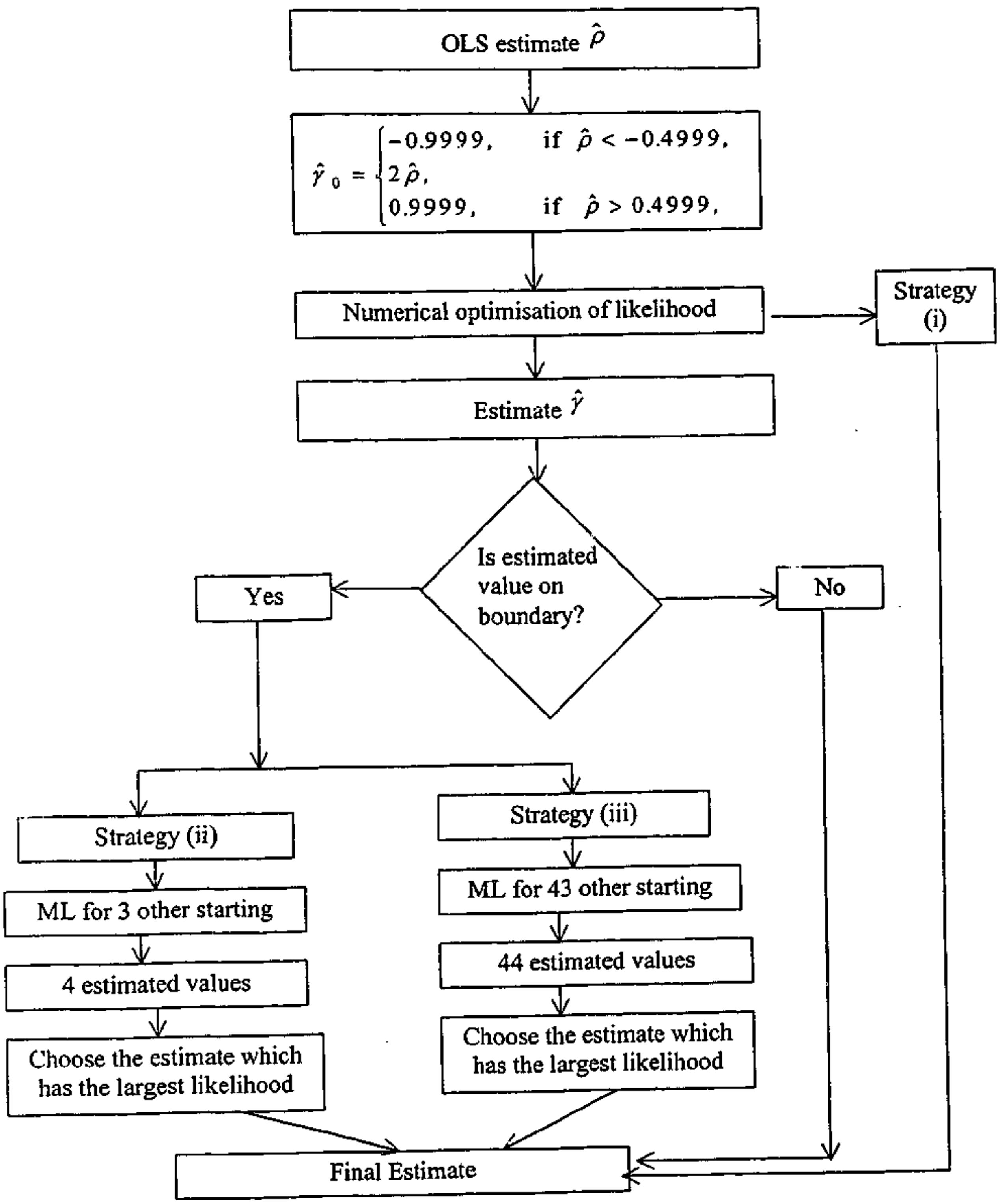
where $\hat{\rho} = \sum_{t=2}^{n} z_t z_{t-1} / \sum_{t=1}^{n} z_t^2$ in which z_t, $t = 1, \cdots, n$, are the OLS residuals from (3.1). In this case, the t test of a regression coefficient is essentially a Wald test. Laskar (1999, p.83) noted problems with the Wald test for the linear regression model with MA(1) errors and was able to overcome these problems by restricting γ to an interval that does not include $\gamma = \pm 1$. We have followed his lead and restricted γ and its estimates to [-0.9999, 0.9999].

The simulation experiment involved generating observations for y from (3.1) and (3.2) in which X is the $n \times 3$ matrix of observations on three regressors, the first being the constant dummy, the second being Australia's quarterly consumer price index commencing 1948(4) and the third being this regressor lagged one quarter. We focused on testing $H_0\colon \beta_j = 0$ against $H_a\colon \beta_j \neq 0$ using a two-sided test. There is an issue of what distribution to use for the calculation of critical values for (3.7). On the one hand we could use the $N(0,1)$ distribution because asymptotically as n tends to infinity, this has the correct distribution. In the related dynamic linear regression model case with well behaved errors there is some evidence that the Student's t distribution is a better approximation than the $N(0,1)$ distribution in small samples. See for example Nankervis and Savin (1987), King and Wu (1991) and Atukorala (1999). This led us to use the Student's t distribution with n degrees of freedom for critical values. (The use of $N(0,1)$ critical values would have resulted in even higher sizes.)

The size results reported below, were calculated using 1000 iterations; for $n =$ 20, 40, 60, 80, 100, and 120; $\gamma = -0.90, -0.75; -0.60, -0.45, -0.30, -0.15, 0, 0.15, 0.30, 0.45, 0.60, 0.75, 0.90$ and $\sigma^2 = 1$. The Newton-Raphson option of the GAUSS constrained optimization algorithm (Aptech Systems, 1998) was used to maximized (3.4) for first three strategies as required. A nominal significance level of 0.05 was used and the t test for maximum likelihood is denoted by $ML(n)$.

Strategy (iv) is based on the simulated annealing algorithm used for finding the global maximum of the likelihood function. This algorithm is a good optimizer for finding a global maximum in the presence of many local maxima. We used SA as an alternative to the other strategies discussed above where multiple starting values was employed. Execution of SA depends on the initial values of different factors. The initial values of different factors we used are discussed in Section 3.3.2. The SA algorithm was coded in the GAUSS programming language.

Figure 3.1 Flow Chart for Strategies (i), (ii) and (iii)



3.3.2 Initial settings used for the SA strategy

In Chapter 2, we discussed the importance of using the SA method for finding the global maximum when multiple local maxima exist in the function being optimised. In this subsection, our main interest is in estimation of the MA(1) parameter γ using simulated annealing. A short description of the starting parameters we used for the SA method is given in the following subsections.

3.3.2.1 Initial temperature (T_0)

One of the important factors of the SA method is the initial temperature. According to Kirkpatrick et al. (1983), a suitable initial temperature should be chosen carefully so that about 80% of all positive transitions (i.e., transitions that maximise the likelihood function) are accepted. In our problem of maximising the likelihood function, we set the initial temperature to be $T_0 = 50$.

3.3.2.2 Temperature reduction factor (r_T)

There is a close relationship between temperature reduction within stages and the number of iterations per stage. An important step is to reduce the temperature and it can be done by multiplying by a constant factor. The most practical and well-known temperature reduction form is $T_{r+1} = \alpha T_r$ where α is a constant close to, but smaller than, 1. Kirkpatrick et al. (1982) suggested that the constant term should be $\alpha = 0.95$. Corana et al. (1987) proposed that the value should be 0.85. In our case, we used $\alpha =$ 0.85.

3.3.2.3 Boundaries for parameter

The boundary values should be determined by the researcher according to the scope of the optimisation problem. In our case we used a lower boundary for γ of -0.9999 and an upper boundary of 0.9999.

3.3.2.4 Number of cycles (n_c)

The number of cycles is denoted by n_c. When the SA method starts to execute, it evaluates $n_c \times N$ (an array of length n_t) functions and then adjusts each element of the step length vector (V) in such a way that approximately half of all functions evaluated are accepted. We used $n_c = 20$ for the number of cycles.

3.3.2.5 Number of iterations before temperature reduction (n_t)

A constant number of iterations (n_t) is used before temperature reduction while implementing the SA method. According to Corana et al. (1987), better results can be obtained by considering the physical background of the SA method and they suggested the value of the number of iterations before temperature reduction should be $n_t = 5$. We used $n_t = 5$ as suggested by Corana et al.

3.3.2.6 Termination criterion (T_c)

The termination criterion (T_c) is an important factor in the SA algorithm. To maintain the quality of the estimated value of the parameter we used a stopping criterion such that the optimal values of the likelihood function for successive stages are constant or their difference is very small. We stop searching if

$|l_{j+1} - l_j| \le T_c$ where l_j is the maximized log-likelihood function at the jth stage. We used a termination criterion of $T_c = 0.00001$.

3.4 Discussion of Results

It was mentioned earlier that numerical optimization methods might have difficulties for finding the optimal value of the likelihood function when estimating the parameter γ. To investigate the difficulties of such problems we compared sizes that result from strategies (i), (ii), (iii) and (iv) for different sample sizes and for different values of the moving average parameter γ in the linear model.

The estimated sizes of the $ML(n)$ test for the four different strategies are reported in Tables 3.1 to 3.10. Note that estimated sizes in the range 0.0365-0.0635 are insignificantly different from the nominal size of 0.05 at the 5% level of significance.

3.4.1 Different sample sizes

An obvious feature is that the majority of the estimated sizes are significantly greater than the nominal size at the 0.05 level. This is particularly true for small samples. For larger samples, the same trend can be seen for negative values of γ but almost all the estimated sizes are not significantly different from the nominal size for positive values of γ for all strategies except for γ values close to the boundary. Typically improvements are gained in estimated sizes through strategy (iii), especially for negative values of γ, for moderate (60, 80) and large (100, 120) samples when the γ value is close to the boundary. However, this is not the case for small samples. For greater accuracy, strategy (iii) is working well to reduce the estimation bias especially for the pile-up effect at the boundary points. We observed that typically strategy (iii) improves accuracy of the estimated sizes and dominates the other strategies by producing smaller sizes. The second improved performance can be seen for strategy (iv). The overall performance of strategy (i) is poor compared to the other strategies. In the next subsection we will discuss and compare results of the estimated sizes for strategies (i) and (ii).

3.4.1.1 Comparison between strategies (i) and (ii)

Tables 3.1 to 3.3 reveal that the majority of the estimated sizes of the $ML(n)$ test for MA(1) errors are significantly above 0.05 for sample sizes $n = 20$ and 40. On the other hand, there is a clear feature that taking greater care in finding the global maximum by using strategy (ii) does improve the test size particularly for small samples and for γ values near the negative and positive boundaries. For example, in Table 3.2 and for testing H_0: $\beta_2 = \beta_{20}$ when $\gamma = -0.45$, the estimated sizes are 0.171 from strategy (i) and 0.140 from strategy (ii) for $n = 40$. In Table 3.3 for testing H_0: $\beta_3 = \beta_{30}$ when $\gamma = 0.60$, the estimated sizes are 0.159 and 0.142 from strategy (i) and (ii) respectively, for $n = 20$. Similar trends may also be noticed for estimated sizes in Tables 3.1 to 3.3 for sample sizes $n = 20$ and 40 with a few exceptions. For example, we see from Table 3.1 that strategy (i) dominates strategy (ii) when γ lies between 0.30 to 0.45. This is because the estimated value very often occurs at the boundary whether the true value is not even close to the boundary value of the parameter. Sometimes the resultant sizes that come from strategy (ii) are very high. It is because the resultant estimate has a non-zero probability of being a boundary value.

For moderate sample sizes, $n = 60$ and 80, in Tables 3.2, 3.5 and 3.8, a trend we observe is that near the boundary, strategy (ii) which involves taking some care in finding the global maximum improves the estimated sizes noticeably as compared to strategy (i). Another obvious feature of the results is that the majority of the estimated sizes are close to the nominal size for positive values of γ which are within the interval (0.15 to 0.90) in Table 3.2, (0.15, to 0.45) in Table 3.5 and ($-$0.15, 0.30) in Table 3.8 except for $\gamma = -0.15$ and $n = 80$. Almost all of the estimated sizes of the above mentioned intervals of γ are approximately the same for all strategies.

The most important feature is that the estimated sizes of the test statistic for different values of γ within the interval ($-$0.30 to 0.95) are insignificantly different from the nominal size. We see that no improvement is found for strategy (ii) over (i) for sample sizes n = 100 and 120 in Table 3.3 for testing H_0: $\beta_1 = \beta_{10}$. The same trend is also observed in Tables 3.6 and 3.9, respectively, for testing H_0: $\beta_2 = \beta_{20}$ for values of γ over the interval ($-$0.15 to 0.60) and for testing H_0: $\beta_3 = \beta_{30}$ for values of γ in the interval ($-$0.15 to 0.45).

On the other hand, significant improvement is noticed for the test sizes for the rest of the γ values in Tables 3.3, 3.6 and 3.9.

The worst scenario is shown in Table 3.3 when $\gamma = -0.75$ where the estimated sizes of the test are around 0.42 under strategy (i). This is because the resultant estimate has a tendency to occur frequently at the boundary where it helps to produce higher estimated sizes. On the other hand, strategy (ii) helps to reduce the estimated sizes noticeably, but the improvement in the estimated sizes is not enough to get close to the nominal size. The estimated sizes from strategy (ii) dominate those of the strategy (i) for negative values of γ except $\gamma = -0.90$ as seen in Table 3.8 for testing H_0: $\beta_3 = \beta_{30}$ when $n = 60$.

3.4.1.2 Comparison between strategies (ii) and (iii)

The overall impression we get through the tables is that many of the estimated sizes are highly significantly different from the nominal size for small and moderate sample sizes. This is true for both strategies (ii) and (iii). In general, no one strategy dominates the other because their estimated sizes are almost the same. However some noticeable improvements are found for strategy (iii) in Table 3.5 for all γ values except $\gamma = -0.45$ and -0.30 when $n = 60$ for testing H_0: $\beta_2 = \beta_{20}$. The opposite scenario can be seen in Table 3.2 where strategy (ii) dominates strategy (iii) for $\gamma = (-0.45 \text{ to } 0.45 \text{ and } 0.90)$ although an appreciable improvement is observed for $\gamma = (-0.75, -0.60)$.

It is clear from Table 3.5 that almost all the estimated sizes obtained from strategy (ii) are equal to the estimated sizes for strategy (iii) when $n = 80$. However improvements are remarkable when the γ values are close to either of the two boundaries. One possible explanation is that the initial estimate from strategy (i) often occurs on a boundary. It gives us the impression that paying special attention to finding the global maximum via strategy (iii), is worthwhile. The same scenario can be seen for $n = 100$ and 120.

3.4.1.3 Comparison between strategies (ii) and (iv)

In this subsection, we compare the results from strategies (ii) and (iv). For $n = 20$, SA unexpectedly produces poorer estimated sizes than strategy (ii). For example, the estimated size of strategy (iv) produces a 0.40 significance level for the $ML(n)$ test when $\gamma = -0.30$ as seen in Table 3.1. On the other hand, strategies (ii) and (iii) produce much better results than strategy (iv) in the case of small samples.

An obvious feature we observe is that the estimated sizes of strategy (iv) are better than strategy (ii) when the γ value is close to a boundary, for example, when $\gamma = -0.90, -0.75, -0.60\ \ 0.75,$ and 0.90. There are few exceptions, when $n = 100$ and $\gamma = -0.75$. This is true for all sample sizes except $n = 60$. In contrast, for the rest of the γ values, we observe that strategy (ii) either dominates strategy (iv) or produces approximately the same sizes where almost all the sizes are barely significantly different from the nominal sizes. From the above discussion we conclude that for the boundary problem, strategy (iv) sometimes performs better than strategy (ii), but overall, strategy (ii) has the best performance by producing smaller estimated sizes.

The results of this chapter suggest that multiple starting values to solve the boundary problem performs well as compared to the use of simulated annealing. This also gives an indication that sometimes a global maximum from strategy (iv) may not produce the most reliable estimated sizes in terms of the hypothesis testing. Theoretically the global maximum of the likelihood may not always produce a consistent estimate while at the same time, one or more local maxima can produce a consistent estimate in the case of multiple maxima, see for example, Kraft and LeCam (1956), LeCam (1979), Bahadur (1958) and Ferguson (1982) and LeCam (1990).

3.4.1.4 Comparison between strategies (iii) and (iv)

In this subsection we compare strategies (iii) and (iv). Strategy (iii), which involves taking extra care in finding a global maximum, can dominate strategy (iv) over an interval of γ values for small samples. This can be seen in Table 3.1 when $n = 20$ for all the γ values and in Table 3.4 for γ in the range (-0.15 to 0.75). The estimated sizes for the other γ values remain the same although the sizes are highly significantly different from the nominal size. An exception is in Table 3.4 for $n = 20$ and $\gamma = -0.75$ when strategy (iv) dominates strategy (iii).

For moderate samples, most of the estimated sizes over the interval for γ of (-0.15 to 0.60) are barely significantly different from the nominal size and the sizes are approximately the same for both strategies. On the other hand, different features are revealed for the rest of the γ values where strategy (iii) dominates strategy (iv) but the improvement is not that much. There are a few exceptions where strategy (iv) dominates strategy (iii) in Table 3.2 when $\gamma = 0$ and for the same γ value in Table 3.7 in the case of $n = 60$.

For large samples, we also observe that the majority of the estimated sizes are barely significantly different from the nominal size especially for positive values of γ for both strategies. Typically improvements are noticeable for negative values of γ where the estimated sizes from strategy (iii) dominate those of strategy (iv) with only one exception when $n = 100$ and $\gamma = -0.60$. Overall we can conclude that strategy (iii) performs reasonably well in respect of all sample sizes and almost all γ values.

3.4.2 For different β values

A feature is that the regressors of the coefficients being tested do make a difference to the results. For example, the best estimated sizes are those for testing $\mathrm{H_0}$: $\beta_3 = \beta_{30}$ for all the strategies. The worst results involve testing $\mathrm{H_0}$: $\beta_1 = \beta_{10}$ for all the strategies.

3.4.3 For different γ values

There is a tendency for the estimated sizes to increase as γ moves away from zero towards the negative boundary or the positive boundary for small samples. Surprisingly, $\gamma = -0.90$ produces much smaller sizes than the other γ values when $n = 20$ and they are significantly different from the nominal size as in Table 3.1. When $n = 40$, for the same γ values, the estimated sizes are barely significantly different from the nominal size and produce more accurate results than the rest. In the case of moderate and large sample sizes, for positive values of γ, the estimated sizes are more accurate than for negative values.

The results suggest that the traditional maximum likelihood based t test using strategy (iii) is more reliable for different γ values than the other strategies with a few exceptions. However the estimated sizes are disappointing in the case of small samples because they are significantly higher than the nominal size for almost all the γ values. An obvious feature we observe from the results is that this strategy can make a significant improvement on estimated size when the γ values are close to the boundary point. There is a trend in terms of accuracy that the sample size increases, and estimated sizes for positive values of γ are more reliable than those for negative values. For large sample sizes and positive values of γ, the majority of the estimated sizes are barely insignificantly different from the nominal size for all strategies. This is true also for moderate samples sizes.

3.5 Conclusion

In this chapter, a Monte Carlo experiment was conducted to investigate the effect of using different strategies for applying maximum likelihood estimation to estimate a nuisance parameter on subsequent hypothesis testing of regression coefficients. We used four strategies, which are (i) one starting value, (ii) if the resultant estimate from strategy (i) is on the boundary, we use an additional three starting values. The same approach was used for strategy (iii) but with forty-three starting values in a standard optimisation computer package. Strategy (iv) is based on the simulated annealing algorithm.

In this study, we used estimated sizes of the test to evaluate the performance of our four strategies. From the result of the simulation study, we observed that the traditional likelihood based test statistic $ML(n)$ can be inaccurate and sometimes provide woefully unacceptable sizes when strategy (i) is used. We observed an improvement in the estimated size of the test for strategy (ii) and (iii) in the case of all sample sizes with very few exceptions. For small samples, strategy (iii) gives better results than strategy (iv), but for moderate and large samples, the estimated sizes for both strategy (iii) and (iv) are almost the same. On the other hand, strategy (ii) improves the estimated sizes but does not dominate strategy (iii) for all the sample sizes. Compared to strategy (iv), strategy (iii) gives better results when the γ values depart from zero towards the positive boundary or are close to boundary values. On the other hand, for positive values of γ, almost all estimated sizes for strategies (iii) and (iv) are approximately the same.

Strategy (iii) involves only attempting to find the global maximum when we are suspicious about our estimate. We are suspicious when it is on the boundary. An interesting result is that strategy (iii) (in particular) generally gives better results than always attempting to find the global maximum. In one sense this is a troubling result. By far the accepted wisdom is that we should always look for the global maximum, but here we find that doing so results in a worse test than if we only look for the global maximum when we are suspicious about our estimate.

Why might this be the case? It could be that some estimates resulting from global maximum located on boundaries are not consistent. Alternatively, our poor results overall suggest a poor procedure and that some attention to improving the test sizes might be needed first before considering whether to advise econometricians that it can be harmful to always look for a global maximum. With some disappointment about the estimated sizes, we will look at this more closely to find an alternate way to get more reliable sizes in terms of accuracy in the next chapter.

Table 3.1 Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ for small n and different strategies for finding a global maximum

n	20				40			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.083	0.082	0.082	0.084	0.067	0.067	0.067	0.068
-0.75	0.160	0.158	0.158	0.160	0.126	0.126	0.126	0.126
-0.60	0.268	0.267	0.267	0.269	0.189	0.181	0.181	0.185
-0.45	0.369	0.365	0.365	0.384	0.189	0.189	0.189	0.195
-0.30	0.368	0.363	0.363	0.403	0.166	0.162	0.162	0.176
-0.15	0.353	0.350	0.350	0.400	0.158	0.153	0.153	0.170
0.00	0.284	0.281	0.281	0.335	0.100	0.099	0.099	0.111
0.15	0.216	0.217	0.217	0.283	0.074	0.074	0.074	0.075
0.30	0.217	0.221	0.221	0.268	0.085	0.085	0.085	0.090
0.45	0.179	0.180	0.180	0.223	0.076	0.076	0.076	0.076
0.60	0.116	0.119	0.119	0.141	0.077	0.077	0.076	0.078
0.75	0.126	0.128	0.127	0.141	0.074	0.074	0.074	0.074
0.90	0.115	0.116	0.116	0.123	0.066	0.066	0.067	0.067

Table 3.2 Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ for moderate n and different strategies for finding a global maximum

n	60				80			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.093	0.093	0.092	0.094	0.164	0.163	0.161	0.162
-0.75	0.179	0.174	0.152	0.155	0.334	0.280	0.269	0.281
-0.60	0.206	0.192	0.155	0.157	0.277	0.170	0.168	0.175
-0.45	0.166	0.135	0.139	0.144	0.161	0.115	0.114	0.116
-0.30	0.125	0.093	0.116	0.119	0.107	0.101	0.101	0.102
-0.15	0.095	0.073	0.095	0.095	0.097	0.097	0.097	0.097
0.00	0.082	0.054	0.095	0.083	0.074	0.074	0.074	0.074
0.15	0.069	0.062	0.069	0.069	0.058	0.058	0.058	0.058
0.30	0.072	0.068	0.072	0.072	0.054	0.054	0.054	0.054
0.45	0.071	0.058	0.064	0.064	0.063	0.058	0.058	0.058
0.60	0.060	0.058	0.052	0.052	0.062	0.053	0.053	0.053
0.75	0.074	0.068	0.064	0.064	0.060	0.052	0.052	0.052
0.90	0.062	0.055	0.060	0.059	0.074	0.066	0.065	0.065

Table 3.3 Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ for large n and different strategies for finding a global maximum

n	100				120			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.256	0.253	0.249	0.251	0.271	0.266	0.262	0.264
-0.75	0.421	0.258	0.312	0.334	0.278	0.242	0.220	0.229
-0.60	0.269	0.159	0.254	0.166	0.169	0.137	0.137	0.140
-0.45	0.131	0.092	0.092	0.092	0.105	0.100	0.100	0.101
-0.30	0.081	0.077	0.077	0.078	0.078	0.078	0.078	0.078
-0.15	0.063	0.063	0.061	0.063	0.065	0.065	0.065	0.065
0.00	0.055	0.055	0.055	0.056	0.056	0.056	0.056	0.056
0.15	0.061	0.061	0.061	0.061	0.060	0.060	0.060	0.060
0.30	0.061	0.061	0.061	0.061	0.065	0.065	0.065	0.065
0.45	0.057	0.059	0.059	0.059	0.062	0.062	0.062	0.062
0.60	0.053	0.056	0.056	0.056	0.048	0.048	0.048	0.048
0.75	0.049	0.052	0.051	0.052	0.046	0.046	0.047	0.047
0.90	0.047	0.048	0.047	0.048	0.065	0.065	0.065	0.065

Table 3.4 Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ for small n and different strategies for finding a global maximum

n	20				40			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.077	0.077	0.077	0.076	0.054	0.054	0.054	0.054
-0.75	0.091	0.091	0.091	0.088	0.105	0.105	0.105	0.107
-0.60	0.113	0.113	0.113	0.116	0.171	0.159	0.159	0.162
-0.45	0.171	0.170	0.170	0.172	0.171	0.140	0.139	0.145
-0.30	0.179	0.177	0.177	0.172	0.128	0.124	0.124	0.137
-0.15	0.184	0.184	0.184	0.208	0.105	0.103	0.103	0.114
0.00	0.172	0.168	0.168	0.197	0.089	0.089	0.089	0.100
0.15	0.181	0.179	0.179	0.197	0.067	0.067	0.067	0.100
0.30	0.165	0.155	0.154	0.177	0.073	0.073	0.073	0.071
0.45	0.159	0.141	0.140	0.159	0.080	0.080	0.080	0.081
0.60	0.149	0.130	0.128	0.140	0.093	0.093	0.091	0.094
0.75	0.168	0.159	0.156	0.171	0.151	0.151	0.141	0.141
0.90	0.126	0.126	0.125	0.109	0.124	0.124	0.119	0.120

Table 3.5 Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ for moderate n and different strategies for finding a global maximum

n	60				80			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.088	0.122	0.088	0.088	0.112	0.111	0.109	0.110
-0.75	0.134	0.217	0.122	0.124	0.264	0.218	0.209	0.220
-0.60	0.150	0.148	0.113	0.116	0.229	0.151	0.152	0.159
-0.45	0.140	0.109	0.121	0.124	0.137	0.097	0.096	0.102
-0.30	0.099	0.079	0.091	0.095	0.071	0.067	0.067	0.068
-0.15	0.071	0.074	0.071	0.071	0.075	0.075	0.075	0.075
0.00	0.063	0.073	0.071	0.064	0.058	0.058	0.058	0.058
0.15	0.072	0.078	0.072	0.072	0.076	0.076	0.076	0.076
0.30	0.068	0.070	0.065	0.061	0.068	0.066	0.066	0.066
0.45	0.100	0.068	0.068	0.068	0.077	0.061	0.061	0.061
0.60	0.154	0.094	0.077	0.078	0.153	0.056	0.055	0.051
0.75	0.210	0.162	0.101	0.101	0.212	0.077	0.076	0.077
0.90	0.168	0.195	0.109	0.109	0.204	0.139	0.125	0.126

Table 3.6 Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ for large n and different strategies for finding a global maximum

n	100				120			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.151	0.148	0.146	0.148	0.236	0.233	0.230	0.233
-0.75	0.320	0.358	0.248	0.261	0.274	0.239	0.222	0.234
-0.60	0.216	0.127	0.124	0.135	0.156	0.127	0.126	0.129
-0.45	0.117	0.076	0.075	0.077	0.092	0.091	0.091	0.101
-0.30	0.089	0.082	0.081	0.083	0.079	0.079	0.079	0.079
-0.15	0.057	0.057	0.054	0.058	0.061	0.061	0.061	0.061
0.00	0.057	0.057	0.056	0.057	0.055	0.055	0.055	0.055
0.15	0.050	0.050	0.050	0.050	0.068	0.068	0.068	0.068
0.30	0.061	0.061	0.060	0.063	0.063	0.063	0.063	0.063
0.45	0.084	0.065	0.065	0.065	0.063	0.063	0.063	0.063
0.60	0.150	0.079	0.077	0.078	0.068	0.067	0.067	0.067
0.75	0.211	0.074	0.074	0.075	0.093	0.091	0.091	0.091
0.90	0.168	0.195	0.109	0.109	0.204	0.139	0.125	0.126

Table 3.7: Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$ for small n and different strategies for finding a global maximum

n	20				40			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.075	0.075	0.075	0.074	0.057	0.057	0.057	0.058
-0.75	0.090	0.090	0.090	0.086	0.105	0.105	0.105	0.105
-0.60	0.111	0.111	0.111	0.112	0.180	0.167	0.167	0.171
-0.45	0.157	0.157	0.157	0.159	0.180	0.141	0.140	0.148
-0.30	0.167	0.164	0.164	0.177	0.124	0.120	0.120	0.133
-0.15	0.166	0.165	0.165	0.185	0.105	0.103	0.103	0.112
0.00	0.164	0.162	0.162	0.190	0.085	0.085	0.085	0.097
0.15	0.174	0.173	0.173	0.201	0.071	0.071	0.071	0.075
0.30	0.168	0.156	0.155	0.178	0.079	0.079	0.079	0.084
0.45	0.169	0.151	0.149	0.167	0.085	0.085	0.085	0.086
0.60	0.159	0.142	0.140	0.151	0.096	0.096	0.094	0.097
0.75	0.172	0.165	0.160	0.179	0.151	0.151	0.141	0.141
0.90	0.119	0.118	0.117	0.127	0.121	0.121	0.114	0.115

Table 3.8: Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$ for moderate n and different strategies for finding a global maximum

n	60				80			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.087	0.090	0.087	0.087	0.164	0.163	0.161	0.162
-0.75	0.127	0.160	0.118	0.120	0.334	0.280	0.269	0.281
-0.60	0.146	0.175	0.113	0.116	0.277	0.170	0.168	0.175
-0.45	0.143	0.133	0.124	0.128	0.161	0.115	0.114	0.116
-0.30	0.096	0.106	0.088	0.093	0.107	0.101	0.101	0.102
-0.15	0.066	0.073	0.066	0.066	0.097	0.097	0.097	0.097
0.00	0.066	0.069	0.066	0.067	0.074	0.074	0.074	0.074
0.15	0.073	0.055	0.073	0.073	0.058	0.058	0.058	0.058
0.30	0.064	0.083	0.061	0.061	0.054	0.054	0.054	0.054
0.45	0.097	0.084	0.068	0.068	0.063	0.058	0.058	0.058
0.60	0.157	0.080	0.080	0.081	0.062	0.053	0.053	0.053
0.75	0.207	0.162	0.103	0.102	0.060	0.052	0.052	0.052
0.90	0.181	0.071	0.123	0.123	0.074	0.066	0.065	0.065

Table 3.9: Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$ for large n and different strategies for finding a global maximum

n	100				120			
γ	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
-0.90	0.256	0.253	0.249	0.251	0.271	0.266	0.262	0.264
-0.75	0.421	0.258	0.312	0.334	0.278	0.242	0.220	0.229
-0.60	0.269	0.159	0.254	0.166	0.169	0.137	0.137	0.140
-0.45	0.131	0.092	0.092	0.092	0.105	0.100	0.100	0.101
-0.30	0.081	0.077	0.077	0.078	0.078	0.078	0.078	0.078
-0.15	0.063	0.063	0.061	0.063	0.065	0.065	0.065	0.065
0.00	0.055	0.055	0.055	0.056	0.056	0.056	0.056	0.056
0.15	0.061	0.061	0.061	0.061	0.060	0.060	0.060	0.060
0.30	0.061	0.061	0.061	0.061	0.065	0.065	0.065	0.065
0.45	0.057	0.059	0.059	0.059	0.062	0.062	0.062	0.062
0.60	0.053	0.056	0.056	0.056	0.048	0.048	0.048	0.048
0.75	0.049	0.052	0.051	0.052	0.046	0.046	0.047	0.047
0.90	0.047	0.048	0.047	0.048	0.065	0.065	0.065	0.065

CHAPTER 4

Improved Tests of Regression Coefficients in the Presence of MA(1) Disturbances

4.1 Introduction

In general, the results of Chapter 3 are somewhat disappointing in the sense that sizes are often significantly greater than the nominal size. This suggests the need to look at other ways of constructing the test statistic that might help with this problem. One obvious approach is to use the maximum marginal likelihood (MML) method to estimate the nuisance parameter γ. As outlined in Chapter 2, there is a growing literature that suggests that the use of the marginal likelihood in place of the concentrated likelihood can help to reduce estimation bias (see for example, Cooper and Thompson (1977), Tunnicliffe Wilson (1989), Rahman and King (1998) and Laskar and King (1998)). We therefore expect nuisance parameters estimated via the marginal likelihood to result in tests with better small sample sizes than those from using estimates from the concentrated likelihood.

We also consider different approximating distributions for the distribution of the test statistic under the null hypothesis.

We have seen in Chapter 3 that the traditional likelihood based test, $(ML(n))$ often produces inaccurate sizes, which are highly significantly different from the nominal size. This situation is true for small samples and to some extent also true for larger samples especially when estimates of the parameter value frequently occur at the boundary. As a result, we consider a modification to the test statistic with the idea of reducing the estimation bias by using the maximum marginal likelihood estimate. Another contribution of this chapter is to consider all estimated parameters (except the variance term) in the count of the loss of degrees of freedom rather than the usual approach of only including estimated regression coefficients in the estimated parameter count.

The aim of this chapter is to investigate these two possible improvements to the size of the test and reassess the importance of taking care in finding the global maximum when applying ML and MML techniques.

The rest of this chapter is outlined as follows. In Section 4.2, we discuss the model and the construction of test statistics using various degrees of freedom based on estimation by both ML and MML. In Section 4.3, we calculate the estimated sizes of tests using different strategies for estimating γ with different degrees of freedom for different values of the moving average parameter γ using Monte Carlo simulation. Section 4.4 contains the discussion of the results for different test statistics for various strategies of ML and MML estimation and for different γ values. Concluding remarks are given in Section 4.5.

4.2 The Model and the Test Statistic

4.2.1 The model and estimation issues

Returning to the general linear regression model with MA(1) errors of Chapter 3, namely

$$y = X\beta + u \tag{4.1}$$

where $u_t = \varepsilon_t + \gamma\varepsilon_{t-1}$, $-1 \le \gamma \le 1$, and $\varepsilon_t \sim IIN(0, \sigma^2)$. Our interest is in testing elements of β which requires an estimate the unknown parameter γ. One possibility is to maximise the concentrated likelihood as outlined in Section 3.2 of Chapter 3. As mentioned in the previous section, an obvious alternative is to maximize the marginal likelihood for γ. According to Tunnicliffe Wilson (1989), the log marginal likelihood for (4.1) is given by

$$l_m(y;\gamma) = -\frac{1}{2}\log|\Sigma(\gamma)| - \frac{1}{2}\log|X'\Sigma^{-1}(\gamma)X| - \frac{m}{2}\log(s^2(\gamma)) \tag{4.2}$$

where $\Sigma(\gamma)$ is defined in (3.3), $m = n - k$, $s^2(\gamma)$ is given by (3.5) and $s^2(\gamma)/n$ is an estimate of σ^2. The MML estimator of γ is found by maximising (4.2) with respect to γ. In the remainder of this chapter, we will use $\hat{\gamma}^{(1)}$ to denote the ML estimator of γ obtained by maximising the concentrated likelihood (3.4) and $\hat{\gamma}^{(2)}$ to denote the MML estimator. The corresponding final estimate of β in each case can be written as

$$\hat{\beta}\left(\hat{\gamma}^{(l)}\right)=\left(X'\Sigma^{-1}\left(\hat{\gamma}^{(l)}\right)X\right)^{-1}X'\Sigma^{-1}\left(\hat{\gamma}^{(l)}\right)y \tag{4.3}$$

where $l=1,2$.

As is the case for the concentrated likelihood, we are unable to maximise the marginal likelihood analytically so have to resort to numerical optimisation. The minor identification problem discussed in Chapter 3 that results from $\sigma^2\Sigma(\gamma)$ and $\sigma_*^2\Sigma(\gamma_*)$ taking exactly the same value whenever $\sigma_*^2=\sigma^2\gamma^2$ and $\gamma_*=1/\gamma$ also effects the marginal likelihood in the same manner in which it did effect the concentrated likelihood. This can be seen from the fact that

$$\begin{aligned}
l_m(y;1/\gamma) &= -\frac{1}{2}log|\Sigma(1/\gamma)|-\frac{1}{2}log|X'\Sigma^{-1}(1/\gamma)X|-\frac{m}{2}log(s^2(1/\gamma))\\
&= -\frac{1}{2}log|1/\gamma^2\Sigma(\gamma)|-\frac{1}{2}log|\gamma^2X'\Sigma^{-1}(\gamma)X|-\frac{m}{2}log(\gamma^2s^2(\gamma))\\
&= -\frac{1}{2}log|\Sigma(\gamma)|+\frac{n}{2}log(\gamma^2)-\frac{1}{2}log(\gamma^2)^k-\frac{1}{2}log|X'\Sigma^{-1}(\gamma)X|\\
&\quad -\frac{m}{2}log(\gamma^2)-\frac{m}{2}log(s^2(\gamma))\\
&= -\frac{1}{2}log|\Sigma(\gamma)|-\frac{1}{2}log|X'\Sigma^{-1}(\gamma)X|-\frac{m}{2}log(s^2(\gamma))\\
&= l_m(y;\gamma).
\end{aligned}$$

4.2.2 The test statistics

Recall from Chapter 3 that our interest is in testing the j th regression coefficient in (4.1), namely testing $H_0:\beta_j=\beta_{j0}$ against $H_1:\beta_j\neq\beta_{j0}$ where $1\leq j\leq k$ and β_{j0} is a known constant. There are now a number of test statistics for consideration and these are of the form

$$t = \frac{\hat{\beta}_j(\hat{\gamma}^{(l)}) - \beta_{j0}}{Se(\hat{\beta}_j(\hat{\gamma}^{(l)}), p)}, \quad l = 1, 2, \tag{4.4}$$

where $\hat{\beta}_j(\hat{\gamma}^{(l)})$ is the j th element of (4.3) and

$$Se(\hat{\beta}_j(\hat{\gamma}^{(l)}), p) = \left[\left(s^2(\hat{\gamma}^{(l)}) / p\right) \left(X' \Sigma^{-1}(\hat{\gamma}^{(l)}) X\right)^{-1}_{jj} \right]^{1/2} \tag{4.5}$$

is the standard error of $\hat{\beta}_j(\hat{\gamma}^{(l)})$ in which $\left(X' \Sigma^{-1}(\hat{\gamma}^{(l)}) X\right)^{-1}_{jj}$ represents the jj th diagonal element of $\left(X' \Sigma^{-1}(\hat{\gamma}^{(l)}) X\right)^{-1}$.

An obvious question that may arise is what should the value of p be? Because $s^2(\hat{\gamma}) / n$ is the MLE of σ^2, $p = n$ is one obvious choice of the value of p as used in Chapter 3. If the value of γ is known then $s^2(\gamma)/(n-k)$ is the unbiased estimator of σ^2 so another option is $p = n - k$. Finally observe that if both β and γ are known,

$$(y - X\beta)' \Sigma^{-1}(\gamma)(y - X\beta) / n \tag{4.6}$$

is an unbiased estimator of σ^2. $s^2(\gamma)$ involves replacing the $k \times 1$ vector β by its estimated value $\hat{\beta}(\gamma)$ and this results in a loss of k degrees of freedom making $p = n - k$ the appropriate denominator. By analogy, therefore, if we replace both β and γ by estimates, this might suggest a loss of more than k degrees of freedom perhaps making $p = n - k - 1$ a more appropriate denominator. In other words, we could summarize the different statistics as (4.4) with $l = 1, 2$ and $p = n,\ n - k,\ n - k - 1$. However the use of $p = n$ with the *MML* estimator does not make sense so $p = n$ should only go with $l = 1$. The critical values come from the Student's t distribution with p degrees of freedom.

4.3 Monte Carlo Experiment

A Monte Carlo experiment was conducted to investigate the importance of taking care in finding the global maximum of the concentrated likelihood and the marginal likelihood. The second aim was to examine whether the use of different degrees of freedom resulted in an improvement in the estimated sizes of tests of regression coefficients in the presence of MA(1) regression disturbances.

4.3.1 Design of the simulation experiment

We reconsidered the four strategies of Chapter 3 in order to re-evaluate their performance for hypothesis testing in the context of ML and MML estimation of the MA(1) parameter γ. The test statistic used based on maximum likelihood estimates are $ML(m)$, $ML(m-1)$ and for marginal likelihood estimates are $MML(m)$ and $MML(m-1)$ with m and $m-1$ degrees of freedom for $n = 20$, 40, 60, 80, 100 and 120. In order to test the null hypothesis H_0: $\beta_j = 0$; against alternative H_1: $\beta_j \neq 0$; we used the nominal size of 5%. The experiment involved 1000 replications. The design matrix, different γ values, the initial estimate of the parameter γ to start with constrained optimisation and the initial values of the different factors for the SA algorithm remain the same as used in Chapter 3.

4.4 Results of the Monte Carlo Experiment

Tables 4.1 to 4.9 contain the Monte Carlo results for testing H_0: $\beta_1 = \beta_{10}$, H_0: $\beta_2 = \beta_{20}$ and H_0: $\beta_3 = \beta_{30}$ respectively for different sample sizes, namely $n = 20$, 40, 60, 80, 100 and 120, different strategies, e.g., strategy (i), (ii), (iii) and (iv), different moving average coefficients, namely $\gamma = -0.90$, -0.75, -0.60, -0.45, -0.30, -0.15, 0, 0.15, 0.30, 0.45, 0.60, 0.75 and 0.90. Estimated sizes in the range (0.0365, 0.0635) are not significantly different from the nominal level of 0.05 at the 5% level of significance. In the following subsections, we will discuss the results in terms of different sample sizes, different strategies, different moving average coefficients, and different test statistics.

The results reflect that the tests based on marginal likelihood estimates of γ produce better estimated sizes than those based on maximum likelihood estimates in terms of accuracy of sizes (closeness to nominal size) with few exceptions. The most striking feature is that the sizes of the $MML(m-1)$ test are best compared to the other tests and the majority of these sizes are insignificantly different from the nominal size. In the next subsection we will discuss and compare results for different test statistics.

4.4.1 Different test statistics

The estimated sizes of the four different tests based on ML and MML are reported in Tables 4.1 to 4.9. When $n = 20$, the results of these tests are highly significantly different from the nominal size. However there are a few exceptions to this. For instance, when values of γ are close to the negative and positive boundaries, the sizes of the $MML(m-1)$ test are insignificantly different from the nominal size for testing H_0: $\beta_1 = \beta_{10}$. The same trend is observed for testing H_0: $\beta_2 = \beta_{20}$ and H_0: $\beta_3 = \beta_{30}$ with the only exceptions being when the values of γ are close to the negative boundary.

In contrast, the sizes of the $MML(m-1)$ test based on strategies (i), (ii), (iii) and (iv) are not significantly different from the nominal size when $n = 40$, and for γ values in the interval (-0.30 to 045) when testing H_0: $\beta_2 = \beta_{20}$ and H_0: $\beta_3 = \beta_{30}$. From the above discussion we conclude that for small samples, the test statistic $MML(m-1)$ performs better than the other test statistics. In other words, the performance of the test based on MML is better than the test based on ML when the sample size is small. On the other hand, the difference between the existing t tests, i.e., $ML(m)$ and $MML(m)$, and the corresponding tests based on corrected degrees of freedom, i.e., $ML(m-1)$ and $MML(m-1)$, is wider when the sample size is smaller. This indicates that for small samples, the test based on corrected degrees of freedom is better than the traditional test. For example, Table 4.4 shows that when $n = 20$, $\gamma = 0.15$ and for strategy (iii) the estimated sizes for test statistics $MML(m-1)$ and $MML(m)$ are 0.088 and 0.97 respectively while for $n = 120$ the estimated sizes become 0.061 and 0.063 respectively in Table 4.6. It is also observed that as the sample size increases, the sizes of the existing t test and tests based on corrected degrees of freedom gradually decrease.

An obvious feature is that as the sample size increases, the estimated sizes of all the tests based on both likelihoods are insignificantly different from the nominal size over a widening interval of γ values. For example, in Table 4.5 when $n = 80$ and γ in the interval $(-0.15 \text{ to } 0.30)$, in Table 4.9 when $n = 100$ for the same γ values and when $n = 120$ and γ in the interval $(-0.15 \text{ to } 0.60)$, the estimated sizes are insignificantly different from the nominal size for both ML and MML tests. Another clear feature is that as the γ value departs from -0.15 towards the negative boundary, the estimated sizes become highly significantly different from the nominal size. Simultaneously, noticeable improvement can be seen from using the marginal likelihood based test with the corrected degrees of freedom, i.e., $MML(m-1)$, through strategy (iii). A possible reason is that using strategy (iii) helps to reduce the estimation bias extensively. A similar feature can be observed for the positive boundary.

The results of the simulation experiment reflect that compared to other test statistics, $MML(m-1)$ produces more accurate results in the sense that the estimated size of the test and the nominal size are very close to each other. The choice of $m-1$ degrees of freedom in the test has a small but worthwhile effect on sizes. The test constructed with these degrees of freedom using both ML and MML estimates shows an appreciable improvement in accuracy. The test based on maximum likelihood with $n-k-1$ degrees of freedom on average cannot even beat the test based on the marginal likelihood with $n-k$ degrees of freedom. As a result, the $MML(m-1)$ test almost always has the most accurate size compared to other tests for small sample sizes as well as large sample sizes and all different values of γ. We recommend the $MML(m-1)$ test for use.

4.4.2 Different sample sizes

In this subsection, we will discuss the effect of estimated sizes of the tests for different sample sizes for testing H_0: $\beta_i = \beta_{i0}$ where $i = 1, 2, 3$. We consider three categories of sample sizes, namely small (n = 20, 40), medium (n = 60, 80) and large (n = 100, 120) sample sizes. For small samples, the sizes of all the tests are significantly different from the nominal size except for $\gamma = -0.75$ and -0.90, where most of the tests are not significantly different from the nominal size. Compared to other tests, the $MML(m-1)$ test gives better results, because it produces smaller estimated sizes. For example, in Table 4.4, the estimated sizes of the two tests, $ML(m)$ and $ML(m-1)$, based on maximum likelihood estimation for strategy (i) are 0.144 and 0.131 respectively, and for the marginal likelihood estimates, these values are 0.091 and 0.079, respectively, when the value of γ is -0.15. Surprisingly, for the value of $\gamma = -0.90$, the estimated sizes of the all tests considered are insignificantly different from the nominal size and the tests based on the marginal likelihood are undersized.

For moderate sample sizes, that is, when $n = 60$ and 80, the estimated sizes of the tests improve slightly compared to those for small sample sizes. The most striking feature is that the estimated sizes of the tests perform very poorly near the boundary except when $n = 60$ and for $\gamma = -0.75$ and -0.90. In this category, the estimated sizes of the tests perform well in the sense that the estimated size is close to the nominal size when γ lies between 0 and 0.30 for $n = 60$ and γ lies between −0.30 and 0.45 when $n = 80$. In the case of large sample sizes, the estimated sizes of the tests perform very well when γ is in the interval −0.15 to 0.30.

The results for different sample sizes show a general trend that as the sample size increases, estimated sizes move closer to the nominal size. In other words, as the sample size increases keeping other factors unchanged, there is a tendency for the estimated size to decrease towards the nominal size.

4.4.1 Different values of the MA(1) parameter

In this subsection, we will discuss the results of the estimated sizes of the test for different values of the moving average parameter γ. The first observable feature is that for small samples, i.e. when $n = 20$ and 40, and for $\gamma = -0.90$ in the case of testing $H_0: \beta_i = \beta_{i0}$, the estimated sizes seem to be more accurate in the sense that they are insignificantly different from the nominal size. The test based on maximum likelihood produces more reasonable results than the other tests, in the sense that the estimated sizes are close to the nominal size 0.05 and the marginal likelihood based tests have a tendency to be slightly undersized. In contrast, when the value of γ departs from zero towards the boundaries except for the value of $\gamma = -0.90$, the sizes gradually increase and they are significantly different from the nominal sizes for testing H_0: $\beta_2 = \beta_{20}$ and H_0: $\beta_3 = \beta_{30}$.

For moderate and large samples and values of γ close to the boundaries, all the tests perform very poorly by producing much higher sizes than the nominal size. Values of γ within the interval $(-0.15, 0.90)$ for testing $H_0: \beta_1 = \beta_{10}$ result in more accurate results and surprisingly the ML and MML based tests provide approximately the same results. The same trend is also observed for testing $H_0: \beta_2 = \beta_{20}$ and $H_0: \beta_3 = \beta_{30}$ when the γ value is in the interval $(-0.30, 0.45)$. The worst scenario revealed in Table 4.2 is for the value of $\gamma = -0.75$ where the estimated sizes of different tests are around 0.42 for strategy (i).

In summary, the majority of the estimated sizes are significantly greater than the nominal size of 0.05 for different values of γ. Surprisingly the estimated sizes are unacceptably high for small samples, particularly, when γ values are -0.60, -0.40 and -0.30; for the ML test as shown in Table 4.1. There is a tendency for the test sizes to increase as γ moves from zero towards either of the boundaries. On the other hand, as the sample size increases for positive values of γ, the estimated sizes give more accurate results. The corrected degrees of freedom based test for marginal likelihood, i.e., $MML(m-1)$, produces the best sizes for different values of γ.

4.4.3 Different strategies

For small samples, the overall performance of strategy (i), i.e., not making any effort to find the global maximum is the worst of the four strategies. For maximum likelihood estimation, the performance of strategy (iii) is best and the estimated sizes of strategy (i) are the worst. In contrast, strategy (iv) dominates all strategies in the case of marginal likelihood based estimates.

As sample size increases, the gap among different strategies tends to decrease gradually and the sizes of the tests improve. For γ in the interval $(-0.15, 0.60)$, the estimated sizes are barely significantly different from the nominal size for moderate sized samples. However, near to the negative boundary, the estimated sizes are woefully significantly different from the nominal level. Strategy (iii) dominates the other strategies except for strategy (iv) in the case of marginal likelihood based estimates. This is also sometimes true for positive values of γ. In Table 4.2, we notice that strategy (ii) gives improved accuracy of the estimated sizes compared to strategy (i) when the estimation is based on marginal likelihood for negative values of γ. The same trend can be seen while comparing strategy (iii) with (ii), and strategy (iii) with (i). For example, the estimated sizes of $MML(m)$ and $MML(m-1)$ from strategy (ii) are 0.093, 0.091 respectively and for strategy (iii) are 0.079 and 0.077, respectively for $\gamma = -0.450$ and $n = 60$ in Table 4.8.

For large samples, the performance of strategy (iv) is not satisfactory compared to other strategies for tests based on maximum likelihood estimates.

The improvement of strategy (iii) combined with the corrected degrees of freedom based test is noticeable and works well particularly when the values of γ are close to the boundary. For the marginal likelihood based test, strategy (iv) produces either the same sizes or a slight improvement over those of strategy (iii) for γ values close to the boundary.

The most significant result is the improvement in sizes that comes from strategy (iv) which dominates all strategies when estimation is based on marginal likelihood. This strategy works well to solve the boundary problem as we mentioned in the previous chapter where the worst performance is observed in the case of maximum likelihood estimation. Strategy (iii) and strategy (iv) produce approximately the same results and the majority of the estimated sizes are insignificantly different from the nominal size for large samples. But for small samples, strategy (iv) produces better sizes when estimation is based on marginal likelihood.

4.5 Conclusion

In Chapter 3, we found unsatisfactory estimated sizes of the test statistics based on maximum likelihood. In this chapter, an attempt was made to improve the estimated sizes by proposing modified test statistics based on corrected degrees of freedom for both the concentrated likelihood and the marginal likelihood estimates. We investigated whether there is any improvement due to modifying the test in context of the estimated sizes for different strategies, sample sizes and moving average parameter values γ, through Monte Carlo simulation.

Our simulation results indicate that the test based on marginal likelihood estimates and $m-1$ degrees of freedom, $MML(m-1)$, performs best irrespective of strategies, samples size, and values of the moving average parameter γ, in terms of accuracy of sizes, i.e., estimated sizes close to the nominal size.

In the context of sample sizes, we observed that there is a close relationship between sample sizes and estimated sizes. In other words, as the sample size increases keeping other factors unchanged, there is a tendency to get more accurate estimated sizes in the sense that the estimated sizes are closer to the nominal size.

Also, the results revealed that there is a tendency for the tests sizes to increase as γ moves from zero towards boundaries where the majority of the estimated sizes are significantly different from the nominal sizes for small samples.

In Chapter 3 we found a rather worrying result that finding the global maximum does not always work in the case of maximum likelihood estimation of γ. To overcome this problem, we investigated the use marginal likelihood for finding the global maximum in this chapter. We found that the performance of the simulated annealing method, which finds the global maximum, gives the best results, compared to the other strategies when estimation is based on marginal likelihood. This is fortunate because looking for the global maximum of a likelihood does seems to be the natural strategy for an econometrician to follow.

The results of our experiment clearly support to use the maximum marginal likelihood method instead of the concentrated likelihood when finding the global maximum while estimating the nuisance parameter. In addition, we conclude that one should use $m-1$ degrees of freedom rather than m degrees of freedom in the test statistic and use critical values from the Student's t distribution with $m-1$ degrees of freedom. This small correction does provide a slight improvement in the resultant sizes.

Table 4.1 Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ with $n = 20$ and 40 for different strategies for maximising the concentrated and marginal likelihood functions

n	20				40			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.058	0.058	0.056	0.058	0.060	0.060	0.060	0.061
$ML(m-1)$	0.053	0.053	0.051	0.052	0.057	0.057	0.057	0.058
$MML(m)$	0.047	0.041	0.040	0.041	0.042	0.043	0.041	0.041
$MML(m-1)$	0.043	0.039	0.038	0.039	0.040	0.041	0.039	0.039
				$\gamma = -0.75$				
$ML(m)$	0.125	0.125	0.122	0.122	0.107	0.106	0.106	0.106
$ML(m-1)$	0.111	0.111	0.108	0.108	0.101	0.099	0.099	0.100
$MML(m)$	0.084	0.072	0.075	0.071	0.074	0.076	0.074	0.075
$MML(m-1)$	0.075	0.066	0.066	0.063	0.073	0.074	0.073	0.074
				$\gamma = -0.60$				
$ML(m)$	0.230	0.230	0.225	0.227	0.173	0.164	0.164	0.168
$ML(m-1)$	0.214	0.214	0.205	0.207	0.171	0.160	0.160	0.163
$MML(m)$	0.153	0.122	0.126	0.123	0.095	0.094	0.098	0.094
$MML(m-1)$	0.144	0.115	0.117	0.114	0.090	0.090	0.094	0.089
				$\gamma = -0.45$				
$ML(m)$	0.337	0.333	0.331	0.346	0.173	0.170	0.170	0.176
$ML(m-1)$	0.330	0.326	0.320	0.334	0.171	0.164	0.164	0.170
$MML(m)$	0.196	0.154	0.159	0.156	0.095	0.088	0.090	0.088
$MML(m-1)$	0.193	0.151	0.153	0.151	0.090	0.085	0.087	0.085
				$\gamma = -0.30$				
$ML(m)$	0.349	0.345	0.339	0.376	0.154	0.149	0.149	0.161
$ML(m-1)$	0.338	0.334	0.326	0.326	0.152	0.147	0.147	0.158
$MML(m)$	0.182	0.157	0.164	0.157	0.080	0.078	0.079	0.080
$MML(m-1)$	0.177	0.153	0.159	0.152	0.078	0.076	0.077	0.078
				$\gamma = -0.15$				
$ML(m)$	0.328	0.326	0.321	0.371	0.151	0.147	0.147	0.162
$ML(m-1)$	0.317	0.315	0.312	0.363	0.145	0.141	0.141	0.156
$MML(m)$	0.140	0.128	0.133	0.128	0.093	0.092	0.092	0.093
$MML(m-1)$	0.136	0.124	0.129	0.124	0.090	0.084	0.089	0.090

Table 4.1 (continued) Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ with $n = 20$ and 40 for different strategies for maximising the concentrated and marginal likelihood functions

n	20				40			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.268	0.267	0.265	0.316	0.095	0.093	0.093	0.105
$ML(m-1)$	0.261	0.260	0.265	0.308	0.091	0.090	0.090	0.102
$MML(m)$	0.124	0.117	0.124	0.117	0.062	0.062	0.062	0.061
$MML(m-1$	0.112	0.104	0.110	0.103	0.058	0.057	0.057	0.057
				γ =0.15				
$ML(m)$	0.201	0.202	0.201	0.269	0.068	0.068	0.068	0.069
$ML(m-1)$	0.198	0.199	0.193	0.261	0.066	0.065	0.065	0.066
$MML(m)$	0.094	0.090	0.097	0.092	0.055	0.055	0.055	0.056
$MML(m-1$	0.087	0.083	0.088	0.083	0.052	0.052	0.052	0.053
				γ =0.30				
$ML(m)$	0.188	0.191	0.187	0.239	0.074	0.074	0.074	0.079
$ML(m-1)$	0.181	0.183	0.179	0.230	0.070	0.070	0.070	0.075
$MML(m)$	0.104	0.098	0.101	0.101	0.062	0.062	0.062	0.061
$MML(m-1$	0.093	0.087	0.089	0.089	0.059	0.059	0.059	0.059
				γ =0.45				
$ML(m)$	0.143	0.144	0.142	0.184	0.067	0.067	0.067	0.067
$ML(m-1)$	0.137	0.137	0.134	0.176	0.063	0.063	0.063	0.063
$MML(m)$	0.085	0.082	0.082	0.085	0.058	0.058	0.058	0.058
$MML(m-1$	0.080	0.077	0.076	0.079	0.058	0.058	0.058	0.058
				γ =0.60				
$ML(m)$	0.095	0.100	0.092	0.115	0.059	0.059	0.058	0.060
$ML(m-1)$	0.083	0.088	0.083	0.106	0.054	0.054	0.053	0.054
$MML(m)$	0.059	0.056	0.057	0.058	0.054	0.054	0.053	0.053
$MML(m-1$	0.053	0.051	0.051	0.052	0.052	0.052	0.051	0.055
				γ =0.75				
$ML(m)$	0.099	0.101	0.098	0.102	0.065	0.065	0.066	0.066
$ML(m-1)$	0.088	0.091	0.090	0.097	0.059	0.059	0.060	0.060
$MML(m)$	0.073	0.071	0.074	0.065	0.065	0.065	0.062	0.062
$MML(m-1$	0.063	0.061	0.064	0.059	0.059	0.059	0.057	0.056
				γ =0.90				
$ML(m)$	0.094	0.095	0.093	0.061	0.056	0.056	0.056	0.056
$ML(m-1)$	0.089	0.090	0.087	0.058	0.051	0.051	0.051	0.051
$MML(m)$	0.068	0.066	0.066	0.041	0.055	0.055	0.054	0.054
$MML(m-1$	0.064	0.061	0.060	0.039	0.050	0.050	0.050	0.050

Table 4.2 Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ with $n = 60$ and 80 for different strategies for maximising the concentrated and marginal likelihood functions

n	60				80			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.091	0.080	0.091	0.093	0.153	0.152	0.150	0.151
$ML(m-1)$	0.089	0.078	0.088	0.090	0.149	0.148	0.146	0.147
$MML(m)$	0.076	0.063	0.070	0.068	0.115	0.107	0.083	0.080
$MML(m-1)$	0.073	0.060	0.066	0.065	0.111	0.103	0.081	0.077
				$\gamma = -0.75$				
$ML(m)$	0.166	0.164	0.142	0.146	0.326	0.274	0.263	0.275
$ML(m-1)$	0.160	0.160	0.137	0.141	0.325	0.272	0.262	0.274
$MML(m)$	0.118	0.107	0.077	0.078	0.226	0.135	0.113	0.109
$MML(m-1)$	0.113	0.103	0.074	0.075	0.223	0.132	0.111	0.106
				$\gamma = -0.60$				
$ML(m)$	0.195	0.180	0.145	0.147	0.272	0.164	0.162	0.170
$ML(m-1)$	0.191	0.177	0.142	0.143	0.268	0.160	0.158	0.165
$MML(m)$	0.135	0.117	0.081	0.076	0.201	0.080	0.078	0.076
$MML(m-1)$	0.132	0.113	0.080	0.075	0.201	0.079	0.076	0.074
				$\gamma = -0.45$				
$ML(m)$	0.155	0.121	0.128	0.133	0.157	0.110	0.109	0.111
$ML(m-1)$	0.148	0.117	0.121	0.126	0.157	0.110	0.109	0.111
$MML(m)$	0.108	0.087	0.074	0.074	0.120	0.076	0.076	0.076
$MML(m-1)$	0.106	0.083	0.073	0.073	0.118	0.073	0.073	0.076
				$\gamma = -0.30$				
$ML(m)$	0.116	0.085	0.108	0.111	0.101	0.095	0.095	0.096
$ML(m-1)$	0.116	0.082	0.108	0.111	0.098	0.092	0.092	0.093
$MML(m)$	0.085	0.067	0.067	0.078	0.075	0.071	0.072	0.071
$MML(m-1)$	0.085	0.066	0.066	0.078	0.073	0.069	0.070	0.069
				$\gamma = -0.15$				
$ML(m)$	0.087	0.063	0.087	0.087	0.087	0.089	0.089	0.089
$ML(m-1)$	0.084	0.062	0.084	0.084	0.088	0.088	0.088	0.086
$MML(m)$	0.060	0.059	0.060	0.060	0.080	0.080	0.080	0.080
$MML(m-1)$	0.060	0.058	0.060	0.060	0.080	0.080	0.080	0.079

Table 4.2 (continued) Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ with $n = 60$ and 80 for different strategies for maximising the concentrated and marginal likelihood functions

n	60				80			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.075	0.050	0.087	0.076	0.067	0.067	0.067	0.067
$ML(m-1)$	0.073	0.050	0.084	0.074	0.065	0.065	0.065	0.064
$MML(m)$	0.066	0.048	0.060	0.060	0.057	0.057	0.057	0.057
$MML(m-1$	0.066	0.044	0.060	0.060	0.057	0.057	0.057	0.057
				γ =0.15				
$ML(m)$	0.065	0.055	0.065	0.065	0.051	0.051	0.051	0.051
$ML(m-1)$	0.065	0.053	0.065	0.064	0.048	0.048	0.048	0.048
$MML(m)$	0.060	0.054	0.060	0.060	0.047	0.047	0.047	0.047
$MML(m-1$	0.058	0.052	0.058	0.058	0.047	0.047	0.047	0.047
				γ =0.30				
$ML(m)$	0.064	0.061	0.064	0.064	0.051	0.051	0.051	0.051
$ML(m-1)$	0.061	0.060	0.061	0.060	0.051	0.051	0.051	0.051
$MML(m)$	0.058	0.060	0.059	0.058	0.050	0.050	0.050	0.050
$MML(m-1$	0.054	0.060	0.054	0.054	0.046	0.046	0.046	0.046
				γ =0.45				
$ML(m)$	0.065	0.053	0.057	0.057	0.060	0.056	0.056	0.056
$ML(m-1)$	0.064	0.055	0.055	0.055	0.057	0.054	0.054	0.054
$MML(m)$	0.064	0.051	0.055	0.055	0.058	0.054	0.054	0.054
$MML(m-1$	0.063	0.048	0.054	0.054	0.056	0.053	0.053	0.052
				γ =0.60				
$ML(m)$	0.050	0.050	0.045	0.045	0.058	0.049	0.049	0.049
$ML(m-1)$	0.049	0.049	0.045	0.045	0.057	0.047	0.047	0.047
$MML(m)$	0.049	0.048	0.045	0.044	0.058	0.048	0.048	0.048
$MML(m-1$	0.047	0.048	0.043	0.043	0.058	0.047	0.047	0.047
				γ =0.75				
$ML(m)$	0.067	0.064	0.060	0.060	0.057	0.049	0.049	0.052
$ML(m-1)$	0.062	0.059	0.056	0.055	0.057	0.049	0.049	0.049
$MML(m)$	0.066	0.061	0.059	0.059	0.057	0.049	0.049	0.049
$MML(m-1$	0.060	0.058	0.054	0.054	0.057	0.049	0.049	0.049
				γ =0.90				
$ML(m)$	0.056	0.050	0.053	0.052	0.070	0.062	0.060	0.060
$ML(m-1)$	0.052	0.048	0.050	0.049	0.068	0.059	0.057	0.056
$MML(m)$	0.056	0.051	0.053	0.053	0.069	0.060	0.058	0.058
$MML(m-1$	0.052	0.048	0.050	0.050	0.066	0.057	0.054	0.054

Table 4.3 Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ with $n = 100$ and 120 for different strategies for maximising the concentrated and marginal likelihood functions

n	100				120			
Strategies	(i)	(ii)	(iii)	SA	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.247	0.244	0.240	0.243	0.267	0.262	0.257	0.259
$ML(m-1)$	0.245	0.242	0.238	0.240	0.264	0.259	0.254	0.256
$MML(m)$	0.179	0.171	0.126	0.124	0.137	0.207	0.133	0.133
$MML(m-1$	0.178	0.170	0.125	0.123	0.136	0.205	0.132	0.132
				$\gamma = -0.75$				
$ML(m)$	0.415	0.254	0.307	0.328	0.275	0.238	0.216	0.225
$ML(m-1)$	0.414	0.251	0.306	0.327	0.274	0.237	0.215	0.224
$MML(m)$	0.289	0.140	0.134	0.128	0.098	0.142	0.100	0.097
$MML(m-1$	0.288	0.140	0.133	0.127	0.098	0.140	0.099	0.097
				$\gamma = -0.60$				
$ML(m)$	0.262	0.151	0.148	0.155	0.168	0.137	0.137	0.140
$ML(m-1)$	0.259	0.147	0.144	0.153	0.165	0.134	0.134	0.137
$MML(m)$	0.200	0.087	0.088	0.086	0.082	0.082	0.082	0.082
$MML(m-1$	0.199	0.086	0.087	0.086	0.081	0.081	0.081	0.081
				$\gamma = -0.45$				
$ML(m)$	0.130	0.091	0.091	0.091	0.100	0.095	0.095	0.096
$ML(m-1)$	0.129	0.090	0.090	0.090	0.099	0.094	0.094	0.095
$MML(m)$	0.106	0.071	0.071	0.071	0.075	0.075	0.075	0.075
$MML(m-1$	0.104	0.069	0.069	0.070	0.072	0.072	0.072	0.072
				$\gamma = -0.30$				
$ML(m)$	0.075	0.071	0.071	0.072	0.077	0.077	0.077	0.077
$ML(m-1)$	0.073	0.069	0.069	0.070	0.075	0.075	0.075	0.075
$MML(m)$	0.060	0.054	0.055	0.055	0.062	0.062	0.062	0.062
$MML(m-1$	0.058	0.053	0.053	0.054	0.062	0.062	0.062	0.062
				$\gamma = -0.15$				
$ML(m)$	0.059	0.059	0.059	0.059	0.063	0.063	0.063	0.063
$ML(m-1)$	0.058	0.058	0.058	0.059	0.061	0.061	0.061	0.061
$MML(m)$	0.053	0.053	0.053	0.054	0.058	0.058	0.058	0.058
$MML(m-1$	0.052	0.052	0.052	0.052	0.058	0.058	0.058	0.058

Table 4.3 (continued) Estimated sizes for testing $H_0: \beta_1 = \beta_{10}$ with $n = 100$ and 120 for different strategies for maximising the concentrated and marginal likelihood functions

n	100				120			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.054	0.054	0.054	0.054	0.055	0.055	0.055	0.055
$ML(m-1)$	0.052	0.052	0.052	0.053	0.053	0.053	0.053	0.053
$MML(m)$	0.047	0.047	0.047	0.047	0.050	0.050	0.050	0.050
$MML(m-1$	0.047	0.047	0.047	0.047	0.049	0.049	0.049	0.049
				γ =0.15				
$ML(m)$	0.058	0.058	0.058	0.058	0.057	0.057	0.057	0.057
$ML(m-1)$	0.057	0.057	0.057	0.058	0.057	0.057	0.057	0.057
$MML(m)$	0.055	0.055	0.055	0.055	0.055	0.055	0.055	0.055
$MML(m-1$	0.055	0.055	0.055	0.055	0.053	0.053	0.053	0.053
				γ =0.30				
$ML(m)$	0.059	0.059	0.059	0.059	0.063	0.063	0.063	0.063
$ML(m-1)$	0.059	0.057	0.059	0.059	0.062	0.062	0.062	0.062
$MML(m)$	0.056	0.055	0.056	0.056	0.062	0.062	0.062	0.062
$MML(m-1$	0.054	0.054	0.054	0.054	0.061	0.061	0.061	0.061
				γ =0.45				
$ML(m)$	0.051	0.054	0.054	0.055	0.058	0.058	0.058	0.058
$ML(m-1)$	0.049	0.052	0.052	0.052	0.057	0.057	0.057	0.057
$MML(m)$	0.049	0.052	0.052	0.052	0.057	0.057	0.057	0.057
$MML(m-1$	0.048	0.051	0.051	0.051	0.057	0.057	0.057	0.057
				γ =0.60				
$ML(m)$	0.053	0.055	0.055	0.056	0.044	0.044	0.044	0.044
$ML(m-1)$	0.052	0.054	0.054	0.055	0.044	0.044	0.044	0.044
$MML(m)$	0.053	0.055	0.055	0.055	0.044	0.044	0.044	0.044
$MML(m-1$	0.051	0.052	0.052	0.053	0.043	0.043	0.043	0.043
				γ =0.75				
$ML(m)$	0.042	0.044	0.044	0.044	0.046	0.046	0.047	0.047
$ML(m-1)$	0.041	0.043	0.043	0.044	0.046	0.046	0.047	0.047
$MML(m)$	0.042	0.044	0.044	0.044	0.047	0.046	0.047	0.047
$MML(m-1$	0.041	0.043	0.043	0.043	0.047	0.046	0.047	0.047
				γ =0.90				
$ML(m)$	0.044	0.045	0.044	0.044	0.063	0.063	0.063	0.063
$ML(m-1)$	0.043	0.044	0.043	0.044	0.061	0.061	0.061	0.061
$MML(m)$	0.044	0.044	0.044	0.044	0.064	0.064	0.064	0.064
$MML(m-1$	0.043	0.043	0.043	0.041	0.063	0.063	0.063	0.063

Table 4.4 Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ with $n = 20$ and 40 for different strategies for maximising the concentrated and marginal likelihood functions

n	20				40			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.058	0.058	0.054	0.053	0.045	0.045	0.045	0.046
$ML(m-1)$	0.051	0.051	0.046	0.045	0.043	0.043	0.043	0.044
$MML(m)$	0.049	0.046	0.045	0.044	0.044	0.044	0.044	0.043
$MML(m-1)$	0.044	0.042	0.039	0.038	0.039	0.041	0.040	0.039
				$\gamma = -0.75$				
$ML(m)$	0.073	0.073	0.072	0.070	0.096	0.093	0.093	0.095
$ML(m-1)$	0.067	0.067	0.063	0.060	0.089	0.087	0.087	0.087
$MML(m)$	0.070	0.065	0.065	0.065	0.074	0.074	0.074	0.072
$MML(m-1)$	0.059	0.055	0.054	0.054	0.069	0.068	0.069	0.065
				$\gamma = -0.60$				
$ML(m)$	0.087	0.087	0.087	0.089	0.159	0.148	0.148	0.151
$ML(m-1)$	0.082	0.082	0.078	0.080	0.152	0.142	0.142	0.144
$MML(m)$	0.071	0.067	0.067	0.066	0.098	0.097	0.100	0.097
$MML(m-1)$	0.062	0.058	0.056	0.056	0.093	0.093	0.096	0.094
				$\gamma = -0.45$				
$ML(m)$	0.137	0.136	0.131	0.131	0.159	0.131	0.130	0.135
$ML(m-1)$	0.129	0.128	0.122	0.122	0.152	0.126	0.125	0.131
$MML(m)$	0.092	0.089	0.090	0.090	0.098	0.090	0.092	0.089
$MML(m-1)$	0.087	0.084	0.085	0.085	0.093	0.084	0.087	0.085
				$\gamma = -0.30$				
$ML(m)$	0.156	0.155	0.150	0.162	0.121	0.117	0.117	0.129
$ML(m-1)$	0.142	0.142	0.130	0.142	0.117	0.114	0.114	0.126
$MML(m)$	0.095	0.091	0.093	0.090	0.069	0.069	0.069	0.069
$MML(m-1)$	0.082	0.078	0.079	0.076	0.067	0.067	0.067	0.067
				$\gamma = -0.15$				
$ML(m)$	0.144	0.143	0.142	0.160	0.097	0.095	0.095	0.105
$ML(m-1)$	0.131	0.130	0.126	0.143	0.093	0.091	0.091	0.101
$MML(m)$	0.091	0.087	0.090	0.087	0.063	0.062	0.062	0.063
$MML(m-1)$	0.079	0.075	0.075	0.073	0.058	0.056	0.056	0.057

Table 4.4 (continued) Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ with $n = 20$ and 40 for different strategies for maximising the concentrated and marginal likelihood functions

n	20				40			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = 0.00$				
$ML(m)$	0.138	0.136	0.132	0.158	0.081	0.081	0.081	0.092
$ML(m-1)$	0.127	0.125	0.119	0.145	0.075	0.075	0.075	0.086
$MML(m)$	0.097	0.095	0.095	0.117	0.069	0.069	0.070	0.068
$MML(m-1$	0.089	0.087	0.085	0.103	0.065	0.065	0.066	0.065
				$\gamma = 0.15$				
$ML(m)$	0.147	0.144	0.140	0.269	0.059	0.059	0.059	0.063
$ML(m-1)$	0.139	0.137	0.126	0.261	0.058	0.058	0.058	0.062
$MML(m)$	0.101	0.096	0.097	0.092	0.052	0.052	0.052	0.053
$MML(m-1$	0.088	0.083	0.083	0.083	0.045	0.045	0.045	0.046
				$\gamma = 0.30$				
$ML(m)$	0.135	0.125	0.124	0.239	0.063	0.063	0.063	0.068
$ML(m-1)$	0.131	0.121	0.117	0.230	0.059	0.059	0.059	0.064
$MML(m)$	0.109	0.096	0.096	0.101	0.063	0.063	0.063	0.063
$MML(m-1$	0.099	0.086	0.086	0.089	0.061	0.061	0.061	0.061
				$\gamma = 0.45$				
$ML(m)$	0.136	0.119	0.114	0.184	0.070	0.070	0.069	0.070
$ML(m-1)$	0.124	0.109	0.105	0.176	0.067	0.067	0.066	0.067
$MML(m)$	0.108	0.083	0.082	0.085	0.068	0.068	0.067	0.066
$MML(m-1$	0.100	0.077	0.073	0.079	0.063	0.063	0.062	0.062
				$\gamma = 0.60$				
$ML(m)$	0.125	0.110	0.105	0.115	0.081	0.081	0.079	0.082
$ML(m-1)$	0.120	0.106	0.098	0.106	0.079	0.079	0.077	0.080
$MML(m)$	0.117	0.099	0.094	0.058	0.079	0.079	0.073	0.074
$MML(m-1$	0.109	0.092	0.083	0.052	0.077	0.077	0.071	0.072
				$\gamma = 0.75$				
$ML(m)$	0.141	0.133	0.130	0.112	0.137	0.137	0.127	0.126
$ML(m-1)$	0.132	0.124	0.119	0.104	0.134	0.134	0.124	0.123
$MML(m)$	0.126	0.114	0.109	0.074	0.129	0.129	0.115	0.115
$MML(m-1$	0.112	0.100	0.093	0.064	0.126	0.126	0.112	0.112
				$\gamma = 0.90$				
$ML(m)$	0.099	0.097	0.094	0.103	0.108	0.108	0.104	0.106
$ML(m-1)$	0.093	0.092	0.083	0.093	0.099	0.099	0.095	0.096
$MML(m)$	0.090	0.082	0.080	0.079	0.104	0.104	0.091	0.092
$MML(m-1$	0.080	0.074	0.069	0.068	0.097	0.097	0.083	0.084

Table 4.5 Estimated sizes for testing $H_0\colon \beta_2 = \beta_{20}$ with $n = 60$ and 80 for different strategies for maximising the concentrated and marginal likelihood functions

n	60				80			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.078	0.113	0.079	0.079	0.108	0.107	0.105	0.106
$ML(m-1)$	0.077	0.109	0.078	0.078	0.106	0.105	0.103	0.103
$MML(m)$	0.076	0.084	0.073	0.073	0.098	0.086	0.079	0.078
$MML(m-1$	0.074	0.079	0.071	0.071	0.096	0.084	0.077	0.075
				$\gamma = -0.75$				
$ML(m)$	0.116	0.205	0.104	0.106	0.255	0.211	0.202	0.213
$ML(m-1)$	0.113	0.201	0.101	0.103	0.253	0.209	0.200	0.211
$MML(m)$	0.099	0.119	0.074	0.073	0.175	0.112	0.101	0.090
$MML(m-1$	0.095	0.116	0.072	0.071	0.172	0.110	0.099	0.088
				$\gamma = -0.60$				
$ML(m)$	0.140	0.136	0.105	0.108	0.220	0.143	0.144	0.151
$ML(m-1)$	0.136	0.135	0.102	0.105	0.217	0.141	0.142	0.148
$MML(m)$	0.106	0.067	0.071	0.070	0.168	0.081	0.083	0.079
$MML(m-1$	0.102	0.066	0.069	0.067	0.166	0.080	0.082	0.078
				$\gamma = -0.45$				
$ML(m)$	0.129	0.103	0.110	0.115	0.133	0.092	0.091	0.097
$ML(m-1)$	0.127	0.101	0.108	0.112	0.130	0.089	0.088	0.094
$MML(m)$	0.100	0.071	0.079	0.078	0.106	0.067	0.067	0.067
$MML(m-1$	0.099	0.070	0.078	0.077	0.106	0.067	0.067	0.065
				$\gamma = -0.30$				
$ML(m)$	0.092	0.073	0.084	0.088	0.069	0.065	0.065	0.065
$ML(m-1)$	0.090	0.070	0.082	0.086	0.067	0.063	0.063	0.063
$MML(m)$	0.069	0.061	0.064	0.064	0.057	0.055	0.056	0.055
$MML(m-1$	0.066	0.065	0.061	0.061	0.054	0.052	0.053	0.052
				$\gamma = -0.15$				
$ML(m)$	0.063	0.063	0.063	0.063	0.067	0.067	0.067	0.067
$ML(m-1)$	0.061	0.059	0.061	0.061	0.062	0.062	0.062	0.062
$MML(m)$	0.055	0.054	0.055	0.055	0.064	0.064	0.064	0.064
$MML(m-1$	0.053	0.053	0.053	0.053	0.062	0.062	0.062	0.062

Table 4.5 (continued) Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ with $n = 60$ and 80 for different strategies for maximising the concentrated and marginal likelihood functions

n	60				80			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.057	0.071	0.063	0.058	0.053	0.053	0.053	0.053
$ML(m-1)$	0.054	0.068	0.061	0.055	0.050	0.050	0.050	0.049
$MML(m)$	0.051	0.066	0.055	0.051	0.050	0.050	0.050	0.050
$MML(m-1$	0.051	0.066	0.053	0.051	0.048	0.048	0.048	0.048
				γ =0.15				
$ML(m)$	0.063	0.060	0.063	0.063	0.071	0.071	0.071	0.071
$ML(m-1)$	0.061	0.060	0.061	0.061	0.068	0.068	0.068	0.068
$MML(m)$	0.064	0.059	0.064	0.064	0.069	0.069	0.069	0.069
$MML(m-1$	0.061	0.057	0.061	0.061	0.068	0.068	0.068	0.068
				γ =0.30				
$ML(m)$	0.062	0.063	0.059	0.059	0.060	0.058	0.058	0.058
$ML(m-1)$	0.058	0.060	0.055	0.055	0.059	0.057	0.057	0.057
$MML(m)$	0.060	0.062	0.057	0.057	0.064	0.062	0.062	0.062
$MML(m-1$	0.058	0.059	0.055	0.055	0.062	0.060	0.060	0.060
				γ =0.45				
$ML(m)$	0.089	0.065	0.058	0.058	0.074	0.057	0.057	0.057
$ML(m-1)$	0.088	0.064	0.057	0.057	0.071	0.054	0.054	0.054
$MML(m)$	0.086	0.059	0.056	0.056	0.076	0.059	0.059	0.059
$MML(m-1$	0.085	0.056	0.053	0.054	0.073	0.056	0.056	0.055
				γ =0.60				
$ML(m)$	0.141	0.086	0.065	0.066	0.130	0.051	0.050	0.051
$ML(m-1)$	0.140	0.081	0.064	0.065	0.128	0.050	0.049	0.050
$MML(m)$	0.141	0.069	0.062	0.063	0.130	0.052	0.051	0.051
$MML(m-1$	0.138	0.065	0.059	0.059	0.128	0.051	0.050	0.050
				γ =0.75				
$ML(m)$	0.199	0.154	0.089	0.090	0.209	0.075	0.074	0.075
$ML(m-1)$	0.196	0.151	0.086	0.087	0.207	0.072	0.071	0.072
$MML(m)$	0.193	0.127	0.083	0.083	0.207	0.074	0.070	0.070
$MML(m-1$	0.190	0.123	0.080	0.080	0.206	0.071	0.066	0.066
				γ =0.90				
$ML(m)$	0.161	0.185	0.100	0.100	0.199	0.134	0.120	0.121
$ML(m-1)$	0.153	0.184	0.096	0.096	0.198	0.133	0.119	0.120
$MML(m)$	0.156	0.156	0.089	0.088	0.189	0.128	0.101	0.102
$MML(m-1$	0.148	0.155	0.086	0.086	0.188	0.126	0.099	0.099

Table 4.6 Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ with $n =$ 100 and 120 for different strategies for maximising the concentrated and marginal likelihood functions

n	100				120			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.145	0.142	0.140	0.142	0.227	0.224	0.221	0.224
$ML(m-1)$	0.143	0.139	0.138	0.141	0.225	0.222	0.219	0.222
$MML(m)$	0.120	0.109	0.089	0.086	0.121	0.172	0.115	0.115
$MML(m-1$	0.119	0.109	0.089	0.086	0.120	0.170	0.114	0.114
				$\gamma = -0.75$				
$ML(m)$	0.315	0.254	0.245	0.257	0.269	0.234	0.217	0.228
$ML(m-1)$	0.311	0.251	0.242	0.254	0.269	0.234	0.216	0.227
$MML(m)$	0.226	0.140	0.113	0.106	0.108	0.139	0.107	0.106
$MML(m-1$	0.225	0.140	0.113	0.106	0.108	0.139	0.107	0.106
				$\gamma = -0.60$				
$ML(m)$	0.209	0.121	0.119	0.130	0.154	0.125	0.124	0.127
$ML(m-1)$	0.205	0.117	0.115	0.127	0.151	0.122	0.121	0.124
$MML(m)$	0.153	0.068	0.068	0.068	0.080	0.079	0.079	0.079
$MML(m-1$	0.151	0.067	0.066	0.066	0.080	0.079	0.079	0.079
				$\gamma = -0.45$				
$ML(m)$	0.113	0.072	0.072	0.073	0.089	0.088	0.088	0.089
$ML(m-1)$	0.112	0.071	0.071	0.073	0.087	0.086	0.086	0.087
$MML(m)$	0.093	0.055	0.055	0.055	0.065	0.065	0.065	0.065
$MML(m-1$	0.093	0.055	0.055	0.055	0.062	0.062	0.062	0.062
				$\gamma = -0.30$				
$ML(m)$	0.083	0.076	0.076	0.079	0.074	0.074	0.074	0.074
$ML(m-1)$	0.081	0.074	0.074	0.076	0.074	0.074	0.074	0.074
$MML(m)$	0.074	0.067	0.067	0.067	0.063	0.063	0.063	0.063
$MML(m-1$	0.073	0.066	0.066	0.066	0.062	0.062	0.062	0.062
				$\gamma = -0.15$				
$ML(m)$	0.052	0.052	0.052	0.053	0.058	0.058	0.058	0.058
$ML(m-1)$	0.050	0.050	0.050	0.051	0.056	0.056	0.056	0.056
$MML(m)$	0.046	0.045	0.046	0.046	0.058	0.058	0.058	0.058
$MML(m-1$	0.045	0.045	0.045	0.045	0.055	0.055	0.055	0.055

Table 4.6 (continued) Estimated sizes for testing $H_0: \beta_2 = \beta_{20}$ with $n = 100$ and 120 for different strategies for maximising the concentrated and marginal likelihood functions

n	100				120			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.051	0.050	0.051	0.053	0.052	0.052	0.052	0.052
$ML(m-1)$	0.048	0.048	0.048	0.048	0.051	0.051	0.051	0.051
$MML(m)$	0.050	0.050	0.050	0.050	0.049	0.049	0.049	0.049
$MML(m-1$	0.047	0.047	0.047	0.047	0.047	0.047	0.047	0.047
				γ =0.15				
$ML(m)$	0.048	0.048	0.048	0.048	0.063	0.063	0.063	0.063
$ML(m-1)$	0.046	0.046	0.046	0.048	0.062	0.062	0.062	0.062
$MML(m)$	0.047	0.047	0.047	0.048	0.063	0.063	0.063	0.063
$MML(m-1$	0.046	0.046	0.046	0.046	0.061	0.061	0.061	0.061
				γ =0.30				
$ML(m)$	0.056	0.056	0.056	0.057	0.057	0.057	0.057	0.057
$ML(m-1)$	0.054	0.054	0.054	0.054	0.054	0.054	0.054	0.054
$MML(m)$	0.056	0.055	0.056	0.057	0.054	0.054	0.054	0.054
$MML(m-1$	0.055	0.055	0.055	0.055	0.054	0.054	0.054	0.054
				γ =0.45				
$ML(m)$	0.082	0.063	0.063	0.063	0.061	0.061	0.061	0.061
$ML(m-1)$	0.082	0.063	0.063	0.063	0.060	0.060	0.060	0.060
$MML(m)$	0.081	0.061	0.062	0.062	0.062	0.062	0.062	0.062
$MML(m-1$	0.080	0.061	0.061	0.062	0.062	0.062	0.062	0.062
				γ =0.60				
$ML(m)$	0.148	0.077	0.076	0.076	0.062	0.061	0.061	0.061
$ML(m-1)$	0.146	0.074	0.074	0.074	0.061	0.060	0.060	0.060
$MML(m)$	0.148	0.077	0.077	0.077	0.061	0.061	0.061	0.061
$MML(m-1$	0.146	0.075	0.075	0.075	0.061	0.061	0.061	0.061
				γ =0.75				
$ML(m)$	0.207	0.069	0.069	0.069	0.090	0.088	0.088	0.088
$ML(m-1)$	0.207	0.069	0.069	0.069	0.090	0.088	0.088	0.088
$MML(m)$	0.204	0.067	0.065	0.064	0.085	0.088	0.084	0.084
$MML(m-1$	0.203	0.067	0.064	0.063	0.085	0.088	0.084	0.084
				γ =0.90				
$ML(m)$	0.274	0.163	0.138	0.140	0.175	0.161	0.133	133.00
$ML(m-1)$	0.273	0.161	0.136	0.139	0.173	0.159	0.132	0.132
$MML(m)$	0.256	0.148	0.103	0.104	0.149	0.159	0.116	0.115
$MML(m-1$	0.255	0.148	0.102	0.102	0.146	0.157	0.115	0.115

Table 4.7 Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$ with $n = 20$ and 40 for different strategies for maximising the concentrated and marginal likelihood functions

n	20				40			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.061	0.061	0.058	0.057	0.046	0.046	0.046	0.047
$ML(m-1)$	0.052	0.052	0.045	0.044	0.043	0.042	0.042	0.043
$MML(m)$	0.049	0.048	0.047	0.046	0.043	0.044	0.044	0.043
$MML(m-1$	0.045	0.043	0.040	0.039	0.041	0.041	0.041	0.040
				$\gamma = -0.75$				
$ML(m)$	0.070	0.070	0.068	0.065	0.093	0.091	0.091	0.090
$ML(m-1)$	0.063	0.063	0.056	0.053	0.088	0.085	0.085	0.086
$MML(m)$	0.068	0.065	0.064	0.064	0.071	0.072	0.072	0.070
$MML(m-1$	0.059	0.057	0.055	0.055	0.068	0.069	0.067	0.065
				$\gamma = -0.60$				
$ML(m)$	0.081	0.081	0.078	0.080	0.164	0.144	0.149	0.151
$ML(m-1)$	0.075	0.075	0.071	0.073	0.155	0.140	0.140	0.142
$MML(m)$	0.069	0.067	0.066	0.066	0.098	0.098	0.100	0.099
$MML(m-1$	0.067	0.064	0.060	0.060	0.095	0.095	0.097	0.095
				$\gamma = -0.45$				
$ML(m)$	0.123	0.123	0.122	0.125	0.164	0.139	0.133	0.139
$ML(m-1)$	0.120	0.120	0.116	0.119	0.155	0.129	0.128	0.136
$MML(m)$	0.095	0.093	0.094	0.093	0.098	0.086	0.089	0.087
$MML(m-1$	0.089	0.087	0.087	0.086	0.095	0.079	0.081	0.078
				$\gamma = -0.30$				
$ML(m)$	0.135	0.135	0.132	0.143	0.119	0.115	0.115	0.127
$ML(m-1)$	0.120	0.120	0.116	0.126	0.117	0.113	0.113	0.124
$MML(m)$	0.086	0.084	0.085	0.084	0.067	0.067	0.067	0.067
$MML(m-1$	0.077	0.075	0.077	0.075	0.066	0.066	0.066	0.066
				$\gamma = -0.15$				
$ML(m)$	0.138	0.137	0.132	0.146	0.089	0.087	0.087	0.099
$ML(m-1)$	0.125	0.124	0.117	0.133	0.086	0.085	0.085	0.097
$MML(m)$	0.088	0.080	0.086	0.082	0.065	0.064	0.064	0.065
$MML(m-1$	0.083	0.074	0.077	0.074	0.063	0.062	0.062	0.063

Table 4.7 (continued) Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$ with $n = 20$ and 40 for different strategies for maximising the concentrated and marginal likelihood functions

n	20				40			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.138	0.137	0.135	0.160	0.074	0.073	0.073	0.084
$ML(m-1)$	0.129	0.129	0.126	0.146	0.068	0.068	0.068	0.080
$MML(m)$	0.101	0.095	0.096	0.094	0.063	0.062	0.064	0.061
$MML(m-1$	0.092	0.089	0.090	0.088	0.059	0.059	0.061	0.059
				γ =0.15				
$ML(m)$	0.146	0.143	0.141	0.169	0.058	0.058	0.058	0.062
$ML(m-1)$	0.138	0.135	0.127	0.155	0.053	0.053	0.053	0.057
$MML(m)$	0.101	0.095	0.097	0.097	0.050	0.050	0.050	0.050
$MML(m-1$	0.091	0.085	0.083	0.084	0.045	0.045	0.045	0.046
				γ =0.30				
$ML(m)$	0.142	0.133	0.132	0.153	0.068	0.068	0.068	0.073
$ML(m-1)$	0.131	0.122	0.117	0.137	0.065	0.065	0.065	0.070
$MML(m)$	0.112	0.102	0.103	0.102	0.069	0.069	0.069	0.069
$MML(m-1$	0.103	0.092	0.092	0.091	0.068	0.068	0.068	0.067
				γ =0.45				
$ML(m)$	0.134	0.116	0.111	0.130	0.076	0.076	0.075	0.076
$ML(m-1)$	0.124	0.106	0.102	0.118	0.073	0.073	0.072	0.073
$MML(m)$	0.115	0.091	0.087	0.089	0.073	0.073	0.072	0.072
$MML(m-1$	0.105	0.080	0.077	0.078	0.069	0.069	0.068	0.068
				γ =0.60				
$ML(m)$	0.138	0.120	0.113	0.125	0.081	0.081	0.078	0.080
$ML(m-1)$	0.121	0.103	0.095	0.107	0.078	0.078	0.075	0.076
$MML(m)$	0.126	0.102	0.099	0.098	0.085	0.085	0.077	0.077
$MML(m-1$	0.112	0.087	0.082	0.081	0.081	0.081	0.072	0.070
				γ =0.75				
$ML(m)$	0.143	0.135	0.125	0.143	0.136	0.136	0.126	0.124
$ML(m-1)$	0.134	0.124	0.117	0.133	0.132	0.132	0.122	0.122
$MML(m)$	0.136	0.124	0.114	0.116	0.131	0.131	0.118	0.118
$MML(m-1$	0.125	0.114	0.103	0.105	0.126	0.121	0.114	0.113
				γ =0.90				
$ML(m)$	0.094	0.097	0.094	0.104	0.108	0.108	0.103	0.103
$ML(m-1)$	0.088	0.092	0.088	0.098	0.103	0.103	0.097	0.097
$MML(m)$	0.086	0.084	0.083	0.083	0.107	0.107	0.096	0.096
$MML(m-1$	0.079	0.079	0.078	0.078	0.103	0.103	0.091	0.091

Table 4.8 Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$ with $n = 60$ and 80 for different strategies for maximising the concentrated and marginal likelihood functions

n	60				80			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.081	0.078	0.081	0.081	0.105	0.104	0.102	0.103
$ML(m-1)$	0.076	0.077	0.077	0.077	0.104	0.103	0.101	0.101
$MML(m)$	0.077	0.065	0.073	0.073	0.095	0.083	0.077	0.076
$MML(m-1$	0.075	0.063	0.072	0.072	0.092	0.080	0.074	0.072
				$\gamma = -0.75$				
$ML(m)$	0.119	0.150	0.105	0.107	0.251	0.208	0.199	0.209
$ML(m-1)$	0.113	0.145	0.100	0.102	0.247	0.205	0.196	0.206
$MML(m)$	0.102	0.103	0.075	0.074	0.173	0.111	0.101	0.091
$MML(m-1$	0.098	0.100	0.075	0.074	0.169	0.109	0.099	0.089
				$\gamma = -0.60$				
$ML(m)$	0.137	0.173	0.106	0.109	0.220	0.143	0.144	0.152
$ML(m-1)$	0.135	0.170	0.104	0.107	0.217	0.141	0.142	0.150
$MML(m)$	0.103	0.124	0.073	0.072	0.172	0.086	0.088	0.084
$MML(m-1$	0.100	0.123	0.071	0.069	0.168	0.083	0.085	0.081
				$\gamma = -0.45$				
$ML(m)$	0.130	0.127	0.110	0.113	0.133	0.091	0.090	0.095
$ML(m-1)$	0.129	0.121	0.108	0.111	0.132	0.090	0.089	0.094
$MML(m)$	0.100	0.093	0.079	0.078	0.109	0.070	0.070	0.070
$MML(m-1$	0.099	0.091	0.077	0.075	0.109	0.070	0.070	0.068
				$\gamma = -0.30$				
$ML(m)$	0.088	0.090	0.079	0.084	0.068	0.064	0.064	0.065
$ML(m-1)$	0.085	0.088	0.076	0.080	0.067	0.063	0.063	0.064
$MML(m)$	0.071	0.075	0.065	0.065	0.060	0.058	0.059	0.058
$MML(m-1$	0.068	0.075	0.061	0.062	0.058	0.056	0.057	0.056
				$\gamma = -0.15$				
$ML(m)$	0.060	0.066	0.060	0.061	0.068	0.068	0.068	0.068
$ML(m-1)$	0.058	0.066	0.058	0.059	0.067	0.067	0.067	0.067
$MML(m)$	0.054	0.058	0.054	0.054	0.064	0.064	0.064	0.064
$MML(m-1$	0.052	0.056	0.052	0.052	0.062	0.062	0.062	0.060

Table 4.8 (continued) Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$ with $n = 60$ and 80 for different strategies for maximising the concentrated and marginal likelihood functions

n	60				80			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.059	0.064	0.060	0.060	0.047	0.047	0.047	0.047
$ML(m-1)$	0.055	0.064	0.058	0.055	0.046	0.046	0.046	0.046
$MML(m)$	0.055	0.061	0.054	0.055	0.044	0.044	0.044	0.044
$MML(m-1$	0.054	0.060	0.052	0.054	0.043	0.043	0.043	0.043
				γ =0.15				
$ML(m)$	0.065	0.049	0.065	0.065	0.073	0.073	0.073	0.073
$ML(m-1)$	0.063	0.049	0.063	0.063	0.071	0.071	0.071	0.070
$MML(m)$	0.066	0.050	0.066	0.066	0.074	0.074	0.074	0.074
$MML(m-1$	0.063	0.049	0.063	0.063	0.067	0.067	0.067	0.066
				γ =0.30				
$ML(m)$	0.056	0.074	0.054	0.054	0.065	0.063	0.063	0.063
$ML(m-1)$	0.054	0.070	0.051	0.051	0.064	0.062	0.062	0.063
$MML(m)$	0.060	0.074	0.057	0.057	0.068	0.066	0.066	0.066
$MML(m-1$	0.056	0.069	0.053	0.052	0.066	0.064	0.064	0.063
				γ =0.45				
$ML(m)$	0.085	0.074	0.055	0.055	0.075	0.058	0.058	0.058
$ML(m-1)$	0.084	0.074	0.053	0.053	0.072	0.055	0.055	0.053
$MML(m)$	0.086	0.072	0.056	0.056	0.075	0.058	0.058	0.058
$MML(m-1$	0.085	0.070	0.054	0.054	0.071	0.054	0.054	0.052
				γ =0.60				
$ML(m)$	0.145	0.075	0.070	0.071	0.133	0.053	0.052	0.053
$ML(m-1)$	0.143	0.073	0.068	0.069	0.132	0.053	0.052	0.053
$MML(m)$	0.145	0.067	0.068	0.068	0.134	0.055	0.054	0.054
$MML(m-1$	0.145	0.065	0.067	0.066	0.133	0.055	0.054	0.053
				γ =0.75				
$ML(m)$	0.197	0.154	0.095	0.094	0.214	0.075	0.072	0.074
$ML(m-1)$	0.191	0.151	0.091	0.091	0.212	0.074	0.071	0.073
$MML(m)$	0.192	0.127	0.087	0.087	0.212	0.074	0.068	0.069
$MML(m-1$	0.186	0.123	0.082	0.082	0.210	0.070	0.068	0.069
				γ =0.90				
$ML(m)$	0.169	0.064	0.114	0.114	0.209	0.136	0.123	0.124
$ML(m-1)$	0.166	0.057	0.113	0.113	0.207	0.136	0.123	0.124
$MML(m)$	0.161	0.058	0.095	0.095	0.198	0.127	0.103	0.104
$MML(m-1$	0.158	0.056	0.094	0.094	0.196	0.126	0.101	0.102

Table 4.9 Estimated sizes for testing $H_0: \beta_3 = \beta_{30}$, with $n = 100$ and 120 for different strategies for maximising the concentrated and marginal likelihood functions

n	100				120			
Strategies	(i)	(ii)	(iii)	SA	(i)	(ii)	(iii)	(iv)
				$\gamma = -0.90$				
$ML(m)$	0.149	0.146	0.144	0.146	0.232	0.229	0.226	0.229
$ML(m-1)$	0.147	0.143	0.142	0.145	0.227	0.224	0.221	0.223
$MML(m)$	0.125	0.113	0.093	0.090	0.125	0.175	0.119	0.119
$MML(m-1$	0.123	0.111	0.091	0.089	0.122	0.173	0.116	0.116
				$\gamma = -0.75$				
$ML(m)$	0.312	0.250	0.240	0.252	0.271	0.236	0.219	0.230
$ML(m-1)$	0.309	0.248	0.238	0.250	0.270	0.235	0.217	0.228
$MML(m)$	0.223	0.139	0.112	0.105	0.108	0.140	0.107	0.106
$MML(m-1$	0.221	0.138	0.110	0.103	0.107	0.140	0.106	0.105
				$\gamma = -0.60$				
$ML(m)$	0.206	0.119	0.117	0.126	0.155	0.124	0.123	0.126
$ML(m-1)$	0.205	0.118	0.116	0.126	0.152	0.127	0.120	0.123
$MML(m)$	0.151	0.065	0.065	0.066	0.082	0.081	0.081	0.081
$MML(m-1$	0.149	0.065	0.062	0.062	0.082	0.081	0.081	0.081
				$\gamma = -0.45$				
$ML(m)$	0.111	0.071	0.071	0.073	0.086	0.085	0.085	0.086
$ML(m-1)$	0.111	0.071	0.071	0.072	0.086	0.085	0.085	0.086
$MML(m)$	0.094	0.056	0.056	0.056	0.062	0.062	0.062	0.062
$MML(m-1$	0.093	0.055	0.055	0.055	0.062	0.062	0.062	0.062
				$\gamma = -0.30$				
$ML(m)$	0.079	0.073	0.073	0.075	0.077	0.077	0.077	0.077
$ML(m-1)$	0.078	0.071	0.072	0.073	0.075	0.075	0.075	0.075
$MML(m)$	0.070	0.064	0.064	0.064	0.065	0.065	0.065	0.065
$MML(m-1$	0.069	0.063	0.063	0.063	0.064	0.064	0.064	0.064
				$\gamma = -0.15$				
$ML(m)$	0.048	0.048	0.048	0.049	0.060	0.060	0.060	0.060
$ML(m-1)$	0.047	0.047	0.047	0.047	0.057	0.057	0.057	0.057
$MML(m)$	0.046	0.046	0.046	0.046	0.059	0.059	0.059	0.059
$MML(m-1$	0.046	0.046	0.046	0.046	0.056	0.056	0.056	0.056

Table 4.9 (continued) Estimated sizes of testing $H_0: \beta_3 = \beta_{30}$ with $n = 100$ and 120 for different strategies for maximising the concentrated and marginal likelihood functions

n	100				120			
Strategies	(i)	(ii)	(iii)	(iv)	(i)	(ii)	(iii)	(iv)
				γ =0.00				
$ML(m)$	0.049	0.049	0.049	0.049	0.053	0.053	0.053	0.053
$ML(m-1)$	0.049	0.049	0.049	0.049	0.051	0.051	0.051	0.051
$MML(m)$	0.048	0.048	0.048	0.049	0.050	0.050	0.050	0.050
$MML(m-1$	0.047	0.047	0.047	0.047	0.050	0.050	0.050	0.050
				γ =0.15				
$ML(m)$	0.049	0.048	0.049	0.049	0.064	0.064	0.064	0.064
$ML(m-1)$	0.045	0.044	0.045	0.045	0.062	0.062	0.062	0.062
$MML(m)$	0.045	0.045	0.045	0.047	0.061	0.061	0.061	0.061
$MML(m-1$	0.045	0.044	0.045	0.045	0.061	0.061	0.061	0.061
				γ =0.30				
$ML(m)$	0.055	0.055	0.055	0.055	0.059	0.059	0.059	0.059
$ML(m-1)$	0.055	0.055	0.055	0.055	0.058	0.058	0.058	0.058
$MML(m)$	0.054	0.054	0.054	0.054	0.060	0.060	0.060	0.060
$MML(m-1$	0.052	0.052	0.052	0.054	0.056	0.056	0.056	0.056
				γ =0.45				
$ML(m)$	0.078	0.059	0.059	0.061	0.063	0.063	0.063	0.063
$ML(m-1)$	0.078	0.059	0.059	0.059	0.063	0.063	0.063	0.063
$MML(m)$	0.078	0.059	0.059	0.059	0.062	0.062	0.062	0.062
$MML(m-1$	0.078	0.059	0.059	0.059	0.062	0.062	0.062	0.062
				γ =0.60				
$ML(m)$	0.142	0.071	0.071	0.071	0.067	0.066	0.066	0.066
$ML(m-1)$	0.142	0.071	0.071	0.071	0.063	0.062	0.062	0.062
$MML(m)$	0.140	0.069	0.069	0.071	0.065	0.065	0.065	0.065
$MML(m-1$	0.140	0.069	0.069	0.069	0.062	0.062	0.062	0.062
				γ =0.75				
$ML(m)$	0.208	0.070	0.070	0.069	0.091	0.088	0.086	0.086
$ML(m-1)$	0.207	0.068	0.069	0.068	0.091	0.088	0.086	0.086
$MML(m)$	0.207	0.069	0.068	0.068	0.085	0.088	0.083	0.083
$MML(m-1$	0.204	0.066	0.064	0.064	0.083	0.087	0.081	0.081
				γ =0.90				
$ML(m)$	0.268	0.162	0.140	0.142	0.188	0.174	0.148	0.147
$ML(m-1)$	0.266	0.159	0.138	0.141	0.186	0.172	0.145	0.145
$MML(m)$	0.249	0.150	0.105	0.106	0.155	0.167	0.121	0.120
$MML(m-1$	0.248	0.148	0.104	0.105	0.154	0.166	0.120	0.120

CHAPTER 5

Forecasting Accuracy and the Issue of Finding the Global Maximum of the Likelihood Function

5.1 Introduction

Forecasting plays an important role in the field of econometrics, statistics and many other branches of science. One standard approach to producing a good forecast is to construct a reliable statistical or econometric model. This typically involves the estimation of unknown parameters of the model. There is a common view that a good forecast can only be obtained when the parameters of the model are estimated in an optimal or near optimal way. In other words, a good estimate of the parameters of the model will give good forecasts.

In the literature, some of which was reviewed in Chapter 2, various methods have been suggested for estimating the parameters of the model. One disadvantage of maximising the likelihood function is that the numerical optimisation technique sometimes ends up with a local maximum rather than a global maximum. In Chapter 3 we found the somewhat disappointing result that finding the global maximum does not always work in the context of sizes of tests of regression coefficients when estimation is based on the maximum likelihood method. In Chapter 4 we investigated the use of the marginal likelihood and in this case, found the global maximum gives encouraging results. This suggests the need to also look at these issues from the point of view of forecasting accuracy.

The aims of this chapter are manyfold. First to look at the effect of accepting a possible local maxima on forecasting performance when estimation of the moving average parameter γ is based on the ML and MML methods. The second is to look at whether there is any difference in the forecasting performance of the model when the first value of the forecasting error used in the recursive formula for calculating forecasting errors is replaced by its estimated value or by zero as suggested by King and McAleer (1987). We investigate these problems in the context of the linear regression model when the error term follows an MA(1) process with parameter γ.

In order to achieve the first aim, we investigate six different strategies for estimating the unknown parameter. The first involves accepting the maximum that comes from maximizing the likelihood from one fixed starting point. The second and third involve the same approach but taking even greater care to find the global maximum by using an additional three and twenty-one different starting values respectively. The fourth and fifth involve taking the best result from an additional three and twenty-one fixed starting points but only when the initial estimated value of γ is on the boundary point. The sixth strategy is based on estimation of γ by grid search.

We use the concentrated likelihood and the marginal likelihood to estimate the moving average parameter through these different strategies and then assess the resultant forecasting performance in terms of accuracy. We compare the forecasting performance of the model for different values of γ in the case of one-step-ahead forecasts (OSAF) and two-step-ahead forecasts (TSAF) by calculating the average mean square forecasted error (AMSFE), average mean absolute error (AMAE) and average absolute mean forecast error (AAMFE) over 1000 simulations.

A surprise result is that TSAF are often more accurate than OSAF. The difference between the two sets of forecasts is a term involving the product of the estimated value of γ and the previous forecasted error. The latter is calculated recursively, and the first value in these recursive calculations may be important. King and McAleer (1987) suggested using zero as the first value to start the recursive calculations of forecasted error. We propose to replace it by a non-zero value rather than zero in such a way that the chosen value gives the minimum forecast error sum of squares.

The rest of this chapter is organized as follows. In Section 5.2, we discuss the model and issues related to forecasting performance for this model. In Section 5.3, we calculate AMSFE, AMAFE and AAMFE in order to evaluate the forecasting performance for different sample sizes and different values of the parameter γ using Monte Carlo simulation. Section 5.4 contains a discussion of results of the Monte Carlo study. Section 5.5 presents some concluding remarks.

5.2 The Model and Various Issues

5.2.1 The Model

Consider the linear regression model with non-spherical disturbances

$$y = X\beta + u \tag{5.1}$$

where y is an $n \times 1$ vector, X is an $n \times k$ matrix of known values and of full column rank, β is a k dimensional vector of unknown parameters. The elements of u are assumed to follow the MA(1) process

$$u_t = \varepsilon_t + \gamma\varepsilon_{t-1},\ -1 \le \gamma \le 1, \text{ where } \varepsilon_t \sim IIN\left(0, \sigma^2\right) \tag{5.2}$$

which implies that $u \sim N\left(0, \sigma^2\Sigma(\gamma)\right)$ where $\Sigma(\gamma)$ is defined in equation (3.3) in Chapter 3.

5.2.2 Prediction for MA(1) error processes

In this subsection we will introduce forecasting equations based on model (5.1) and (5.2) using estimated parameters of the model for different strategies for finding the global maximum, different likelihood functions and different starting values. For the next observation, model (5.1) can be written as

$$y_{n+1} = x'_{n+1}\beta + u_{n+1} \tag{5.3}$$

where y_{n+1} is the next value of y, x_{n+1} is the $k \times 1$ vector of observations on the regressors at time $n+1$ and u_{n+1} is the associated disturbance term. In the presence of moving average disturbances, the predicted value of y_{n+1} can be written as

$$\hat{y}_{n+1}^{(l)} = x'_{n+1}\hat{\beta}\left(\hat{\gamma}^{(l)}\right) + \hat{\gamma}^{(l)}\hat{\varepsilon}_n^{(l)} \tag{5.4}$$

where $\hat{y}_{n+1}^{(l)}$ is the forecast value of y at time $n+1$, $\hat{\beta}\left(\hat{\gamma}^{(l)}\right)$ is the estimated value of β which is defined by equation (4.3) in Chapter 4, $\hat{\varepsilon}_n^{(l)}$ is the estimated forecasted error component and $\hat{\gamma}^{(l)}$ is the estimated value of the MA(1) parameter. $l = 1, 2$ in which $l = 1$ indicates the estimated value comes from maximising the concentrated likelihood and $l = 2$ indicates the estimated value comes from maximising the marginal likelihood as in Chapter 4. Following King and McAleer (1987), the prediction error $\hat{\varepsilon}_n^{(l)}$ can be obtained via the recursive procedure starting with

$$\hat{\varepsilon}_0^{(l)} = 0$$

and then calculating

$$\hat{y}_t^{(l)} = x'_t\hat{\beta}\left(\hat{\gamma}^{(l)}\right) + \hat{\gamma}^{(l)}\hat{\varepsilon}_{t-1}^{(l)}, \tag{5.5}$$

followed by

$$\hat{\varepsilon}_t^{(l)} = y_t - \hat{y}_t^{(l)} \tag{5.6}$$

for $t = 1, 2, \cdots, n$ and $l = 1, 2$.

In the context of a regression with MA(1) errors, the TSAF and the forecasts of longer lead times are the same. This is due to the fact that the second term in (5.4) no longer appears in the prediction formula. The TSAF of model (5.1) can be expressed as

$$\hat{y}_{n+2}^{(l)} = x'_{n+2}\hat{\beta}\left(\hat{\gamma}^{(l)}\right) \tag{5.7}$$

where $\hat{y}_{n+2}$ is the forecast value of y at time $n+2$ and x_{n+2} is the $k \times 1$ vector of observations on the independent variables at time $n+2$. The estimated $\hat{\beta}\left(\hat{\gamma}^{(l)}\right)$ is the same as in (5.4). Equation (5.7) shows that the two-step-ahead forecast formula does not include the second term of (5.4), which is the product of the estimate of γ and the forecasted error.

In the next subsection we will discuss some related issues of forecasting performance.

5.2.3 Different issues related to the forecasting performance

5.2.3.1 The first issue

As we have observed many times in this thesis, numerical optimisation algorithms can have problems converging to the global maximum and sometimes may end up with an unexpected local maximum instead of the global maximum. In Chapter 3, we discussed the consequences of accepting local maxima in hypothesis testing and introduced different strategies to improve the accuracy of test sizes. In this subsection, we will discuss the same problem with six different strategies for estimating the moving average parameter of the model with the aim of looking at their affect on the forecasting performance. The strategies considered in this chapter are (i) use of one starting value, which is based on OLS to start a standard optimisation routine and accepting the estimates as a final estimate. This strategy is the same as strategy (i) defined in Chapter 3.

One might have some doubt whether the resultant estimates from strategy (i) are those from a global maximum or a local maximum. With that thought in mind, we used an additional two different strategies namely, strategy (ii) and strategy (iii) for finding the global maximum based on multiple starting values of the parameter. Strategy (ii) is based on optimisation repeated with an additional three different starting values and choosing the outcome which provides the largest maximised likelihood function. Strategy (iii) follows the same procedure except those three starting values are replaced by 21 different starting values. Strategy (ii) and strategy (iii) are specially designed to improve the accuracy of the forecasting performance by searching for the global maximum.

The main drawback of estimation of the MA(1) error model is that there is a tendency that the estimated value of γ occurs either at the positive boundary or the negative boundary of the γ space, which we discussed in detail in Chapter 3. For the reasons outlined there, we used an additional two strategies namely strategy (iv) and strategy (v), which use the best results from strategy (ii) and strategy (iii) but only when the final estimate of strategy (i) is at the boundary. In Figure 5.1 and Figure 5.2, the above mentioned strategies are presented in a concise form. We also used grid search which we call strategy (vi) and which we will discuss in the next subsection.

5.2.3.2. Grid search

This section discusses finding the global maximum for estimating the γ parameter by using the grid search technique. We can maximize the likelihood function with respect to different values of γ by a trial and error method. A grid search evaluates the function at grid points that cover the entire range of possible values of γ. This technique inspects the results and repeats the process with a finer grid over a selected zone, which is centred at the γ value for the largest value of the likelihood function...

We applied the grid search approach by evaluating the likelihood function over the parameter range l and u where l is the lower limit and u is the upper limit of the parameter. The steps involved are as follows. Beginning from the lower value l, generate a sequence of grid points such that the difference between any two successive elements of the sequence is equal to ϕ, where ϕ is a small number. Calculate the value of the likelihood function for each corresponding γ value; store all the likelihood function values and select the γ value that gives the largest value of the likelihood function as the estimate. Repeat the process with a finer grid using a smaller increment over a smaller selected region. For example, let $\hat{\gamma}^*$ is the selected value of γ, set two further limits of $\hat{\gamma}^* \mp \delta$ and generate a new sequence of grid points such that the difference between any two successive elements of the sequence is equal to ϕ^* where ϕ^* is a new value such that $\phi^* < \phi$. Using the same procedure discussed above, find the γ value say $\hat{\gamma}^{**}$ that has the largest value of the likelihood. Again set two further limits of $\hat{\gamma}^{**} \mp \delta^*$ and generate a new sequence of grid points such that the difference between any two successive elements of the sequence is equal to ϕ^{**} where ϕ^{**} is a new value such that $\phi^{**} < \phi^*$. Repeat the grid search until changing the δ and ϕ values does not change the largest value of the likelihood within a prescribed level of tolerance. Finally, the grid search converges to the global maximum thus giving the estimated value of the parameter γ.

5.2.3.3 The second issue related to estimation

The joint effect of the estimated parameter and the forecast error has an impact with respect to producing good quality OSAF. On the other hand, TSAF are free from the joint effect of the parameter estimate and forecasted errors.

In Section 5.2.2, we mentioned that (5.4) can provide best forecasts once we know the true forecast error $\varepsilon_n^{(l)}$. Unfortunately, the true forecast errors are unknown and need to be estimated. But problems may occur when estimating $\varepsilon_n^{(l)}$. Normally, we estimate the forecast error starting with an initial value of zero, namely $\hat{\varepsilon}_0 = 0$ and then calculate the rest of the estimated error recursively according to King and McAleer (1987) as in equations (5.5) and (5.6). In order to understand why the value of the estimated forecast error may cause trouble for the forecasting performance note that

$$\begin{aligned}
\hat{\varepsilon}_n^{(l)} &= \left(y_n - x_n{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) - \hat{\gamma}^{(l)}\hat{\varepsilon}_{n-1}^{(l)} \right) \\
&= y_n - x_n{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) - \hat{\gamma}^{(l)}\left(y_{n-1} - x_{n-1}{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) - \hat{\gamma}^{(l)}\hat{\varepsilon}_{n-2}^{(l)} \right) \\
&= y_n - x_n{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) - \hat{\gamma}^{(l)}\left(y_{n-1} - x_{n-1}{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) \right) \\
&\qquad + \hat{\gamma}^{(l)2}\left(y_{n-2} - x_{n-2}{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) - \hat{\gamma}\,\hat{\varepsilon}_{n-3}^{(l)} \right) \\
&= \sum_{i=1}^{n-1}\left(-\hat{\gamma}^{(l)}\right)^i \left(y_{n-i} - x_{n-i}{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) \right) + \left(-\hat{\gamma}^{(l)}\right)^n \hat{\varepsilon}_0^{(l)}.
\end{aligned} \tag{5.8}$$

It is obvious that $\hat{\varepsilon}_n^{(l)}$ depends, among other things, on the first element of the estimated forecasted error $\hat{\varepsilon}_0^{(l)}$. As we discussed in Chapter 3, the estimated value of $\hat{\gamma}^{(l)}$ often occurs close to –1 or 1 in which case the term $\hat{\gamma}^{(l)n}\hat{\varepsilon}_0^{(l)}$ has quite an effect on $\hat{\varepsilon}_n^{(l)}$ and setting $\hat{\varepsilon}_0^{(l)}$ to an arbitrary value such as zero might not be appropriate. To rectify this we suggest estimating the first term of the forecast error by minimising the error sum of squares. We have

$$\begin{aligned}\sum_{t=1}^{n}\hat{\varepsilon}_t^{(l)2} &= \sum_{t=1}^{n}\left(y_t - x_t{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) - \hat{\gamma}^{(l)}\hat{\varepsilon}_{t-1}^{(l)}\right)^2 \\ &= \sum_{t=1}^{n}\left(\sum_{i=1}^{t-1}\left(-\hat{\gamma}^{(l)}\right)^i\left(y_{t-i} - x_{t-i}{}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right)\right) + \left(-\hat{\gamma}^{(l)}\right)^t\hat{\varepsilon}_0^{(l)}\right)^2\end{aligned} \tag{5.9}$$

which can be minimised to obtain the estimated value $\hat{\varepsilon}_0^{(l)}$ by using an optimisation technique. This is then substituted into equation (5.5) and with (5.6) is used to calculate the successive forecast errors. Finally the one-step-ahead forecasts can be obtained as

$$\tilde{y}_{n+1}^{(l)} = x_{n+1}'\hat{\beta}\left(\hat{\gamma}^{(l)}\right) + \hat{\gamma}^{(l)}\hat{\varepsilon}_n^{(l)}. \tag{5.10}$$

Equation (5.10) shows that one step-ahead forecasts depend on the combined effect of $\hat{\varepsilon}_n^{(l)}$ and $\hat{\gamma}^{(l)}$ and equation (5.7), the two step-ahead forecast, does not have the combined effect of $\hat{\varepsilon}_n^{(l)}$ and $\hat{\gamma}^{(l)}$. It indicates that there may be a possibility of better performance from TSAF than OSAF. In the next section, we will investigate the forecasting performance of the model by Monte Carlo methods.

Monte Carlo Experiment

In this section, we outline a Monte Carlo study conducted to assess whether there is any improvement in the forecasting performance using our proposed strategies. We used average mean square forecasting error (AMSFE), average mean absolute forecast errors (AMFE) and average absolute mean forecast error (AAMFE) to assess the forecasting performance.

5.3.1 Design of the simulation experiment

We investigated the forecasting performance for the six different strategies outlined earlier. The additional starting values for strategies (ii) and (iv) where $\gamma = -0.5, 0, 0.5$; while those for strategies (iii) and (v) were $\gamma = \pm.95, \pm.90, \pm.80 \cdots \pm.10, 0$. For this study, the following design matrices were used:

X_1 A constant or intercept plus two white noise regressors generated as $x_{it} = v_{it}$, where $v_{it} \sim IN(0,1)$, $i = 2, 3$.

X_2 A constant or intercept plus two autoregressive regressors generated as

$$x_{it} = 0.8x_{it-1} + v_{it}, \text{ where } v_{it} \sim IN(0,1),\ i = 2, 3. \tag{5.13}$$

X_3 A constant or intercept plus five autoregressive regressors generated from (5.13).

X_4 A constant or intercept plus two random walk regressors generated as

$$x_{it} = x_{it-1} + v_{it}, \text{ where } v_{it} \sim IN(0,1),\ i = 2, 3.$$

X_5 A constant or intercept plus two explosive regressors generated as

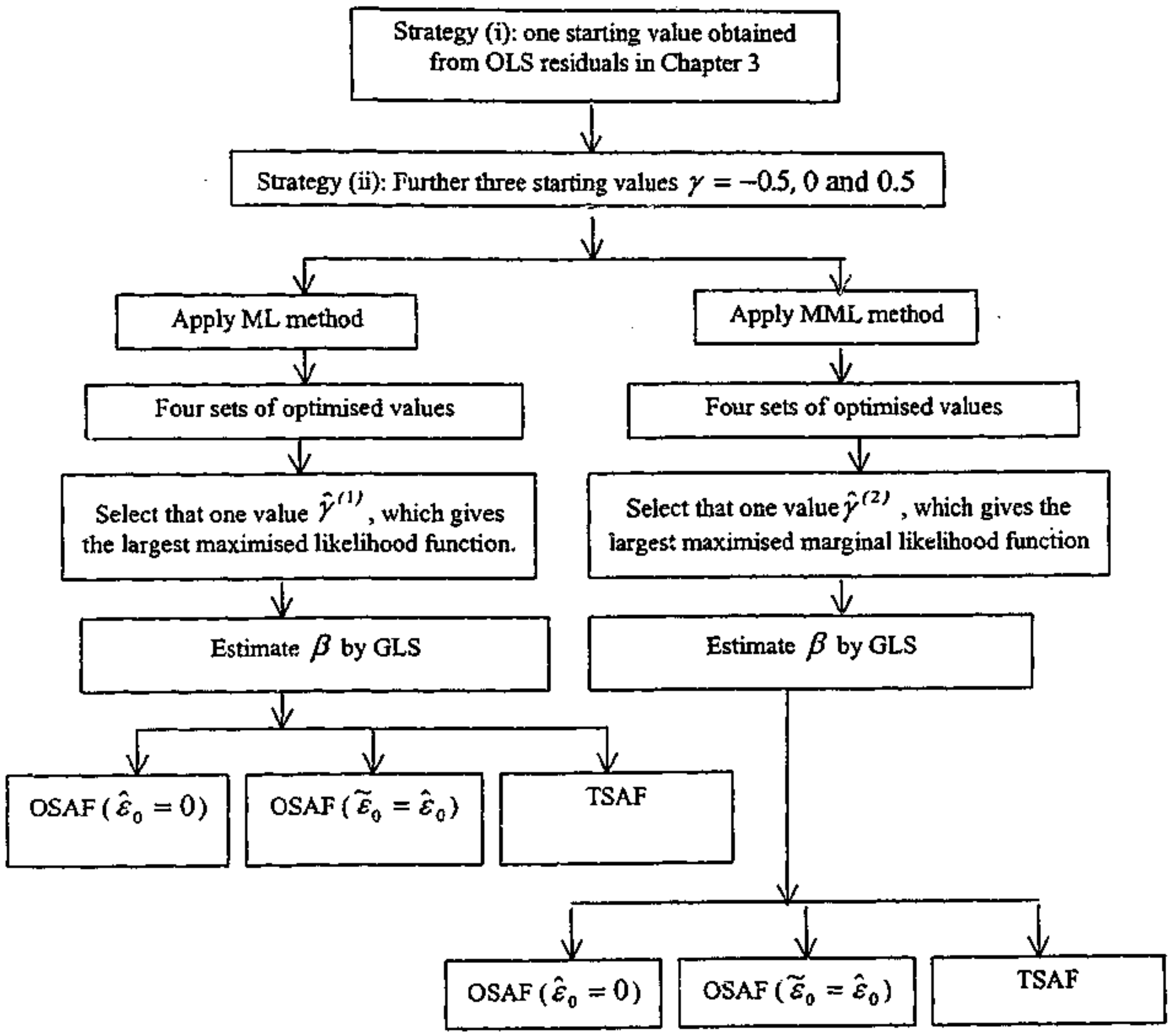
$$x_{it} = 1.02x_{it-1} + v_{it}, \text{ where } v_{it} \sim IN(0,1),\ i = 2, 3.$$

The above design matrices represent a range of different economic time series behaviour. The y's are generated using (5.1) and (5.2) and the values of $\gamma = \pm0.90, \pm0.75, \pm0.60, \pm0.45, \pm0.30, \pm0.15, 0$. After the generation of the true model, the parameters of the models were estimated using ML and MML via the different strategies, then one-step-ahead forecasts and two-step-ahead forecasts were calculated.

We generated 400 data points and threw away 100 data points at the beginning of each sample to avoid the problem of initialization and then divided the remaining sample into two parts of length n_1 and n_2. The former is for estimation and later is a sample for prediction. We used prediction sample size, $n_2 = 100$ and the samples for estimation were taken to be of size $n_1 = 30, 60$ and 120 observations, which meant we discarded the last 170, 140 and 80 observations, respectively, from the generated data.

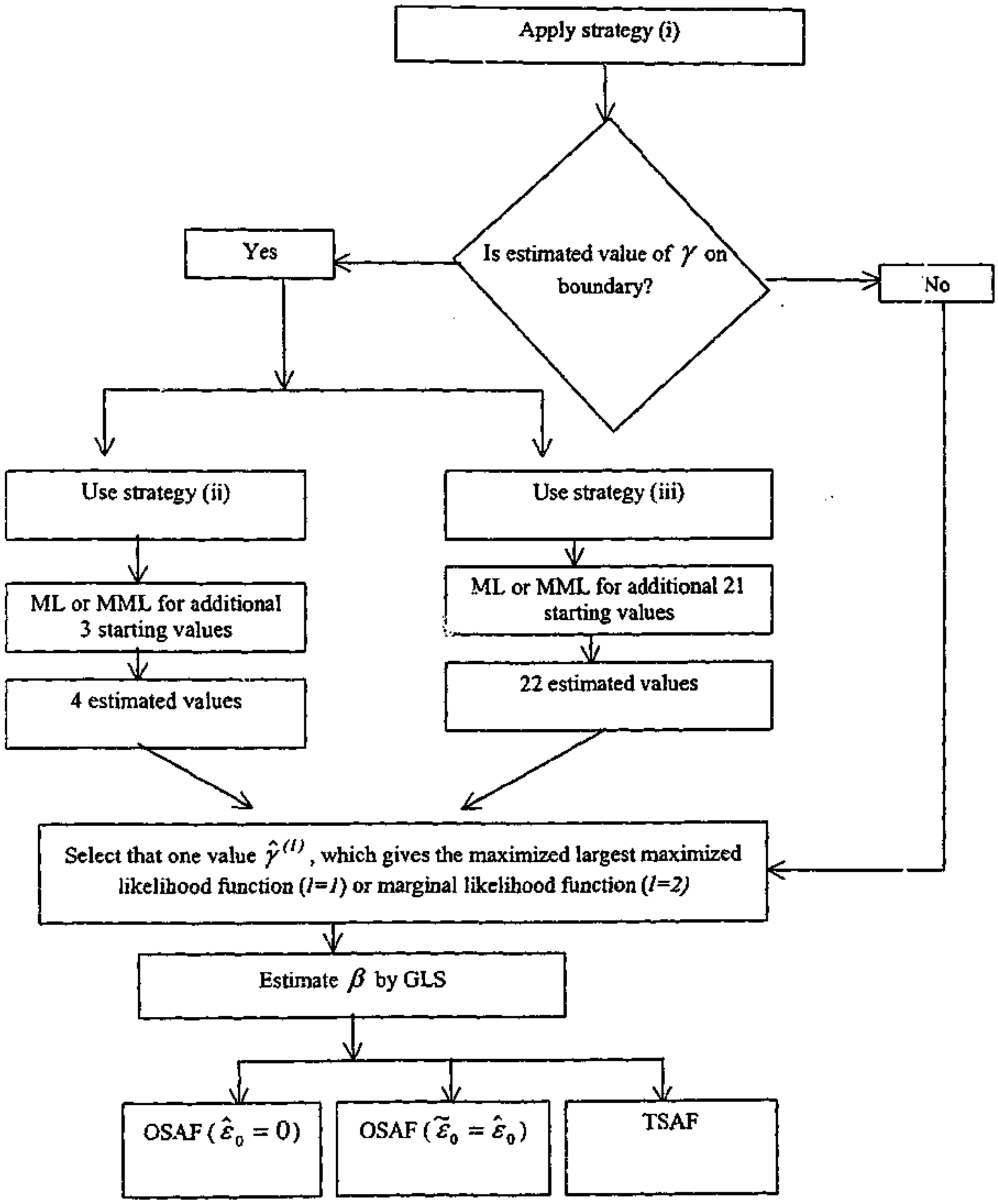
To perform OSAF, we used the final estimated value of $\hat{\gamma}$ which come from different strategies by estimating the regression coefficient vector $\hat{\beta}\left(\hat{\gamma}^{(l)}\right)$ from the first n_1 observations. Keeping these two estimated values, i.e, $\hat{\gamma}$ and $\hat{\beta}\left(\hat{\gamma}^{(l)}\right)$, unchanged, we estimated the forecasted error according to King and McAleer (1987) starting with $\hat{\varepsilon}_0^{(l)} = 0$ by substituting the estimated value of the forecasted error in equation (5.5) and (5.6) respectively. The experiment involves a total of 1000 (N) repetitions or iterations. The ML or MML estimation was carried out by using the Gauss constrained optimization (Aptech system, 1996) with the Newton-Raphson method used to maximise the likelihood. Because of difficulties with computational time, we implemented strategy (vi) that is, grid search for all design matrices in case OSAF and for X_2 only in case of TSAF. This was done to make the project manageable in terms of time.

Figure 5.1 Flow Chart for Finding Estimate Using Strategy (ii) for Forecasting



Note: Strategy (iii) follows the same steps, the only difference is that instead of 3 additional starting values, it uses 21 extra starting values as mentioned in strategy (v).

Figure 5.2 Flow Chart for Strategies (iv) and (v)



5.3.1.1 Working procedure for grid search

We used –0.9999 and 0.9999 for the lower and upper limits of the parameter, between where, the estimated value of the parameter is assumed to lie and the spacing between two initial grid points was set at 0.000999, which gives us 2000 grid points. For each grid point, we calculated the values of the objective functions (i.e., concentrated likelihood and marginal likelihood) and stored their values. The point that gives the maximum value of the objective function is the first selected value of γ, say γ^*. We set two further limits of with $\delta = 0.01$, and calculated 1000 grid points (so the difference between two successive points becomes much smaller than for the previous grid of points) and repeated the process mentioned earlier. Let γ^{**} be the second selected value. For finer searches, we chose $\delta^* = 0.001$ and $\delta^{**} = 0.0001$ to estimate γ^{***} and γ^{****} respectively. We chose the level of accuracy as 0.00001. The process was repeated until the difference between the likelihoods for two successive values for γ met this level of accuracy, i.e., the process continued until $\left|\hat{l}_{n+1} - \hat{l}_n\right| < 0.00001$, where $\hat{l}_{n+1}$ is the likelihood for the $(n+1)$th estimated value of γ, which is our final estimate. Through this very thorough grid search, we are very confident we have found the global maximum of the likelihood function.

5.3.1.2 Prediction accuracy

Our aim was to measure the forecasting accuracy and compare the results, which were obtained using different methods and different strategies for the above model. In order to measure the accuracy of forecasts we used mean square forecast error (MSFE), mean absolute forecast error (MAFE) and absolute mean forecast error (AMFE), these being averaged over the 1000 iterations to produce the average MSFE (AMSFE), average MAFE (AMAFE) and average AMFE (AAMFE) respectively.

5.3.1.3 Average mean square forecast error

The mean square forecast error is the sum of the squared forecast errors for each of the observations divided by the number of forecasts made, $n_2 = 100$, and in the case of OSAF is defined as

$$\mathrm{MSFE} = \frac{1}{n_2}\sum_{i=1}^{n_2}\left(\hat{\varepsilon}_{n+i}^{(l)}\right)^2$$

where $\hat{\varepsilon}_{n+i}^{(l)} = y_{n+i} - \hat{y}_{n+i}^{(l)}$ is the forecast error. Because the experiment was repeated $N = 1000$ times, producing a total of N MSFE's, the average mean square forecast error is calculated as

$$\mathrm{AMSFE} = \frac{1}{N}\sum_{s=1}^{N}\mathrm{MSFE}_s\,.$$

AMSFE is a criterion widely used to evaluate the forecasting performance. The perception is that the lower the value of AMSFE is, the better the forecasting procedure is.

5.3.1.4 Average mean absolute forecast error

The MAFE involves obtaining the average of the mean absolute forecast error defined in the case of OSAF by

$$\text{MAFE} = \frac{1}{n_2}\sum_{l=1}^{n_2}\left|\hat{\varepsilon}_{n+l}^{(l)}\right|$$

There are N=1000 replications so the average mean absolute forecast error is calculated as

$$\text{AMAFE} = \frac{1}{N}\sum_{s=1}^{N}\text{MAFE}_s \,.$$

MAFE ignores the signs of the errors and measures the average absolute size of the forecasting error. It is a measure that is appropriate if the cost of making an error is proportional to the absolute size of the forecast error. The smaller the value of AMAFE, the better the forecasting performances is.

5.3.1.5 Average absolute mean forecast error

The AMFE involves obtaining the average of the absolute mean error defined in the case of OSAF by

$$\text{AMFE} = \frac{1}{n_2}\left|\sum_{l=1}^{n_2}\hat{\varepsilon}_{n+l}^{(l)}\right|.$$

The experiment was repeated N=1000 times, which produces a total of N AMFE's, and average absolute mean forecast error is calculated as

$$\text{AAMFE} = \frac{1}{N}\sum_{s=1}^{N}\text{AMFE}_s$$

Bias in forecasts is revealed by looking at the sum of forecast errors. A value close to zero indicates unbiasedness. AMFE is the average absolute value of this sum. The larger AMFE is, the greater the potential for a problem of biased forecasts from the forecasting model. Similar formulae apply in the case of TSAF.

5.4 Discussion of Results

The results of the six different strategies, two estimation methods namely ML and MML, three different sample sizes and thirteen different values of the parameter γ, are reported in Tables 5.1 to 5.13, for design matrix X_1; Tables 5.14 to 5.26 for design matrix X_2; Tables 5.27 to 5.39 for design matrix X_3 and Tables 5.40 to 5.52 for design matrix X_4. Our main concern is to see the effect of accepting local maxima on forecasting performance and to assess the improvements in terms of accuracy in the context of OSAF and TSAF when estimation comes from different strategies based on ML and MML.

5.4.1 Comparison between ML and MML

An obvious feature of the simulation results is that the predictions based on MML estimates are clearly more accurate on average than those based on ML estimates. In some cases, the difference between the performance of the two approaches is extremely large particularly for non-stationary design matrices, OSAF and small and moderate sample sizes. An extreme example is that of strategy (iv) for X_4 with $n = 30$ and $\gamma = 0.60$ where the AMSFE for ML is 126 times higher than that of MML. On the other hand, for stationary design matrices, differences in performance for small samples are much smaller and for larger sample sizes the differences are very small for both OSAF and TSAF with a few exceptions.

For stationary design matrices, TSAF for ML and MML methods provide more accurate results compared to OSAF with the only exceptions occurring for larger sample sizes and γ values in the interval [0.60, 0.90]. This is true for all strategies except strategy (i). On the other hand, almost equal performances are observed for larger n_1 values, nonstationary regressors (X_3 and X_4) and larger γ values. Overall the comparison between the forecasting performances of the ML and MML estimators reveals that the MML estimator gives a better performance than does the ML estimator keeping other factors unchanged. In the next subsection we will compare the forecasting performances between different strategies.

5.4.2 Comparison between the strategies

The results obtained using the six strategies have differences as well as similarities in their behavior. Thus, we shall make a comparison between different strategies in order to find the best possible strategy for use in forecasting. First of all, we compare the forecasting performance of strategy (i) and strategy (ii). Our results show that overall, the forecasting performance of strategy (ii) is better than that of strategy (i). In particular, we observe that the strategies based on the maximum marginal likelihood estimator show an on average better performance for all strategies than do those using the maximum likelihood. Noticeable improvement is observed for strategy (ii) over strategy (i) in the case of marginal likelihood for small samples. Similar trends are also observed for moderate and large samples where both likelihoods show better performance for strategy (ii) compared to strategy (i).

There are very few improvements that can be seen from strategy (iii) over strategy (ii). The same pattern is observed for strategy (iii) over strategy (iv) and strategy (iv) over strategy (v). None of the strategies are dominating the others. Forecasts for strategy (vi) based on MML give the best performance over all other strategies when used for one-step-ahead forecasts for small sample sizes. The above feature is not true for maximum likelihood based forecasts. The forecasting performances of the other strategies show a significant improvement over that of strategy (i) in the case of two-step-ahead forecasts with a few exceptions.

Overall, we see that estimation of parameter by any strategy incorporated with maximum marginal likelihood estimates gives a better performance than for maximum likelihood estimates. The forecasting performance for strategy (vi) based on both likelihoods is better than for other strategies for small sample sizes. The second best forecasting performance occurs for strategy (v) and the worst performance can be seen for strategy (i).

5.4.3 Comparison between forecasts when initial term of the forecast error is estimated or set to zero

In this subsection, the forecasting performance when the initial forecast error is estimated is compared with that when the initial forecast error is set to zero. The results show that for different strategies, different sample sizes and for different design matrices, estimating the initial forecast error gives a slightly better forecasting performance.

In the next subsection we will compare the forecasting performance for different sample sizes keeping other factors fixed.

5.4.4 Comparison over sample sizes

The results very clearly indicate that the forecasting performance almost always improves as the sample size n_1 increases, the only exception being for $\gamma = -0.90$ and design matrix X_2. As the sample size increases, the forecasting performance based on the method of maximum likelihood gradually improves for all the strategies and gives the same results as for maximum marginal likelihood. Another important feature is that for large samples, the forecasting accuracy from OSAF is approximately the same as for TSAF for different strategies. On the other hand, for some positive values of γ, OSAF are better than TSAF particularly for design matrix X_4 in the case of moderate and large sample sizes. In the next subsection we will look at the forecasting performances for different design matrices holding other factors constant.

5.4.5 Comparison between different design matrices

The simulation results indicate that the forecasting performance depends on the design matrix, particularly for small samples. The best results occur for the design matrices X_1 and X_2, which contain stationary regressors while design matrices X_3 and X_4, which contain nonstationary regressors, yield the worst forecasts. The results also reveal that for design matrices X_3 and X_4 when $n = 30$, the forecasting performance is unacceptably bad particularly for negative γ values. This is also true in the case of X_4 when $n = 60$. We conclude from these results that one cannot forecast with great confidence from the linear regression model with MA(1) errors involving nonstationary regressors and sample sizes less than 100. In the next subsection, we will look at the forecasting performances for different measures of forecasting performance keeping other factors constant.

5.4.6 Different measures of forecasting performance

Now we turn to comparing the forecasting performance for different measures of forecasting accuracy, namely AMSFE, AMAFE and AAMFE. The results show that AMSFE, AMAFE and AAMFE tend to be much larger for negative γ values as compared to positive γ values. The largest values of AMSFE, AMAFE and AAMFE observed for a given design matrix and value of n, typically occur at $\gamma = -0.30$ when $n = 30$, $\gamma = -0.60$ when $n = 60$ and $\gamma = -0.90$ when $n = 120$. Almost without exception, strategy (i), which involves no effort to find the global maximum, has the highest AMSFE, AMAFE and AAMFE. Some exceptions occur for the unreliable cases of X_3 and X_4 and $n = 30$.

Almost always the smallest AMSFE, AMAFE and AAMFE occur for all the strategies based on the MML estimator and the best performance occurs for strategy (vi). Typically, the differences between the AMSFE, AMAFE and

AAMFE of the four strategies other than strategy (i) and (vi) are not great. In the next subsection we will compare the forecasting performances based on one-step-ahead and two-step-ahead forecasts keeping other factors unchanged.

5.4.7 Comparison between OSAF and TSAF

Now, we turn to compare the prediction performance between OSAF and TSAF. From the results, we see that TSAF are clearly more accurate on average than OSAF for different sample sizes, design matrices and values of γ. For design matrix X_4 with $n = 30$ and $\gamma = -0.15$, we see that the AMSFE for OSAF is 497 times higher than that of TSAF in the case of ML estimates. One reason might be the joint effect of the estimated parameter and the forecasted error makes the forecasting performance worse in the case of OSAF. On the other hand, TSAF give a more accurate forecasting performance because they are free from this joint effect. As the sample size increases, the performance of OSAF gradually improves and sometimes they provide an approximately similar performance to that of TSAF. Finally the results of the simulation study shows that one-step-ahead forecasts are not worth using for small samples in econometric data series for both likelihoods. For larger samples, we find that the forecasting performance of OSAF is better than that of TSAF in the case of stationary design matrices. On the other hand, in the case of nonstationary design matrices, TSAF are more accurate compared to OSAF.

5.5 Conclusion

In this chapter, we carefully looked at the estimation of unknown parameters by maximizing a likelihood function and then substituting these estimated parameter values in the model for forecasting. We compared the forecasting performance for two methods of estimation, six different strategies for finding the global maximum, three different sample sizes, thirteen different values of the moving average parameter for one-step-ahead and two-step-ahead forecasts in turn and evaluated the performances of the estimators using AMSFE, AMAFE and AAMFE criteria. We looked at the consequences of accepting local maxima when estimating the unknown parameter and investigated whether taking extra care to find the global maximum of the likelihood function improves the forecasting performance. We also looked at the forecasting performance when the initial term of the forecasting error is estimated (in such a way that the estimated value gives the minimum error sum of squares) and compared it with that when the initial term is set to zero in the recursive calculations.

Our simulation results suggest that forecasting based on the MML estimator is much better than that based on the ML estimator. Furthermore, MML provides the best forecasts for different values of γ, different strategies and different sample sizes in the context of stationary and non-stationary design matrices. Strategy (vi), i.e., grid search, combined with the MML estimation almost always shows the best performance over the other strategies. We have found that the second best forecasting performance result comes from using the MML estimator by further searching for the global maximum whenever the initial estimate of the moving average parameter γ, is on the boundary. Estimating the first term of the forecasted error does not do much to improve forecast accuracy because any improvement is very small.

The main and most significant contribution of this chapter is that TSAF combined with MML estimates give a better performance than OSAF for small and moderate samples for all design matrices. On the other hand for larger sample sizes, OSAF provide better performances in the case of stationary and non-stationary design matrices. We found that overall, as the sample size increases, the forecasting performance for different strategies based on ML and MML estimators gives improved performances in the case of OSAF and TSAF.

In view of the findings of Chapter 3, searching for the global maximum does not always provide the satisfactory estimated sizes when estimation is based on maximum likelihood. As an alternative method, we used the MML method and found it gives more satisfactory estimated sizes in Chapter 4. Finally we can conclude that strategy (vi), i.e. grid search combined with the MML method, gives the best forecasting performance.

Table 5.1 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = -0.90$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	2.6292	2.139	1.6791	1.1425	2.0080	1.6215
			(ii)	2.6464	2.012	1.6934	1.0006	2.0175	1.5172
			(iii)	2.6254	2.012	1.6771	1.0002	2.0066	1.5171
			(iv)	2.6588	2.011	1.7012	0.9992	2.0234	1.5159
			(v)	2.6254	2.012	1.6771	0.9990	2.0066	1.5161
			(vi)	2.5462	1.944	1.1043	0.5547	2.1982	1.5962
Estimated	OSAF	30	(i)	4.4620	2.987	2.6988	1.5315	2.9223	2.0009
			(ii)	4.6677	2.900	3.1327	1.4598	3.2806	1.9641
			(iii)	4.4620	2.900	2.6988	1.4598	2.9223	1.9641
			(iv)	4.6677	2.900	3.1327	1.4598	3.2806	1.9641
			(v)	4.4667	2.902	2.5828	1.3326	2.8650	1.8476
			(vi)	2.5446	1.943	1.0776	0.5378	2.0849	1.5380
Zero	TSAF	30	(i)	1.8705	1.877	0.0346	0.0354	1.0911	1.0931
			(ii)	1.8693	1.879	0.0346	0.0355	1.0907	1.0936
			(iii)	1.8705	1.879	0.0345	0.0355	1.0911	1.0936
			(iv)	1.8691	1.879	0.0346	0.0355	1.0907	1.0936
			(v)	1.8705	1.879	0.0345	0.0355	1.0911	1.0936
Zero	OSAF	60	(i)	1.5951	1.436	1.3498	0.9982	1.6536	1.4163
			(ii)	1.5796	1.354	1.3274	0.8292	1.6374	1.2953
			(iii)	1.5795	1.354	1.3246	0.8292	1.6363	1.2953
			(iv)	1.5635	1.322	1.2990	0.7518	1.6153	1.2455
			(v)	1.5620	1.322	1.2949	0.7508	1.6131	1.2448
			(vi)	1.5183	1.287	0.7682	0.3890	1.6626	1.2545
Estimated	OSAF	60	(i)	1.5942	1.428	1.1861	0.8749	1.5062	1.3053
			(ii)	1.5786	1.343	1.1639	0.7206	1.4913	1.2003
			(iii)	1.5783	1.343	1.1625	0.7206	1.4907	1.2003
			(iv)	1.5622	1.309	1.1403	0.6632	1.4733	1.1646
			(v)	1.5604	1.308	1.3027	0.7427	1.6210	1.2354
			(vi)	1.5194	1.287	0.7387	0.3747	1.5361	1.1952
Zero	TSAF	60	(i)	1.8512	1.852	0.0253	0.0255	1.0863	1.0868
			(ii)	1.8512	1.852	0.0253	0.0254	1.0863	1.0867
			(iii)	1.8512	1.852	0.0253	0.0254	1.0863	1.0867
			(iv)	1.8513	1.851	0.0252	0.0253	1.0864	1.0867
			(v)	1.8513	1.851	0.0252	0.0254	1.0864	1.0867
Zero	OSAF	120	(i)	1.1669	1.131	0.8913	0.6167	1.2496	1.1001
			(ii)	1.1674	1.107	0.8734	0.4764	1.2404	1.0234
			(iii)	1.1654	1.107	0.8637	0.4764	1.2350	1.0234
			(iv)	1.1616	1.096	0.8404	0.3760	1.2215	0.9665
			(v)	1.1596	1.096	0.8307	0.3760	1.2161	0.9665
			(vi)	1.1466	1.066	0.8373	0.3688	1.1711	0.9437
Estimated	OSAF	120	(i)	1.1640	1.116	0.6586	0.4586	1.0837	0.9844
			(ii)	1.1632	1.086	0.6502	0.3509	1.0819	0.9319
			(iii)	1.1613	1.086	0.6396	0.3509	1.0760	0.9319
			(iv)	1.1562	1.071	0.6274	0.2910	1.0698	0.9030
			(v)	1.1541	1.071	0.9333	0.3953	1.2980	0.9831
			(vi)	1.1475	1.067	0.8205	0.3775	1.2737	0.9721
Zero	TSAF	120	(i)	1.8265	1.826	0.0156	0.0157	1.0779	1.0777
			(ii)	1.8265	1.826	0.0155	0.0156	1.0779	1.0777
			(iii)	1.8264	1.826	0.0155	0.0156	1.0778	1.0777
			(iv)	1.8264	1.826	0.0156	0.0156	1.0779	1.0777
			(v)	1.8264	1.826	0.0156	0.0156	1.0778	1.0777

Table 5.2 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = -0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	5.1646	3.2271	2.6138	1.4612	2.9315	1.950
			(ii)	5.1495	2.7537	2.6127	1.1925	2.9306	1.716
			(iii)	5.0913	2.7533	2.5858	1.1924	2.9082	1.716
			(iv)	5.2345	2.7532	2.6625	1.1907	2.9721	1.714
			(v)	5.0913	2.7496	2.5858	1.1869	2.9082	1.711
			(vi)	4.9506	2.6328	1.9905	1.1799	2.0232	1.725
Estimated	OSAF	30	(i)	5.1638	3.2231	2.5915	1.4405	2.8935	1.920
			(ii)	5.1487	2.7492	2.5881	1.1821	2.8888	1.694
			(iii)	5.0905	2.7488	2.5577	1.1809	2.8646	1.694
			(iv)	5.2338	2.7419	2.6323	1.1776	2.9267	1.692
			(v)	5.0906	2.7455	2.6214	1.2024	2.9539	1.725
			(vi)	4.9491	2.6321	0.9813	1.1955	2.9313	1.693
Zero	TSAF	30	(i)	1.6017	1.6022	0.0544	0.0546	1.0106	1.010
			(ii)	1.6012	1.6011	0.0544	0.0547	1.0105	1.010
			(iii)	1.6012	1.6011	0.0544	0.0547	1.0105	1.010
			(iv)	1.6011	1.6012	0.0544	0.0546	1.0104	1.010
			(v)	1.6012	1.6011	0.0544	0.0547	1.0105	1.010
Zero	OSAF	60	(i)	2.5866	1.9387	1.5387	0.9379	1.9452	1.463
			(ii)	2.3183	1.4504	1.3399	0.5666	1.7792	1.159
			(iii)	2.2723	1.4636	1.3016	0.5636	1.7489	1.157
			(iv)	2.2865	1.4088	1.3018	0.5141	1.7497	1.117
			(v)	2.2203	1.4084	1.2543	0.5112	1.7112	1.115
			(vi)	2.1936	1.3732	1.1815	0.5103	1.7034	1.103
Estimated	OSAF	60	(i)	2.5833	1.9333	1.4940	0.9159	1.9045	1.441
			(ii)	2.3137	1.4561	1.3167	0.5575	1.7547	1.147
			(iii)	2.2675	1.4557	1.2809	0.5552	1.7261	1.145
			(iv)	2.2818	1.3995	1.2822	0.4992	1.7271	1.100
			(v)	2.2151	1.3991	1.2968	0.5130	1.7520	1.113
			(vi)	2.1932	1.3731	1.1796	0.5067	1.6957	1.092
Zero	TSAF	60	(i)	1.8512	1.8523	0.0253	0.0255	1.0863	1.086
			(ii)	1.8512	1.8521	0.0253	0.0254	1.0863	1.086
			(iii)	1.8512	1.8521	0.0253	0.0254	1.0863	1.086
			(iv)	1.8513	1.8519	0.0252	0.0253	1.0864	1.086
			(v)	1.8513	1.8519	0.0252	0.0254	1.0864	1.086
Zero	OSAF	120	(i)	1.6303	1.4885	0.7840	0.5993	1.3280	1.193
			(ii)	1.3261	1.0947	0.4915	0.2062	1.0956	0.883
			(iii)	1.3165	1.0947	0.4801	0.2062	1.0875	0.883
			(iv)	1.2597	1.0683	0.4097	0.1636	1.0292	0.852
			(v)	1.2476	1.0683	0.3942	0.1636	1.0182	0.852
			(vi)	1.2305	1.0533	0.3939	0.1568	1.1639	0.851
Estimated	OSAF	120	(i)	1.6233	1.4806	0.7833	0.6147	1.3258	1.202
			(ii)	1.3172	1.0838	0.4693	0.2064	1.0748	0.880
			(iii)	1.3076	1.0838	0.4603	0.2064	1.0680	0.880
			(iv)	1.2492	1.0572	0.3952	0.1636	1.0171	0.848
			(v)	1.2371	1.0572	0.4189	0.1637	1.0384	0.850
			(vi)	1.2307	1.0533	0.3935	0.1554	1.1486	0.843
Zero	TSAF	120	(i)	1.5776	1.5775	0.0299	0.0299	1.0017	1.001
			(ii)	1.5744	1.5732	0.0295	0.0295	1.0007	1.000
			(iii)	1.5743	1.5732	0.0295	0.0295	1.0007	1.000
			(iv)	1.5740	1.5729	0.0295	0.0295	1.0006	1.000
			(v)	1.5738	1.5729	0.0295	0.0295	1.0005	1.000

Table 5.3 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = -0.60$

Initial value	Step	Sample		AMSFE		AAMFE		AMAE	
of $\hat{\varepsilon}_0$		size	Strategy	ML	MML	ML	MML	ML	MML
Zero	OSAF	30	(i)	8.0720	4.0789	3.1852	1.4393	3.5679	1.998
			(ii)	7.9946	2.9397	3.1572	1.0274	3.5399	1.617
			(iii)	7.8946	2.9374	3.1029	1.0271	3.4918	1.616
			(iv)	8.3202	2.9364	3.2921	1.0251	3.6616	1.615
			(v)	7.8946	2.9341	3.1029	1.0249	3.4918	1.614
			(vi)	7.2242	2.8062	1.1003	1.0168	3.3543	1.603
Estimated	OSAF	30	(i)	8.0787	4.0782	3.2271	1.4530	3.5986	2.002
			(ii)	7.9929	2.9373	3.1932	1.0392	3.5662	1.623
			(iii)	7.8929	2.9350	3.1445	1.0386	3.5228	1.622
			(iv)	8.3186	2.9340	3.3372	1.0359	3.6935	1.619
			(v)	7.8929	2.9317	3.2305	1.0472	3.6182	1.635
			(vi)	7.8248	2.8072	3.0915	1.0318	3.6042	1.618
Zero	TSAF	30	(i)	1.4312	1.4210	0.0718	0.0704	0.9556	0.952
			(ii)	1.4306	1.4178	0.0716	0.0703	0.9556	0.951
			(iii)	1.4303	1.4178	0.0717	0.0703	0.9554	0.951
			(iv)	1.4317	1.4178	0.0717	0.0703	0.9561	0.951
			(v)	1.4303	1.4178	0.0717	0.0703	0.9554	0.951
Zero	OSAF	60	(i)	3.4619	2.5379	1.3393	0.8647	1.8918	1.473
			(ii)	2.6186	1.3848	0.9591	0.3396	1.5464	0.999
			(iii)	2.5016	1.3855	0.9012	0.3404	1.4955	1.000
			(iv)	2.7201	1.3209	0.9955	0.3087	1.5812	0.971
			(v)	2.4971	1.3216	0.8958	0.3096	1.4910	0.971
			(vi)	2.6015	1.2966	0.2615	0.0418	1.5850	0.966
Estimated	OSAF	60	(i)	2.5833	1.9333	0.9959	0.9159	1.9045	1.441
			(ii)	2.3137	1.4561	0.9567	0.5575	1.7547	1.147
			(iii)	2.2675	1.4557	0.9557	0.5552	1.7261	1.145
			(iv)	2.2818	1.3995	0.9560	0.4992	1.7271	1.100
			(v)	2.2151	1.3991	0.8668	0.5130	1.7520	1.113
			(vi)	2.2012	1.2967	0.8665	0.4990	1.5560	1.109
Zero	TSAF	60	(i)	1.3958	1.3932	0.0591	0.0589	0.9438	0.943
			(ii)	1.3897	1.3858	0.0592	0.0589	0.9415	0.940
			(iii)	1.3893	1.3858	0.0591	0.0588	0.9414	0.940
			(iv)	1.3904	1.3856	0.0593	0.0589	0.9417	0.940
			(v)	1.3893	1.3857	0.0591	0.0588	0.9414	0.940
Zero	OSAF	120	(i)	1.6437	1.5738	0.4876	0.4321	1.1395	1.096
			(ii)	1.1440	1.0428	0.1864	0.1183	0.8775	0.824
			(iii)	1.1168	1.0428	0.1735	0.1183	0.8669	0.824
			(iv)	1.1298	1.0411	0.1896	0.1182	0.8776	0.823
			(v)	1.0961	1.0411	0.1679	0.1182	0.8608	0.823
			(vi)	1.1187	1.0361	0.1688	0.1062	0.8893	0.821
Estimated	OSAF	120	(i)	1.6384	1.5683	0.5404	0.4863	1.1833	1.142
			(ii)	1.1384	1.0371	0.1953	0.1216	0.8805	0.824
			(iii)	1.1111	1.0371	0.1824	0.1216	0.8697	0.824
			(iv)	1.1236	1.0350	0.1963	0.1195	0.8797	0.823
			(v)	1.0898	1.0350	0.1809	0.1205	0.8703	0.824
			(vi)	1.1192	1.0361	0.1847	0.1215	0.8845	0.820
Zero	TSAF	120	(i)	1.3842	1.3838	0.0434	0.0433	0.9387	0.938
			(ii)	1.3783	1.3777	0.0430	0.0429	0.9366	0.936
			(iii)	1.3782	1.3777	0.0430	0.0429	0.9366	0.936
			(iv)	1.3782	1.3777	0.0430	0.0429	0.9366	0.936
			(v)	1.3781	1.3777	0.0430	0.0429	0.9365	0.936

Table 5.4 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = -0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	2.3916	1.8863	0.7152	0.4776	1.3506	1.1365
			(ii)	1.6581	1.0919	0.4173	0.1699	1.0748	0.8507
			(iii)	1.5959	1.0919	0.4028	0.1699	1.0610	0.8507
			(iv)	1.7712	1.0932	0.4739	0.1703	1.1271	0.8517
			(v)	1.5952	1.0894	0.4025	0.1672	1.0607	0.8488
			(vi)	1.6761	1.0632	0.4130	0.1643	1.0678	0.8376
Estimated	OSAF	30	(i)	2.3895	1.8846	0.7333	0.4830	1.3686	1.1440
			(ii)	1.6561	1.0899	0.4320	0.1686	1.0861	0.8488
			(iii)	1.5936	1.0899	0.4177	0.1686	1.0725	0.8488
			(iv)	1.7692	1.0912	0.4888	0.1685	1.1382	0.8489
			(v)	1.5928	1.0875	0.4218	0.1684	1.0787	0.8510
			(vi)	1.6723	1.0622	0.4220	0.1645	1.0654	0.8322
Zero	TSAF	30	(i)	1.2851	1.2701	0.0964	0.0952	0.9088	0.9040
			(ii)	1.2835	1.2661	0.0959	0.0948	0.9082	0.9023
			(iii)	1.2812	1.2661	0.0959	0.0948	0.9075	0.9023
			(iv)	1.2849	1.2661	0.0962	0.0948	0.9087	0.9023
			(v)	1.2812	1.2661	0.0959	0.0948	0.9075	0.9023
Zero	OSAF	60	(i)	2.3916	1.8863	0.7152	0.4776	1.3506	1.1365
			(ii)	1.6581	1.0919	0.4173	0.1699	1.0748	0.8507
			(iii)	1.5959	1.0919	0.4028	0.1699	1.0610	0.8507
			(iv)	1.7712	1.0932	0.4739	0.1703	1.1271	0.8517
			(v)	1.5952	1.0894	0.4025	0.1672	1.0607	0.8488
			(vi)	1.7217	1.0876	0.0891	0.0220	1.1296	0.8569
Estimated	OSAF	60	(i)	2.3895	1.8846	0.7333	0.4830	1.3686	1.1440
			(ii)	1.6561	1.0899	0.4320	0.1686	1.0861	0.8488
			(iii)	1.5936	1.0899	0.4177	0.1686	1.0725	0.8488
			(iv)	1.7692	1.0912	0.4888	0.1685	1.1382	0.8489
			(v)	1.5928	1.0875	0.4218	0.1718	1.0787	0.8510
			(vi)	1.7223	1.0878	0.0649	0.0164	1.1214	0.8508
Zero	TSAF	60	(i)	1.2303	1.2285	0.0774	0.0772	0.8878	0.8871
			(ii)	1.2256	1.2236	0.0768	0.0765	0.8864	0.8856
			(iii)	1.2254	1.2236	0.0768	0.0765	0.8863	0.8856
			(iv)	1.2261	1.2236	0.0769	0.0766	0.8866	0.8856
			(v)	1.2254	1.2236	0.0768	0.0766	0.8863	0.8856
Zero	OSAF	120	(i)	1.2889	1.2819	0.2350	0.2278	0.9222	0.9188
			(ii)	1.0295	1.0224	0.1211	0.1137	0.8179	0.8143
			(iii)	1.0295	1.0224	0.1211	0.1137	0.8179	0.8143
			(iv)	1.0295	1.0224	0.1211	0.1137	0.8179	0.8143
			(v)	1.0295	1.0224	0.1211	0.1137	0.8179	0.8143
			(vi)	1.0254	1.0190	0.0075	0.0026	0.8188	0.8128
Estimated	OSAF	120	(i)	1.2859	1.2791	0.2488	0.2415	0.9333	0.9298
			(ii)	1.0267	1.0197	0.1211	0.1135	0.8168	0.8132
			(iii)	1.0267	1.0197	0.1211	0.1135	0.8168	0.8132
			(iv)	1.0267	1.0197	0.1211	0.1135	0.8168	0.8132
			(v)	1.0267	1.0197	0.1219	0.1135	0.8174	0.8132
			(vi)	1.0255	1.0190	0.0068	0.0026	0.8181	0.8128
Zero	TSAF	120	(i)	1.2130	1.2129	0.0609	0.0609	0.8801	0.8800
			(ii)	1.2111	1.2109	0.0607	0.0607	0.8794	0.8793
			(iii)	1.2111	1.2109	0.0607	0.0607	0.8794	0.8793
			(iv)	1.2111	1.2109	0.0607	0.0607	0.8794	0.8793
			(v)	1.2111	1.2109	0.0607	0.0607	0.8794	0.8793

Table 5.5 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = -0.30$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	9.1042	3.2263	2.2670	0.6920	2.8470	1.356
			(ii)	8.8764	2.3120	2.2135	0.5010	2.8012	1.161
			(iii)	7.9872	2.2681	2.0183	0.5008	2.6034	1.157
			(iv)	9.9274	2.3099	2.4848	0.4988	3.0576	1.160
			(v)	7.9861	2.2660	2.0183	0.4986	2.6031	1.155
			(vi)	8.6761	2.2332	2.4054	0.4473	2.8712	1.132
Estimated	OSAF	30	(i)	9.1044	3.2148	2.2685	0.6847	2.8580	1.356
			(ii)	8.8708	2.3073	2.2084	0.5028	2.8032	1.165
			(iii)	7.9876	2.2693	2.0132	0.5026	2.6060	1.160
			(iv)	9.9208	2.3052	2.4854	0.5011	3.0655	1.164
			(v)	7.9856	2.2670	2.0197	0.5026	2.6139	1.161
			(vi)	7.5745	2.2112	2.0051	0.4324	2.5121	1.131
Zero	TSAF	30	(i)	1.1613	1.1490	0.1238	0.1234	0.8663	0.862
			(ii)	1.1587	1.1462	0.1235	0.1234	0.8653	0.861
			(iii)	1.1575	1.1462	0.1236	0.1234	0.8649	0.861
			(iv)	1.1617	1.1462	0.1236	0.1234	0.8664	0.861
			(v)	1.1575	1.1461	0.1236	0.1234	0.8649	0.861
Zero	OSAF	60	(i)	1.3658	1.2743	0.2543	0.2158	0.9371	0.902
			(ii)	1.1972	1.0901	0.1968	0.1539	0.8831	0.844
			(iii)	1.1902	1.0901	0.1943	0.1539	0.8807	0.844
			(iv)	1.2717	1.0611	0.2270	0.1441	0.9107	0.835
			(v)	1.1902	1.0611	0.1943	0.1441	0.8807	0.835
			(vi)	1.2567	1.0587	0.0285	0.0021	0.9173	0.834
Estimated	OSAF	60	(i)	1.3654	1.2740	0.2563	0.2143	0.9394	0.903
			(ii)	1.1967	1.0896	0.1990	0.1563	0.8847	0.847
			(iii)	1.1897	1.0896	0.1973	0.1563	0.8828	0.847
			(iv)	1.2712	1.0609	0.2333	0.1441	0.9155	0.836
			(v)	1.1896	1.0609	0.2035	0.1450	0.8874	0.836
			(vi)	1.2575	1.0588	0.0347	0.0014	0.9144	0.834
Zero	TSAF	60	(i)	1.1215	1.1211	0.0933	0.0934	0.8500	0.849
			(ii)	1.1197	1.1195	0.0932	0.0932	0.8492	0.849
			(iii)	1.1197	1.1195	0.0932	0.0932	0.8492	0.849
			(iv)	1.1198	1.1195	0.0932	0.0932	0.8492	0.849
			(v)	1.1197	1.1195	0.0932	0.0932	0.8492	0.849
Zero	OSAF	120	(i)	1.0518	1.0485	0.1211	0.1170	0.8253	0.823
			(ii)	1.0266	1.0233	0.1120	0.1078	0.8166	0.814
			(iii)	1.0266	1.0233	0.1120	0.1078	0.8166	0.814
			(iv)	1.0266	1.0233	0.1120	0.1078	0.8166	0.814
			(v)	1.0266	1.0233	0.1120	0.1078	0.8166	0.814
			(vi)	1.0255	1.0222	0.0033	0.0000	0.8178	0.814
Estimated	OSAF	120	(i)	1.0507	1.0474	0.1234	0.1184	0.8267	0.824
			(ii)	1.0255	1.0222	0.1129	0.1078	0.8167	0.814
			(iii)	1.0255	1.0222	0.1129	0.1079	0.8167	0.814
			(iv)	1.0255	1.0222	0.1129	0.1078	0.8167	0.814
			(v)	1.0255	1.0222	0.1128	0.1079	0.8166	0.814
			(vi)	1.0255	1.0222	0.0033	0.0000	0.8178	0.814
Zero	TSAF	120	(i)	1.1043	1.1043	0.0743	0.0743	0.8432	0.843
			(ii)	1.1041	1.1041	0.0742	0.0742	0.8431	0.843
			(iii)	1.1041	1.1041	0.0742	0.0742	0.8431	0.843
			(iv)	1.1041	1.1041	0.0742	0.0742	0.8431	0.843
			(v)	1.1041	1.1041	0.0742	0.0742	0.8431	0.843

Table 5.6 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = -0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	6.621	2.2666	1.2800	0.4005	1.9446	1.0908
			(ii)	6.797	1.6430	1.3120	0.2802	1.9692	0.9726
			(iii)	6.058	1.5747	1.1799	0.2802	1.8349	0.9662
			(iv)	6.071	1.6058	1.5840	0.2878	2.2322	0.9769
			(v)	6.058	1.5747	1.1799	0.2802	1.8349	0.9662
			(vi)	5.586	1.5705	0.9982	0.2800	1.2345	0.9235
Estimated	OSAF	30	(i)	11.504	1.2166	2.5139	0.1722	3.1561	0.8982
			(ii)	3.488	1.2166	0.7448	0.1722	1.4261	0.8982
			(iii)	3.488	1.2166	0.7448	0.1722	1.4261	0.8982
			(iv)	3.488	1.2166	0.7448	0.1722	1.4261	0.8982
			(v)	3.485	1.2166	0.9269	0.1722	1.5363	0.8982
			(vi)	2.692	1.0021	0.8962	0.1601	1.1542	0.8255
Zero	TSAF	30	(i)	1.096	1.0885	0.1472	0.1471	0.8487	0.8458
			(ii)	1.096	1.0877	0.1472	0.1470	0.8486	0.8456
			(iii)	1.094	1.0877	0.1472	0.1469	0.8481	0.8456
			(iv)	1.099	1.0876	0.1475	0.1468	0.8500	0.8455
			(v)	1.094	1.0877	0.1472	0.1469	0.8481	0.8456
Zero	OSAF	60	(i)	1.202	1.0574	0.1637	0.1382	0.8548	0.8313
			(ii)	1.202	1.0574	0.1637	0.1382	0.8548	0.8313
			(iii)	1.202	1.0574	0.1637	0.1382	0.8548	0.8313
			(iv)	1.212	1.0574	0.1704	0.1382	0.8602	0.8313
			(v)	1.202	1.0574	0.1637	0.1382	0.8548	0.8313
			(vi)	1.197	1.0552	0.0143	0.0062	0.8594	0.8304
Estimated	OSAF	60	(i)	1.202	[illegible]	0.1660	0.1383	0.8566	0.8314
			(ii)	1.202	1.0573	0.1660	0.1383	0.8566	0.8314
			(iii)	1.202	1.0573	0.1660	0.1383	0.8566	0.8314
			(iv)	1.212	1.0573	0.1717	0.1383	0.8613	0.8314
			(v)	1.202	1.0573	0.1678	0.1383	0.8581	0.8314
			(vi)	1.197	1.0552	0.0143	0.0062	0.8594	0.8304
Zero	TSAF	60	(i)	1.054	1.0535	0.1121	0.1120	0.8258	0.8256
			(ii)	1.054	1.0535	0.1121	0.1120	0.8258	0.8256
			(iii)	1.054	1.0535	0.1121	0.1120	0.8258	0.8256
			(iv)	1.054	1.0535	0.1121	0.1120	0.8258	0.8256
			(v)	1.054	1.0535	0.1121	0.1120	0.8258	0.8256
Zero	OSAF	120	(i)	1.014	1.0138	0.1124	0.1106	0.8118	0.8113
			(ii)	1.014	1.0138	0.1124	0.1106	0.8118	0.8113
			(iii)	1.014	1.0138	0.1124	0.1106	0.8118	0.8113
			(iv)	1.014	1.0138	0.1124	0.1106	0.8118	0.8113
			(v)	1.014	1.0138	0.1124	0.1106	0.8118	0.8113
			(vi)	1.014	1.0140	0.0037	0.0036	0.8118	0.8113
Estimated	OSAF	120	(i)	1.014	1.0137	0.1124	0.1107	0.8118	0.8113
			(ii)	1.014	1.0137	0.1124	0.1107	0.8118	0.8113
			(iii)	1.014	1.0137	0.1124	0.1107	0.8118	0.8113
			(iv)	1.014	1.0137	0.1124	0.1107	0.8118	0.8113
			(v)	1.014	1.0137	0.1124	0.1107	0.8118	0.8113
			(vi)	1.014	1.0140	0.0037	0.0036	0.8118	0.8113
Zero	TSAF	120	(i)	1.029	1.0292	0.0931	0.0931	0.8154	0.8154
			(ii)	1.029	1.0292	0.0931	0.0931	0.8154	0.8154
			(iii)	1.029	1.0292	0.0931	0.0931	0.8154	0.8154
			(iv)	1.029	1.0292	0.0931	0.0931	0.8154	0.8154
			(v)	1.029	1.0292	0.0931	0.0931	0.8154	0.8154

Table 5.7 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = 0$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	4.210	1.7543	0.7037	0.2627	1.4000	0.9582
			(ii)	4.445	1.6260	0.7657	0.2363	1.4536	0.9348
			(iii)	4.149	1.6171	0.6961	0.2363	1.3893	0.9330
			(iv)	5.781	1.6970	1.0200	0.2363	1.6978	0.9408
			(v)	4.087	1.5533	0.6961	0.2362	1.3855	0.9290
			(vi)	5.460	1.5467	0.7174	0.2266	1.6659	0.9379
Estimated	OSAF	30	(i)	4.201	1.7490	0.7060	0.2640	1.4048	0.9598
			(ii)	4.449	1.6207	0.7680	0.2372	1.4600	0.9359
			(iii)	4.153	1.6117	0.6974	0.2372	1.3945	0.9342
			(iv)	5.789	1.7406	1.0158	0.2372	1.7009	0.9441
			(v)	4.093	1.5648	0.7032	0.2399	1.3966	0.9332
			(vi)	4.463	1.5505	0.7768	0.2457	1.6593	0.9259
Zero	TSAF	30	(i)	1.055	1.0518	0.1698	0.1699	0.8370	0.8355
			(ii)	1.056	1.0517	0.1699	0.1698	0.8375	0.8355
			(iii)	1.055	1.0516	0.1699	0.1698	0.8370	0.8355
			(iv)	1.058	1.0518	0.1699	0.1698	0.8381	0.8356
			(v)	1.055	1.0516	0.1698	0.1698	0.8370	0.8355
Zero	OSAF	60	(i)	1.053	1.0476	0.1363	0.1332	0.8304	0.8278
			(ii)	1.053	1.0476	0.1363	0.1332	0.8304	0.8278
			(iii)	1.053	1.0476	0.1363	0.1332	0.8304	0.8278
			(iv)	1.084	1.0476	0.1438	0.1332	0.8372	0.8278
			(v)	1.053	1.0476	0.1363	0.1332	0.8304	0.8278
			(vi)	1.078	1.0455	0.1378	0.1329	0.8359	0.8268
Estimated	OSAF	60	(i)	1.053	1.0478	0.1363	0.1333	0.8305	0.8278
			(ii)	1.053	1.0478	0.1363	0.1333	0.8305	0.8278
			(iii)	1.053	1.0478	0.1363	0.1333	0.8305	0.8278
			(iv)	1.085	1.0478	0.1424	0.1333	0.8363	0.8278
			(v)	1.053	1.0478	0.1363	0.1333	0.8305	0.8278
			(vi)	1.078	1.0455	0.1368	0.1328	0.8352	0.8268
Zero	TSAF	60	(i)	1.020	1.0206	0.1311	0.1310	0.8168	0.8168
			(ii)	1.020	1.0206	0.1311	0.1310	0.8168	0.8168
			(iii)	1.020	1.0206	0.1311	0.1310	0.8168	0.8168
			(iv)	1.020	1.0206	0.1311	0.1310	0.8169	0.8168
			(v)	1.020	1.0206	0.1311	0.1310	0.8168	0.8168
Zero	OSAF	120	(i)	1.015	1.0154	0.1098	0.1089	0.8117	0.8114
			(ii)	1.015	1.0154	0.1098	0.1089	0.8117	0.8114
			(iii)	1.015	1.0154	0.1098	0.1089	0.8117	0.8114
			(iv)	1.015	1.0154	0.1098	0.1089	0.8117	0.8114
			(v)	1.015	1.0154	0.1098	0.1089	0.8117	0.8114
			(vi)	1.015	1.0153	0.1090	0.1032	0.8117	0.8113
Estimated	OSAF	120	(i)	1.015	1.0155	0.1098	0.1089	0.8117	0.8114
			(ii)	1.015	1.0155	0.1098	0.1089	0.8117	0.8114
			(iii)	1.015	1.0155	0.1098	0.1089	0.8117	0.8114
			(iv)	1.015	1.0155	0.1098	0.1089	0.8117	0.8114
			(v)	1.015	1.0155	0.1098	0.1089	0.8117	0.8114
			(vi)	1.015	1.0155	0.0033	0.0032	0.8117	0.8113
Zero	TSAF	120	(i)	1.005	1.0059	0.1081	0.1081	0.8072	0.8072
			(ii)	1.005	1.0059	0.1081	0.1081	0.8072	0.8072
			(iii)	1.005	1.0059	0.1081	0.1081	0.8072	0.8072
			(iv)	1.005	1.0059	0.1081	0.1081	0.8072	0.8072
			(v)	1.005	1.0059	0.1081	0.1081	0.8072	0.8072

Table 5.8 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = 0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.4125	2.0162	0.4350	0.1936	1.2086	0.955
			(ii)	3.5423	1.7029	0.5203	0.1967	1.2602	0.935
			(iii)	2.9966	1.5624	0.4351	0.1844	1.1754	0.913
			(iv)	5.0928	1.7533	0.6821	0.1963	1.4408	0.939
			(v)	2.9347	1.4865	0.4350	0.1843	1.1677	0.905
			(vi)	4.8167	1.3910	0.4492	0.0041	1.4158	0.904
Estimated	OSAF	30	(i)	1.1532	1.1342	0.1799	0.1726	0.8849	0.876
			(ii)	1.1532	1.1342	0.1799	0.1726	0.8849	0.876
			(iii)	1.1532	1.1342	0.1799	0.1726	0.8849	0.876
			(iv)	1.1532	1.1342	0.1799	0.1726	0.8849	0.876
			(v)	1.1532	1.1342	0.1799	0.1726	0.8849	0.876
			(vi)	1.0400	1.0365	0.1687	0.1683	0.8243	0.822
Zero	TSAF	30	(i)	1.0731	1.0699	0.1940	0.1936	0.8495	0.848
			(ii)	1.0738	1.0696	0.1942	0.1935	0.8497	0.848
			(iii)	1.0727	1.0696	0.1940	0.1936	0.8494	0.848
			(iv)	1.0752	1.0699	0.1939	0.1934	0.8503	0.848
			(v)	1.0727	1.0696	0.1940	0.1936	0.8494	0.848
Zero	OSAF	60	(i)	1.0408	1.0373	0.1273	0.1261	0.8245	0.823
			(ii)	1.0408	1.0373	0.1273	0.1261	0.8245	0.823
			(iii)	1.0408	1.0373	0.1273	0.1261	0.8245	0.823
			(iv)	1.0408	1.0373	0.1273	0.1261	0.8245	0.823
			(v)	1.0408	1.0373	0.1273	0.1261	0.8245	0.823
			(vi)	1.0400	1.0365	0.1287	0.1187	0.8243	0.822
Estimated	OSAF	60	(i)	1.0409	1.0374	0.1273	0.1261	0.8246	0.823
			(ii)	1.0409	1.0374	0.1273	0.1261	0.8246	0.823
			(iii)	1.0409	1.0374	0.1273	0.1261	0.8246	0.823
			(iv)	1.0409	1.0374	0.1273	0.1261	0.8246	0.823
			(v)	1.0409	1.0374	0.1273	0.1261	0.8246	0.823
			(vi)	1.0400	1.0365	0.1287	0.1287	0.8243	0.822
Zero	TSAF	60	(i)	1.0303	1.0302	0.1435	0.1435	0.8230	0.822
			(ii)	1.0303	1.0302	0.1435	0.1435	0.8230	0.822
			(iii)	1.0303	1.0302	0.1435	0.1435	0.8230	0.822
			(iv)	1.0303	1.0302	0.1435	0.1435	0.8230	0.822
			(v)	1.0303	1.0302	0.1435	0.1435	0.8230	0.822
Zero	OSAF	120	(i)	1.0179	1.0175	0.1088	0.1084	0.8125	0.812
			(ii)	1.0179	1.0175	0.1088	0.1084	0.8125	0.812
			(iii)	1.0179	1.0175	0.1088	0.1084	0.8125	0.812
			(iv)	1.0179	1.0175	0.1088	0.1084	0.8125	0.812
			(v)	1.0179	1.0175	0.1088	0.1084	0.8125	0.812
			(vi)	1.0168	1.0165	0.1016	0.1016	0.8121	0.812
Estimated	OSAF	120	(i)	1.0177	1.0173	0.1087	0.1083	0.8125	0.812
			(ii)	1.0177	1.0173	0.1087	0.1083	0.8125	0.812
			(iii)	1.0177	1.0173	0.1087	0.1083	0.8125	0.812
			(iv)	1.0177	1.0173	0.1087	0.1083	0.8125	0.812
			(v)	1.0177	1.0173	0.1087	0.1083	0.8125	0.812
			(vi)	1.0168	1.0165	0.1016	0.1014	0.8121	0.811
Zero	TSAF	120	(i)	1.0251	1.0251	0.1248	0.1248	0.8171	0.817
			(ii)	1.0251	1.0251	0.1248	0.1248	0.8171	0.817
			(iii)	1.0251	1.0251	0.1248	0.1248	0.8171	0.817
			(iv)	1.0251	1.0251	0.1248	0.1248	0.8171	0.817
			(v)	1.0251	1.0251	0.1248	0.1248	0.8171	0.817

Table 5.9 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = 0.30$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.3196	2.4925	0.2657	0.1838	1.1590	1.037
			(ii)	3.4126	1.7518	0.3561	0.1847	1.1787	0.952
			(iii)	2.5877	1.6864	0.2666	0.1848	1.0699	0.942
			(iv)	4.1609	1.8741	0.4618	0.1845	1.2756	0.963
			(v)	2.4642	1.6582	0.2666	0.1848	1.0566	0.938
			(vi)	4.1304	1.6377	0.4524	0.1758	1.4305	0.935
Estimated	OSAF	30	(i)	3.2689	2.4211	0.2646	0.1835	1.1528	1.031
			(ii)	3.4011	1.7686	0.3555	0.1843	1.1766	0.952
			(iii)	2.5957	1.6966	0.2655	0.1845	1.0673	0.942
			(iv)	4.1683	1.9019	0.4619	0.1842	1.2762	0.964
			(v)	2.5783	1.7542	0.2643	0.1847	1.0603	0.943
			(vi)	2.9096	1.6080	0.2709	0.1744	1.1551	0.936
Zero	TSAF	30	(i)	1.1310	1.1283	0.2169	0.2170	0.8743	0.873
			(ii)	1.1304	1.1271	0.2168	0.2170	0.8742	0.873
			(iii)	1.1297	1.1270	0.2168	0.2170	0.8738	0.873
			(iv)	1.1315	1.1275	0.2170	0.2169	0.8747	0.873
			(v)	1.1297	1.1270	0.2168	0.2170	0.8738	0.873
Zero	OSAF	60	(i)	1.3145	1.2966	0.1315	0.1314	0.8626	0.858
			(ii)	1.0654	1.0474	0.1316	0.1316	0.8306	0.826
			(iii)	1.0654	1.0474	0.1316	0.1316	0.8306	0.826
			(iv)	1.0753	1.0474	0.1317	0.1316	0.8319	0.826
			(v)	1.0517	1.0474	0.1317	0.1316	0.8285	0.826
			(vi)	1.0841	1.0471	0.1356	0.1316	0.8320	0.826
Estimated	OSAF	60	(i)	1.3128	1.2988	0.1314	0.1314	0.8631	0.859
			(ii)	1.0610	1.0469	0.1316	0.1315	0.8301	0.826
			(iii)	1.0610	1.0469	0.1316	0.1315	0.8301	0.826
			(iv)	1.0873	1.0469	0.1316	0.1315	0.8328	0.826
			(v)	1.0513	1.0469	0.1316	0.1315	0.8284	0.826
			(vi)	1.0890	1.0471	0.1324	0.1316	0.8331	0.826
Zero	TSAF	60	(i)	1.0731	1.0699	0.1940	0.1936	0.8495	0.848
			(ii)	1.0738	1.0696	0.1942	0.1935	0.8497	0.848
			(iii)	1.0727	1.0696	0.1940	0.1936	0.8494	0.848
			(iv)	1.0752	1.0699	0.1939	[illegible]	0.8503	0.848
			(v)	1.0727	1.0696	0.1940	0.1936	0.8494	0.848
Zero	OSAF	120	(i)	1.0135	1.0129	0.1090	0.1090	0.8110	0.810
			(ii)	1.0135	1.0129	0.1090	0.1090	0.8110	0.810
			(iii)	1.0135	1.0129	0.1090	0.1090	0.8110	0.810
			(iv)	1.0135	1.0129	0.1090	0.1090	0.8110	0.810
			(v)	1.0135	1.0129	0.1090	0.1090	0.8110	0.810
			(vi)	1.0168	1.0115	0.1116	0.1065	0.8121	0.814
Estimated	OSAF	120	(i)	1.0128	1.0122	0.1090	0.1090	0.8106	0.810
			(ii)	1.0128	1.0122	0.1090	0.1090	0.8106	0.810
			(iii)	1.0128	1.0122	0.1090	0.1090	0.8106	0.810
			(iv)	1.0128	1.0122	0.1090	0.1090	0.8106	0.810
			(v)	1.0128	1.0122	0.1090	0.1090	0.8106	0.810
			(vi)	1.0104	1.0098	0.1060	0.1050	0.8085	0.809
Zero	TSAF	120	(i)	1.1004	1.1003	0.1703	0.1703	0.8543	0.854
			(ii)	1.0991	1.0990	0.1702	0.1702	0.8537	0.853
			(iii)	1.0991	1.0990	0.1702	0.1702	0.8537	0.853
			(iv)	1.0992	1.0990	0.1702	0.1702	0.8538	0.853
			(v)	1.0992	1.0990	0.1702	0.1702	0.8538	0.853

Table 5.10 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = 0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.1521	2.7425	0.1996	0.1703	1.1841	1.107
			(ii)	3.1060	2.0040	0.3140	0.1712	1.1787	0.997
			(iii)	2.3390	1.9124	0.2007	0.1716	1.0620	0.982
			(iv)	3.4171	1.8822	0.3674	0.1711	1.2063	0.982
			(v)	2.0838	1.7156	0.2107	0.1716	1.0224	0.957
			(vi)	2.0547	1.5954	0.2037	0.1512	1.0127	0.941
Estimated	OSAF	30	(i)	3.1933	3.1586	0.1897	0.1899	1.3099	1.300
			(ii)	1.7034	1.6676	0.1914	0.1919	1.0299	1.019
			(iii)	1.7034	1.6676	0.1914	0.1919	1.0299	1.019
			(iv)	1.7034	1.6676	0.1914	0.1919	1.0299	1.019
			(v)	1.6496	1.6028	0.1912	0.1918	1.0191	1.006
			(vi)	1.2151	1.4736	0.1358	0.1759	1.1811	0.969
Zero	TSAF	30	(i)	1.2220	1.2205	0.2413	0.2413	0.9146	0.914
			(ii)	1.2209	1.2180	0.2412	0.2410	0.9145	0.913
			(iii)	1.2196	1.2178	0.2411	0.2411	0.9140	0.913
			(iv)	1.2209	1.2180	0.2413	0.2410	0.9146	0.913
			(v)	1.2195	1.2178	0.2411	0.2411	0.9140	0.913
Zero	OSAF	60	(i)	2.3806	2.3373	0.1344	0.1351	0.9917	0.984
			(ii)	1.1920	1.1814	0.1354	0.1361	0.8503	0.847
			(iii)	1.1921	1.1814	0.1353	0.1361	0.8503	0.847
			(iv)	1.1874	1.1784	0.1354	0.1361	0.8491	0.846
			(v)	1.1874	1.1784	0.1354	0.1361	0.8491	0.846
			(vi)	1.1617	1.1427	0.1101	0.1314	0.8434	0.840
Estimated	OSAF	60	(i)	2.1325	2.0720	0.1347	0.1353	0.9686	0.959
			(ii)	1.1513	1.1402	0.1356	0.1363	0.8464	0.842
			(iii)	1.1513	1.1402	0.1356	0.1363	0.8464	0.842
			(iv)	1.1444	1.1354	0.1356	0.1363	0.8448	0.841
			(v)	1.1591	1.1501	0.1356	0.1363	0.8461	0.843
			(vi)	1.1411	1.1321	0.1201	0.1114	0.8423	0.841
Zero	TSAF	60	(i)	1.2135	1.2133	0.1969	0.1969	0.9002	0.900
			(ii)	1.2105	1.2103	0.1974	0.1974	0.8993	0.899
			(iii)	1.2105	1.2103	0.1974	0.1974	0.8993	0.899
			(iv)	1.2105	1.2103	0.1974	0.1974	0.8993	0.899
			(v)	1.2105	1.2103	0.1974	0.1974	0.8993	0.899
Zero	OSAF	120	(i)	1.4799	1.4789	0.1120	0.1123	0.8676	0.867
			(ii)	1.0312	1.0301	0.1121	0.1123	0.8178	0.817
			(iii)	1.0312	1.0301	0.1121	0.1123	0.8178	0.817
			(iv)	1.0312	1.0301	0.1121	0.1123	0.8178	0.817
			(v)	1.0312	1.0301	0.1121	0.1123	0.8178	0.817
			(vi)	1.0286	1.0276	0.0070	0.0070	0.8168	0.816
Estimated	OSAF	120	(i)	1.5983	1.5972	0.1119	0.1122	0.8748	0.874
			(ii)	1.0278	1.0267	0.1119	0.1122	0.8165	0.816
			(iii)	1.0278	1.0267	0.1119	0.1122	0.8165	0.816
			(iv)	1.0278	1.0267	0.1119	0.1122	0.8165	0.816
			(v)	1.0278	1.0267	0.1119	0.1122	0.8165	0.816
			(vi)	1.0286	1.0276	0.1117	0.1117	0.8168	0.816
Zero	TSAF	120	(i)	1.2050	1.2050	0.1627	0.1627	0.8903	0.890
			(ii)	1.2026	1.2026	0.1627	0.1627	0.8894	0.889
			(iii)	1.2026	1.2026	0.1627	0.1627	0.8894	0.889
			(iv)	1.2026	1.2026	0.1627	0.1627	0.8894	0.889
			(v)	1.2026	1.2026	0.1627	0.1627	0.8894	0.889

Table 5.11 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = 0.60$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.9301	3.4938	0.2048	0.1864	1.3418	1.272
			(ii)	3.6652	2.3960	0.2635	0.1885	1.2052	1.086
			(iii)	2.6892	2.3478	0.2069	0.1886	1.1371	1.077
			(iv)	3.5595	2.1691	0.2896	0.1886	1.1902	1.050
			(v)	2.4267	2.1202	0.2072	0.1889	1.0890	1.041
			(vi)	2.3924	2.1190	0.1974	0.1848	1.0832	1.040
Estimated	OSAF	30	(i)	3.9653	3.5113	0.2032	0.1858	1.3435	1.274
			(ii)	3.6363	2.3497	0.2618	0.1878	1.1992	1.079
			(iii)	2.6640	2.3087	0.2053	0.1879	1.1317	1.071
			(iv)	3.5406	2.1500	0.2884	0.1880	1.1861	1.046
			(v)	2.5105	2.2054	0.2060	0.1884	1.1637	1.055
			(vi)	2.3613	2.0645	0.2043	0.1746	1.1549	1.039
Zero	TSAF	30	(i)	1.3743	1.3732	0.2699	0.2699	0.9730	0.972
			(ii)	1.3738	1.3718	0.2702	0.2697	0.9728	0.972
			(iii)	1.3729	1.3717	0.2697	0.2697	0.9725	0.972
			(iv)	1.3736	1.3717	0.2703	0.2697	0.9729	0.972
			(v)	1.3728	1.3716	0.2697	0.2697	0.9724	0.972
Zero	OSAF	60	(i)	3.0835	2.8978	0.1266	0.1275	1.1644	1.136
			(ii)	1.3979	1.2731	0.1289	0.1300	0.8871	0.868
			(iii)	1.3884	1.2731	0.1289	0.1300	0.8853	0.868
			(iv)	1.2164	1.1507	0.1289	0.1300	0.8650	0.851
			(v)	1.1854	1.1507	0.1289	0.1300	0.8598	0.851
			(vi)	1.1809	1.1245	0.0045	0.0046	0.8580	0.845
Estimated	OSAF	60	(i)	2.8554	2.6799	0.1265	0.1274	1.1411	1.115
			(ii)	1.3615	1.2312	0.1288	0.1299	0.8795	0.861
			(iii)	1.3444	1.2312	0.1288	0.1299	0.8768	0.861
			(iv)	1.1863	1.1220	0.1288	0.1298	0.8592	0.845
			(v)	1.1654	1.1363	0.1287	0.1298	0.8557	0.847
			(vi)	1.1737	1.1158	0.0045	0.0046	0.8573	0.844
Zero	TSAF	60	(i)	1.3472	1.3470	0.2093	0.2092	0.9489	0.948
			(ii)	1.3415	1.3414	0.2087	0.2087	0.9470	0.946
			(iii)	1.3416	1.3414	0.2087	0.2087	0.9470	0.946
			(iv)	1.3415	1.3414	0.2087	0.2087	0.9469	0.946
			(v)	1.3415	1.3414	0.2087	0.2087	0.9469	0.946
Zero	OSAF	120	(i)	3.0068	2.9991	0.1020	0.1024	1.0878	1.085
			(ii)	1.1134	1.0706	0.1034	0.1038	0.8323	0.826
			(iii)	1.1116	1.0706	0.1034	0.1038	0.8317	0.826
			(iv)	1.0851	1.0642	0.1034	0.1039	0.8295	0.824
			(v)	1.0759	1.0642	0.1034	0.1039	0.8272	0.824
			(vi)	1.0643	1.0468	0.0000	0.0001	0.8250	0.820
Estimated	OSAF	120	(i)	2.9319	2.9176	0.1019	0.1023	1.0826	1.079
			(ii)	1.1079	1.0706	0.1032	0.1037	0.8312	0.825
			(iii)	1.1062	1.0706	0.1032	0.1037	0.8306	0.825
			(iv)	1.0874	1.0606	0.1032	0.1037	0.8287	0.822
			(v)	1.0878	1.0725	0.1032	0.1037	0.8275	0.823
			(vi)	1.0606	1.0426	0.0000	0.0001	0.8242	0.819
Zero	TSAF	120	(i)	1.3639	1.3639	0.1664	0.1664	0.9474	0.947
			(ii)	1.3545	1.3545	0.1666	0.1666	0.9442	0.944
			(iii)	1.3545	1.3545	0.1666	0.1666	0.9442	0.944
			(iv)	1.3546	1.3545	0.1666	0.1666	0.9442	0.944
			(v)	1.3545	1.3545	0.1666	0.1666	0.9442	0.944

Table 5.12 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = 0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.3809	3.0919	0.1566	0.1579	1.3553	1.303
			(ii)	2.7806	2.4436	0.1648	0.1602	1.2285	1.162
			(iii)	2.7683	2.4224	0.1653	0.1603	1.2263	1.157
			(iv)	2.5411	2.2226	0.1838	0.1606	1.1533	1.108
			(v)	2.2882	2.1682	0.1661	0.1611	1.1272	1.096
			(vi)	2.7104	2.3992	0.0271	0.0013	1.1903	1.142
Estimated	OSAF	30	(i)	3.3064	3.0882	0.1565	0.1578	1.3415	1.298
			(ii)	2.7244	2.4633	0.1648	0.1601	1.2135	1.157
			(iii)	2.7144	2.4433	0.1653	0.1602	1.2117	1.152
			(iv)	2.4827	2.1759	0.1829	0.1604	1.1330	1.091
			(v)	2.3911	2.2757	0.1666	0.1609	1.1412	1.112
			(vi)	2.5445	2.2347	0.0266	0.0013	1.1571	1.109
Zero	TSAF	30	(i)	1.5740	1.5733	0.2839	0.2839	1.0407	1.040
			(ii)	1.5732	1.5723	0.2840	0.2839	1.0403	1.040
			(iii)	1.5733	1.5724	0.2841	0.2840	1.0404	1.040
			(iv)	1.5734	1.5722	0.2841	0.2839	1.0404	1.040
			(v)	1.5732	1.5722	0.2841	0.2840	1.0404	1.040
Zero	OSAF	60	(i)	2.9530	2.7729	0.1203	0.1214	1.2337	1.198
			(ii)	1.7199	1.5621	0.1221	0.1234	0.9828	0.950
			(iii)	1.7176	1.5468	0.1221	0.1234	0.9820	0.946
			(iv)	1.5050	1.3713	0.1224	0.1236	0.9321	0.906
			(v)	1.4956	1.3408	0.1224	0.1236	0.9296	0.899
			(vi)	1.4338	1.3201	0.0062	0.0062	0.9160	0.891
Estimated	OSAF	60	(i)	2.6240	2.4844	0.1202	0.1213	1.1821	1.153
			(ii)	1.6347	1.4976	0.1220	0.1234	0.9644	0.933
			(iii)	1.6322	1.4818	0.1220	0.1234	0.9636	0.930
			(iv)	1.4216	1.3175	0.1223	0.1235	0.9140	0.892
			(v)	1.4689	1.3415	0.1224	0.1236	0.9213	0.895
			(vi)	1.4172	1.3155	0.1211	0.1209	0.9189	0.886
Zero	TSAF	60	(i)	1.5420	1.5419	0.2184	0.2185	1.0141	1.014
			(ii)	1.5374	1.5373	0.2177	0.2178	1.0124	1.012
			(iii)	1.5374	1.5374	0.2177	0.2177	1.0124	1.012
			(iv)	1.5372	1.5372	0.2177	0.2176	1.0124	1.012
			(v)	1.5373	1.5372	0.2176	0.2176	1.0124	1.012
Zero	OSAF	120	(i)	3.2606	3.1748	0.1006	0.1012	1.2222	1.203
			(ii)	1.3792	1.1537	0.1027	0.1035	0.8923	0.843
			(iii)	1.3536	1.1537	0.1026	0.1035	0.8885	0.843
			(iv)	1.1827	1.0651	0.1029	0.1036	0.8533	0.827
			(v)	1.1571	1.0652	0.1029	0.1036	0.8494	0.827
			(vi)	1.2458	1.0781	0.0066	0.0066	0.8631	0.828
Estimated	OSAF	120	(i)	3.0814	3.0024	0.1005	0.1011	1.1944	1.176
			(ii)	1.3459	1.1499	0.1027	0.1035	0.8850	0.840
			(iii)	1.3294	1.1499	0.1027	0.1035	0.8821	0.840
			(iv)	1.1700	1.0600	0.1030	0.1036	0.8497	0.823
			(v)	1.1340	1.0491	0.1030	0.1036	0.8431	0.822
			(vi)	1.2541	1.0870	0.0066	0.0066	0.8630	0.828
Zero	TSAF	120	(i)	1.5625	1.5624	0.1828	0.1828	1.0140	1.013
			(ii)	1.5530	1.5528	0.1827	0.1826	1.0110	1.010
			(iii)	1.5530	1.5528	0.1827	0.1826	1.0110	1.010
			(iv)	1.5529	1.5527	0.1826	0.1826	1.0110	1.010
			(v)	1.5529	1.5527	0.1826	0.1826	1.0110	1.010

Table 5.13 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = 0.90$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	2.6905	2.5655	0.1640	0.1652	1.2830	1.251
			(ii)	2.5652	2.3856	0.1770	0.1663	1.2530	1.205
			(iii)	2.5538	2.3794	0.1775	0.1666	1.2496	1.203
			(iv)	2.2979	2.1727	0.1776	0.1671	1.1789	1.148
			(v)	2.2791	2.1632	0.1784	0.1676	1.1734	1.145
			(vi)	1.4677	1.4677	0.1429	0.1329	1.8145	0.990
Estimated	OSAF	30	(i)	2.6487	2.6050	0.1491	0.1615	1.2380	1.221
			(ii)	2.5465	1.8165	0.1477	0.1564	1.7359	1.078
			(iii)	2.3365	1.8023	0.1477	0.1643	1.7359	1.073
			(iv)	2.2691	1.8165	0.1493	0.1564	1.2628	1.078
			(v)	2.2381	1.5215	0.1493	0.1645	1.2628	0.985
			(vi)	2.1914	1.4895	0.1332	0.1233	1.1483	0.953
Zero	TSAF	30	(i)	1.8089	1.8093	0.3161	0.3162	1.1171	1.117
			(ii)	1.8091	1.8093	0.3161	0.3162	1.1172	1.117
			(iii)	1.8092	1.8093	0.3161	0.3162	1.1172	1.117
			(iv)	1.8091	1.8093	0.3161	0.3162	1.1172	1.117
			(v)	1.8092	1.8094	0.3161	0.3162	1.1172	1.117
Zero	OSAF	60	(i)	1.7880	1.7881	0.2454	0.2453	1.0948	1.094
			(ii)	1.7875	1.7875	0.2452	0.2452	1.0946	1.094
			(iii)	1.7875	1.7875	0.2452	0.2452	1.0946	1.094
			(iv)	1.7871	1.7871	0.2451	0.2451	1.0945	1.094
			(v)	1.7871	1.7872	0.2451	0.2451	1.0945	1.094
			(vi)	1.6837	1.5749	0.0073	0.0075	1.0158	0.983
Estimated	OSAF	60	(i)	1.6078	1.5727	0.1255	0.1263	1.0029	0.991
			(ii)	1.5015	1.4716	0.1266	0.1273	0.9695	0.960
			(iii)	1.5009	1.4681	0.1266	0.1274	0.9693	0.959
			(iv)	1.3459	1.2976	0.1272	0.1282	0.9220	0.906
			(v)	1.6164	1.5071	0.1272	0.1283	0.9961	0.964
			(vi)	1.4998	1.4302	0.2273	0.2175	0.9694	0.948
Zero	TSAF	60	(i)	1.7882	1.7881	0.2454	0.2453	1.0948	1.094
			(ii)	1.7875	1.7875	0.2452	0.2452	1.0946	1.094
			(iii)	1.7875	1.7875	0.2452	0.2452	1.0946	1.094
			(iv)	1.7871	1.7871	0.2451	0.2451	1.0945	1.094
			(v)	1.7871	1.7872	0.2451	0.2451	1.0945	1.094
Zero	OSAF	120	(i)	1.7836	1.7834	0.2099	0.2099	1.0855	1.085
			(ii)	1.7832	1.7824	0.2097	0.2098	1.0855	1.085
			(iii)	1.7831	1.7824	0.2097	0.2098	1.0855	1.085
			(iv)	1.7824	1.7818	0.2098	0.2099	1.0853	1.085
			(v)	1.7823	1.7818	0.2098	0.2099	1.0853	1.085
			(vi)	1.5235	1.0204	0.2449	0.1563	0.7342	0.641
Estimated	OSAF	120	(i)	1.8180	1.6997	0.1071	0.1078	1.0458	1.012
			(ii)	1.6403	1.4354	0.1077	0.1089	0.9953	0.937
			(iii)	1.6383	1.4354	0.1077	0.1089	0.9946	0.937
			(iv)	1.4330	1.2009	0.1084	0.1095	0.9357	0.870
			(v)	1.7501	1.3200	0.1084	0.1095	1.0224	0.903
			(vi)	1.2914	1.2895	0.1733	0.1833	1.1483	0.953
Zero	TSAF	120	(i)	1.7836	1.7834	0.2099	0.2099	1.0855	1.085
			(ii)	1.7832	1.7824	0.2097	0.2098	1.0855	1.085
			(iii)	1.7831	1.7824	0.2097	0.2098	1.0855	1.085
			(iv)	1.7824	1.7818	0.2098	0.2099	1.0853	1.085
			(v)	1.4998	1.4312	0.1973	0.1175	0.9694	0.948

Table 5 14 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = -0.90$

Initial value	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	2.341	1.9509	1.7385	1.1659	2.0519	1.6419
			(ii)	2.341	1.8682	1.7385	1.0643	2.0519	1.5670
			(iii)	2.341	1.8692	1.7385	1.0631	2.0519	1.5669
			(iv)	2.343	1.8681	1.7415	1.0623	2.0538	1.5657
			(v)	2.341	1.8691	1.7385	1.0611	2.0519	1.5655
			(vi)	2.276	1.8357	1.5118	1.0544	2.0503	1.5515
Estimated	OSAF	30	(i)	2.340	1.9469	1.6780	1.1333	1.9869	1.6022
			(ii)	2.340	1.8634	1.6780	1.0295	1.9869	1.5253
			(iii)	2.340	1.8644	1.6780	1.0291	1.9869	1.5255
			(iv)	2.342	1.8631	1.6811	1.0284	1.9887	1.5247
			(v)	2.340	1.8643	1.8250	1.1050	2.1208	1.6015
			(vi)	2.318	1.1309	1.6674	1.0015	1.1791	1.1788
Zero	TSAF	30	(i)	1.849	1.8510	0.0340	0.0347	1.0845	1.0850
			(ii)	1.849	1.8511	0.0340	0.0348	1.0845	1.0850
			(iii)	1.849	1.8512	0.0340	0.0348	1.0845	1.0851
			(iv)	1.849	1.8511	0.0339	0.0348	1.0845	1.0850
			(v)	1.849	1.8512	0.0340	0.0348	1.0845	1.0851
			(vi)	1.849	1.8511	0.0339	0.0348	1.0845	1.0850
Zero	OSAF	60	(i)	1.786	1.6121	1.4028	1.0777	1.7089	1.4926
			(ii)	1.765	1.4871	1.3762	0.8580	1.6875	1.3349
			(iii)	1.764	1.4869	1.3719	0.8574	1.6849	1.3346
			(iv)	1.754	1.4404	1.3605	0.7737	1.6749	1.2751
			(v)	1.749	1.4380	1.3536	0.7698	1.6705	1.2721
			(vi)	1.576	1.3357	1.2118	0.7544	1.0503	1.2515
Estimated	OSAF	60	(i)	1.785	1.6048	1.2570	0.9657	1.5798	1.3920
			(ii)	1.764	1.4763	1.2356	0.7737	1.5643	1.2577
			(iii)	1.763	1.4761	1.2336	0.7727	1.5629	1.2571
			(iv)	1.753	1.4279	1.2230	0.6975	1.5546	1.2051
			(v)	1.748	1.4256	1.3468	0.7622	1.6743	1.2671
			(vi)	1.698	1.2230	1.3257	0.7421	1.6601	1.1211
Zero	TSAF	60	(i)	1.839	1.8392	0.0250	0.0250	1.0828	1.0829
			(ii)	1.839	1.8391	0.0250	0.0252	1.0828	1.0829
			(iii)	1.839	1.8391	0.0250	0.0251	1.0828	1.0829
			(iv)	1.839	1.8391	0.0250	0.0251	1.0828	1.0829
			(v)	1.839	1.8391	0.0250	0.0251	1.0828	1.0828
			(vi)	1.839	1.8391	0.0250	0.0251	1.0828	1.0829
Zero	OSAF	120	(i)	1.168	1.1323	0.9228	0.6471	1.2696	1.1220
			(ii)	1.167	1.1125	0.9112	0.4925	1.2651	1.0379
			(iii)	1.164	1.1125	0.9008	0.4925	1.2589	1.0379
			(iv)	1.163	1.0985	0.8806	0.3975	1.2491	0.9822
			(v)	1.161	1.0985	0.8702	0.3975	1.2429	0.9822
			(vi)	1.147	1.0696	0.7521	0.2369	1.4417	1.0216
Estimated	OSAF	120	(i)	1.164	1.1190	0.7231	0.5028	1.1246	1.0118
			(ii)	1.163	1.0928	0.7132	0.3831	1.1209	0.9547
			(iii)	1.160	1.0928	0.7050	0.3831	1.1165	0.9547
			(iv)	1.159	1.0741	0.6863	0.3072	1.1074	0.9134
			(v)	1.155	1.0741	0.9949	0.4131	1.3459	0.9946
			(vi)	1.142	1.0615	0.9752	0.4236	1.2359	0.9716
Zero	TSAF	120	(i)	1.821	1.8213	0.0160	0.0159	1.0764	1.0764
			(ii)	1.821	1.8213	0.0160	0.0159	1.0764	1.0764
			(iii)	1.821	1.8213	0.0159	0.0159	1.0764	1.0764
			(iv)	1.821	1.8213	0.0159	0.0159	1.0764	1.0764
			(v)	1.826	1.8260	0.0156	0.0156	1.0778	1.0777
			(vi)	1.821	1.8213	0.0159	0.0159	1.0764	1.0764

Table 5 .15 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = -0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	5.178	3.3466	2.6749	1.5253	3.0270	2.0369
			(ii)	5.181	2.9608	2.6793	1.2994	3.0298	1.8390
			(iii)	5.176	2.9589	2.6745	1.2994	3.0263	1.8384
			(iv)	5.238	2.9608	2.7134	1.2994	3.0574	1.8390
			(v)	5.176	2.9589	2.6745	1.2994	3.0263	1.8384
			(vi)	5.000	2.8685	3.9086	1.3998	3.2352	1.2278
Estimated	OSAF	30	(i)	5.178	3.3435	2.6683	1.5154	3.0055	2.0196
			(ii)	5.181	2.9575	2.6717	1.2941	3.0080	1.8275
			(iii)	5.175	2.9556	2.6677	1.2941	3.0048	1.8269
			(iv)	5.238	2.9575	2.7006	1.2941	3.0318	1.8275
			(v)	5.175	2.9553	2.7710	1.3295	3.1081	1.8673
			(vi)	5.231	2.4856	2.7682	1.1273	3.1697	1.7130
Zero	TSAF	30	(i)	1.568	1.5701	0.0540	0.0544	1.0001	1.0007
			(ii)	1.568	1.5702	0.0540	0.0544	1.0001	1.0008
			(iii)	1.568	1.5702	0.0540	0.0544	1.0001	1.0008
			(iv)	1.568	1.5702	0.0540	0.0544	1.0001	1.0008
			(v)	1.568	1.5702	0.0540	0.0544	1.0001	1.0008
			(vi)	1.568	1.5702	0.0540	0.0544	1.0010	1.0008
Zero	OSAF	60	(i)	3.095	2.2933	1.6954	1.0703	2.1206	1.6004
			(ii)	2.850	1.6419	1.5459	0.6326	1.9866	1.2259
			(iii)	2.827	1.6419	1.5306	0.6326	1.9736	1.2259
			(iv)	2.777	1.5767	1.5021	0.5632	1.9465	1.1684
			(v)	2.716	1.5535	1.4649	0.5532	1.9162	1.1591
			(vi)	2.212	1.3432	1.9086	0.5998	1.2120	1.1441
Estimated	OSAF	60	(i)	3.091	2.2888	1.6995	1.0746	2.1058	1.5947
			(ii)	2.845	1.6337	1.5460	0.6274	1.9710	1.2148
			(iii)	2.822	1.6337	1.5306	0.6274	1.9582	1.2148
			(iv)	2.771	1.5677	1.4926	0.5554	1.9263	1.1552
			(v)	2.710	1.5445	1.5166	0.5540	1.9604	1.1573
			(vi)	2.669	1.5320	1.5373	0.5977	1.4046	1.1041
Zero	TSAF	60	(i)	1.563	1.5631	0.0402	0.0403	0.9987	0.9986
			(ii)	1.563	1.5623	0.0401	0.0401	0.9985	0.9982
			(iii)	1.562	1.5623	0.0401	0.0401	0.9985	0.9982
			(iv)	1.562	1.5622	0.0401	0.0401	0.9985	0.9982
			(v)	1.562	1.5622	0.0401	0.0401	0.9984	0.9982
			(vi)	1.562	1.5622	0.0401	0.0401	0.9984	0.9982
Zero	OSAF	120	(i)	1.692	1.5336	0.8702	0.6623	1.3969	1.2441
			(ii)	1.355	1.1055	0.5366	0.1971	1.1283	0.8766
			(iii)	1.340	1.1055	0.5159	0.1971	1.1129	0.8766
			(iv)	1.284	1.0725	0.4341	0.1658	1.0477	0.8515
			(v)	1.269	1.0725	0.4135	0.1658	1.0323	0.8515
			(vi)	1.254	1.0573	0.2217	0.1338	1.0941	0.8524
Estimated	OSAF	120	(i)	1.685	1.5255	0.8739	0.6726	1.3916	1.2427
			(ii)	1.345	1.0949	0.5235	0.1917	1.1141	0.8692
			(iii)	1.330	1.0949	0.5045	0.1917	1.0993	0.8691
			(iv)	1.274	1.0614	0.4208	0.1582	1.0352	0.8432
			(v)	1.259	1.0614	0.4483	0.1677	1.0614	0.8501
			(vi)	1.254	1.0576	0.4563	0.1628	1.0674	0.8464
Zero	TSAF	120	(i)	1.564	1.5648	0.0303	0.0303	0.9973	0.9972
			(ii)	1.564	1.5642	0.0300	0.0299	0.9971	0.9970
			(iii)	1.564	1.5642	0.0300	0.0299	0.9971	0.9970
			(iv)	1.564	1.5642	0.0299	0.0299	0.9971	0.9970
			(v)	1.564	1.5642	0.0299	0.0299	0.9971	0.9970
			(vi)	1.564	1.5642	0.0299	0.0299	0.9971	0.9970

Table 5.16 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = -0.60$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	9.0597	4.4189	3.4741	1.5848	3.9061	2.167
			(ii)	9.0014	3.7492	3.4585	1.3076	3.8908	1.911
			(iii)	8.9511	3.7487	3.4332	1.3075	3.8676	1.911
			(iv)	9.2613	3.7497	3.5596	1.3083	3.9832	1.911
			(v)	8.9511	3.7487	3.4332	1.3075	3.8676	1.911
			(vi)	8.7463	3.5901	3.1278	1.9186	3.6890	1.797
Estimated	OSAF	30	(i)	9.0592	4.4173	3.5002	1.5928	3.9193	2.172
			(ii)	9.0009	3.7474	3.4826	1.3165	3.9033	1.920
			(iii)	8.9505	3.7468	3.4607	1.3164	3.8822	1.919
			(iv)	8.8609	3.7478	3.5787	1.3173	3.7925	1.920
			(v)	8.8492	3.7464	3.5920	1.3418	3.7028	1.944
			(vi)	8.8405	3.7417	3.5586	1.2108	3.6131	1.812
Zero	TSAF	30	(i)	1.3805	1.3806	0.0719	0.0713	0.9398	0.939
			(ii)	1.3805	1.3806	0.0718	0.0712	0.9398	0.939
			(iii)	1.3805	1.3806	0.0718	0.0712	0.9398	0.939
			(iv)	1.3805	1.3806	0.0719	0.0712	0.9398	0.939
			(v)	1.3805	1.3806	0.0718	0.0712	0.9398	0.939
			(vi)	1.3805	1.3806	0.0719	0.0712	0.9398	0.939
Zero	OSAF	60	(i)	4.3370	3.0484	1.6641	1.0162	2.2007	1.624
			(ii)	3.4595	1.4189	1.2544	0.3313	1.8256	0.993
			(iii)	3.3281	1.4189	1.1977	0.3313	1.7734	0.993
			(iv)	3.4463	1.3970	1.2542	0.3237	1.8244	0.986
			(v)	3.2770	1.3963	1.1763	0.3236	1.7530	0.986
			(vi)	4.2772	2.3223	1.7783	0.9346	2.1466	1.657
Estimated	OSAF	60	(i)	4.3329	3.0438	1.7119	1.0447	2.2390	1.647
			(ii)	3.4549	1.4138	1.2853	0.3313	1.8429	0.988
			(iii)	3.3234	1.4138	1.2241	0.3313	1.7884	0.988
			(iv)	3.4418	1.3931	1.2890	0.3255	1.8451	0.982
			(v)	3.2721	1.3917	1.2436	0.3293	1.8080	0.987
			(vi)	4.2872	2.3773	1.7463	1.3363	2.0466	1.667
Zero	TSAF	60	(i)	1.3716	1.3715	0.0603	0.0601	0.9358	0.935
			(ii)	1.3703	1.3697	0.0603	0.0599	0.9353	0.935
			(iii)	1.3702	1.3697	0.0603	0.0599	0.9353	0.935
			(iv)	1.3702	1.3696	0.0604	0.0599	0.9353	0.935
			(v)	1.3702	1.3696	0.0603	0.0599	0.9353	0.935
			(vi)	1.3702	1.3696	0.0604	0.0599	0.9353	0.935
Zero	OSAF	120	(i)	1.7506	1.6681	0.5666	0.4928	1.2105	1.153
			(ii)	1.1993	1.0572	0.2100	0.1225	0.8994	0.828
			(iii)	1.1638	1.0572	0.2007	0.1225	0.8906	0.828
			(iv)	1.1823	1.0543	0.2007	0.1195	0.8904	0.826
			(v)	1.1447	1.0543	0.1906	0.1195	0.8808	0.826
			(vi)	1.1377	1.0492	0.1490	0.1117	0.9025	0.825
Estimated	OSAF	120	(i)	1.7453	1.6629	0.6063	0.5380	1.2411	1.189
			(ii)	1.1937	1.0516	0.2203	0.1256	0.9030	0.829
			(iii)	1.1578	1.0516	0.2099	0.1256	0.8935	0.829
			(iv)	1.1761	1.0486	0.2089	0.1228	0.8932	0.827
			(v)	1.1381	1.0486	0.2099	0.1233	0.8935	0.827
			(vi)	1.1238	1.0491	0.1451	0.1201	0.8952	0.825
Zero	TSAF	120	(i)	1.3695	1.3695	0.0443	0.0442	0.9340	0.934
			(ii)	1.3686	1.3686	0.0440	0.0439	0.9336	0.933
			(iii)	1.3686	1.3686	0.0440	0.0439	0.9336	0.933
			(iv)	1.3686	1.3686	0.0440	0.0439	0.9336	0.933
			(v)	1.3686	1.3686	0.0440	0.0439	0.9336	0.933
			(vi)	1.3686	1.3686	0.0440	0.0439	0.9336	0.933

Table 5.17 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = -0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	14.3846	5.698	3.9956	1.4263	4.5606	2.0938
			(ii)	14.5728	4.446	4.0396	1.0946	4.6065	1.7635
			(iii)	14.1118	4.446	3.9393	1.0946	4.5046	1.7635
			(iv)	15.1657	4.467	4.2441	1.1007	4.8040	1.7694
			(v)	14.1118	4.446	3.9393	1.0946	4.5046	1.7635
			(vi)	14.2627	4.227	3.2113	0.8204	3.2445	1.1169
Estimated	OSAF	30	(i)	14.3838	5.697	4.0489	1.4440	4.6193	2.1103
			(ii)	14.5717	4.446	4.1002	1.1122	4.6676	1.7774
			(iii)	14.1109	4.445	3.9969	1.1121	4.5645	1.7773
			(iv)	15.1647	4.466	4.3093	1.1170	4.8679	1.7825
			(v)	14.1089	4.445	4.1139	1.1417	4.6517	1.8022
			(vi)	9.2009	1.640	3.1753	0.8104	3.6492	0.8169
Zero	TSAF	30	(i)	1.2402	1.237	0.0962	0.0947	0.8936	0.8925
			(ii)	1.2402	1.237	0.0961	0.0945	0.8937	0.8924
			(iii)	1.2400	1.237	0.0961	0.0945	0.8936	0.8924
			(iv)	1.2405	1.237	0.0963	0.0945	0.8937	0.8924
			(v)	1.2400	1.237	0.0961	0.0945	0.8936	0.8924
			(vi)	1.2405	1.237	0.0963	0.0945	0.8937	0.8924
Zero	OSAF	60	(i)	2.8130	1.919	0.8549	0.4799	1.4909	1.1425
			(ii)	2.0589	1.155	0.5708	0.1920	1.2240	0.8738
			(iii)	2.0402	1.155	0.5537	0.1920	1.2103	0.8738
			(iv)	2.1981	1.144	0.6135	0.1868	1.2659	0.8694
			(v)	2.0249	1.143	0.5468	0.1865	1.2046	0.8691
			(vi)	2.6270	1.336	0.6995	0.1892	1.3401	0.8715
Estimated	OSAF	60	(i)	2.8112	1.918	0.8801	0.4816	1.5141	1.1492
			(ii)	2.0567	1.153	0.5932	0.1895	1.2406	0.8721
			(iii)	2.0381	1.153	0.5765	0.1895	1.2266	0.8721
			(iv)	2.1959	1.142	0.6390	0.1842	1.2866	0.8677
			(v)	2.0230	1.142	0.5655	0.1849	1.2189	0.8679
			(vi)	2.0000	1.312	0.4525	0.1712	1.2412	0.8256
Zero	TSAF	60	(i)	1.2107	1.210	0.0785	0.0783	0.8805	0.8805
			(ii)	1.2098	1.209	0.0782	0.0781	0.8803	0.8802
			(iii)	1.2098	1.209	0.0781	0.0781	0.8803	0.8802
			(iv)	1.2099	1.209	0.0782	0.0781	0.8803	0.8802
			(v)	1.2098	1.209	0.0781	0.0781	0.8803	0.8802
			(vi)	1.2099	1.209	0.0782	0.0781	0.8803	0.8802
Zero	OSAF	120	(i)	1.3472	1.340	0.2631	0.2532	0.9459	0.9416
			(ii)	1.0364	1.025	0.1294	0.1168	0.8229	0.8162
			(iii)	1.0364	1.025	0.1294	0.1168	0.8229	0.8162
			(iv)	1.0416	1.025	0.1327	0.1168	0.8257	0.8162
			(v)	1.0364	1.025	0.1294	0.1168	0.8229	0.8162
			(vi)	1.0377	1.022	0.1022	0.1002	0.8236	0.8149
Estimated	OSAF	120	(i)	1.3441	1.337	0.2901	0.2806	0.9671	0.9629
			(ii)	1.0335	1.022	0.1290	0.1167	0.8215	0.8152
			(iii)	1.0335	1.022	0.1290	0.1167	0.8215	0.8152
			(iv)	1.0388	1.022	0.1335	0.1167	0.8255	0.8152
			(v)	1.0335	1.022	0.1282	0.1167	0.8210	0.8152
			(vi)	1.0324	1.012	0.1278	0.1159	0.8210	0.8004
Zero	TSAF	120	(i)	1.2041	1.204	0.0624	0.0624	0.8771	0.8771
			(ii)	1.2039	1.203	0.0622	0.0622	0.8770	0.8770
			(iii)	1.2039	1.203	0.0622	0.0622	0.8770	0.8770
			(iv)	1.2039	1.203	0.0622	0.0622	0.8770	0.8770
			(v)	1.2039	1.203	0.0622	0.0622	0.8770	0.8770
			(vi)	1.2039	1.203	0.0622	0.0622	0.8770	0.8770

Table 5.18 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = -0.30$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	16.038	4.3591	3.7483	0.8306	4.3611	1.5383
			(ii)	15.754	2.9796	3.6976	0.6299	4.3132	1.3223
			(iii)	15.596	2.9794	3.5494	0.6298	4.1566	1.3223
			(iv)	15.439	2.9796	3.5412	0.6299	4.1499	1.3223
			(v)	15.096	2.9794	3.5494	0.6298	4.1486	1.3223
			(vi)	17.754	2.8551	3.5436	0.3489	4.1433	1.3223
Estimated	OSAF	30	(i)	16.032	4.3506	3.7557	0.8278	4.3767	1.5395
			(ii)	15.753	2.9744	3.7080	0.6337	4.3317	1.3282
			(iii)	15.096	2.9742	3.5601	0.6336	4.1745	1.3282
			(iv)	15.239	2.9744	3.4122	0.6337	4.1921	1.3282
			(v)	15.094	2.9854	3.6101	0.6358	4.2142	1.3302
			(vi)	17.141	2.8551	3.8622	0.5489	4.0131	1.2201
Zero	TSAF	30	(i)	1.132	1.1309	0.1241	0.1237	0.8569	0.8558
			(ii)	1.131	1.1306	0.1236	0.1236	0.8565	0.8557
			(iii)	1.132	1.1306	0.1240	0.1236	0.8566	0.8557
			(iv)	1.132	1.1306	0.1236	0.1236	0.8569	0.8557
			(v)	1.132	1.1306	0.1240	0.1236	0.8566	0.8557
			(vi)	1.132	1.1306	0.1236	0.1236	0.8569	0.8557
Zero	OSAF	60	(i)	1.954	1.4546	0.4089	0.2531	1.0900	0.9422
			(ii)	1.598	1.1015	0.3368	0.1600	1.0137	0.8500
			(iii)	1.538	1.1015	0.3075	0.1600	0.9881	0.8500
			(iv)	1.857	1.0759	0.4130	0.1537	1.0878	0.8441
			(v)	1.499	1.0759	0.2992	0.1538	0.9791	0.8441
			(vi)	2.466	1.0212	0.4259	0.1345	1.1776	0.8025
Estimated	OSAF	60	(i)	1.954	1.4543	0.4074	0.2533	1.0899	0.9414
			(ii)	1.598	1.1012	0.3356	0.1594	1.0151	0.8495
			(iii)	1.538	1.1012	0.3084	0.1594	0.9905	0.8496
			(iv)	1.857	1.0758	0.4205	0.1538	1.0938	0.8441
			(v)	1.499	1.0754	0.3090	0.1542	0.9888	0.8445
			(vi)	2.387	1.0015	0.4542	0.1215	1.1219	0.8332
Zero	TSAF	60	(i)	1.112	1.1121	0.0975	0.0974	0.8468	0.8467
			(ii)	1.112	1.1119	0.0975	0.0974	0.8466	0.8465
			(iii)	1.112	1.1119	0.0975	0.0974	0.8466	0.8465
			(iv)	1.112	1.1119	0.0976	0.0974	0.8467	0.8465
			(v)	1.112	1.1119	0.0975	0.0974	0.8466	0.8465
			(vi)	1.112	1.1119	0.0975	0.0974	0.8466	0.8465
Zero	OSAF	120	(i)	1.067	1.0634	0.1346	0.1289	0.8368	0.8344
			(ii)	1.028	1.0242	0.1142	0.1085	0.8176	0.8152
			(iii)	1.028	1.0242	0.1142	0.1085	0.8176	0.8152
			(iv)	1.028	1.0242	0.1142	0.1085	0.8176	0.8152
			(v)	1.028	1.0242	0.1142	0.1085	0.8176	0.8152
			(vi)	1.028	1.0131	0.1040	0.1013	0.8091	0.8037
Estimated	OSAF	120	(i)	1.066	1.0623	0.1351	0.1294	0.8370	0.8346
			(ii)	1.023	0.1142	0.1084	0.8172	0.8148	0.8148
			(iii)	1.027	1.0232	0.1142	0.1084	0.8172	0.8148
			(iv)	1.027	1.0232	0.1142	0.1084	0.8172	0.8148
			(v)	1.027	1.0232	0.1142	0.1084	0.8172	0.8148
			(vi)	1.027	1.0232	0.1142	0.1084	0.8172	0.8148
Zero	TSAF	120	(i)	1.097	1.0970	0.0747	0.0747	0.8401	0.8401
			(ii)	1.097	1.0970	0.0746	0.0746	0.8401	0.8401
			(iii)	1.097	1.0970	0.0746	0.0746	0.8401	0.8401
			(iv)	1.097	1.0970	0.0746	0.0746	0.8401	0.8401
			(v)	1.097	1.0970	0.0746	0.0746	0.8401	0.8401
			(vi)	1.097	1.0970	0.0746	0.0746	0.8401	0.8401

Table 5.19 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_1 design matrix when $\gamma = -0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	15.115	2.8664	2.8316	0.5323	3.5266	1.2494
			(ii)	15.849	2.3954	2.9782	0.4212	3.6784	1.1314
			(iii)	14.743	2.3901	2.7517	0.4200	3.4474	1.1300
			(iv)	14.637	2.3954	3.7788	0.4212	3.4164	1.1314
			(v)	14.711	2.3901	2.7517	0.4200	3.4446	1.1300
			(vi)	18.988	2.2683	2.1811	0.4280	3.3704	1.1248
Estimated	OSAF	30	(i)	15.113	2.8830	2.8806	0.5427	3.5695	1.2567
			(ii)	15.827	2.4142	3.0273	0.4296	3.7175	1.1380
			(iii)	14.721	2.4088	2.7957	0.4296	3.6841	1.1368
			(iv)	14.715	2.4142	3.8354	0.4296	3.6507	1.1380
			(v)	14.707	2.3922	2.8342	0.4305	3.5154	1.1369
			(vi)	18.603	2.1800	2.1634	0.4220	3.2783	1.1012
Zero	TSAF	30	(i)	1.089	1.0837	0.1478	0.1468	0.8461	0.8439
			(ii)	1.089	1.0837	0.1476	0.1468	0.8462	0.8439
			(iii)	1.089	1.0837	0.1477	0.1468	0.8460	0.8439
			(iv)	1.092	1.0837	0.1482	0.1468	0.8474	0.8439
			(v)	1.089	1.0837	0.1477	0.1468	0.8460	0.8439
			(vi)	1.089	.0837	0.1477	0.1468	0.8460	0.8439
Zero	OSAF	60	(i)	1.538	1.2685	0.2432	0.1828	0.9326	0.8749
			(ii)	1.353	1.0782	0.2264	0.1497	0.9146	0.8417
			(iii)	1.311	1.0782	0.2060	0.1497	0.8958	0.8417
			(iv)	1.628	1.0782	0.2847	0.1497	0.9751	0.8417
			(v)	1.311	1.0782	0.2060	0.1497	0.8958	0.8417
			(vi)	2.201	1.2512	0.2754	0.1575	0.8995	0.8471
Estimated	OSAF	60	(i)	1.538	1.2687	0.2401	[illegible]	0.9312	0.8728
			(ii)	1.354	1.0785	0.2246	0.1495	0.9139	0.8415
			(iii)	1.311	1.0785	0.2054	0.1495	0.8963	0.8415
			(iv)	1.628	1.0785	0.2865	0.1494	0.9760	0.8415
			(v)	1.311	1.0785	0.2056	0.1494	0.8973	0.8415
			(vi)	1.300	1.0765	0.2046	0.1441	0.8965	0.8405
Zero	TSAF	60	(i)	1.053	1.0528	0.1175	0.1174	0.8261	0.8260
			(ii)	1.052	1.0526	0.1176	0.1175	0.8260	0.8259
			(iii)	1.052	1.0526	0.1176	0.1175	0.8260	0.8259
			(iv)	1.053	1.0526	0.1175	0.1175	0.8261	0.8259
			(v)	1.052	1.0526	0.1176	0.1175	0.8260	0.8259
			(vi)	1.052	1.0526	0.1176	0.1175	0.8260	0.8259
Zero	OSAF	120	(i)	1.019	1.0170	0.1168	0.1126	0.8146	0.8129
			(ii)	1.019	1.0170	0.1168	0.1126	0.8146	0.8129
			(iii)	1.019	1.0170	0.1168	0.1126	0.8146	0.8129
			(iv)	1.019	1.0170	0.1168	0.1126	0.8146	0.8129
			(v)	1.019	1.0170	0.1168	0.1126	0.8146	0.8129
			(vi)	1.020	1.0174	0.0036	0.0035	0.8145	0.8129
Estimated	OSAF	120	(i)	1.019	1.0170	0.1168	0.1126	0.8145	0.8128
			(ii)	1.019	1.0170	0.1168	0.1126	0.8145	0.8128
			(iii)	1.019	1.0170	0.1168	0.1126	0.8145	0.8128
			(iv)	1.019	1.0170	0.1168	0.1126	0.8145	0.8128
			(v)	1.019	1.0170	0.1168	0.1126	0.8145	0.8128
			(vi)	1.597	1.0743	0.2716	0.1475	0.9675	0.8399
Zero	TSAF	120	(i)	1.027	1.0270	0.0945	0.0945	0.8147	0.8147
			(ii)	1.027	1.0270	0.0945	0.0945	0.8147	0.8147
			(iii)	1.027	1.0270	0.0945	0.0945	0.8147	0.8147
			(iv)	1.027	1.0270	0.0945	0.0945	0.8147	0.8147
			(v)	1.027	1.0270	0.0945	0.0945	0.8147	0.8147
			(vi)	1.027	1.0270	0.0945	0.0945	0.8147	0.8147

Table 5.20 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = 0$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	10.8122	2.0841	1.7961	0.2808	2.5125	0.996
			(ii)	12.3687	2.0096	2.0503	0.2662	2.7685	0.982
			(iii)	10.7673	2.0014	1.7919	0.2642	2.5043	0.980
			(iv)	10.1679	1.9932	1.5335	0.2662	2.4401	0.982
			(v)	9.5645	1.9850	1.3919	0.2642	2.3759	0.980
			(vi)	15.6980	1.8730	1.1954	0.2356	2.5789	0.985
Estimated	OSAF	30	(i)	10.8047	2.0764	1.7969	0.2763	2.5123	0.994
			(ii)	12.3629	2.0028	2.0486	0.2636	2.7674	0.980
			(iii)	10.7606	1.9945	1.7900	0.2614	2.5027	0.978
			(iv)	9.8583	1.9862	1.5314	0.2592	2.5027	0.976
			(v)	9.5560	1.9779	1.2728	0.2570	2.5027	0.974
			(vi)	15.4980	1.7730	1.1843	0.2361	2.4765	0.973
Zero	TSAF	30	(i)	1.0720	1.0662	0.1705	0.1708	0.8425	0.840
			(ii)	1.0723	1.0663	0.1705	0.1708	0.8426	0.840
			(iii)	1.0721	1.0662	0.1705	0.1708	0.8425	0.840
			(iv)	1.0758	1.0663	0.1709	0.1708	0.8439	0.840
			(v)	1.0721	1.0662	0.1705	0.1708	0.8425	0.840
			(vi)	1.0721	1.0662	0.1705	0.1708	0.8425	0.840
Zero	OSAF	60	(i)	1.1541	1.1338	0.1608	0.1508	0.8518	0.843
			(ii)	1.1541	1.1338	0.1608	0.1508	0.8518	0.843
			(iii)	1.1541	1.1338	0.1608	0.1508	0.8518	0.843
			(iv)	1.1623	1.1338	0.1625	0.1508	0.8537	0.843
			(v)	1.1541	1.1338	0.1608	0.1508	0.8518	0.843
			(vi)	2.1459	1.1238	0.2599	0.1506	2.0499	0.841
Estimated	OSAF	60	(i)	1.1543	1.1340	0.1626	0.1512	0.8528	0.843
			(ii)	1.1543	1.1339	0.1626	0.1512	0.8528	0.843
			(iii)	1.1543	1.1340	0.1626	0.1512	0.8528	0.843
			(iv)	1.1624	1.1339	0.1653	0.1512	0.8552	0.843
			(v)	1.1543	1.1340	0.1610	0.1508	0.8519	0.843
			(vi)	2.1533	1.1556	0.2805	0.1509	1.0858	0.843
Zero	TSAF	60	(i)	1.0294	1.0292	0.1367	0.1366	0.8211	0.821
			(ii)	1.0294	1.0292	0.1367	0.1366	0.8211	0.821
			(iii)	1.0294	1.0292	0.1367	0.1366	0.8211	0.821
			(iv)	1.0294	1.0292	0.1367	0.1366	0.8210	0.821
			(v)	1.0294	1.0292	0.1367	0.1366	0.8211	0.821
			(vi)	1.0294	1.0292	0.1367	0.1366	0.8211	0.821
Zero	OSAF	120	(i)	1.0208	1.0191	0.1133	0.1105	0.8141	0.813
			(ii)	1.0208	1.0191	0.1133	0.1105	0.8141	0.813
			(iii)	1.0208	1.0191	0.1133	0.1105	0.8141	0.813
			(iv)	1.0208	1.0191	0.1133	0.1105	0.8141	0.813
			(v)	1.0208	1.0191	0.1133	0.1105	0.8141	0.813
			(vi)	2.0210	1.0193	0.0025	0.0024	0.8142	0.813
Estimated	OSAF	120	(i)	1.0209	1.0192	0.1133	0.1105	0.8142	0.813
			(ii)	1.0209	1.0192	0.1133	0.1105	0.8142	0.813
			(iii)	1.0209	1.0192	0.1133	0.1105	0.8142	0.813
			(iv)	1.0209	1.0192	0.1133	0.1105	0.8142	0.813
			(v)	1.0209	1.0192	0.1133	0.1105	0.8142	0.813
			(vi)	1.0201	1.0193	0.1025	0.1024	0.8142	0.813
Zero	TSAF	120	(i)	1.0089	1.0089	0.1096	0.1096	0.8088	0.808
			(ii)	1.0089	1.0089	0.1096	0.1096	0.8088	0.808
			(iii)	1.0089	1.0089	0.1096	0.1096	0.8088	0.808
			(iv)	1.0089	1.0089	0.1096	0.1096	0.8088	0.808
			(v)	1.0089	1.0089	0.1096	0.1096	0.8088	0.808
			(vi)	1.0089	1.0089	0.1096	0.1096	0.8088	0.808

Table 5.21 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = 0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	7.4980	1.926	0.9479	0.2041	1.7295	0.9607
			(ii)	9.3430	1.936	1.2362	0.2068	2.0154	0.9602
			(iii)	7.3919	1.848	0.9980	0.2055	1.7172	0.9516
			(iv)	7.4408	1.806	0.9598	0.2036	1.7109	0.9477
			(v)	7.3286	1.756	0.9479	0.2041	1.7100	0.9410
			(vi)	12.9641	1.521	1.9938	0.2036	2.3021	0.8742
Estimated	OSAF	30	(i)	7.5146	1.920	0.9418	0.2019	1.7302	0.9617
			(ii)	9.3318	1.911	1.2353	0.2042	2.0197	0.9618
			(iii)	7.2957	1.824	0.9919	0.2020	1.7170	0.9528
			(iv)	7.2596	1.722	0.9597	0.2041	1.5748	0.9474
			(v)	7.2235	1.711	0.9587	0.2004	1.4248	0.9400
			(vi)	12.1360	1.518	2.9592	0.1943	2.3952	0.8542
Zero	TSAF	30	(i)	1.0979	1.094	0.1954	0.1947	0.6580	0.8564
			(ii)	1.0992	1.094	0.1957	0.1946	0.8585	0.8564
			(iii)	1.0978	1.093	0.1953	0.1946	0.8579	0.8563
			(iv)	1.1012	1.094	0.1957	0.1946	0.8592	0.8564
			(v)	1.0978	1.093	0.1953	0.1946	0.8579	0.8563
			(vi)	1.0978	1.093	0.1953	0.1946	0.8579	0.8563
Zero	OSAF	60	(i)	1.0536	1.045	0.1393	0.1334	0.8304	0.8265
			(ii)	1.0536	1.045	0.1393	0.1334	0.8304	0.8265
			(iii)	1.0536	1.045	0.1393	0.1334	0.8304	0.8265
			(iv)	1.0744	1.045	0.1475	0.1334	0.8377	0.8265
			(v)	1.0536	1.045	0.1393	0.1334	0.8304	0.8265
			(vi)	1.0342	1.028	0.1285	0.1305	0.6754	0.8015
Estimated	OSAF	60	(i)	1.0539	1.045	0.1393	0.1334	0.8305	0.8267
			(ii)	1.0539	1.045	0.1393	0.1334	0.8305	0.8267
			(iii)	1.0539	1.045	0.1393	0.1334	0.8305	0.8267
			(iv)	1.0746	1.045	0.1463	0.1334	0.8365	0.8267
			(v)	1.0539	1.045	0.1393	0.1334	0.8305	0.8267
			(vi)	1.0717	1.043	0.1400	0.1323	0.8331	0.8256
Zero	TSAF	60	(i)	1.0491	1.048	0.1513	0.1512	0.8309	0.8308
			(ii)	1.0491	1.048	0.1513	0.1512	0.8309	0.8308
			(iii)	1.0491	1.048	0.1513	0.1512	0.8309	0.8308
			(iv)	1.0491	1.048	0.1512	0.1512	0.8309	0.8308
			(v)	1.0491	1.048	0.1513	0.1512	0.8309	0.8308
			(vi)	1.0491	1.048	0.1513	0.1512	0.8309	0.8308
Zero	OSAF	120	(i)	1.0202	1.019	0.1127	0.1108	0.8137	0.8132
			(ii)	1.0202	1.019	0.1127	0.1108	0.8137	0.8132
			(iii)	1.0202	1.019	0.1127	0.1108	0.8137	0.8132
			(iv)	1.0202	1.019	0.1127	0.1108	0.8137	0.8132
			(v)	1.0202	1.019	0.1127	0.1108	0.8137	0.8132
			(vi)	1.0193	1.018	0.1122	0.1022	0.8133	0.8128
Estimated	OSAF	120	(i)	1.0200	1.019	0.1127	0.1107	0.8137	0.8132
			(ii)	1.0200	1.019	0.1127	0.1107	0.8137	0.8132
			(iii)	1.0200	1.019	0.1127	0.1107	0.8137	0.8132
			(iv)	1.0200	1.019	0.1127	0.1107	0.8137	0.8132
			(v)	1.0200	1.019	0.1127	0.1107	0.8137	0.8132
			(vi)	1.0193	1.018	0.1123	0.1104	0.8133	0.8128
Zero	TSAF	120	(i)	1.0491	1.048	0.1513	0.1512	0.8309	0.8308
			(ii)	1.0328	1.032	0.1276	0.1277	0.8206	0.8206
			(iii)	1.0328	1.032	0.1276	0.1277	0.8206	0.8206
			(iv)	1.0328	1.032	0.1276	0.1277	0.8206	0.8206
			(v)	1.0328	1.032	0.1276	0.1277	0.8206	0.8206
			(vi)	1.0328	1.032	0.1276	0.1277	0.8206	0.8206

Table 5.22 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = 0.30$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	5.3049	1.977	0.6643	0.1871	1.4584	0.9741
			(ii)	7.3667	1.705	0.9758	0.1875	1.7365	0.9454
			(iii)	5.2830	1.678	0.8647	0.1875	1.4217	0.9382
			(iv)	5.1993	1.659	0.8536	0.1873	1.4299	0.9388
			(v)	4.9332	1.642	0.7647	0.1874	1.4144	0.9298
			(vi)	11.8242	1.390	2.7577	0.1672	2.9972	0.9208
Estimated	OSAF	30	(i)	5.2757	1.945	0.6722	0.1870	1.4603	0.9700
			(ii)	7.3638	1.696	0.6871	0.1873	1.7459	0.9434
			(iii)	4.9913	1.656	0.6726	0.1873	1.5279	0.9369
			(iv)	4.6188	1.653	0.6581	0.1872	1.4299	0.9380
			(v)	4.5463	1.641	0.6750	0.1871	1.4093	0.9349
			(vi)	11.5289	1.174	2.6632	0.1422	2.3971	0.9123
Zero	TSAF	30	(i)	1.1716	1.168	0.2210	0.2202	0.8909	0.8897
			(ii)	1.1715	1.168	0.2214	0.2203	0.8907	0.8895
			(iii)	1.1708	1.168	0.2210	0.2203	0.8906	0.8895
			(iv)	1.1747	1.168	0.2218	0.2202	0.8919	0.8895
			(v)	1.1708	1.168	0.2210	0.2203	0.8906	0.8895
			(vi)	1.1708	1.168	0.2210	0.2203	0.8906	0.8895
Zero	OSAF	60	(i)	1.1723	1.167	0.1409	0.1373	0.8460	0.8434
			(ii)	1.0550	1.050	0.1405	0.1370	0.8313	0.8287
			(iii)	1.0550	1.050	0.1405	0.1370	0.8313	0.8287
			(iv)	1.0550	1.050	0.1405	0.1370	0.8313	0.8287
			(v)	1.0550	1.050	0.1405	0.1370	0.8313	0.8287
			(vi)	1.3210	1.053	0.1577	0.1016	0.9006	0.8279
Estimated	OSAF	60	(i)	1.1714	1.166	0.1409	0.1373	0.8461	0.8435
			(ii)	1.0543	1.0492	0.1406	0.1370	0.8310	0.8284
			(iii)	1.0543	1.049	0.1406	0.1370	0.8310	0.8284
			(iv)	1.0543	1.049	0.1406	0.1370	0.8310	0.8284
			(v)	1.0543	1.049	0.1406	0.1370	0.8310	0.8284
			(vi)	1.0530	1.048	0.1394	0.1358	0.8304	0.8279
Zero	TSAF	60	(i)	1.1251	1.125	0.1777	0.1778	0.8640	0.8640
			(ii)	1.1225	1.122	0.1771	0.1772	0.8632	0.8632
			(iii)	1.1225	1.122	0.1771	0.1772	0.8632	0.8632
			(iv)	1.1225	1.122	0.1771	0.1772	0.8632	0.8632
			(v)	1.1225	1.122	0.1771	0.1772	0.8632	0.8632
			(vi)	1.1225	1.122	0.1771	0.1772	0.8632	0.8632
Zero	OSAF	120	(i)	1.0143	1.013	0.1123	0.1110	0.8121	0.8117
			(ii)	1.0143	1.013	0.1123	0.1110	0.8121	0.8117
			(iii)	1.0143	1.013	0.1123	0.1110	0.8121	0.8117
			(iv)	1.0143	1.013	0.1123	0.1110	0.8121	0.8117
			(v)	1.0143	1.013	0.1123	0.1110	0.8121	0.8117
			(vi)	1.0113	1.010	0.0047	0.0046	0.8106	0.8102
Estimated	OSAF	120	(i)	1.0135	1.012	0.1123	0.1110	0.8117	0.8113
			(ii)	1.0135	1.012	0.1123	0.1110	0.8117	0.8113
			(iii)	1.0135	1.012	0.1123	0.1110	0.8117	0.8113
			(iv)	1.0135	1.012	0.1123	0.1110	0.8117	0.8113
			(v)	1.0135	1.012	0.1123	0.1110	0.8117	0.8113
			(vi)	1.0113	1.010	0.1117	0.1104	0.8106	0.8102
Zero	TSAF	120	(i)	1.0950	1.095	0.1447	0.1447	0.8474	0.8474
			(ii)	1.0950	1.095	0.1447	0.1447	0.8474	0.8474
			(iii)	1.0950	1.095	0.1447	0.1447	0.8474	0.8474
			(iv)	1.0950	1.095	0.1447	0.1447	0.8474	0.8474
			(v)	1.0950	1.095	0.1447	0.1447	0.8474	0.8474
			(vi)	1.0950	1.095	0.1447	0.1447	0.8474	0.8474

Table 5.23 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = 0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	4.1908	2.439	0.3713	0.1791	1.2694	1.0553
			(ii)	6.1720	1.896	0.6071	0.1697	1.4419	0.9825
			(iii)	3.7752	1.828	0.3971	0.1699	1.2057	0.9724
			(iv)	3.7378	1.770	0.3871	0.1698	1.2045	0.9647
			(v)	3.6660	1.699	0.3771	0.1700	1.1907	0.9541
			(vi)	8.0343	1.531	0.8179	0.1321	1.7050	0.8791
Estimated	OSAF	30	(i)	4.1696	2.429	0.3776	0.1798	1.2694	1.0514
			(ii)	6.1624	1.872	0.6183	0.1699	1.4451	0.9775
			(iii)	3.7513	1.787	0.3935	0.1643	1.2059	0.9656
			(iv)	3.3402	1.586	0.3687	0.1623	1.1667	0.9640
			(v)	3.1929	1.505	0.3960	0.1617	0.9975	0.9538
			(vi)	2.9953	1.425	0.3893	0.1615	0.9883	0.7738
Zero	TSAF	30	(i)	1.2899	1.287	0.2456	0.2454	0.9393	0.9384
			(ii)	1.2872	1.283	0.2453	0.2453	0.9384	0.9373
			(iii)	1.2856	1.283	0.2453	0.2452	0.9379	0.9371
			(iv)	1.2888	1.283	0.2454	0.2453	0.9389	0.9373
			(v)	1.2856	1.283	0.2453	0.2452	0.9379	0.9370
			(vi)	1.2856	1.283	0.2453	0.2452	0.9379	0.9370
Zero	OSAF	60	(i)	1.6828	1.679	0.1397	0.1376	0.9188	0.9170
			(ii)	1.1472	1.143	0.1404	0.1384	0.8458	0.8440
			(iii)	1.1472	1.143	0.1404	0.1384	0.8458	0.8440
			(iv)	1.1472	1.143	0.1404	0.1384	0.8458	0.8440
			(v)	1.1472	1.143	0.1404	0.1384	0.8458	0.8440
			(vi)	1.1343	1.031	0.1317	0.1321	0.8405	0.8398
Estimated	OSAF	60	(i)	1.6588	1.655	0.1399	0.1378	0.9155	0.9137
			(ii)	1.1194	1.115	0.1406	0.1385	0.8429	0.8411
			(iii)	1.1194	1.115	0.1406	0.1385	0.8429	0.8411
			(iv)	1.1194	1.115	0.1406	0.1385	0.8429	0.8411
			(v)	1.1265	1.122	0.1406	0.1385	0.8437	0.8419
			(vi)	1.0950	1.091	0.1387	0.1366	0.8401	0.8383
Zero	TSAF	60	(i)	1.2520	1.251	0.1999	0.1999	0.9150	0.9149
			(ii)	1.2476	1.247	0.2004	0.2004	0.9137	0.9136
			(iii)	1.2476	1.247	0.2004	0.2004	0.9137	0.9136
			(iv)	1.2476	1.247	0.2004	0.2004	0.9137	0.9136
			(v)	1.2476	1.247	0.2004	0.2004	0.9137	0.9136
			(vi)	1.2476	1.247	0.2004	0.2004	0.9137	0.9136
Zero	OSAF	120	(i)	1.4362	1.435	0.1141	0.1133	0.8628	0.8624
			(ii)	1.0314	1.030	0.1141	0.1133	0.8181	0.8177
			(iii)	1.0314	1.030	0.1141	0.1133	0.8181	0.8177
			(iv)	1.0314	1.030	0.1141	0.1133	0.8181	0.8177
			(v)	1.0314	1.030	0.1141	0.1133	0.8181	0.8177
			(vi)	1.0290	1.028	0.1126	0.1116	0.8172	0.8168
Estimated	OSAF	120	(i)	1.4730	1.472	0.1140	0.1132	0.8660	0.8656
			(ii)	1.0282	1.027	0.1140	0.1132	0.8169	0.8165
			(iii)	1.0282	1.027	0.1140	0.1132	0.8169	0.8165
			(iv)	1.0282	1.027	0.1140	0.1132	0.8169	0.8165
			(v)	1.0282	1.027	0.1140	0.1132	0.8169	0.8165
			(vi)	1.0290	1.028	0.1135	0.1127	0.8172	0.8168
Zero	TSAF	120	(i)	1.2351	1.235	0.1641	0.1641	0.9006	0.9006
			(ii)	1.2248	1.224	0.1639	0.1639	0.8976	0.8976
			(iii)	1.2248	1.224	0.1639	0.1639	0.8976	0.8976
			(iv)	1.2248	1.224	0.1639	0.1639	0.8976	0.8976
			(v)	1.2248	1.224	0.1639	0.1639	0.8976	0.8976
			(vi)	1.2248	1.224	0.1639	0.1639	0.8976	0.8976

Table 5.24 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = 0.60$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.0655	2.734	0.2350	0.1690	1.2122	1.1417
			(ii)	2.8011	[illegible]	0.2813	0.1707	1.1531	1.0276
			(iii)	2.4488	2.089	0.2366	0.1706	1.1034	1.0272
			(iv)	3.5742	1.884	0.3608	0.1705	1.1968	0.9961
			(v)	2.2031	1.874	0.2369	0.1708	1.0639	0.9937
			(vi)	3.4941	1.726	0.3305	0.1704	1.1904	0.9042
Estimated	OSAF	30	(i)	3.1228	2.797	0.2382	0.1692	1.2187	1.1470
			(ii)	2.7940	2.081	0.2840	0.1703	1.1525	1.0261
			(iii)	2.4430	2.080	0.2398	0.1708	1.1030	1.0257
			(iv)	3.5403	1.864	0.3620	0.1710	1.1940	0.9943
			(v)	2.2536	1.931	0.2421	0.1709	1.0808	1.0077
			(vi)	3.1504	1.620	0.3201	0.1603	1.1713	0.8787
Zero	TSAF	30	(i)	1.4466	1.446	0.2752	0.2750	0.9975	0.9968
			(ii)	1.4452	1.442	0.2754	0.2750	0.9964	0.9955
			(iii)	1.4449	1.442	0.2751	0.2749	0.9963	0.9955
			(iv)	1.4454	1.442	0.2756	0.2749	0.9965	0.9953
			(v)	1.4446	1.442	0.2751	0.2749	0.9961	0.9953
			(vi)	1.4446	1.442	0.2751	0.2749	0.9961	0.9953
Zero	OSAF	60	(i)	2.3374	2.317	0.1382	0.1373	1.0530	1.0478
			(ii)	1.3111	1.256	0.1366	0.1356	0.8727	0.8640
			(iii)	1.3028	1.256	0.1365	0.1356	0.8713	0.8640
			(iv)	1.1868	1.169	0.1365	0.1356	0.8558	0.8509
			(v)	1.1806	1.168	0.1365	0.1356	0.8542	0.8506
			(vi)	1.1741	1.168	0.1405	0.1219	0.8504	0.8504
Estimated	OSAF	60	(i)	1.6588	1.655	0.1399	0.1378	0.9155	0.9137
			(ii)	1.1194	1.115	0.1406	0.1385	0.8429	0.8411
			(iii)	1.1194	1.115	0.1406	0.1385	0.8429	0.8411
			(iv)	1.1194	1.115	0.1406	0.1385	0.8429	0.8411
			(v)	1.1265	1.122	0.1406	0.1385	0.8437	0.8419
			(vi)	1.0508	1.034	0.1049	0.1047	0.8227	0.8188
Zero	TSAF	60	(i)	2.2273	2.208	0.1380	0.1371	1.0435	1.0384
			(ii)	1.2568	1.218	0.1364	0.1354	0.8648	0.8586
			(iii)	1.2473	1.218	0.1363	0.1354	0.8633	0.8586
			(iv)	1.1532	1.134	0.1363	0.1354	0.8504	0.8454
			(v)	1.1528	1.141	0.1363	0.1354	0.8502	0.8468
			(vi)	1.1528	1.141	0.1363	0.1354	0.8502	0.8468
Zero	OSAF	120	(i)	2.3421	2.332	0.1052	0.1050	1.0021	0.9995
			(ii)	1.0730	1.048	0.1053	0.1050	0.8277	0.8222
			(iii)	1.0568	1.048	0.1052	0.1050	0.8242	0.8222
			(iv)	1.0657	1.040	0.1053	0.1050	0.8261	0.8207
			(v)	1.0495	1.040	0.1053	0.1050	0.8226	0.8207
			(vi)	1.0572	1.034	0.1061	0.1000	0.8238	0.8188
Estimated	OSAF	120	(i)	2.3434	2.335	0.1051	0.1048	1.0042	1.0020
			(ii)	1.0715	1.046	0.1051	0.1048	0.8264	0.8206
			(iii)	1.0525	1.046	0.1051	0.1048	0.8223	0.8206
			(iv)	1.0663	1.034	0.1051	0.1048	0.8252	0.8187
			(v)	1.0532	1.034	0.1051	0.1048	0.8218	0.8187
			(vi)	1.0080	1.008	0.1023	0.1010	0.7779	0.7898
Zero	TSAF	120	(i)	1.4243	1.424	0.1699	0.1698	0.9658	0.9658
			(ii)	1.3833	1.383	0.1683	0.1683	0.9542	0.9541
			(iii)	1.3832	1.383	0.1683	0.1683	0.9542	0.9541
			(iv)	1.3833	1.383	0.1683	0.1682	0.9542	0.9541
			(v)	1.3832	1.383	0.1683	0.1682	0.9542	0.9541
			(vi)	1.3832	1.383	0.1683	0.1682	0.9542	0.9541

Table 5.25 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = 0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.9323	2.837	0.2686	0.1598	1.3648	1.2537
			(ii)	4.0991	2.444	0.3224	0.1613	1.3300	1.1606
			(iii)	3.5027	2.427	0.2706	0.1616	1.2641	1.1574
			(iv)	5.3615	2.052	0.4001	0.1619	1.3342	1.0752
			(v)	3.1392	2.012	0.2711	0.1624	1.1852	1.0680
			(vi)	12.1559	1.954	2.9225	0.1501	3.2580	1.0019
Estimated	OSAF	30	(i)	3.8520	2.743	0.2679	0.1597	1.3447	1.2306
			(ii)	4.0263	2.397	0.3231	0.1612	1.3099	1.1426
			(iii)	3.4395	2.374	0.2698	0.1615	1.2449	1.1392
			(iv)	5.2703	1.990	0.4009	0.1618	1.3109	1.0550
			(v)	3.1830	2.052	0.2703	0.1623	1.1919	1.0726
			(vi)	12.1997	1.844	2.8852	0.1425	2.9779	1.002
Zero	TSAF	30	(i)	1.6662	1.663	0.2878	0.2873	1.0687	1.0678
			(ii)	1.6655	1.660	0.2877	0.2872	1.0684	1.0667
			(iii)	1.6639	1.661	0.2877	0.2872	1.0679	1.0670
			(iv)	1.6655	1.660	0.2876	0.2871	1.0684	1.0666
			(v)	1.6638	1.661	0.2876	0.2872	1.0679	1.0669
			(vi)	1.6638	1.661	0.2876	0.2872	1.0679	1.0669
Zero	OSAF	60	(i)	2.2594	2.20	0.1279	0.1279	1.0986	1.0859
			(ii)	1.5399	1.516	0.1277	0.1275	0.9454	0.939
			(iii)	1.528	1.504	0.1276	0.1276	0.9429	0.9357
			(iv)	1.3912	1.320	0.1277	0.1277	0.9088	0.8929
			(v)	1.3679	1.308	0.1278	0.1278	0.9035	0.8894
			(vi)	1.3231	1.254	0.1262	0.1071	0.8334	0.7839
Estimated	OSAF	60	(i)	1.4626	1.440	0.1295	0.1296	0.9640	0.9570
			(ii)	1.3760	1.379	0.1304	0.1304	0.9363	0.9371
			(iii)	1.3759	1.378	0.1305	0.1306	0.9362	0.9367
			(iv)	1.2579	1.234	0.1307	0.1309	0.8974	0.8898
			(v)	1.4854	1.424	0.1308	0.1312	0.9631	0.9448
			(vi)	1.4307	1.333	0.1262	0.1263	0.9488	0.9272
Zero	TSAF	60	(i)	1.5961	1.596	0.2281	0.2281	1.0330	1.0329
			(ii)	1.5824	1.582	0.2248	0.2247	1.0281	1.0281
			(iii)	1.5824	1.582	0.2248	0.2248	1.0281	1.0281
			(iv)	1.5823	1.582	0.2248	0.2245	1.0281	1.0280
			(v)	1.5822	1.582	0.2247	0.2246	1.0281	1.0280
			(vi)	1.5822	1.582	0.2247	0.2246	1.0281	1.0280
Zero	OSAF	120	(i)	2.7565	2.713	0.1060	0.1060	1.1338	1.1243
			(ii)	1.2321	1.142	0.1041	0.1043	0.8605	0.8413
			(iii)	1.2222	1.142	0.1041	0.1043	0.8582	0.8413
			(iv)	1.1162	1.062	0.1043	0.1044	0.8390	0.8266
			(v)	1.1063	1.062	0.1044	0.1044	0.8367	0.8266
			(vi)	1.1076	1.057	0.0053	0.0053	0.8350	0.8232
Estimated	OSAF	120	(i)	2.5936	2.561	0.1060	0.1060	1.1093	1.1017
			(ii)	1.2195	1.138	0.1041	0.1044	0.8551	0.8373
			(iii)	1.2022	1.138	0.1042	0.1044	0.8517	0.8373
			(iv)	1.1085	1.054	0.1044	0.1045	0.8352	0.8221
			(v)	1.0858	1.046	0.1044	0.1045	0.8303	0.8210
			(vi)	1.1613	1.000	0.1138	0.1023	0.8476	0.8192
Zero	TSAF	120	(i)	1.6342	1.634	0.1899	0.1898	1.0369	1.0368
			(ii)	1.5817	1.581	0.1841	0.1841	1.0203	1.0201
			(iii)	1.5817	1.581	0.1841	0.1841	1.0203	1.0201
			(iv)	1.5809	1.580	0.1841	0.1841	1.0201	1.0200
			(v)	1.5809	1.580	0.1841	0.1841	1.0201	1.0200
			(vi)	1.5809	1.580	0.1841	0.1841	1.0201	1.0200

Table 5 .26 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = 0.90$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	2.4244	2.0774	0.2338	0.167	1.2066	1.1421
			(ii)	2.6099	2.0001	0.2614	0.168	1.2194	1.1197
			(iii)	2.3402	1.9820	0.2349	0.168	1.1813	1.1144
			(iv)	2.5030	1.8224	0.2879	0.169	1.1911	1.0683
			(v)	2.1549	1.7996	0.2360	0.169	1.1288	1.0616
			(vi)	7.9052	1.7184	3.6075	0.167	3.1258	1.0191
Estimated	OSAF	30	(i)	2.0394	1.6862	0.2341	0.167	1.1006	1.0340
			(ii)	2.2284	1.6224	0.2654	0.168	1.1183	1.0153
			(iii)	1.9723	1.6091	0.2351	0.168	1.0810	1.0117
			(iv)	2.1695	1.4970	0.2908	0.169	1.1005	0.9781
			(v)	2.1450	1.7868	0.2351	0.169	1.1255	1.0583
			(vi)	7.8726	1.3012	3.5453	0.163	3.1212	0.9941
Zero	TSAF	30	(i)	1.8908	1.8896	0.3202	0.319	1.1409	1.1406
			(ii)	1.8920	1.8901	0.3201	0.319	1.1412	1.1406
			(iii)	1.8915	1.8903	0.3204	0.320	1.1412	1.1408
			(iv)	1.8915	1.8899	0.3200	0.319	1.1412	1.1407
			(v)	1.8913	1.8901	0.3204	0.320	1.1412	1.1408
			(vi)	1.8913	1.8901	0.3204	0.320	1.1412	1.1408
Zero	OSAF	60	(i)	1.9756	1.9288	0.1295	0.129	1.1046	1.0910
			(ii)	1.7835	1.8257	0.1305	0.130	1.0504	1.0600
			(iii)	1.7827	1.8193	0.1305	0.130	1.0501	1.0578
			(iv)	1.5423	1.5101	0.1307	0.130	0.9781	0.9682
			(v)	1.5401	1.5028	0.1308	0.131	0.9773	0.9657
			(vi)	2.9952	2.6518	0.3195	0.143	1.7580	1.5119
Estimated	OSAF	60	(i)	1.4626	1.4409	0.1295	0.129	0.9640	0.9570
			(ii)	1.3760	1.3795	0.1304	0.130	0.9363	0.9371
			(iii)	1.3759	1.3785	0.1305	0.130	0.9362	0.9367
			(iv)	1.2579	1.2348	0.1307	0.130	0.8974	0.8898
			(v)	1.4854	1.4241	0.1308	0.131	0.9631	0.9448
			(vi)	2.9612	2.7515	0.2998	0.130	1.6165	1.4637
Zero	TSAF	60	(i)	1.8300	1.8300	0.2504	0.250	1.1079	1.1079
			(ii)	1.8296	1.8299	0.2505	0.250	1.1075	1.1076
			(iii)	1.8296	1.8300	0.2505	0.250	1.1075	1.1077
			(iv)	1.8287	1.8285	0.2497	0.249	1.1072	1.1071
			(v)	1.8289	1.8289	0.2497	0.249	1.1072	1.1073
			(vi)	1.8289	1.8289	0.2497	0.249	1.1072	1.1073
Zero	OSAF	120	(i)	2.1376	1.9974	0.1097	0.110	1.1326	1.0917
			(ii)	1.8690	1.6116	0.1103	0.110	1.0613	0.9883
			(iii)	1.8636	1.6116	0.1103	0.110	1.0597	0.9883
			(iv)	1.5748	1.3040	0.1105	0.111	0.9785	0.9014
			(v)	1.5693	1.3040	0.1105	0.111	0.9769	0.9014
			(vi)	1.5011	1.2331	0.1033	0.103	0.9782	0.8803
Estimated	OSAF	120	(i)	1.6309	1.5570	0.1098	0.110	0.9971	0.9754
			(ii)	1.4651	1.3377	0.1104	0.110	0.9497	0.9119
			(iii)	1.4628	1.3377	0.1104	0.110	0.9490	0.9119
			(iv)	1.2946	1.1459	0.1107	0.111	0.8984	0.8538
			(v)	1.5186	1.2416	0.1107	0.111	0.9613	0.8815
			(vi)	1.4811	1.1989	0.9733	0.110	0.9572	0.7905
Zero	TSAF	120	(i)	1.8159	1.8143	0.2140	0.214	1.0957	1.0953
			(ii)	1.8107	1.8072	0.2138	0.213	1.0944	1.0932
			(iii)	1.8107	1.8072	0.2138	0.213	1.0944	1.0932
			(iv)	1.8058	1.8042	0.2131	0.212	1.0928	1.0923
			(v)	1.8057	1.8042	0.2131	0.212	1.0927	1.0923
			(vi)	1.8057	1.8042	0.2131	0.212	1.0927	1.0923

Table 5.27 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = -0.90$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	14.467	10.5014	7.1720	5.0882	7.2811	5.3763
			(ii)	14.473	8.7219	7.1790	4.2004	7.2867	4.5486
			(iii)	14.467	8.7369	7.1720	4.2040	7.2811	4.5529
			(iv)	14.474	8.7268	7.1880	4.1942	7.2942	4.5426
			(v)	14.467	8.7208	7.1720	4.1888	7.2811	4.5384
			(vi)	13.733	8.3022	6.8094	3.3241	7.1991	4.4910
Estimated	OSAF	30	(i)	14.466	10.4989	7.1551	5.0662	7.2495	5.3438
			(ii)	14.473	8.7181	7.1616	4.1831	7.2548	4.5225
			(iii)	14.466	8.7331	7.1551	4.1871	7.2495	4.5271
			(iv)	14.474	8.7227	7.1685	4.1841	7.2595	4.5236
			(v)	14.466	8.7167	7.2203	4.2232	7.3250	4.5667
			(vi)	13.738	8.3061	7.1949	4.1273	7.2519	4.4639
Zero	TSAF	30	(i)	1.864	1.8679	0.1022	0.1046	1.0934	1.0945
			(ii)	1.864	1.8679	0.1020	0.1054	1.0934	1.0944
			(iii)	1.864	1.8682	0.1022	0.1054	1.0934	1.0946
			(iv)	1.864	1.8678	0.1019	0.1053	1.0933	1.0945
			(v)	1.864	1.8682	0.1022	0.1054	1.0934	1.0946
Zero	OSAF	60	(i)	6.948	4.7991	2.0231	1.3783	2.4194	1.8594
			(ii)	6.989	4.3146	2.0309	1.2312	2.4262	1.7323
			(iii)	6.948	4.3004	2.0231	1.2266	2.4194	1.7282
			(iv)	6.998	4.2309	2.0315	1.2158	2.4280	1.7156
			(v)	6.948	4.2166	2.0231	1.2111	2.4194	1.7115
			(vi)	4.549	2.3404	1.4092	0.9939	1.9060	1.0457
Estimated	OSAF	60	(i)	6.948	4.7912	1.9327	1.3231	2.3059	1.7836
			(ii)	6.989	4.3045	1.9406	1.1878	2.3129	1.6704
			(iii)	6.947	4.2902	1.9317	1.1840	2.3054	1.6671
			(iv)	6.998	4.2201	1.9436	1.1702	2.3155	1.6555
			(v)	6.948	4.2057	2.0253	1.2132	2.4238	1.7109
			(vi)	4.445	2.4148	1.3892	0.7892	1.8455	1.0321
Zero	TSAF	60	(i)	1.844	1.8448	0.0677	0.0701	1.0859	1.0863
			(ii)	1.844	1.8446	0.0678	0.0695	1.0859	1.0862
			(iii)	1.844	1.8446	0.0677	0.0695	1.0859	1.0862
			(iv)	1.844	1.8446	0.0678	0.0695	1.0859	1.0862
			(v)	1.844	1.8446	0.0677	0.0695	1.0859	1.0862
Zero	OSAF	120	(i)	5.782	4.1507	3.1580	2.2730	3.3473	2.5468
			(ii)	5.373	3.2795	3.0119	1.8279	3.2077	2.1385
			(iii)	5.354	3.2795	3.0021	1.8279	3.1984	2.1385
			(iv)	4.825	2.4235	2.7767	1.4196	2.9870	1.7606
			(v)	4.805	2.4235	2.7669	1.4196	2.9777	1.7606
			(vi)	4.549	2.3404	6.4092	2.1939	6.9060	3.0457
Estimated	OSAF	120	(i)	5.776	4.1352	3.1629	2.2810	3.3287	2.5352
			(ii)	5.366	3.2602	3.0131	1.8409	3.1887	2.1334
			(iii)	5.346	3.2602	3.0042	1.8409	3.1799	2.1334
			(iv)	4.816	2.3989	2.7748	1.4230	2.9655	1.7511
			(v)	4.795	2.3987	2.8483	1.4267	3.0590	1.7646
			(vi)	4.553	2.3403	6.6041	2.6739	6.8773	3.0481
Zero	TSAF	120	(i)	1.820	1.8207	0.0676	0.0675	1.0782	1.0783
			(ii)	1.820	1.8207	0.0675	0.0667	1.0781	1.0781
			(iii)	1.820	1.8207	0.0675	0.0667	1.0781	1.0781
			(iv)	1.820	1.8207	0.0671	0.0656	1.0781	1.0781
			(v)	1.820	1.8207	0.0671	0.0656	1.0781	1.0781

Table 5.28 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = -0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	20.416	12.9416	8.2457	5.0992	8.3851	5.4511
			(ii)	20.483	10.0732	8.2804	3.8904	8.4180	4.3192
			(iii)	20.376	10.0701	8.2263	3.8908	8.3667	4.3187
			(iv)	20.866	10.1192	8.3955	3.9075	8.5288	4.3349
			(v)	20.376	10.0701	8.2263	3.8908	8.3667	4.3187
			(vi)	19.701	9.9562	5.9963	2.8562	7.9589	4.2297
Estimated	OSAF	30	(i)	20.415	12.9390	8.2745	5.1205	8.4033	5.4663
			(ii)	20.481	10.0702	8.3079	3.9071	8.4349	4.3305
			(iii)	20.374	10.0672	8.2566	3.9076	8.3864	4.3301
			(iv)	20.865	10.1162	8.4247	3.9251	8.5450	4.3470
			(v)	20.374	10.0672	8.2914	3.9206	8.4341	4.3439
			(vi)	19.700	9.5914	5.9843	2.7453	7.6543	4.0266
Zero	TSAF	30	(i)	1.592	1.5957	0.1210	0.1252	1.0136	1.0155
			(ii)	1.592	1.5957	0.1210	0.1249	1.0136	1.0154
			(iii)	1.592	1.5957	0.1210	0.1251	1.0136	1.0155
			(iv)	1.592	1.5957	0.1209	0.1250	1.0135	1.0154
			(v)	1.592	1.5957	0.1210	0.1251	1.0136	1.0155
Zero	OSAF	60	(i)	13.311	5.6644	2.7184	1.2872	3.2329	1.8514
			(ii)	12.929	4.0260	2.6624	0.9449	3.1800	1.5225
			(iii)	12.639	4.0259	2.5952	0.9449	3.1158	1.5225
			(iv)	12.924	3.9689	2.6575	0.9332	3.1790	1.5106
			(v)	12.570	3.9512	2.5742	0.9262	3.0974	1.5050
			(vi)	11.911	3.7237	4.2183	1.0925	5.2219	2.0616
Estimated	OSAF	60	(i)	13.310	5.6594	2.7190	1.2739	3.2187	1.8323
			(ii)	12.927	4.0193	2.6688	0.9308	3.1693	1.5068
			(iii)	12.638	4.0192	2.5991	0.9308	3.1031	1.5068
			(iv)	12.922	3.9623	2.6622	0.9196	3.1653	1.4965
			(v)	12.567	3.9457	2.6266	0.9367	3.1713	1.5126
			(vi)	11.913	3.7247	2.5066	0.9255	2.2120	1.3377
Zero	TSAF	60	(i)	1.574	1.5725	0.0991	0.0950	1.0059	1.0052
			(ii)	1.574	1.5724	0.0990	0.0944	1.0059	1.0050
			(iii)	1.574	1.5724	0.0986	0.0944	1.0058	1.0050
			(iv)	1.574	1.5723	0.0991	0.0943	1.0059	1.0050
			(v)	1.574	1.5723	0.0987	0.0944	1.0059	1.0050
Zero	OSAF	120	(i)	10.076	8.2330	2.9726	2.3450	3.2998	2.7364
			(ii)	4.100	1.7721	1.5565	0.7027	1.9542	1.1830
			(iii)	4.028	1.7721	1.5350	0.7027	1.9337	1.1830
			(iv)	3.732	1.3688	1.4168	0.5574	1.8212	1.0472
			(v)	3.593	1.3688	1.3630	0.5574	1.7788	1.0472
			(vi)	3.468	1.3315	1.2932	0.4848	1.7576	1.0446
Estimated	OSAF	120	(i)	10.064	8.2214	3.0586	2.4073	3.3705	2.7850
			(ii)	4.089	1.7597	1.5782	0.7038	1.9649	1.1778
			(iii)	4.016	1.7597	1.5557	0.7038	1.9429	1.1778
			(iv)	3.719	1.3559	1.4402	0.5544	1.8349	1.0394
			(v)	3.580	1.3561	1.4049	0.5551	1.8068	1.0416
			(vi)	3.470	1.3313	1.3359	0.4804	1.7647	1.0311
Zero	TSAF	120	(i)	1.563	1.5632	0.1106	0.1094	1.0026	1.0025
			(ii)	1.563	1.5630	0.0988	0.0962	1.0012	1.0010
			(iii)	1.563	1.5630	0.0988	0.0962	1.0012	1.0010
			(iv)	1.563	1.5629	0.0987	0.0962	1.0012	1.0010
			(v)	1.563	1.5629	0.0986	0.0962	1.0012	1.0010

Table 5.29 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = -0.60$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	31.099	15.6378	9.6443	4.9262	9.8521	5.3453
			(ii)	30.971	11.9706	9.6137	3.7210	9.8193	4.1962
			(iii)	30.820	11.9669	9.5566	3.7206	9.7665	4.1958
			(iv)	30.922	11.9632	9.8751	3.7210	9.7688	4.1963
			(v)	30.820	11.9595	9.5566	3.7206	9.7665	4.1958
			(vi)	30.113	11.9558	9.0665	3.6358	9.3528	4.1137
Estimated	OSAF	30	(i)	31.096	15.6362	9.6871	4.9268	9.8919	5.3453
			(ii)	30.969	11.9686	9.6652	3.7291	9.8674	4.2059
			(iii)	30.818	11.9649	9.6007	3.7287	9.8076	4.2054
			(iv)	30.920	11.9686	9.9309	3.7291	10.1178	4.2059
			(v)	30.817	11.9653	9.6240	3.7240	9.8539	4.2070
			(vi)	30.112	11.4251	9.5729	3.7943	9.2990	4.4220
Zero	TSAF	30	(i)	1.422	1.4221	0.1501	0.1506	0.9623	0.9622
			(ii)	1.422	1.4220	0.1505	0.1507	0.9624	0.9620
			(iii)	1.422	1.4220	0.1504	0.1507	0.9624	0.9620
			(iv)	1.422	1.4220	0.1506	0.1507	0.9625	0.9620
			(v)	1.422	1.4220	0.1504	0.1507	0.9624	0.9620
Zero	OSAF	60	(i)	20.235	6.7715	2.8897	1.0641	3.5199	1.6815
			(ii)	16.831	3.1617	2.4807	0.6410	3.0916	1.2469
			(iii)	16.469	3.1615	2.4271	0.6410	3.0393	1.2469
			(iv)	17.467	3.0979	2.5637	0.6307	3.1783	1.2372
			(v)	16.466	3.0977	2.4234	0.6307	3.0362	1.2372
			(vi)	15.035	1.9161	1.2285	0.6204	1.9992	1.0142
Estimated	OSAF	60	(i)	20.235	6.7643	2.9533	1.0676	3.5708	1.6882
			(ii)	16.831	3.1535	2.5395	0.6421	3.1307	1.2466
			(iii)	16.469	3.1533	2.4862	0.6421	3.0793	1.2466
			(iv)	16.078	3.0897	2.6336	0.6330	3.2263	1.2382
			(v)	15.195	3.0892	2.5575	0.6466	3.1538	1.2503
			(vi)	14.312	2.9379	2.4294	0.5696	3.1447	1.1913
Zero	TSAF	60	(i)	1.395	1.3916	0.1384	0.1323	0.9526	0.9503
			(ii)	1.393	1.3903	0.1351	0.1314	0.9517	0.9497
			(iii)	1.393	1.3903	0.1349	0.1314	0.9516	0.9497
			(iv)	1.394	1.3902	0.1363	0.1314	0.9520	0.9497
			(v)	1.393	1.3902	0.1350	0.1314	0.9516	0.9497
Zero	OSAF	120	(i)	3.784	3.9476	0.7204	0.6885	1.2500	1.2325
			(ii)	1.148	1.0216	0.3630	0.3064	0.9054	0.8638
			(iii)	1.148	1.0216	0.3630	0.3064	0.9054	0.8638
			(iv)	1.151	1.0216	0.3670	0.3064	0.9086	0.8638
			(v)	1.148	1.0216	0.3630	0.3064	0.9054	0.8638
			(vi)	1.290	1.0187	0.2458	0.2058	0.8605	0.8579
Estimated	OSAF	120	(i)	9.011	8.8682	2.0245	1.9001	2.4663	2.3698
			(ii)	1.381	1.0527	0.5484	0.3597	1.0445	0.8914
			(iii)	1.370	1.0527	0.5371	0.3597	1.0342	0.8914
			(iv)	1.308	1.0527	0.5344	0.3597	1.0308	0.8914
			(v)	1.288	1.0527	0.5229	0.3593	1.0208	0.8910
			(vi)	1.290	1.0486	0.6511	0.3634	1.1580	0.8805
Zero	TSAF	120	(i)	1.366	1.3664	0.1467	0.1465	0.9444	0.9444
			(ii)	1.366	1.3660	0.1346	0.1348	0.9422	0.9423
			(iii)	1.366	1.3660	0.1345	0.1348	0.9422	0.9423
			(iv)	1.366	1.3660	0.1344	0.1348	0.9422	0.9423
			(v)	1.366	1.3660	0.1344	0.1348	0.9422	0.942

Table 5.30 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = -0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	45.211	15.9993	10.6286	3.8287	10.9513	4.3927
			(ii)	45.887	9.9430	10.6995	2.4298	11.0190	3.0344
			(iii)	44.579	9.9425	10.4625	2.4297	10.7900	3.0343
			(iv)	44.271	9.9420	10.2255	2.4296	10.5610	3.0342
			(v)	40.963	9.9425	10.2215	2.4297	10.7900	3.0343
			(vi)	40.855	9.5154	10.2175	1.1596	10.0190	3.0507
Estimated	OSAF	30	(i)	45.208	15.9982	10.7298	3.8563	11.0440	4.4174
			(ii)	45.883	9.9425	10.7975	2.4394	11.1098	3.0415
			(iii)	44.575	9.9420	10.5628	2.4393	10.8824	3.0414
			(iv)	44.267	9.9415	10.3281	2.4653	10.6550	3.0665
			(v)	40.259	9.9415	10.0934	2.4502	10.4276	3.0529
			(vi)	40.051	9.5210	10.0587	2.5896	10.2002	3.0468
Zero	TSAF	30	(i)	1.310	1.3047	0.1918	0.1855	0.9321	0.9291
			(ii)	1.310	1.3036	0.1927	0.1836	0.9325	0.9283
			(iii)	1.310	1.3036	0.1920	0.1836	0.9321	0.9283
			(iv)	1.312	1.3036	0.1942	0.1836	0.9332	0.9283
			(v)	1.310	1.3036	0.1920	0.1836	0.9321	0.9283
Zero	OSAF	60	(i)	14.466	3.8330	1.8102	0.6290	2.4657	1.2541
			(ii)	11.050	1.9299	1.4542	0.4041	2.0903	1.0156
			(iii)	10.438	1.9298	1.3901	0.4041	2.0171	1.0156
			(iv)	10.505	1.9414	1.6864	0.4089	2.1137	1.0198
			(v)	10.438	1.9298	1.3901	0.4041	2.0171	1.0156
			(vi)	11.503	1.8820	1.4332	0.3177	2.6425	1.0146
Estimated	OSAF	60	(i)	14.464	3.8310	1.8382	0.6390	2.4873	1.2610
			(ii)	11.048	1.9274	1.4689	0.4009	2.1000	1.0118
			(iii)	10.435	1.9273	1.4088	0.4009	2.0288	1.0118
			(iv)	10.503	1.9389	1.7071	0.4063	2.0974	1.0166
			(v)	10.435	1.9276	1.4075	0.3960	2.0362	1.0087
			(vi)	11.497	1.8795	1.5766	0.3946	2.3375	1.0168
Zero	TSAF	60	(i)	1.246	1.2436	0.1654	0.1612	0.9052	0.9036
			(ii)	1.245	1.2431	0.1639	0.1607	0.9045	0.9033
			(iii)	1.245	1.2431	0.1639	0.1607	0.9044	0.9033
			(iv)	1.245	1.2431	0.1641	0.1607	0.9047	0.9033
			(v)	1.245	1.2431	0.1639	0.1607	0.9044	0.9033
Zero	OSAF	120	(i)	1.347	1.3403	0.2631	0.2532	0.9459	0.9416
			(ii)	1.036	1.0256	0.1294	0.1168	0.8229	0.8162
			(iii)	1.036	1.0256	0.1294	0.1168	0.8229	0.8162
			(iv)	1.041	1.0256	0.1327	0.1168	0.8257	0.8162
			(v)	1.036	1.0256	0.1294	0.1168	0.8229	0.8162
			(vi)	1.137	1.0185	0.1421	0.1117	0.9178	0.8114
Estimated	OSAF	120	(i)	3.780	3.9435	0.7491	0.7204	1.2771	1.2635
			(ii)	1.145	1.0190	0.3654	0.3064	0.9065	0.8630
			(iii)	1.145	1.0190	0.3654	0.3064	0.9065	0.8630
			(iv)	1.148	1.0190	0.3699	0.3064	0.9102	0.8630
			(v)	1.145	1.0190	0.3683	0.3064	0.9090	0.8630
			(vi)	1.138	1.0185	0.3631	0.3029	0.9084	0.8614
Zero	TSAF	120	(i)	1.201	1.2016	0.1665	0.1663	0.8898	0.8898
			(ii)	1.201	1.2015	0.1634	0.1631	0.8890	0.8890
			(iii)	1.201	1.2015	0.1634	0.1631	0.8890	0.8890
			(iv)	1.201	1.2015	0.1635	0.1631	0.8890	0.8890
			(v)	1.201	1.2015	0.1634	0.1631	0.8890	0.8890

Table 5.31 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = -0.30$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	49.022	13.6756	9.4257	2.4850	9.8519	3.1200
			(ii)	48.946	8.7914	9.4039	1.6384	9.8320	2.2825
			(iii)	46.783	8.7899	8.9809	1.6383	9.4170	2.2824
			(iv)	55.749	8.7914	10.6531	1.6384	11.0531	2.2825
			(v)	46.783	8.7899	8.9809	1.6383	9.4170	2.2824
			(vi)	52.419	8.5154	9.3567	1.5473	11.6352	2.1994
Estimated	OSAF	30	(i)	49.019	13.6663	9.4979	2.5068	9.9219	3.1398
			(ii)	48.943	8.7906	9.4775	1.6484	9.9052	2.2908
			(iii)	46.780	8.7890	9.0474	1.6483	9.4834	2.2907
			(iv)	55.745	8.7906	10.7451	1.6484	11.1443	2.2908
			(v)	46.779	8.7893	9.0808	1.6489	9.5112	2.2900
			(vi)	52.425	8.3167	9.3561	1.4410	11.6161	2.0026
Zero	TSAF	30	(i)	1.229	1.2282	0.2309	0.2284	0.9129	0.9118
			(ii)	1.229	1.2267	0.2305	0.2271	0.9129	0.9111
			(iii)	1.229	1.2267	0.2312	0.2271	0.9128	0.9111
			(iv)	1.231	1.2267	0.2323	0.2271	0.9142	0.9111
			(v)	1.229	1.2267	0.2312	0.2271	0.9128	0.9111
Zero	OSAF	60	(i)	6.230	2.2616	0.7909	0.3693	1.4356	1.0043
			(ii)	5.262	1.3177	0.6803	0.2945	1.3166	0.9215
			(iii)	4.514	1.3177	0.6363	0.2945	1.2693	0.9215
			(iv)	6.913	1.3177	0.8233	0.2945	1.4701	0.9215
			(v)	4.506	1.3177	0.6342	0.2945	1.2671	0.9215
			(vi)	4.442	1.3048	0.6172	0.2813	1.2061	0.9211
Estimated	OSAF	60	(i)	6.229	2.2611	0.8053	0.3693	1.4446	1.0036
			(ii)	5.262	1.3171	0.6925	0.2965	1.3230	0.9219
			(iii)	4.513	1.3171	0.6485	0.2965	1.2745	0.9219
			(iv)	6.912	1.3171	0.8368	0.2965	1.4807	0.9219
			(v)	4.505	1.3171	0.6532	0.2980	1.2778	0.9226
			(vi)	4.416	1.3051	0.6080	0.2958	1.2056	0.9112
Zero	TSAF	60	(i)	1.108	1.1081	0.2285	0.2284	0.8674	0.8671
			(ii)	1.108	1.1081	0.2284	0.2283	0.8673	0.8671
			(iii)	1.108	1.1081	0.2284	0.2283	0.8673	0.8671
			(iv)	1.108	1.1081	0.2284	0.2283	0.8674	0.8671
			(v)	1.108	1.1081	0.2284	0.2283	0.8673	0.8671
Zero	OSAF	120	(i)	1.292	1.0204	0.3383	0.2957	0.8940	0.8610
			(ii)	1.289	1.0172	0.3366	0.2940	0.8925	0.8595
			(iii)	1.289	1.0172	0.3366	0.2940	0.8925	0.8595
			(iv)	1.289	1.0172	0.3366	0.2940	0.8925	0.8595
			(v)	1.289	1.0172	0.3366	0.2940	0.8925	0.8595
			(vi)	1.256	1.0165	0.2482	0.2730	0.7916	0.8578
Estimated	OSAF	120	(i)	1.291	1.0195	0.3418	0.2974	0.8968	0.8621
			(ii)	1.288	1.0163	0.3384	0.2940	0.8939	0.8592
			(iii)	1.288	1.0163	0.3384	0.2940	0.8939	0.8592
			(iv)	1.288	1.0163	0.3384	0.2940	0.8938	0.8592
			(v)	1.288	1.0163	0.3379	0.2940	0.8933	0.8592
			(vi)	1.256	1.0165	0.3327	0.2904	0.8704	0.8578
Zero	TSAF	120	(i)	1.093	1.0933	0.2018	0.2015	0.8604	0.8603
			(ii)	1.093	1.0933	0.2019	0.2016	0.8604	0.8604
			(iii)	1.093	1.0933	0.2019	0.2016	0.8604	0.8604
			(iv)	1.093	1.0933	0.2019	0.2016	0.8604	0.8604
			(v)	1.093	1.0933	0.2019	0.2016	0.8604	0.8604

Table 5.32 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = -0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	41.7958	7.065	6.8543	1.2837	7.4096	1.9515
			(ii)	44.7685	4.784	7.3717	0.8775	7.9270	1.5605
			(iii)	41.6070	4.775	6.7963	0.8777	7.3547	1.5590
			(iv)	57.6594	4.726	9.3246	0.8663	9.8430	1.5495
			(v)	41.6070	4.717	6.7963	0.8665	7.3547	1.5480
			(vi)	54.3667	4.566	3.5394	0.8224	10.1524	1.5792
Estimated	OSAF	30	(i)	41.7762	7.078	6.9020	1.2910	7.4576	1.9592
			(ii)	44.7450	4.778	7.4253	0.8842	7.9818	1.5666
			(iii)	41.5875	4.772	6.8430	0.8844	7.4019	1.5655
			(iv)	57.6345	4.719	9.3937	0.8723	9.9153	1.5548
			(v)	41.5964	4.713	6.8868	0.8725	7.4401	1.5552
			(vi)	54.3687	4.533	9.4428	0.8817	10.1073	1.5750
Zero	TSAF	30	(i)	1.2325	1.220	0.2639	0.2567	0.9235	0.9174
			(ii)	1.2314	1.220	0.2649	0.2565	0.9233	0.9174
			(iii)	1.2323	1.220	0.2641	0.2566	0.9234	0.9173
			(iv)	1.2371	1.220	0.2667	0.2565	0.9259	0.9174
			(v)	1.2323	1.220	0.2641	0.2566	0.9234	0.9173
Zero	OSAF	60	(i)	3.4186	1.351	0.4781	0.2983	1.0984	0.9159
			(ii)	2.4967	1.167	0.4362	0.2827	1.0579	0.9006
			(iii)	2.4863	1.167	0.4344	0.2827	1.0560	0.9006
			(iv)	3.6252	1.167	0.4990	0.2827	1.1292	0.9006
			(v)	2.4863	1.167	0.4344	0.2827	1.0560	0.9006
			(vi)	3.4519	1.166	0.0557	0.0217	1.1403	0.8985
Estimated	OSAF	60	(i)	3.4192	1.351	0.4797	0.2957	1.0993	0.9140
			(ii)	2.4970	1.168	0.4424	0.2827	1.0625	0.9008
			(iii)	2.4867	1.168	0.4397	0.2827	1.0600	0.9008
			(iv)	3.6254	1.168	0.5059	0.2827	1.1351	0.9008
			(v)	2.4865	1.168	0.4441	0.2827	1.0630	0.9008
			(vi)	3.4512	1.166	0.4714	0.2784	1.1370	0.8985
Zero	TSAF	60	(i)	1.1087	1.108	0.2285	0.2284	0.8674	0.8671
			(ii)	1.1086	1.108	0.2284	0.2283	0.8673	0.8671
			(iii)	1.1086	1.108	0.2284	0.2283	0.8673	0.8671
			(iv)	1.1088	1.108	0.2284	0.2283	0.8674	0.8671
			(v)	1.1086	1.108	0.2284	0.2283	0.8673	0.8671
Zero	OSAF	120	(i)	1.0162	1.013	0.3018	0.2897	0.8596	0.8543
			(ii)	1.0162	1.013	0.3018	0.2897	0.8596	0.8543
			(iii)	1.0162	1.013	0.3018	0.2897	0.8596	0.8543
			(iv)	1.0162	1.013	0.3018	0.2897	0.8596	0.8543
			(v)	1.0162	1.013	0.3018	0.2897	0.8596	0.8543
			(vi)	1.0165	1.013	0.3018	0.2772	0.8585	0.8533
Estimated	OSAF	120	(i)	1.0161	1.013	0.3019	0.2898	0.8596	0.8543
			(ii)	1.0161	1.013	0.3019	0.2898	0.8596	0.8543
			(iii)	1.0161	1.013	0.3019	0.2898	0.8596	0.8543
			(iv)	1.0161	1.013	0.3019	0.2898	0.8596	0.8543
			(v)	1.0161	1.013	0.3019	0.2898	0.8596	0.8543
			(vi)	1.0165	1.013	0.2986	0.2867	0.8585	0.8533
Zero	TSAF	120	(i)	1.0237	1.023	0.2422	0.2420	0.8428	0.8428
			(ii)	1.0237	1.023	0.2422	0.2420	0.8428	0.8428
			(iii)	1.0237	1.023	0.2422	0.2420	0.8428	0.8428
			(iv)	1.0237	1.023	0.2422	0.2420	0.8428	0.8428
			(v)	1.0237	1.023	0.2422	0.2420	0.8428	0.8428

Table 5.33 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = 0$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	34.206	6.1399	4.6100	0.8295	5.2309	1.5163
			(ii)	39.069	3.9950	5.1542	0.6806	5.7644	1.3597
			(iii)	34.113	3.9226	4.5905	0.6578	5.2089	1.3376
			(iv)	39.156	3.9591	4.9691	0.6701	5.5486	1.3517
			(v)	34.110	3.8867	4.5905	0.6473	5.2085	1.3296
			(vi)	49.266	3.6912	5.1954	0.6335	6.1989	1.2853
Estimated	OSAF	30	(i)	34.197	6.1411	4.6297	0.8306	5.2519	1.5187
			(ii)	39.059	3.9885	5.1766	0.6811	5.7894	1.3610
			(iii)	34.104	3.9161	4.6082	0.6583	5.2283	1.3390
			(iv)	32.200	3.9533	4.0025	0.6355	4.5862	1.3525
			(v)	34.115	3.8947	4.6150	0.6487	5.2365	1.3316
			(vi)	49.264	3.6887	5.1919	0.6328	6.1823	1.2523
Zero	TSAF	30	(i)	1.254	1.2456	0.2963	0.2910	0.9379	0.9335
			(ii)	1.256	1.2455	0.2972	0.2909	0.9387	0.9334
			(iii)	1.254	1.2453	0.2964	0.2907	0.9380	0.9333
			(iv)	1.262	1.2455	0.2986	0.2909	0.9414	0.9334
			(v)	1.254	1.2453	0.2964	0.2907	0.9380	0.9333
Zero	OSAF	60	(i)	1.766	1.1509	0.3000	0.2572	0.9423	0.8932
			(ii)	1.906	1.1509	0.3100	0.2572	0.9521	0.8932
			(iii)	1.766	1.1509	0.3000	0.2572	0.9423	0.8932
			(iv)	2.382	1.1509	0.3455	0.2572	0.9893	0.8932
			(v)	1.766	1.1509	0.3000	0.2572	0.9423	0.8932
			(vi)	2.252	1.1491	0.1243	0.0069	1.0179	0.8911
Estimated	OSAF	60	(i)	1.766	1.1511	0.2992	0.2573	0.9436	0.8933
			(ii)	1.906	1.1511	0.3083	0.2573	0.9529	0.8933
			(iii)	1.766	1.1511	0.2992	0.2573	0.9436	0.8933
			(iv)	2.382	1.1511	0.3423	0.2573	0.9889	0.8933
			(v)	1.766	1.1511	0.3030	0.2573	0.9436	0.8933
			(vi)	2.256	1.1491	0.3818	0.2527	1.0134	0.8911
Zero	TSAF	60	(i)	1.114	1.1143	0.2510	0.2507	0.8786	0.8785
			(ii)	1.114	1.1143	0.2510	0.2507	0.8786	0.8785
			(iii)	1.114	1.1143	0.2510	0.2507	0.8786	0.8785
			(iv)	1.114	1.1143	0.2512	0.2507	0.8788	0.8785
			(v)	1.114	1.1143	0.2510	0.2507	0.8786	0.8785
Zero	OSAF	120	(i)	1.016	1.0149	0.2813	0.2735	0.8533	0.8500
			(ii)	1.016	1.0149	0.2813	0.2735	0.8533	0.8500
			(iii)	1.016	1.0149	0.2813	0.2735	0.3533	0.8500
			(iv)	1.016	1.0149	0.2813	0.2735	0.8533	0.8500
			(v)	1.016	1.0149	0.2813	0.2735	0.8533	0.8500
			(vi)	1.016	1.0151	0.2804	0.2728	0.8522	0.8490
Estimated	OSAF	120	(i)	1.016	1.0150	0.2813	0.2735	0.8534	0.8501
			(ii)	1.016	1.0150	0.2813	0.2735	0.8534	0.8501
			(iii)	1.016	1.0150	0.2813	0.2735	0.8534	0.8501
			(iv)	1.016	1.0150	0.2813	0.2735	0.8534	0.8501
			(v)	1.016	1.0150	0.2813	0.2735	0.8534	0.8501
			(vi)	1.016	1.0151	0.2778	0.2700	0.8522	0.8490
Zero	TSAF	120	(i)	1.004	1.0048	0.2720	0.2719	0.8454	0.8454
			(ii)	1.004	1.0048	0.2720	0.2719	0.8454	0.8454
			(iii)	1.004	1.0048	0.2720	0.2719	0.8454	0.8454
			(iv)	1.004	1.0048	0.2720	0.2719	0.8454	0.8454
			(v)	1.004	1.0048	0.2720	0.2719	0.8454	0.8454

Table 5.34 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = 0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	34.2062	6.139	4.6100	0.8295	5.2309	1.5163
			(ii)	39.0696	3.995	5.1542	0.6806	5.7644	1.3597
			(iii)	34.1132	3.922	4.5905	0.6578	5.2089	1.3376
			(iv)	52.2106	3.959	6.9691	0.6701	7.5486	1.3517
			(v)	34.1109	3.886	4.5905	0.6473	5.2085	1.3296
			(vi)	34.5822	3.715	0.9082	0.5732	5.1504	1.2093
Estimated	OSAF	30	(i)	15.8970	3.247	2.0765	0.4439	2.7655	1.1567
			(ii)	21.8502	3.371	2.8903	0.4440	3.5608	1.1665
			(iii)	15.7666	3.176	2.0526	0.4440	2.7342	1.1483
			(iv)	36.2368	3.261	4.5169	0.4441	5.1737	1.1539
			(v)	15.7483	3.093	2.0517	0.4448	2.7306	1.1394
			(vi)	34.5498	3.049	3.4125	0.4357	5.1387	1.1229
Zero	TSAF	30	(i)	1.3184	1.308	0.3240	0.3222	0.9673	0.9632
			(ii)	1.3198	1.310	0.3253	0.3224	0.9682	0.9636
			(iii)	1.3176	1.307	0.3239	0.3221	0.9670	0.9630
			(iv)	1.3248	1.310	0.3272	0.3223	0.9704	0.9636
			(v)	1.3175	1.307	0.3238	0.3220	0.9670	0.9630
Zero	OSAF	60	(i)	1.2675	1.123	0.2710	0.2500	0.8900	0.8790
			(ii)	1.5006	1.123	0.2930	0.2500	0.9100	0.8790
			(iii)	1.2675	1.123	0.2710	0.2500	0.8900	0.8790
			(iv)	1.5006	1.123	0.2930	0.2500	0.9100	0.8790
			(v)	1.2675	1.123	0.2710	0.2500	0.8900	0.8790
			(vi)	1.4680	1.123	0.0311	0.0078	0.9097	0.8777
Estimated	OSAF	60	(i)	1.2677	1.124	0.2729	0.2500	0.8984	0.8791
			(ii)	1.5007	1.124	0.2966	0.2500	0.9219	0.8791
			(iii)	1.2677	1.124	0.2729	0.2500	0.8984	0.8791
			(iv)	1.5007	1.124	0.2966	0.2500	0.9219	0.8791
			(v)	1.2676	1.124	0.2732	0.2500	0.8987	0.8791
			(vi)	1.4669	1.123	0.2814	0.2462	0.9109	0.8777
Zero	TSAF	60	(i)	1.1509	1.150	0.2838	0.2839	0.8974	0.8973
			(ii)	1.1509	1.150	0.2838	0.2839	0.8974	0.8973
			(iii)	1.1509	1.150	0.2838	0.2839	0.8974	0.8973
			(iv)	1.1509	1.150	0.2838	0.2839	0.8974	0.8973
			(v)	1.1509	1.150	0.2838	0.2839	0.8974	0.8973
Zero	OSAF	120	(i)	1.0158	1.015	0.2841	0.2785	0.8520	0.8499
			(ii)	1.0158	1.015	0.2841	0.2785	0.8520	0.8499
			(iii)	1.0158	1.015	0.2841	0.2785	0.8520	0.8499
			(iv)	1.0158	1.015	0.2841	0.2785	0.8520	0.8499
			(v)	1.0158	1.015	0.2841	0.2785	0.8520	0.8499
			(vi)	1.0149	1.014	0.0106	0.0103	0.8504	0.8484
Estimated	OSAF	120	(i)	1.0156	1.014	0.2841	0.2785	0.8520	0.8499
			(ii)	1.0156	1.014	0.2841	0.2785	0.8520	0.8499
			(iii)	1.0156	1.014	0.2841	0.2785	0.8520	0.8499
			(iv)	1.0156	1.014	0.2841	0.2785	0.8520	0.8499
			(v)	1.0156	1.014	0.2841	0.2785	0.8520	0.8499
			(vi)	1.0143	1.013	0.2838	0.2780	0.8499	0.8356
Zero	TSAF	120	(i)	1.0280	1.028	0.3187	0.3187	0.8674	0.8673
			(ii)	1.0280	1.028	0.3187	0.3187	0.8674	0.8673
			(iii)	1.0280	1.028	0.3187	0.3187	0.8674	0.8673
			(iv)	1.0280	1.028	0.3187	0.3187	0.8674	0.8673
			(v)	1.0251	1.025	0.1248	0.1248	0.8171	0.8171

Table 5.35 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = 0.30$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	11.882	1.9547	1.4458	0.2802	2.1870	1.0366
			(ii)	17.596	1.7154	2.1054	0.2805	2.8146	1.0066
			(iii)	11.641	1.6775	1.4461	0.2805	2.1564	1.0014
			(iv)	27.823	1.6645	3.1332	0.2804	3.8626	0.9989
			(v)	11.589	1.6266	1.4461	0.2805	2.1496	0.9937
			(vi)	26.472	1.5955	2.0029	0.2716	2.9400	0.9826
Estimated	OSAF	30	(i)	11.888	1.9487	1.4507	0.2805	2.1889	1.0335
			(ii)	17.603	1.7054	2.1201	0.2808	2.8255	1.0043
			(iii)	11.651	1.6776	1.4510	0.2808	2.1598	1.0000
			(iv)	27.819	1.6598	3.1593	0.2807	3.8835	0.9980
			(v)	11.667	1.5046	1.4454	0.2800	2.1522	0.9918
			(vi)	16.430	1.4598	2.0003	0.2635	2.8430	0.9715
Zero	TSAF	30	(i)	1.431	1.4249	0.3516	0.3494	1.0106	1.0082
			(ii)	1.431	1.4220	0.3524	0.3490	1.0112	1.0071
			(iii)	1.428	1.4224	0.3511	0.3489	1.0098	1.0073
			(iv)	1.435	1.4220	0.3552	0.3489	1.0128	1.0071
			(v)	1.428	1.4224	0.3511	0.3488	1.0098	1.0073
Zero	OSAF	60	(i)	1.297	1.2865	0.2589	0.2500	0.9026	0.8965
			(ii)	1.147	1.1359	0.2583	0.2494	0.8873	0.8811
			(iii)	1.147	1.1359	0.2583	0.2494	0.8873	0.8811
			(iv)	1.132	1.1210	0.2582	0.2493	0.8850	0.8788
			(v)	1.132	1.1210	0.2582	0.2493	0.8850	0.8788
			(vi)	1.131	1.1208	0.2538	0.2451	0.8838	0.8777
Estimated	OSAF	60	(i)	1.349	1.3381	0.2590	0.2501	0.9037	0.8976
			(ii)	1.141	1.1307	0.2584	0.2495	0.8866	0.8804
			(iii)	1.141	1.1307	0.2584	0.2495	0.8866	0.8804
			(iv)	1.131	1.1202	0.2583	0.2494	0.8847	0.8786
			(v)	1.131	1.1202	0.2583	0.2494	0.8847	0.8786
			(vi)	1.131	1.1208	0.2538	0.2451	0.8838	0.8777
Zero	TSAF	60	(i)	1.246	1.2465	0.3243	0.3235	0.9435	0.9433
			(ii)	1.244	1.2437	0.3226	0.3217	0.9419	0.9417
			(iii)	1.244	1.2437	0.3226	0.3217	0.9419	0.9417
			(iv)	1.244	1.2436	0.3224	0.3215	0.9418	0.9416
			(v)	1.244	1.2436	0.3224	0.3215	0.9418	0.9416
Zero	OSAF	120	(i)	1.011	1.0104	0.2755	0.2718	0.8473	0.8459
			(ii)	1.011	1.0104	0.2755	0.2718	0.8473	0.8459
			(iii)	1.011	1.0104	0.2755	0.2718	0.8473	0.8459
			(iv)	1.011	1.0104	0.2755	0.2718	0.8473	0.8459
			(v)	1.011	1.0104	0.2755	0.2718	0.8473	0.8459
			(vi)	1.008	1.0071	0.0078	0.2685	0.8448	0.8434
Estimated	OSAF	120	(i)	1.010	1.0097	0.2754	0.2716	0.8470	0.8455
			(ii)	1.010	1.0097	0.2754	0.2716	0.8470	0.8455
			(iii)	1.010	1.0097	0.2754	0.2716	0.8470	0.8455
			(iv)	1.010	1.0097	0.2754	0.2716	0.8470	0.8455
			(v)	1.010	1.0097	0.2754	0.2716	0.8470	0.8455
			(vi)	1.007	1.0071	0.2723	0.2685	0.8448	0.8434
Zero	TSAF	120	(i)	1.092	1.0928	0.3524	0.3525	0.9032	0.9032
			(ii)	1.092	1.0928	0.3524	0.3525	0.9032	0.9032
			(iii)	1.092	1.0928	0.3524	0.3525	0.9032	0.9032
			(iv)	1.092	1.0928	0.3524	0.3525	0.9032	0.9032
			(v)	1.092	1.0928	0.3524	0.3525	0.9032	0.9032

Table 5.36 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = 0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	8.6220	2.545	0.8876	0.3165	1.7110	1.1438
			(ii)	5.8512	1.982	0.6508	0.2943	1.2543	1.0544
			(iii)	8.0746	1.942	0.8904	0.2943	1.6316	1.0493
			(iv)	9.8400	1.884	0.9492	0.2944	1.6783	1.0397
			(v)	5.0094	1.844	0.8905	0.2944	1.1216	1.0347
			(vi)	20.0168	1.759	2.8915	0.2504	2.8451	1.0272
Estimated	OSAF	30	(i)	8.6186	2.556	0.8880	0.3157	1.7066	1.1399
			(ii)	5.8326	1.962	0.5088	0.2935	1.2495	1.0481
			(iii)	8.0427	2.009	0.8510	0.2934	1.6254	1.0417
			(iv)	9.8453	2.895	0.9579	0.2936	1.7845	1.0370
			(v)	5.0042	1.866	0.4934	0.2930	1.1239	1.0347
			(vi)	4.9897	1.733	0.3937	0.2479	1.0042	1.0145
Zero	TSAF	30	(i)	1.6220	1.615	0.4104	0.4081	1.0883	1.0862
			(ii)	1.6121	1.602	0.4083	0.4063	1.0856	1.0823
			(iii)	1.6059	1.601	0.4077	0.4061	1.0836	1.0820
			(iv)	1.6156	1.602	0.4094	0.4063	1.0869	1.0822
			(v)	1.6055	1.601	0.4078	0.4061	1.0836	1.0819
Zero	OSAF	60	(i)	1.8053	1.800	0.2520	0.2475	0.9661	0.9632
			(ii)	1.1803	1.174	0.2472	0.2428	0.8888	0.8856
			(iii)	1.1803	1.174	0.2472	0.2428	0.8888	0.8856
			(iv)	1.1760	1.171	0.2473	0.2428	0.8877	0.8849
			(v)	1.1760	1.171	0.2473	0.2428	0.8877	0.8849
			(vi)	1.1510	1.146	0.0044	0.0041	0.8843	0.8815
Estimated	OSAF	60	(i)	1.7910	1.786	0.2521	0.2476	0.9662	0.9633
			(ii)	1.1704	1.163	0.2473	0.2429	0.8880	0.8847
			(iii)	1.1704	1.163	0.2473	0.2429	0.8880	0.8847
			(iv)	1.1643	1.159	0.2473	0.2429	0.8866	0.8837
			(v)	1.1650	1.160	0.2473	0.2429	0.8866	0.8837
			(vi)	1.1512	1.146	0.2432	0.2388	0.8843	0.8815
Zero	TSAF	60	(i)	1.4116	1.411	0.3638	0.3643	1.0112	1.0112
			(ii)	1.3932	1.392	0.3526	0.3532	1.0022	1.0022
			(iii)	1.3932	1.392	0.3526	0.3532	1.0022	1.0022
			(iv)	1.3931	1.392	0.3527	0.3532	1.0022	1.0022
			(v)	1.3931	1.392	0.3527	0.3532	1.0022	1.0022
Zero	OSAF	120	(i)	1.2517	1.251	0.2786	0.2760	0.8813	0.8804
			(ii)	1.0255	1.025	0.2696	0.2669	0.8511	0.8502
			(iii)	1.0255	1.025	0.2696	0.2669	0.8511	0.8502
			(iv)	1.0255	1.025	0.2696	0.2669	0.8511	0.8502
			(v)	1.0255	1.025	0.2696	0.2669	0.8511	0.8502
			(vi)	1.0232	1.022	0.0202	0.0202	0.8494	0.8485
Estimated	OSAF	120	(i)	1.2475	1.247	0.2785	0.2758	0.8804	0.8794
			(ii)	1.0227	1.022	0.2695	0.2668	0.8500	0.8491
			(iii)	1.0227	1.022	0.2695	0.2668	0.8500	0.8491
			(iv)	1.0227	1.022	0.2695	0.2668	0.8500	0.8491
			(v)	1.0227	1.022	0.2695	0.2668	0.8500	0.8491
			(vi)	1.0232	1.022	0.2669	0.2642	0.8494	0.8485
Zero	TSAF	120	(i)	1.2313	1.231	0.4061	0.4059	0.9766	0.9766
			(ii)	1.2208	1.220	0.3864	0.3862	0.9593	0.9592
			(iii)	1.2208	1.220	0.3864	0.3862	0.9593	0.9592
			(iv)	1.2208	1.220	0.3864	0.3862	0.9593	0.9592
			(v)	1.2208	1.220	0.3864	0.3862	0.9593	0.9592

Table 5.37 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = 0.60$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	6.2646	2.787	0.5300	0.2660	1.4473	1.1945
			(ii)	7.5672	2.218	0.7104	0.2681	1.5431	1.0925
			(iii)	5.6109	2.157	0.5321	0.2681	1.3400	1.0821
			(iv)	8.0324	1.962	0.8881	0.2683	1.6855	1.0567
			(v)	5.4227	1.894	0.5323	0.2684	1.3066	1.0447
			(vi)	10.7773	1.169	1.0949	0.2413	1.9588	1.0069
Estimated	OSAF	30	(i)	6.3433	2.868	0.5284	0.2661	1.4543	1.2023
			(ii)	7.5532	2.217	0.7101	0.2682	1.5417	1.0916
			(iii)	5.5966	2.152	0.5305	0.2681	1.3375	1.0807
			(iv)	8.0159	1.954	0.8861	0.2684	1.6823	1.0559
			(v)	5.4557	1.926	0.5312	0.2685	1.3163	1.0545
			(vi)	10.6804	1.017	0.9151	0.2369	1.9497	1.0025
Zero	TSAF	30	(i)	1.8329	1.825	0.4309	0.4301	1.1549	1.1527
			(ii)	1.8221	1.813	0.4294	0.4290	1.1515	1.1487
			(iii)	1.8196	1.812	0.4295	0.4288	1.1509	1.1485
			(iv)	1.8217	1.812	0.4291	0.4287	1.1512	1.1484
			(v)	1.8187	1.811	0.4291	0.4286	1.1505	1.1483
Zero	OSAF	60	(i)	2.2223	2.218	0.2468	0.2441	1.0672	1.0652
			(ii)	1.3060	1.241	0.2360	0.2335	0.9025	0.8933
			(iii)	1.3070	1.241	0.2362	0.2335	0.9029	0.8933
			(iv)	1.1776	1.174	0.2361	0.2335	0.8853	0.8836
			(v)	1.1787	1.174	0.2364	0.2335	0.8857	0.8836
			(vi)	1.2097	1.207	0.0064	0.0060	0.8856	0.8842
Estimated	OSAF	60	(i)	2.3391	2.335	0.2466	0.2439	1.0863	1.0844
			(ii)	1.2788	1.217	0.2357	0.2332	0.8979	0.8898
			(iii)	1.2799	1.217	0.2359	0.2332	0.8983	0.8898
			(iv)	1.1569	1.153	0.2359	0.2332	0.8813	0.8795
			(v)	1.1687	1.164	0.2361	0.2332	0.8833	0.8812
			(vi)	1.2038	1.202	0.2328	0.2303	0.8851	0.8838
Zero	TSAF	60	(i)	1.5679	1.567	0.4049	0.4045	1.0721	1.0719
			(ii)	1.5304	1.531	0.3753	0.3757	1.0509	1.0513
			(iii)	1.5315	1.531	0.3759	0.3757	1.0515	1.0513
			(iv)	1.5300	1.530	0.3754	0.3756	1.0508	1.0511
			(v)	1.5311	1.530	0.3759	0.3756	1.0513	1.0511
Zero	OSAF	120	(i)	2.6333	2.620	0.3014	0.3002	1.0713	1.0686
			(ii)	1.0518	1.063	0.2536	0.2522	0.8540	0.8534
			(iii)	1.0330	1.063	0.2535	0.2521	0.8501	0.8534
			(iv)	1.0518	1.033	0.2536	0.2521	0.8540	0.8497
			(v)	1.0330	1.033	0.2535	0.2521	0.8501	0.8497
			(vi)	1.0646	1.027	0.2083	0.2088	0.8539	0.8469
Estimated	OSAF	120	(i)	2.3116	2.296	0.3013	0.3001	1.0458	1.0427
			(ii)	1.0469	1.049	0.2535	0.2520	0.8526	0.8507
			(iii)	1.0274	1.049	0.2534	0.2520	0.8483	0.8507
			(iv)	1.0469	1.027	0.2535	0.2520	0.8526	0.8478
			(v)	1.0274	1.027	0.2534	0.2520	0.8483	0.8478
			(vi)	1.0703	1.027	0.2506	0.2490	0.8546	0.8469
Zero	TSAF	120	(i)	1.4455	1.445	0.5121	0.5124	1.1099	1.1099
			(ii)	1.3828	1.382	0.4046	0.4043	1.0224	1.0223
			(iii)	1.3826	1.382	0.4041	0.4043	1.0223	1.0223
			(iv)	1.3828	1.382	0.4046	0.4041	1.0224	1.0223
			(v)	1.3826	1.382	0.4041	0.4041	1.0223	1.0223

Table 5.38 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = 0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	6.346	4.0516	0.5375	0.3350	1.5582	1.3602
			(ii)	11.195	2.3851	0.8393	0.2488	1.7769	1.1797
			(iii)	5.968	2.3784	0.5403	0.2491	1.4771	1.1779
			(iv)	12.327	2.0753	0.9425	0.2499	1.8017	1.1117
			(v)	5.613	2.0465	0.5417	0.2503	1.4003	1.1062
			(vi)	18.612	2.0035	1.1258	0.6023	2.6294	1.0032
Estimated	OSAF	30	(i)	6.332	4.0179	0.5402	0.3356	1.5526	1.3489
			(ii)	11.166	2.3008	0.8437	0.2487	1.7671	1.1543
			(iii)	5.946	2.2925	0.5430	0.2490	1.4668	1.1521
			(iv)	12.280	2.0169	0.9488	0.2497	1.7905	1.0930
			(v)	5.720	2.1493	0.5450	0.2502	1.4201	1.1228
			(vi)	18.300	1.8250	0.9976	0.2396	2.6086	1.0037
Zero	TSAF	30	(i)	2.068	2.0602	0.4468	0.4443	1.2254	1.2225
			(ii)	2.063	2.0525	0.4476	0.4431	1.2241	1.2203
			(iii)	2.063	2.0533	0.4465	0.4430	1.2237	1.2204
			(iv)	2.062	2.0512	0.4474	0.4428	1.2241	1.2198
			(v)	2.063	2.0520	0.4463	0.4428	1.2236	1.2199
Zero	OSAF	60	(i)	2.032	2.0665	0.2251	0.2234	1.0846	1.0913
			(ii)	1.440	1.4678	0.2196	0.2183	0.9505	0.9556
			(iii)	1.437	1.4557	0.2196	0.2183	0.9495	0.9527
			(iv)	1.307	1.3136	0.2195	0.2179	0.9180	0.9188
			(v)	1.299	1.2982	0.2193	0.2178	0.9158	0.9150
			(vi)	2.309	2.3768	0.0065	0.0065	1.0552	1.0609
Estimated	OSAF	60	(i)	1.929	1.9564	0.2249	0.2233	1.0659	1.0711
			(ii)	1.399	1.4044	0.2196	0.2182	0.9393	0.9402
			(iii)	1.393	1.3915	0.2196	0.2183	0.9378	0.9371
			(iv)	1.259	1.2641	0.2194	0.2179	0.9061	0.9066
			(v)	1.294	1.2927	0.2192	0.2178	0.9126	0.9116
			(vi)	2.404	2.4329	0.2160	0.2144	1.0678	1.0709
Zero	TSAF	60	(i)	1.767	1.7662	0.4027	0.4017	1.1303	1.1296
			(ii)	1.746	1.7454	0.3852	0.3846	1.1187	1.1182
			(iii)	1.746	1.7458	0.3852	0.3847	1.1188	1.1184
			(iv)	1.746	1.7450	0.3845	0.3837	1.1186	1.1181
			(v)	1.745	1.7447	0.3840	0.3834	1.1183	1.1178
Zero	OSAF	120	(i)	2.976	2.9625	0.3263	0.3259	1.1871	1.1842
			(ii)	1.194	1.1825	0.2441	0.2443	0.8780	0.8748
			(iii)	1.183	1.1825	0.2438	0.2443	0.8757	0.8748
			(iv)	1.082	1.0575	0.2425	0.2415	0.8563	0.8510
			(v)	1.071	1.0575	0.2423	0.2415	0.8540	0.8510
			(vi)	1.665	1.4207	0.0023	0.0023	0.9063	0.8827
Estimated	OSAF	120	(i)	2.530	2.5229	0.3262	0.3259	1.1323	1.1303
			(ii)	1.167	1.1544	0.2441	0.2444	0.8721	0.8692
			(iii)	1.152	1.1544	0.2438	0.2444	0.8692	0.8692
			(iv)	1.071	1.0506	0.2426	0.2416	0.8518	0.8473
			(v)	1.045	1.0405	0.2423	0.2416	0.8474	0.8457
			(vi)	1.679	1.4192	0.2397	0.2387	0.9058	0.8819
Zero	TSAF	120	(i)	1.655	1.6555	0.6100	0.6100	1.2066	1.2065
			(ii)	1.584	1.5844	0.4295	0.4310	1.0899	1.0909
			(iii)	1.584	1.5844	0.4289	0.4310	1.0894	1.0909
			(iv)	1.583	1.5829	0.4257	0.4246	1.0880	1.0874
			(v)	1.583	1.5829	0.4251	0.4246	1.0875	1.0870

Table 5.39 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_3 design matrix when $\gamma = 0.90$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	3.531	2.0996	0.4291	0.2506	1.3346	1.1669
			(ii)	4.325	1.9879	0.4965	0.2513	1.3802	1.1362
			(iii)	3.439	1.9778	0.4300	0.2515	1.3090	1.1334
			(iv)	4.166	1.8328	0.4980	0.2529	1.3352	1.0922
			(v)	3.277	1.8214	0.4317	0.2534	1.2629	1.0890
			(vi)	19.407	1.5682	1.9166	0.2321	3.5614	1.0064
Estimated	OSAF	30	(i)	3.164	1.7058	0.4310	0.2506	1.2357	1.0590
			(ii)	3.978	1.6390	0.4989	0.2513	1.2870	1.0397
			(iii)	3.107	1.6375	0.4319	0.2515	1.2190	1.0392
			(iv)	3.869	1.8019	0.5005	0.2529	1.2552	1.0255
			(v)	3.308	1.8545	0.4334	0.2534	1.2730	1.0979
			(vi)	18.636	1.3783	1.6817	0.2235	3.3151	1.0093
Zero	TSAF	30	(i)	2.280	2.2795	0.4822	0.4818	1.2924	1.2918
			(ii)	2.282	2.2809	0.4818	0.4811	1.2929	1.2922
			(iii)	2.282	2.2811	0.4817	0.4811	1.2928	1.2922
			(iv)	2.283	2.2815	0.4819	0.4813	1.2930	1.2925
			(v)	2.283	2.2816	0.4818	0.4814	1.2930	1.2925
Zero	OSAF	60	(i)	1.975	1.9288	0.1295	0.1297	1.1046	1.0910
			(ii)	1.783	1.8257	0.1305	0.1305	1.0504	1.0600
			(iii)	1.782	1.8193	0.1305	0.1306	1.0501	1.0578
			(iv)	1.542	1.5101	0.1307	0.1309	0.9781	0.9682
			(v)	1.540	1.5028	0.1308	0.1312	0.9773	0.9657
			(vi)	2.724	1.6658	0.2139	0.1438	1.8379	0.8285
Estimated	OSAF	60	(i)	1.431	1.4351	0.1941	0.1938	0.9633	0.9645
			(ii)	1.353	1.3635	0.1950	0.1938	0.9384	0.9409
			(iii)	1.353	1.3632	0.1949	0.1938	0.9382	0.9407
			(iv)	1.230	1.2285	0.1946	0.1942	0.8992	0.8984
			(v)	1.432	1.4291	0.1945	0.1942	0.9584	0.9571
			(vi)	2.421	1.5858	0.2539	0.2038	0.9979	0.9870
Zero	TSAF	60	(i)	1.973	1.9731	0.3738	0.3739	1.1799	1.1798
			(ii)	1.970	1.9704	0.3729	0.3715	1.1784	1.1781
			(iii)	1.970	1.9705	0.3727	0.3715	1.1783	1.1780
			(iv)	1.970	1.9698	0.3698	0.3697	1.1775	1.1774
			(v)	1.970	1.9698	0.3696	0.3697	1.1775	1.1774
Zero	OSAF	120	(i)	2.102	2.0383	0.2500	0.2492	1.1365	1.1204
			(ii)	1.711	1.7484	0.2280	0.2310	1.0348	1.0441
			(iii)	1.708	1.7484	0.2281	0.2310	1.0338	1.0441
			(iv)	1.376	1.2707	0.2129	0.2123	0.9413	0.9113
			(v)	1.373	1.2707	0.2129	0.2123	0.9403	0.9113
			(vi)	2.469	1.9314	0.0049	0.0056	2.0539	1.6564
Estimated	OSAF	120	(i)	1.530	1.5040	0.2499	0.2491	0.9927	0.9850
			(ii)	1.353	1.3636	0.2279	0.2309	0.9382	0.9415
			(iii)	1.353	1.3636	0.2280	0.2309	0.9382	0.9415
			(iv)	1.282	1.1309	0.2128	0.2122	0.8833	0.8678
			(v)	1.324	1.2277	0.2129	0.2122	0.9236	0.8961
			(vi)	2.382	1.1892	0.2233	0.2121	0.9533	0.6568
Zero	TSAF	120	(i)	1.833	1.8332	0.4939	0.4913	1.1912	1.1899
			(ii)	1.825	1.8257	0.4451	0.4510	1.1671	1.1695
			(iii)	1.825	1.8257	0.4452	0.4510	1.1671	1.1695
			(iv)	1.818	1.8178	0.4101	0.4076	1.1519	1.1506
			(v)	1.818	1.8178	0.4102	0.4076	1.1519	1.1506

Table 5.40 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_2 design matrix when $\gamma = -0.90$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	415.290	287.2989	14.7023	10.8071	16.1352	11.9757
			(ii)	415.806	235.7341	14.7271	9.1396	16.1598	10.1825
			(iii)	415.290	235.7159	14.7023	9.1391	16.1352	10.1820
			(iv)	416.529	235.7344	14.7630	9.1396	16.1954	10.1825
			(v)	415.290	235.7161	14.7023	9.1391	16.1352	10.1820
			(vi)	376.859	213.5615	9.1443	8.0042	15.4345	9.7371
Estimated	OSAF	30	(i)	415.288	287.2873	14.6966	10.8099	16.1021	11.9618
			(ii)	415.805	235.7174	14.7214	9.1232	16.1267	10.1641
			(iii)	415.288	235.6994	14.6966	9.1227	16.1021	10.1636
			(iv)	416.527	235.7177	14.7564	9.1232	16.1618	10.1641
			(v)	415.289	235.6971	14.7001	9.1337	16.1353	10.1876
			(vi)	376.854	213.5570	14.1654	9.0020	15.4010	9.7137
Zero	TSAF	30	(i)	2.305	2.3689	0.5624	0.5992	1.3080	1.3349
			(ii)	2.305	2.3733	0.5625	0.6024	1.3081	1.3372
			(iii)	2.305	2.3733	0.5624	0.6024	1.3080	1.3372
			(iv)	2.303	2.3733	0.5609	0.6024	1.3069	1.3372
			(v)	2.305	2.3733	0.5624	0.6024	1.3080	1.3372
Zero	OSAF	60	(i)	529.022	327.2419	19.9814	13.6542	20.1037	13.8045
			(ii)	527.793	292.4316	19.9151	12.5378	20.0383	12.6968
			(iii)	525.282	292.3532	19.8410	12.5370	19.9644	12.6960
			(iv)	527.141	284.8694	19.8826	12.3133	20.0064	12.4719
			(v)	524.630	283.8677	19.8084	12.2834	19.9324	12.4419
			(vi)	476.859	258.1472	17.1535	10.7947	19.0462	11.9519
Estimated	OSAF	60	(i)	529.020	327.2488	19.9810	13.6648	20.0743	13.7959
			(ii)	527.791	292.4396	19.9161	12.5555	20.0104	12.6953
			(iii)	525.281	292.3609	19.8429	12.5547	19.9374	12.6945
			(iv)	527.139	284.8731	19.8850	12.3309	19.9797	12.4704
			(v)	524.636	283.8755	19.7841	12.3122	19.9188	12.4711
			(vi)	474.954	248.2223	16.1431	10.3022	18.1232	11.3112
Zero	TSAF	60	(i)	2.153	2.1636	0.6570	0.6657	1.3173	1.3230
			(ii)	2.153	2.1626	0.6560	0.6633	1.3169	1.3221
			(iii)	2.154	2.1626	0.6567	0.6633	1.3174	1.3221
			(iv)	2.153	2.1624	0.6555	0.6638	1.3167	1.3219
			(v)	2.154	2.1624	0.6562	0.6637	1.3172	1.3219
Zero	OSAF	120	(i)	15.706	10.7317	3.7502	2.7693	3.9874	3.0561
			(ii)	15.361	7.7516	3.6876	2.1786	3.9274	2.4914
			(iii)	15.302	7.7511	3.6685	2.1786	3.9103	2.4913
			(iv)	15.194	6.0234	3.6607	1.8499	3.9014	2.1769
			(v)	15.135	6.0228	3.6415	1.8498	3.8843	2.1768
			(vi)	10.859	5.1472	2.1535	1.7947	2.0462	1.9519
Estimated	OSAF	120	(i)	15.703	10.7141	3.7577	2.7521	3.9452	3.0069
			(ii)	15.358	7.7276	3.6921	2.1507	3.8838	2.4431
			(iii)	15.300	7.7271	3.6742	2.1506	3.8674	2.4430
			(iv)	15.191	5.9962	3.6610	1.8155	3.8554	2.1271
			(v)	15.132	5.9946	3.7470	1.8673	3.9726	2.1787
			(vi)	13.802	5.5454	1.7310	1.2745	3.7617	2.0894
Zero	TSAF	120	(i)	1.827	1.8270	0.1089	0.1080	1.0834	1.0833
			(ii)	1.827	1.8271	0.1090	0.1091	1.0835	1.0834
			(iii)	1.827	1.8271	0.1089	0.1091	1.0835	1.0834
			(iv)	1.827	1.8270	0.1089	0.1091	1.0835	1.0834
			(v)	1.827	1.8270	0.1088	0.1091	1.0835	1.0834

Table 5.41 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = -0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAFE ML	AMAFE MML
Zero	OSAF	30	(i)	786.168	495.6083	21.1944	13.8461	22.6228	14.9280
			(ii)	784.154	379.5085	21.1824	10.8185	22.6057	11.6972
			(iii)	783.863	379.4439	21.1518	10.8177	22.5753	11.6965
			(iv)	790.665	379.5089	21.4063	10.8185	22.8292	11.6972
			(v)	783.863	379.4443	21.1518	10.8178	22.5753	11.6965
			(vi)	715.159	343.5661	20.1228	9.5244	21.8284	11.2056
Estimated	OSAF	30	(i)	786.153	495.5887	21.2813	13.8902	22.6763	14.9519
			(ii)	784.140	379.4824	21.2659	10.8447	22.6575	11.7135
			(iii)	783.848	379.4174	21.2361	10.8439	22.6279	11.7128
			(iv)	790.651	379.4829	21.4889	10.8447	22.8807	11.7136
			(v)	783.856	379.4216	21.2154	10.8391	22.6486	11.7201
			(vi)	715.093	343.5408	20.1245	9.5151	21.8055	11.1986
Zero	TSAF	30	(i)	2.338	2.4053	0.7901	0.8143	1.3885	1.4089
			(ii)	2.337	2.4190	0.7893	0.8205	1.3880	1.4132
			(iii)	2.337	2.4190	0.7901	0.8205	1.3883	1.4132
			(iv)	2.332	2.4190	0.7882	0.8205	1.3864	1.4132
			(v)	2.337	2.4190	0.7901	0.8205	1.3883	1.4132
Zero	OSAF	60	(i)	981.828	405.4571	25.5036	12.4933	25.6462	12.6590
			(ii)	925.802	261.2874	24.3510	9.4076	24.4980	9.5783
			(iii)	915.928	261.2584	24.0997	9.4072	24.2449	9.5779
			(iv)	918.955	258.5121	24.1742	9.3637	24.3210	9.5337
			(v)	905.251	254.4298	23.8191	9.2569	23.9644	9.4276
			(vi)	831.562	234.7358	20.2115	8.6696	23.2090	9.1408
Estimated	OSAF	60	(i)	981.812	405.4301	25.5905	12.5091	25.7132	12.6622
			(ii)	925.787	261.2632	24.4415	9.4201	24.5708	9.5817
			(iii)	915.912	261.2340	24.1873	9.4197	24.3135	9.5813
			(iv)	918.940	258.4877	24.2640	9.3769	24.3941	9.5376
			(v)	905.247	254.4147	23.8397	9.2459	24.0038	9.4177
			(vi)	823.321	224.8562	20.1201	8.4021	22.1216	8.1218
Zero	TSAF	60	(i)	2.222	2.2075	0.9567	0.9383	1.4534	1.4428
			(ii)	2.215	2.1954	0.9503	0.9303	1.4496	1.4361
			(iii)	2.215	2.1954	0.9517	0.9303	1.4498	1.4361
			(iv)	2.213	2.1950	0.9475	0.9302	1.4479	1.4360
			(v)	2.214	2.1941	0.9496	0.9294	1.4485	1.4353
Zero	OSAF	120	(i)	45.568	28.1570	5.2596	3.4697	5.5335	3.7863
			(ii)	34.109	5.6666	4.2023	1.3589	4.4992	1.7243
			(iii)	32.337	5.6666	4.0081	1.3589	4.3067	1.7243
			(iv)	30.008	3.1294	3.8051	1.1151	4.1094	1.4821
			(v)	28.046	3.1293	3.5885	1.1151	3.8956	1.4821
			(vi)	28.209	2.9389	3.2524	1.0087	3.0222	0.9587
Estimated	OSAF	120	(i)	45.554	28.1320	5.3993	3.5456	5.6368	3.8394
			(ii)	34.094	5.6404	4.2872	1.3697	4.5592	1.7269
			(iii)	32.322	5.6404	4.0922	1.3697	4.3662	1.7269
			(iv)	29.994	3.0984	3.8883	1.1219	4.1660	1.4818
			(v)	28.033	3.0983	3.6596	1.1251	3.9545	1.4857
			(vi)	27.209	2.9389	2.8524	1.0087	2.7222	0.9587
Zero	TSAF	120	(i)	1.594	1.5939	0.2344	0.2316	1.0325	1.0320
			(ii)	1.593	1.5921	0.2300	0.2244	1.0312	1.0301
			(iii)	1.593	1.5921	0.2296	0.2244	1.0312	1.0301
			(iv)	1.593	1.5922	0.2297	0.2251	1.0311	1.0303
			(v)	1.593	1.5922	0.2292	0.2251	1.0311	1.0303

Table 5.42 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = -0.60$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	1339.0160	666.8769	27.3612	14.8021	29.0131	15.7799
			(ii)	1316.9830	428.2572	27.0279	10.9622	28.6640	11.7341
			(iii)	1311.6200	428.1576	26.9382	10.9614	28.5747	11.7332
			(iv)	1365.3030	427.9450	27.9213	10.9508	29.6043	11.7226
			(v)	1311.6200	427.8442	26.9382	10.9500	28.5747	11.7218
			(vi)	1235.0800	387.9871	24.7572	9.6551	26.3072	10.9582
Estimated	OSAF	30	(i)	1338.9960	666.8354	27.4643	14.8707	29.1045	15.8288
			(ii)	1316.9630	428.2079	27.1324	11.0049	28.7564	11.7680
			(iii)	1311.6010	428.1086	27.0391	11.0041	28.6641	11.7671
			(iv)	1365.2840	427.8961	28.0296	10.9937	29.6955	11.7567
			(v)	1311.6150	427.7961	26.9926	10.9800	28.6663	11.7494
			(vi)	1234.9480	387.9642	24.3916	9.6151	26.3567	10.9645
Zero	TSAF	30	(i)	2.7095	2.7211	1.0601	1.0724	1.5587	1.5658
			(ii)	2.7066	2.6993	1.0587	1.0644	1.5571	1.5581
			(iii)	2.7058	2.6993	1.0586	1.0644	1.5570	1.5581
			(iv)	2.7047	2.6993	1.0585	1.0645	1.5563	1.5581
			(v)	2.7058	2.6993	1.0586	1.0645	1.5570	1.5581
Zero	OSAF	60	(i)	1221.3470	474.0401	23.7341	11.0820	23.9057	11.2692
			(ii)	966.9406	133.9338	19.6867	5.4036	19.8611	5.5987
			(iii)	931.8440	133.9133	19.1017	5.4035	19.2769	5.5986
			(iv)	1001.5720	126.2411	20.3727	5.3295	20.5453	5.5244
			(v)	924.2878	126.2142	.3.9299	5.3293	19.1043	5.5242
			(vi)	906.4593	115.2288	17.0161	4.8306	18.6582	4.3536
Estimated	OSAF	60	(i)	1221.326	473.9949	23.8192	11.0955	23.9869	11.2806
			(ii)	966.9212	133.8940	19.7526	5.4136	19.9219	5.6040
			(iii)	931.8265	133.8734	19.1656	5.4135	19.3357	5.6039
			(iv)	1001.550	126.2026	20.4397	5.3390	20.6072	5.5295
			(v)	924.2795	126.1962	18.9743	5.3444	19.1616	5.5356
			(vi)	906.2593	112.0080	17.0151	4.8205	18.2542	3.9536
Zero	TSAF	60	(i)	2.5535	2.4939	1.2660	1.2307	1.6593	1.6311
			(ii)	2.5112	2.4414	1.2436	1.2069	1.6408	1.6093
			(iii)	2.5058	2.4414	1.2400	1.2069	1.6379	1.6093
			(iv)	2.5129	2.4414	1.2454	1.2069	1.6420	1.6093
			(v)	2.5050	2.4414	1.2396	1.2069	1.6374	1.6093
Zero	OSAF	120	(i)	49.1098	39.9043	3.9160	3.1295	4.2336	3.4751
			(ii)	11.7598	2.0928	1.7464	0.9245	2.0984	1.2997
			(iii)	11.0954	2.0928	1.6857	0.9245	2.0348	1.2997
			(iv)	12.7675	1.7224	1.8018	0.8981	2.1518	1.2734
			(v)	10.7909	1.7225	1.6569	0.8981	2.0062	1.2735
			(vi)	11.5781	1.4884	1.2458	0.8207	2.1536	1.2263
Estimated	OSAF	120	(i)	49.0883	39.8932	3.9960	3.1832	4.3122	3.5317
			(ii)	11.7531	2.0874	1.7701	0.9222	2.1139	1.2960
			(iii)	11.0794	2.0875	1.7080	0.9222	2.0503	1.2960
			(iv)	12.7601	1.7173	1.8221	0.8973	2.1648	1.2711
			(v)	10.7740	1.7173	1.6794	0.8969	2.0239	1.2708
			(vi)	11.6781	1.2884	1.5597	0.8707	2.0136	1.2463
Zero	TSAF	120	(i)	1.4334	1.4331	0.3488	0.3471	1.0137	1.0133
			(ii)	1.4295	1.4292	0.3407	0.3396	1.0096	1.0093
			(iii)	1.4295	1.4292	0.3412	0.3396	1.0097	1.0093
			(iv)	1.4295	1.4292	0.3409	0.3396	1.0097	1.0093
			(v)	1.4296	1.4292	0.3415	0.3396	1.0098	1.0093

Table 5.43 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = -0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	2063.8880	764.4447	32.1762	13.0655	33.8187	13.945
			(ii)	2019.3050	387.2031	31.6885	8.0829	33.3496	8.778
			(iii)	1987.4500	387.1210	31.2251	8.0823	32.8204	8.777
			(iv)	111.8560	387.2041	33.1711	8.0829	34.8630	8.778
			(v)	1987.4500	387.1214	31.2251	8.0823	32.8204	8.777
			(vi)	1910.3600	351.2550	30.2009	7.3950	33.3279	7.440
Estimated	OSAF	30	(i)	2063.8540	764.3950	32.2970	13.0896	33.9348	13.971
			(ii)	2019.2720	387.1499	31.8078	8.0952	32.4628	8.788
			(iii)	1987.4170	387.0673	31.3438	8.0946	32.9327	8.788
			(iv)	2111.8190	387.1510	33.3022	8.0953	34.9844	8.788
			(v)	1987.4450	387.0820	31.2869	8.0926	32.9103	8.788
			(vi)	1910.0890	350.2096	30.0912	7.3861	33.1434	7.400
Zero	TSAF	30	(i)	3.4135	3.2594	1.3863	1.3488	1.7969	1.756
			(ii)	3.4019	3.2213	1.3829	1.3399	1.7933	1.747
			(iii)	3.3930	3.2213	1.3837	1.3399	1.7924	1.747
			(iv)	3.4122	3.2213	1.3849	1.3399	1.7964	1.747
			(v)	3.3930	3.2213	1.3837	1.3399	1.7924	1.747
Zero	OSAF	60	(i)	981.5634	334.1945	16.1092	7.0062	16.3071	7.207
			(ii)	599.8766	19.2687	11.5898	3.1526	11.7827	3.354
			(iii)	553.4246	19.2683	10.9975	3.1526	11.1903	3.354
			(iv)	665.1002	19.2687	12.4224	3.1526	12.6160	3.354
			(v)	542.5760	19.2683	10.8379	3.1526	11.0309	3.354
			(vi)	542.4117	17.9874	10.0116	2.1048	10.1003	3.274
Estimated	OSAF	60	(i)	981.5425	334.1918	16.1442	7.0154	16.3439	7.218
			(ii)	599.8657	19.2708	11.6044	3.1545	11.7953	3.354
			(iii)	553.4116	19.2704	11.0088	3.1546	11.1996	3.354
			(iv)	665.0879	19.2708	12.4388	3.1545	12.6312	3.354
			(v)	542.5678	19.2703	10.8537	3.1569	11.0496	3.357
			(vi)	541.3772	16.8253	10.0056	2.0017	10.0050	2.998
Zero	TSAF	60	(i)	3.0025	2.9306	1.5655	1.5385	1.8806	1.855
			(ii)	2.9172	2.8691	1.5319	1.5121	1.8503	1.833
			(iii)	2.9137	2.8691	1.5311	1.5121	1.8490	1.833
			(iv)	2.9211	2.8691	1.5333	1.5121	1.8512	1.833
			(v)	2.9152	2.8691	1.5317	1.5121	1.8496	1.833
Zero	OSAF	120	(i)	23.1591	18.2763	1.8775	1.5129	2.2268	1.976
			(ii)	4.1857	1.6049	1.0689	0.8803	1.4211	1.247
			(iii)	3.5914	1.6049	1.0379	0.8803	1.3904	1.247
			(iv)	4.5426	1.6049	1.0938	0.8803	1.4458	1.247
			(v)	3.5914	1.6049	1.0379	0.8803	1.3904	1.247
			(vi)	4.3479	1.3242	1.1121	0.7153	1.4405	1.201
Estimated	OSAF	120	(i)	23.1535	18.2714	1.8952	1.6253	2.2439	1.989
			(ii)	4.1819	1.6019	1.0765	0.8814	1.4260	1.248
			(iii)	3.5876	1.6019	1.0449	0.8815	1.3947	1.248
			(iv)	4.5386	1.6019	1.1022	0.8814	1.4514	1.248
			(v)	3.5875	1.6019	1.0465	0.8813	1.3973	1.247
			(vi)	3.2469	1.2763	1.0626	0.8015	1.3801	1.200
Zero	TSAF	120	(i)	1.3254	1.3251	0.4735	0.4728	1.0216	1.021
			(ii)	1.3237	1.3232	0.4720	0.4710	1.0204	1.019
			(iii)	1.3236	1.3232	0.4717	0.4710	1.0203	1.019
			(iv)	1.3237	1.3232	0.4718	0.4710	1.0203	1.019
			(v)	1.3236	1.3232	0.4717	0.4710	1.0203	1.019

Table 5.44 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = -030$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	1930.622	372.5553	26.6432	7.4611	28.0316	8.0268
			(ii)	1866.668	147.5764	26.1092	4.8497	27.4440	5.2755
			(iii)	1784.477	147.8054	24.9871	4.8510	26.296C	5.2892
			(iv)	2116.298	147.5766	29.6594	4.8497	31.1213	5.2755
			(v)	1784.477	147.8056	24.9871	4.3510	26.2960	5.2892
			(vi)	1794.730	134.1554	25.9941	3.2332	26.7503	4.9047
Estimated	OSAF	30	(i)	1930.589	372.5227	26.7474	7.4856	28.1315	8.0474
			(ii)	1866.635	147.5684	26.2067	4.8624	27.5375	5.2856
			(iii)	1784.445	147.7750	25.0852	4.8637	26.3890	5.2985
			(iv)	2116.259	147.5687	29.7705	4.8625	31.2282	5.2856
			(v)	1784.481	147.7696	25.0198	4.8591	26.3579	5.2957
			(vi)	1914.361	134.1379	27.9994	3.1394	27.9543	4.1070
Zero	TSAF	30	(i)	4.044	3.9695	1.5949	1.5822	1.9678	1.9586
			(ii)	4.028	3.8967	1.5954	1.5676	1.9678	1.9431
			(iii)	4.006	3.9085	1.5901	1.5712	1.9623	1.9467
			(iv)	4.057	3.8967	1.6016	1.5676	1.9756	1.9431
			(v)	4.006	3.9085	1.5901	1.5712	1.9623	1.9467
Zero	OSAF	60	(i)	415.174	131.4279	8.0449	4.3299	8.2452	4.5354
			(ii)	287.711	21.0322	6.6741	3.0316	6.8813	3.2388
			(iii)	270.492	21.0324	6.5040	3.0317	6.7076	3.2389
			(iv)	403.428	21.0322	7.9440	3.0316	8.1588	3.2388
			(v)	270.492	21.0324	6.5040	3.0317	6.7076	3.2389
			(vi)	365.788	19.5216	7.8223	2.4131	7.8396	2.1582
Estimated	OSAF	60	(i)	415.167	131.4249	8.0620	4.3348	8.2618	4.5408
			(ii)	287.708	21.0319	6.6889	3.0315	6.8939	3.2389
			(iii)	270.489	21.0321	6.5165	3.0315	6.7190	3.2389
			(iv)	403.427	21.0319	7.9519	3.0315	8.1644	3.2389
			(v)	270.487	21.0323	6.5141	3.0302	6.7188	3.2382
			(vi)	362.446	18.9962	5.8223	2.4131	6.9462	1.9657
Zero	TSAF	60	(i)	3.876	3.8819	1.9185	1.9130	2.1791	2.1766
			(ii)	3.861	3.8673	1.9103	1.9069	2.1729	2.1716
			(iii)	3.857	3.8673	1.9093	1.9069	2.1719	2.1716
			(iv)	3.862	3.8673	1.9102	1.9069	2.1733	2.1716
			(v)	3.857	3.8673	1.9093	1.9069	2.1719	2.1716
Zero	OSAF	120	(i)	1.753	1.4464	0.8998	0.8365	1.2671	1.2140
			(ii)	1.485	1.4260	0.8783	0.8287	1.2461	1.2065
			(iii)	1.485	1.4260	0.8784	0.8287	1.2461	1.2065
			(iv)	1.485	1.4260	0.8783	0.8287	1.2461	1.2065
			(v)	1.485	1.4260	0.8784	0.8287	1.2461	1.2065
			(vi)	1.485	1.4260	0.8784	0.8287	1.2461	1.2065
Estimated	OSAF	120	(i)	1.751	1.4453	0.9014	0.8362	1.2681	1.2134
			(ii)	1.483	1.4249	0.8785	0.8287	1.2459	1.2062
			(iii)	1.483	1.4249	0.8785	0.8288	1.2459	1.2062
			(iv)	1.483	1.4249	0.8785	0.8287	1.2459	1.2062
			(v)	1.483	1.4249	0.8785	0.8288	1.2459	1.2062
			(vi)	1.468	1.4119	0.0054	0.0056	1.2306	1.1920
Zero	TSAF	120	(i)	1.286	1.2864	0.5788	0.5786	1.0527	1.0527
			(ii)	1.286	1.2865	0.5787	0.5789	1.0526	1.0528
			(iii)	1.286	1.2865	0.5787	0.5789	1.0526	1.0528
			(iv)	1.286	1.2865	0.5787	0.5789	1.0526	1.0528
			(v)	1.286	1.2865	0.5787	0.5789	1.0526	1.0528

Table 5.45 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = -0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	2014.601	237.0345	23.4688	5.0792	24.6194	5.4836
			(ii)	2107.810	141.1000	24.1753	3.9457	25.3928	4.3312
			(iii)	1970.578	141.0893	22.9301	3.9481	24.0482	4.3334
			(iv)	2687.068	141.1000	29.6486	3.9457	31.1148	4.3312
			(v)	1970.578	141.0893	22.9301	3.9481	24.0482	4.3334
			(vi)	2430.573	128.1962	27.7886	2.3559	29.7700	3.1944
Estimated	OSAF	30	(i)	2014.555	237.0156	23.5450	5.0838	24.6929	5.4832
			(ii)	2107.779	141.0838	24.2532	3.9482	25.4717	4.3307
			(iii)	1970.548	141.0728	23.0037	3.9506	24.1208	4.3330
			(iv)	2687.026	141.0838	29.7413	3.9482	31.2116	4.3307
			(v)	1970.566	141.0886	22.9731	3.9471	24.1115	4.3326
			(vi)	2430.324	127.1870	27.7831	2.3558	29.7859	3.1965
Zero	TSAF	30	(i)	5.330	5.1521	2.0027	1.9696	2.3394	2.3039
			(ii)	5.346	5.1455	1.9984	1.9672	2.3376	2.3015
			(iii)	5.327	5.1475	2.0001	1.9680	2.3374	2.3023
			(iv)	5.402	5.1455	2.0112	1.9672	2.3509	2.3015
			(v)	5.327	5.1475	2.0001	1.9680	2.3374	2.3023
Zero	OSAF	60	(i)	183.895	35.8066	4.4461	2.8447	4.6488	3.0597
			(ii)	148.847	7.3417	4.1824	2.6757	4.3852	2.8907
			(iii)	148.847	7.3418	4.1824	2.6757	4.3852	2.8907
			(iv)	213.131	7.3417	4.8525	2.6757	5.0584	2.8907
			(v)	148.847	7.3418	4.1824	2.6757	4.3852	2.8907
			(vi)	193.346	7.1012	4.5723	2.1350	4.9017	2.8263
Estimated	OSAF	60	(i)	183.890	35.8057	4.4529	2.8446	4.6556	3.0595
			(ii)	148.842	7.3408	4.1883	2.6758	4.3911	2.8907
			(iii)	148.842	7.3409	4.1883	2.6758	4.3911	2.8907
			(iv)	213.124	7.3408	4.8630	2.6758	5.0697	2.8907
			(v)	148.844	7.3409	4.1878	2.6758	4.3910	2.8907
			(vi)	193.346	7.1012	4.5723	2.1350	4.9017	2.8263
Zero	TSAF	60	(i)	4.646	4.6301	2.1757	2.1710	2.4094	2.4059
			(ii)	4.641	4.6236	2.1754	2.1697	2.4090	2.4046
			(iii)	4.641	4.6236	2.1754	2.1697	2.4090	2.4046
			(iv)	4.650	4.6236	2.1771	2.1697	2.4104	2.4046
			(v)	4.641	4.6236	2.1754	2.1697	2.4090	2.4046
Zero	OSAF	120	(i)	1.409	1.3723	0.8253	0.7918	1.2001	1.1741
			(ii)	1.409	1.3723	0.8253	0.7918	1.2001	1.1741
			(iii)	1.409	1.3723	0.8253	0.7918	1.2001	1.1741
			(iv)	1.409	1.3723	0.8253	0.7918	1.2001	1.1741
			(v)	1.409	1.3723	0.8253	0.7918	1.2001	1.1741
			(vi)	1.402	1.3521	0.0119	0.0101	1.1998	1.1712
Estimated	OSAF	120	(i)	1.409	1.3722	0.8253	0.7919	1.2001	1.1741
			(ii)	1.409	1.3722	0.8253	0.7918	1.2001	1.1741
			(iii)	1.409	1.3722	0.8253	0.7919	1.2001	1.1741
			(iv)	1.409	1.3722	0.8253	0.7918	1.2001	1.1741
			(v)	1.409	1.3722	0.8253	0.7919	1.2001	1.1741
			(vi)	1.398	1.3624	0.0108	0.0099	1.1866	1.1613
Zero	TSAF	120	(i)	1.268	1.2687	0.6692	0.6693	1.0879	1.0879
			(ii)	1.268	1.2687	0.6692	0.6693	1.0879	1.0879
			(iii)	1.268	1.2687	0.6692	0.6693	1.0879	1.0879
			(iv)	1.268	1.2687	0.6692	0.6693	1.0879	1.0879
			(v)	1.268	1.2687	0.6692	0.6693	1.0879	1.0879

Table 5.46 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = 0$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	1344.659	101.9056	14.6070	3.2230	15.4786	3.6090
			(ii)	1427.210	110.2126	15.3438	3.2047	16.2941	3.5883
			(iii)	1324.523	108.5697	14.3189	3.0771	15.1885	3.4594
			(iv)	1229.836	112.9268	13.7940	3.2047	15.7555	3.3405
			(v)	1129.149	102.2839	12.8691	3.0771	15.1885	3.2594
			(vi)	1116.462	101.6410	11.2442	3.0625	14.7781	3.1899
Estimated	OSAF	30	(i)	1344.631	101.9021	14.5481	3.2252	15.5175	3.5096
			(ii)	1427.180	110.1986	15.2922	3.2075	16.1440	3.4412
			(iii)	1324.496	107.5586	14.1982	3.0793	15.0118	3.3728
			(iv)	1228.812	101.9186	13.1042	3.2075	14.8796	3.2144
			(v)	1119.128	100.2786	12.0103	3.0817	13.7474	3.1560
			(vi)	1016.444	35.6386	10.9163	3.0632	12.6152	3.0976
Zero	TSAF	30	(i)	6.504	6.3464	2.2483	2.2226	2.5761	2.5477
			(ii)	6.519	6.3604	2.2532	2.2252	2.5806	2.5512
			(iii)	6.498	6.3467	2.2483	2.2210	2.5758	2.5471
			(iv)	6.624	6.3604	2.2631	2.2252	2.5932	2.5512
			(v)	6.498	6.3467	2.2483	2.2210	2.5758	2.5471
Zero	OSAF	60	(i)	29.775	6.0705	2.9197	2.5446	3.1262	2.7555
			(ii)	29.775	6.0705	2.9197	2.5446	3.1262	2.7555
			(iii)	29.775	6.0705	2.9197	2.5446	3.1262	2.7555
			(iv)	37.431	6.0700	3.0210	2.5446	3.2298	2.7555
			(v)	29.775	6.0705	2.9197	2.5446	3.1262	2.7555
			(vi)	28.312	5.8826	2.2753	2.0437	2.8551	2.5954
Estimated	OSAF	60	(i)	29.773	6.0702	2.9211	2.5447	3.1279	2.7555
			(ii)	29.773	6.0701	2.9211	2.5446	3.1279	2.7555
			(iii)	29.773	6.0702	2.9211	2.5447	3.1279	2.7555
			(iv)	37.429	6.0701	3.0228	2.5446	3.2318	2.7555
			(v)	29.772	6.0702	2.9225	2.5447	3.1295	2.7555
			(vi)	28.012	5.4221	2.2753	2.0437	2.8551	2.5954
Zero	TSAF	60	(i)	5.582	5.5564	2.5020	2.4954	2.7018	2.6970
			(ii)	5.582	5.5564	2.5020	2.4954	2.7018	2.6970
			(iii)	5.582	5.5564	2.5020	2.4954	2.7018	2.6970
			(iv)	5.580	5.5564	2.5012	2.4954	2.7011	2.6970
			(v)	5.582	5.5564	2.5020	2.4954	2.7018	2.6970
Zero	OSAF	120	(i)	1.430	1.4047	0.8376	0.8138	1.2119	1.1933
			(ii)	1.430	1.4047	0.8376	0.8138	1.2119	1.1933
			(iii)	1.430	1.4047	0.8376	0.8138	1.2119	1.1933
			(iv)	1.430	1.4047	0.8376	0.8138	1.2119	1.1933
			(v)	1.430	1.4047	0.8376	0.8138	1.2119	1.1933
			(vi)	1.421	1.3223	0.7041	0.5712	1.1198	1.1081
Estimated	OSAF	120	(i)	1.430	1.4048	0.8376	0.8138	1.2119	1.1934
			(ii)	1.430	1.4047	0.8376	0.8138	1.2119	1.1934
			(iii)	1.430	1.4048	0.8376	0.8138	1.2119	1.1934
			(iv)	1.430	1.4047	0.8376	0.8138	1.2119	1.1934
			(v)	1.430	1.4048	0.8376	0.8138	1.2119	1.1934
			(vi)	1.417	1.3924	0.7413	0.6401	1.1980	1.1800
Zero	TSAF	120	(i)	1.380	1.3801	0.8097	0.8098	1.1849	1.1850
			(ii)	1.380	1.3801	0.8097	0.8098	1.1849	1.1850
			(iii)	1.380	1.3801	0.8097	0.8098	1.1849	1.1850
			(iv)	1.380	1.3801	0.8097	0.8098	1.1849	1.1850
			(v)	1.380	1.3801	0.8097	0.8098	1.1849	1.1850

Table 5.47 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = 0.15$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	521.448	36.1973	7.2631	2.5403	7.8895	2.9340
			(ii)	657.857	34.5251	8.3773	2.5182	9.0167	2.9035
			(iii)	521.325	34.4380	7.2600	2.5191	7.8756	2.8974
			(iv)	1107.437	34.3062	12.2825	2.5171	13.0367	2.8853
			(v)	521.252	34.2191	7.2601	2.5180	7.8692	2.8792
			(vi)	522.974	31.5520	7.6552	2.0986	7.5248	2.8217
Estimated	OSAF	30	(i)	521.458	36.2111	7.2651	2.5419	7.8981	2.9362
			(ii)	657.852	34.5556	8.3837	2.5196	9.0299	2.9065
			(iii)	521.330	34.4470	7.2621	2.5205	7.8837	2.8991
			(iv)	1107.443	34.3447	12.3055	2.5185	13.0685	2.8883
			(v)	521.269	34.2369	7.2696	2.5220	7.8835	2.8834
			(vi)	1002.935	31.5957	10.6561	2.0981	12.5173	2.8225
Zero	TSAF	30	(i)	7.906	7.8399	2.5197	2.5096	2.8392	2.8278
			(ii)	7.916	7.8338	2.5140	2.5057	2.8346	2.8252
			(iii)	7.903	7.8296	2.5128	2.5036	2.8341	2.8231
			(iv)	7.920	7.8298	2.5178	2.5033	2.8385	2.8232
			(v)	7.903	7.8256	2.5130	2.5011	2.8342	2.8211
Zero	OSAF	60	(i)	6.548	5.9140	2.6319	2.4921	2.8389	2.7053
			(ii)	6.548	5.9141	2.6319	2.4921	2.8389	2.7053
			(iii)	6.548	5.9140	2.6319	2.4921	2.8389	2.7053
			(iv)	6.548	5.9141	2.6319	2.4921	2.8389	2.7053
			(v)	6.548	5.9140	2.6319	2.4921	2.8389	2.7053
			(vi)	6.330	5.7214	2.4627	2.2571	2.7756	2.6457
Estimated	OSAF	60	(i)	6.548	5.9142	2.6319	2.4920	2.8390	2.7054
			(ii)	6.548	5.9142	2.6319	2.4921	2.8390	2.7054
			(iii)	6.548	5.9142	2.6319	2.4920	2.8390	2.7054
			(iv)	6.548	5.9142	2.6319	2.4921	2.8390	2.7054
			(v)	6.548	5.9142	2.6319	2.4920	2.8390	2.7054
			(vi)	6.320	5.6113	2.3521	2.2342	2.5621	2.4113
Zero	TSAF	60	(i)	7.078	7.0699	2.8387	2.8325	3.0288	3.0235
			(ii)	7.078	7.0698	2.8387	2.8325	3.0288	3.0235
			(iii)	7.078	7.0699	2.8387	2.8325	3.0288	3.0235
			(iv)	7.078	7.0698	2.8387	2.8325	3.0288	3.0235
			(v)	7.078	7.0699	2.8387	2.8325	3.0288	3.0235
Zero	OSAF	120	(i)	1.403	1.3866	0.8054	0.7886	1.1884	1.1757
			(ii)	1.403	1.3866	0.8054	0.7886	1.1884	1.1757
			(iii)	1.403	1.3866	0.8054	0.7886	1.1884	1.1757
			(iv)	1.403	1.3866	0.8054	0.7886	1.1884	1.1757
			(v)	1.403	1.3866	0.8054	0.7886	1.1884	1.1757
			(vi)	1.400	1.3821	0.7067	0.6038	1.1642	1.1696
Estimated	OSAF	120	(i)	1.403	1.3865	0.8054	0.7886	1.1884	1.1757
			(ii)	1.403	1.3865	0.8054	0.7886	1.1884	1.1757
			(iii)	1.403	1.3865	0.8054	0.7886	1.1884	1.1757
			(iv)	1.403	1.3865	0.8054	0.7886	1.1884	1.1757
			(v)	1.403	1.3865	0.8054	0.7886	1.1884	1.1757
			(vi)	1.390	1.3743	0.0087	0.0081	1.1753	1.1629
Zero	TSAF	120	(i)	1.504	1.5046	0.9071	0.9070	1.2663	1.2663
			(ii)	1.504	1.5046	0.9071	0.9070	1.2663	1.2663
			(iii)	1.504	1.5046	0.9071	0.9070	1.2663	1.2663
			(iv)	1.504	1.5046	0.9071	0.9070	1.2663	1.2663
			(v)	1.504	1.5046	0.9071	0.9070	1.2663	1.2663

Table 5.48 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = 0.30$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	478.862	7.7195	5.7363	2.2595	6.2551	2.6474
			(ii)	974.053	7.4778	8.6491	2.2539	9.1800	2.6196
			(iii)	477.272	7.4944	5.7239	2.2560	6.2127	2.6217
			(iv)	1353.147	7.4742	11.6642	2.2538	12.4002	2.6189
			(v)	477.242	7.4908	5.7241	2.2559	6.2093	2.6210
			(vi)	1224.449	7.2644	1.2474	0.0357	11.8894	2.5667
Estimated	OSAF	30	(i)	478.845	7.7004	0.6259	2.2594	6.2563	2.6453
			(ii)	974.061	7.4838	0.7974	2.2537	9.1922	2.6201
			(iii)	477.285	7.5004	0.6328	2.2559	6.2168	2.6222
			(iv)	1353.150	7.4815	1.2891	2.2537	12.4231	2.6197
			(v)	477.320	7.5696	0.6377	2.2551	6.2120	2.6274
			(vi)	1224.302	7.2029	1.2338	0.0350	11.8968	2.5617
Zero	TSAF	30	(i)	9.523	9.4466	2.8772	2.8750	3.1833	3.1786
			(ii)	9.521	9.3244	2.8818	2.8628	3.1881	3.1670
			(iii)	9.421	9.3340	2.8682	2.8639	3.1747	3.1681
			(iv)	9.536	9.3191	2.8858	2.8623	3.1931	3.1665
			(v)	9.417	9.3288	2.8682	2.8634	3.1746	3.1676
Zero	OSAF	60	(i)	6.022	5.6960	2.4622	2.3808	2.6824	2.6049
			(ii)	5.653	5.3261	2.4306	2.3490	2.6506	2.5730
			(iii)	5.653	5.3260	2.4306	2.3490	2.6506	2.5730
			(iv)	5.653	5.3261	2.4306	2.3490	2.6506	2.5730
			(v)	5.653	5.3260	2.4306	2.3490	2.6506	2.5730
			(vi)	5.479	5.1646	2.1023	2.1018	2.5927	2.5173
Estimated	OSAF	60	(i)	6.062	5.7353	2.1451	2.3811	2.6829	2.6055
			(ii)	5.651	5.3242	2.1049	2.3492	2.6504	2.5728
			(iii)	5.651	5.3241	2.1049	2.3492	2.6504	2.5727
			(iv)	5.651	5.3242	2.1049	2.3492	2.6504	2.5728
			(v)	5.651	5.3241	2.1049	2.3492	2.6504	2.5727
			(vi)	5.377	5.0216	2.1023	2.1018	2.5927	2.5173
Zero	TSAF	60	(i)	9.173	9.1825	3.1105	3.1105	3.3047	3.3050
			(ii)	8.115	8.1237	3.0429	3.0428	3.2374	3.2376
			(iii)	8.115	8.1237	3.0430	3.0428	3.2374	3.2376
			(iv)	8.115	8.1237	3.0429	3.0428	3.2374	3.2376
			(v)	8.115	8.1237	3.0430	3.0428	3.2374	3.2376
Zero	OSAF	120	(i)	1.402	1.3889	0.7969	0.7846	1.1825	1.1732
			(ii)	1.402	1.3890	0.7969	0.7846	1.1825	1.1732
			(iii)	1.402	1.3889	0.7969	0.7846	1.1825	1.1732
			(iv)	1.402	1.3890	0.7969	0.7846	1.1825	1.1732
			(v)	1.402	1.3889	0.7969	0.7846	1.1825	1.1732
			(vi)	1.397	1.3653	0.6988	0.6764	1.1765	1.1544
Estimated	OSAF	120	(i)	1.401	1.3884	0.8073	0.7845	1.1821	1.1728
			(ii)	1.401	1.3884	0.8073	0.7845	1.1821	1.1729
			(iii)	1.401	1.3884	0.8073	0.7845	1.1821	1.1728
			(iv)	1.401	1.3884	0.8073	0.7845	1.1821	1.1729
			(v)	1.401	1.3884	0.8073	0.7845	1.1821	1.1728
			(vi)	1.385	1.3732	0.8068	0.7073	1.1684	1.1594
Zero	TSAF	120	(i)	1.705	1.7050	1.0156	1.0151	1.3694	1.3690
			(ii)	1.705	1.7050	1.0156	1.0151	1.3694	1.3690
			(iii)	1.705	1.7050	1.0156	1.0151	1.3694	1.3690
			(iv)	1.705	1.7050	1.0156	1.0151	1.3694	1.3690
			(v)	1.705	1.7050	1.0156	1.0151	1.3694	1.3690

Table 5.49 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = 0.45$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	111.745	6.9122	2.9787	2.0024	3.4883	2.4783
			(ii)	198.529	6.3754	3.7665	1.9979	4.3473	2.4223
			(iii)	111.138	6.3620	2.9758	1.9983	3.4276	2.4199
			(iv)	382.952	6.1557	5.2312	2.8832	5.8117	2.4066
			(v)	110.979	6.1423	2.9764	2.0004	3.4150	2.4042
			(vi)	347.427	5.9873	4.3959	2.1274	5.6075	2.3571
Estimated	OSAF	30	(i)	111.760	6.9501	2.9800	2.0025	3.4877	2.4795
			(ii)	198.521	6.3749	3.7698	1.9980	4.3483	2.4209
			(iii)	111.123	6.3655	2.9770	1.9984	3.4250	2.4192
			(iv)	382.952	6.1697	5.2405	2.0001	5.8196	2.4060
			(v)	111.004	6.2122	2.9737	2.0005	3.4146	2.4101
			(vi)	347.319	5.9939	4.4027	2.1274	4.6101	2.3569
Zero	TSAF	30	(i)	10.616	10.5463	2.9094	2.8968	3.2741	3.2640
			(ii)	10.319	10.2546	2.8796	2.8662	3.2417	3.2309
			(iii)	10.319	10.2549	2.8787	2.8667	3.2415	3.2313
			(iv)	10.325	10.2577	2.8801	2.8683	3.2435	3.2314
			(v)	10.320	10.2581	2.8780	2.8688	3.2407	3.2318
Zero	OSAF	60	(i)	6.731	6.5851	2.4179	2.3721	2.6809	2.6381
			(ii)	17.766	5.0417	2.4253	2.2596	2.6573	2.4938
			(iii)	5.223	5.0416	2.3080	2.2596	2.5401	2.4937
			(iv)	17.766	5.0413	2.4261	2.2604	2.6577	2.4941
			(v)	5.222	5.0412	2.3088	2.2604	2.5405	2.4941
			(vi)	5.390	4.8618	2.0443	2.0697	2.5953	2.4388
Estimated	OSAF	60	(i)	6.731	6.6000	2.4179	2.3721	2.6807	2.6389
			(ii)	17.741	5.0163	2.4247	2.2596	2.6554	2.4923
			(iii)	5.197	5.0162	2.3080	2.2596	2.5387	2.4923
			(iv)	17.740	5.0152	2.4255	2.2604	2.6556	2.4926
			(v)	5.209	5.0280	2.3088	2.2604	2.5393	2.4929
			(vi)	4.867	3.3843	2.0443	2.0186	2.4953	2.2311
Zero	TSAF	60	(i)	12.890	12.9223	3.5201	3.5283	3.7158	3.7230
			(ii)	9.297	9.3007	3.2802	3.2820	3.4814	3.4823
			(iii)	9.293	9.3007	3.2787	3.2820	3.4800	3.4823
			(iv)	9.301	9.3038	3.2816	3.2835	3.4828	3.4837
			(v)	9.296	9.3038	3.2801	3.2835	3.4814	3.4837
Zero	OSAF	120	(i)	1.484	1.4769	0.8073	0.7990	1.2023	1.1962
			(ii)	1.397	1.3893	0.8033	0.7951	1.1868	1.1806
			(iii)	1.397	1.3893	0.8033	0.7951	1.1868	1.1806
			(iv)	1.397	1.3893	0.8033	0.7951	1.1868	1.1806
			(v)	1.397	1.3893	0.8033	0.7951	1.1868	1.1806
			(vi)	1.397	1.3890	0.8020	0.7941	1.1865	1.1800
Estimated	OSAF	120	(i)	1.600	1.5927	0.8072	0.7990	1.2160	1.2098
			(ii)	1.394	1.3864	0.8033	0.7951	1.1857	1.1795
			(iii)	1.394	1.3864	0.8033	0.7951	1.1857	1.1795
			(iv)	1.394	1.3864	0.8033	0.7951	1.1857	1.1795
			(v)	1.394	1.3864	0.8033	0.7951	1.1857	1.1795
			(vi)	1.385	1.3775	0.0617	0.0613	1.1732	1.1673
Zero	TSAF	120	(i)	1.999	1.9992	1.1681	1.1679	1.5219	1.5218
			(ii)	1.980	1.9807	1.1568	1.1566	1.5116	1.5115
			(iii)	1.980	1.9807	1.1568	1.1566	1.5116	1.5115
			(iv)	1.980	1.9807	1.1568	1.1566	1.5116	1.5115
			(v)	1.980	1.9807	1.1568	1.1566	1.5116	1.5115

Table 5.50 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = 0.60$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	39.555	7.3928	2.4318	2.0745	2.9470	2.5966
			(ii)	323.152	6.4837	4.0719	2.0710	4.5273	2.4996
			(iii)	38.731	6.4882	2.4318	2.0716	2.8585	2.5007
			(iv)	396.256	6.2925	4.8372	2.0711	5.2741	2.4772
			(v)	38.467	6.2319	2.4329	2.0728	2.8298	2.4705
			(vi)	358.646	6.1128	4.2949	2.1244	5.0955	2.4299
Estimated	OSAF	30	(i)	39.585	7.4082	2.4310	2.0744	2.9513	2.6006
			(ii)	323.124	6.4530	4.0703	2.0709	4.5245	2.4957
			(iii)	38.704	6.4574	2.4310	2.0715	2.8541	2.4968
			(iv)	396.255	6.2867	4.8391	2.0710	5.2784	2.4769
			(v)	38.516	6.2998	2.4331	2.0727	2.8394	2.4828
			(vi)	358.590	6.0071	4.2949	2.1245	5.0905	2.4896
Zero	TSAF	30	(i)	13.192	13.1409	3.3509	3.3516	3.7070	3.7039
			(ii)	12.742	12.6048	3.3050	3.2940	3.6650	3.6476
			(iii)	12.650	12.6120	3.2954	3.2944	3.6528	3.6484
			(iv)	12.755	12.5953	3.3047	3.2920	3.6649	3.6451
			(v)	12.648	12.5891	3.2923	3.2909	3.6491	3.6445
Zero	OSAF	60	(i)	7.931	7.8116	2.5965	2.5684	2.9080	2.8806
			(ii)	5.439	5.3388	2.3371	2.3087	2.5837	2.5585
			(iii)	5.438	5.3387	2.3370	2.3087	2.5837	2.5585
			(iv)	5.362	5.2310	2.3352	2.3052	2.5752	2.5453
			(v)	5.362	5.2309	2.3352	2.3052	2.5752	2.5453
			(vi)	5.188	5.0589	2.1559	2.1515	2.5177	2.4882
Estimated	OSAF	60	(i)	7.914	7.7938	2.5966	2.5685	2.9124	2.8849
			(ii)	5.412	5.3012	2.3372	2.3089	2.5799	2.5535
			(iii)	5.412	5.3011	2.3372	2.3089	2.5799	2.5535
			(iv)	5.344	5.2134	2.3354	2.3054	2.5720	2.5422
			(v)	5.360	5.2251	2.3353	2.3054	2.5743	2.5438
			(vi)	5.188	5.0589	2.1559	2.1515	2.5177	2.4882
Zero	TSAF	60	(i)	18.135	18.1340	4.2891	4.2925	4.4910	4.4928
			(ii)	11.773	11.7919	3.7067	3.7138	3.9172	3.9222
			(iii)	11.773	11.7919	3.7067	3.7138	3.9172	3.9222
			(iv)	11.746	11.7398	3.7023	3.7044	3.9127	3.9133
			(v)	11.746	11.7398	3.7023	3.7044	3.9127	3.9133
Zero	OSAF	120	(i)	2.401	2.4060	0.8044	0.7993	1.3279	1.3259
			(ii)	1.442	1.4074	0.7910	0.7855	1.1901	1.1826
			(iii)	1.438	1.4073	0.7909	0.7855	1.1894	1.1826
			(iv)	1.414	1.4058	0.7912	0.7858	1.1868	1.1824
			(v)	1.410	1.4058	0.7911	0.7858	1.1862	1.1824
			(vi)	1.411	1.3998	0.7843	0.7382	1.1852	1.1693
Estimated	OSAF	120	(i)	2.426	2.4359	0.8044	0.7993	1.3325	1.3311
			(ii)	1.434	1.4024	0.7909	0.7854	1.1883	1.1807
			(iii)	1.426	1.4024	0.7908	0.7854	1.1868	1.1807
			(iv)	1.412	1.3997	0.7911	0.7858	1.1856	1.1803
			(v)	1.404	1.3997	0.7910	0.7858	1.1841	1.1803
			(vi)	1.405	1.3886	0.7039	0.7038	1.1738	1.1679
Zero	TSAF	120	(i)	2.420	2.4207	1.3084	1.3079	1.6864	1.6861
			(ii)	2.321	2.3202	1.2585	1.2580	1.6376	1.6371
			(iii)	2.320	2.3202	1.2582	1.2580	1.6374	1.6371
			(iv)	2.321	2.3209	1.2590	1.2586	1.6379	1.6376
			(v)	2.320	2.3209	1.2587	1.2586	1.6377	1.6376

Table 5.51 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = 0.75$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	236.0916	23.2400	3.317	2.3247	3.8397	2.8466
			(ii)	304.1220	7.1119	3.748	2.1181	4.2401	2.5804
			(iii)	261.4522	7.1813	3.461	2.1283	3.9306	2.5897
			(iv)	392.5131	6.7734	4.110	2.1189	4.5612	2.5374
			(v)	261.1300	6.8301	3.464	2.1333	3.8914	2.5480
			(vi)	355.4910	6.6673	0.472	0.1954	4.4302	2.4986
Estimated	OSAF	30	(i)	236.0232	23.1351	0.195	2.3253	3.8247	2.8315
			(ii)	304.0733	7.0425	0.120	2.1181	4.2280	2.5655
			(iii)	261.4067	7.1157	0.326	2.1282	3.9170	2.5752
			(iv)	392.4673	6.7120	0.486	2.1189	4.5506	2.5256
			(v)	261.1767	6.8646	0.334	2.1332	3.8961	2.5532
			(vi)	355.4115	11.0908	0.471	0.1953	4.4165	2.5504
Zero	TSAF	30	(i)	16.7081	16.7040	3.786	3.7801	4.1555	4.1510
			(ii)	16.3156	16.2220	3.736	3.7302	4.1070	4.0983
			(iii)	16.3231	16.2784	3.738	3.7343	4.1083	4.1026
			(iv)	16.2105	16.1349	3.718	3.7141	4.0921	4.0857
			(v)	16.2303	16.1912	3.722	3.7203	4.0943	4.0913
Zero	OSAF	60	(i)	6.7517	6.7024	2.396	2.3828	2.7161	2.7059
			(ii)	5.0558	4.9772	2.149	2.1282	2.4372	2.4181
			(iii)	5.0489	4.9702	2.150	2.1288	2.4372	2.4181
			(iv)	4.7834	4.6984	2.130	2.1130	2.3986	2.3806
			(v)	4.7759	4.6908	2.133	2.1157	2.3994	2.3813
			(vi)	4.6387	4.5620	0.137	0.1366	2.3449	2.3280
Estimated	OSAF	60	(i)	6.5774	6.5290	0.132	2.3828	2.6993	2.6890
			(ii)	4.9806	4.9082	0.151	2.1282	2.4244	2.4054
			(iii)	4.9746	4.9021	0.150	2.1288	2.4245	2.4054
			(iv)	4.7482	4.6714	0.141	2.1131	2.3886	2.3718
			(v)	4.7771	4.6961	0.141	2.1158	2.3951	2.3776
			(vi)	4.6387	4.5620	0.137	0.1366	2.3449	2.3280
Zero	TSAF	60	(i)	16.8045	16.8126	4.327	4.3292	4.5370	4.5403
			(ii)	12.2900	12.2790	3.766	3.7533	3.9896	3.9791
			(iii)	12.2945	12.2832	3.767	3.7542	3.9905	3.9800
			(iv)	11.9884	11.9560	3.718	3.7138	3.9459	3.9420
			(v)	11.9981	11.9653	3.723	3.7179	3.9491	3.9450
Zero	OSAF	120	(i)	2.7273	2.7187	0.812	0.8104	1.3986	1.3961
			(ii)	1.5356	1.5201	0.785	0.7834	1.2022	1.1986
			(iii)	1.5356	1.5201	0.785	0.7834	1.2022	1.1986
			(iv)	1.4292	1.4186	0.786	0.7839	1.1840	1.1808
			(v)	1.4292	1.4186	0.786	0.7839	1.1840	1.1808
			(vi)	1.3961	1.3870	0.010	0.0102	1.1647	1.1614
Estimated	OSAF	120	(i)	2.6593	2.6550	0.012	0.8104	1.3876	1.3859
			(ii)	1.5245	1.5155	0.011	0.7834	1.1973	1.1946
			(iii)	1.5245	1.5155	0.011	0.7834	1.1973	1.1946
			(iv)	1.4111	1.4034	0.010	0.7840	1.1786	1.1757
			(v)	1.4107	1.4015	0.010	0.7840	1.1789	1.1755
			(vi)	1.3961	1.3870	0.010	0.0102	1.1647	1.1614
Zero	TSAF	120	(i)	2.8538	2.8547	1.469	1.4699	1.8573	1.8579
			(ii)	2.7009	2.7015	1.375	1.3764	1.7725	1.7730
			(iii)	2.7009	2.7015	1.375	1.3764	1.7725	1.7730
			(iv)	2.7002	2.7004	1.375	1.3759	1.7722	1.7724
			(v)	2.7002	2.7004	1.375	1.3759	1.7722	1.7724

Table 5.52 AMSFE, AMAFE and AAMFE of one and two step ahead forecasts using different strategies, sample sizes, X_4 design matrix when $\gamma = 0.90$

Initial value of $\hat{\varepsilon}_0$	Step	Sample size	Strategy	AMSFE ML	AMSFE MML	AAMFE ML	AAMFE MML	AMAE ML	AMAE MML
Zero	OSAF	30	(i)	109.685	6.3835	2.7270	1.9712	3.1955	2.4447
			(ii)	251.844	6.3302	3.3650	1.9884	3.8353	2.4438
			(iii)	148.889	6.3406	2.8142	1.9913	3.2756	2.4474
			(iv)	283.247	6.1582	3.5044	1.9947	3.9487	2.4216
			(v)	148.709	6.1728	2.8251	1.9985	3.2570	2.4263
			(vi)	256.918	6.2014	3.0500	1.0400	3.8542	2.4054
Estimated	OSAF	30	(i)	109.301	5.9833	2.7274	1.9711	3.1272	2.3728
			(ii)	251.485	5.9631	3.3650	1.9883	3.7711	2.3789
			(iii)	148.541	5.9751	2.8140	1.9912	3.2118	2.3829
			(iv)	282.965	7.1463	3.5045	1.9946	3.8996	2.4042
			(v)	148.784	6.2369	2.8217	1.9983	3.2680	2.4383
			(vi)	256.620	7.1608	0.0506	0.0395	3.8107	2.3819
Zero	TSAF	30	(i)	17.225	17.2190	3.7672	3.7618	4.1674	4.1602
			(ii)	17.230	17.1228	3.7720	3.7632	4.1701	4.1578
			(iii)	17.170	17.1400	3.7641	3.7644	4.1626	4.1608
			(iv)	17.135	17.0599	3.7619	3.7526	4.1625	4.1502
			(v)	17.103	17.0767	3.7562	3.7543	4.1578	4.1534
Zero	OSAF	60	(i)	5.087	5.0492	2.0820	2.0759	2.4333	2.4260
			(ii)	4.847	4.7704	2.0423	2.0374	2.3804	2.3683
			(iii)	4.842	4.7672	2.0410	2.0377	2.3792	2.3683
			(iv)	4.542	4.4610	2.0123	2.0010	2.3245	2.3088
			(v)	4.538	4.4653	2.0120	2.0054	2.3237	2.3121
			(vi)	4.404	4.3297	2.0189	2.0172	2.2761	2.2585
Estimated	OSAF	60	(i)	4.648	4.6255	2.0820	2.0759	2.3596	2.3544
			(ii)	4.492	4.4392	2.0424	2.0375	2.3176	2.3087
			(iii)	4.488	4.4371	2.0411	2.0378	2.3164	2.3088
			(iv)	4.288	4.2318	2.0124	2.0012	2.2762	2.2640
			(v)	4.518	4.4540	2.0121	2.0055	2.3186	2.3060
			(vi)	4.304	4.2206	2.0181	2.0162	2.2756	2.2575
Zero	TSAF	60	(i)	13.606	13.5815	4.0073	4.0032	4.2436	4.2416
			(ii)	13.103	13.0030	3.9007	3.8966	4.1465	4.1402
			(iii)	13.080	12.9864	3.8964	3.8954	4.1424	4.1389
			(iv)	12.597	12.5219	3.8157	3.8022	4.0642	4.0523
			(v)	12.572	12.5240	3.8127	3.8067	4.0610	4.0562
Zero	OSAF	120	(i)	2.191	2.1652	0.8065	0.8058	1.3721	1.3655
			(ii)	1.981	1.9914	0.8010	0.8022	1.3258	1.3274
			(iii)	1.981	1.9914	0.8010	0.8022	1.3258	1.3274
			(iv)	1.729	1.6388	0.8086	0.8092	1.2714	1.2522
			(v)	1.729	1.6388	0.8086	0.8092	1.2715	1.2522
			(vi)	1.700	1.4321	0.8073	0.8092	1.2656	1.2431
Estimated	OSAF	120	(i)	1.762	1.7523	0.8064	0.8058	1.2702	1.2676
			(ii)	1.638	1.6561	0.8009	0.8021	1.2428	1.2460
			(iii)	1.637	1.6561	0.8009	0.8021	1.2426	1.2460
			(iv)	1.521	1.4846	0.8084	0.8091	1.2183	1.2103
			(v)	1.671	1.5742	0.8085	0.8091	1.2542	1.2317
			(vi)	1.594	1.5304	0.8023	0.8039	1.2265	1.2114
Zero	TSAF	120	(i)	3.224	3.2235	1.5602	1.5592	1.9774	1.9768
			(ii)	3.203	3.2058	1.5390	1.5425	1.9617	1.9637
			(iii)	3.203	3.2058	1.5391	1.5425	1.9618	1.9637
			(iv)	3.198	3.1984	1.5418	1.5420	1.9612	1.9614
			(v)	3.198	3.1984	1.5419	1.5420	1.9613	1.9614

CHAPTER 6

Some Issues on the Estimation of the Covariance Matrix in the General Linear Regression Model

6.1 Introduction

As we discussed in Chapter 2, Godambe (1960) used the score of the estimating equation to show an optimum property of regular maximum likelihood estimation. Godambe and Thompson (1974) modified the score of the estimating equation for regular maximum likelihood estimation as a bias correction. Mahmood (2000) proposed a method of estimation for the covariance matrix parameters of the general linear regression model using a bias corrected version of Mak's (1993) algorithm based on the marginal likelihood. He demonstrated that the LS method and the marginal likelihood method produce equivalent estimates thus allowing the normality assumption to be dropped for marginal likelihood based estimation of the general linear regression model.

We use a different approach in which the general linear model is transformed and the LS method is applied to the transformed model in the context of an estimating equation. We transform the linear model by pre-multiplying by the n th root of the determinant of lower triangular Cholesky decomposition of the covariance matrix times the inverse of the same matrix.

Our aim is to explore whether there is equivalence between the LS method and the maximum likelihood method. We will compare the estimating equation for our LS method based on the transformed model with that from the concentrated likelihood function.

A related contribution is to express the estimating equation as a polynomial of infinite degree in the case of a non-stationary AR(1) error process and to use different mathematical expressions to find the roots of an approximating polynomial up to degree four. A special case is when the design matrix is a single vector, for then the estimating equation turns into a finite degree polynomial, i.e., a fifth degree polynomial for this non-stationary AR(1) process. For the general case, we suggest a way to find the roots of the polynomial and select an appropriate root as a final estimated value of the parameter of the model under consideration. Finally we derive an iterative formula for estimating the parameters based on the least squares method.

This chapter is organized as follows. In Section 6.2 we consider the model, the likelihood, concentrated likelihood and marginal likelihood for the general linear regression model. In Section 6.3 we present the score of the concentrated likelihood as given by Mahmood (2000) and an alternative version. The score of the marginal likelihood is discussed in Section 6.4. Applying the least squares method to estimate the parameter in Section 6.5, we show that it results in the same estimates as those from the estimating equations of the concentrated likelihood. In Section 6.6 we discuss estimating the covariance matrix parameter in the linear regression model with a non-stationary AR(1) error process. In Section 6.7, we discuss how to find roots of the estimating equations when they are approximated by polynomials of different degrees. The chapter ends with some concluding remarks in Section 6.8.

6.2 The Model

Consider the linear model

$$y = X\beta + u \tag{6.1}$$

where y is an $n \times 1$ vector, X is an $n \times k$ matrix of known values and of full column rank, β is a k-dimensional vector of unknown parameters and $u \sim N(0, \Sigma(\gamma))$ with $\Sigma(\gamma) \neq \sigma^2 I$ where $\Sigma(\gamma)$ is a general positive definite matrix. In the context of (6.1), the most problematic part is the estimation of the parameter γ. Therefore our initial emphasis is on estimating the unknown parameter γ when the parameters of interest are β and σ^2. In this case γ is an unknown nuisance parameter. Once γ is estimated, β and σ^2 can easily be estimated. The likelihood and the loglikelihood for model (6.1) are respectively

$$L(y;\gamma,\sigma^2,\beta) = (2\pi)^{-\frac{n}{2}}\sigma^{-n}|\Sigma(\gamma)|^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}(y - X\beta)'\Sigma(\gamma)^{-1}(y - X\beta)\right\} \tag{6.2}$$

and

$$l(y;\gamma,\sigma^2,\beta) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2}\log|\Sigma(\gamma)| \\ -\frac{1}{2\sigma^2}(y - X\beta)'\,\Sigma(\gamma)^{-1}(y - X\beta). \tag{6.3}$$

The above expressions are also known as the classical likelihood and classical log likelihood. The concentrated log likelihood for the parameter γ of the model (6.1) can be written as

$$l_c(y;\gamma) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\left(\frac{s^2(\gamma)}{n}\right) - \frac{1}{2}\log|\Sigma(\gamma)| - \frac{n}{2} \tag{6.4}$$

where

$$s^2(\gamma) = \left(y - X\hat{\beta}(\gamma)\right)'\,\Sigma(\gamma)^{-1}\left(y - X\hat{\beta}(\gamma)\right) \tag{6.5}$$

and

$$\hat{\beta}(\gamma) = \left(X'\Sigma(\gamma)^{-1}X\right)^{-1}X'\Sigma(\gamma)^{-1}y. \tag{6.6}$$

Another useful method for estimating γ as discussed in Chapter 2 is to maximize its marginal likelihood function. According to Tunnicliffe Wilson (1989), the marginal likelihood function for γ in (6.1) is

$$L_m(y;\gamma) = |\Sigma(\gamma)|^{-\frac{1}{2}}|X'\Sigma(\gamma)X|^{-\frac{1}{2}}s^2(\gamma)^{-\frac{m}{2}} \tag{6.7}$$

in which $m = n - k$. The log of the marginal likelihood is

$$l_m(y;\gamma) = -\frac{1}{2}\log|\Sigma(\gamma)| - \frac{1}{2}\log\left|X'\Sigma(\gamma)^{-1}X\right| - \frac{m}{2}\log s^2(\gamma). \tag{6.8}$$

6.3 Score of the Concentrated Likelihood

The concentrated likelihood given by equation (6.4) can also be written as

$$\begin{aligned} l_c(y;\gamma) &= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\left(s^2(\gamma)\right) + \frac{n}{2}\log n - \frac{1}{2}\log\left|L(\gamma)L(\gamma)'\right| - \frac{n}{2} \\ &= -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\left(s^2(\gamma)\right) + \frac{n}{2}\log n - \log\left|L'(\gamma)\right| - \frac{n}{2} \end{aligned} \tag{6.9}$$

where $L(\gamma)$ is the lower triangular matrix of $\Sigma(\gamma)$ defined by $\Sigma(\gamma) = L(\gamma)L(\gamma)'$.

The estimated value of the parameter γ can be obtained by setting the derivative of the concentrated likelihood equal to zero and solving for the parameter value from that equation. Differentiating equation (6.9) with respect to γ and equating to zero, according to Mahmood (2000), we have

$$\frac{\partial l_c(y;\gamma)}{\partial\gamma} = -\frac{1}{2}tr\left[\Sigma^{-1}(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\right] + \frac{n}{2}\frac{y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y}{e(\gamma)'e(\gamma)} = 0 \tag{6.10}$$

where

$$\Delta(\gamma) = \left[\Sigma^{-1}(\gamma) - \Sigma^{-1}(\gamma)X\left(X'\Sigma^{-1}(\gamma)X\right)^{-1}X'\Sigma^{-1}(\gamma)\right] \tag{6.11}$$

and

$$e(\gamma)'e(\gamma) = y'\Delta(\gamma)y. \tag{6.12}$$

If we replace $\Sigma(\gamma)$ with $\Sigma(\gamma) = L(\gamma)L(\gamma)'$ in equation (6.10), we have the following estimating equation of the concentrated likelihood

$$\frac{\partial l_c(y;\gamma)}{\partial\gamma} = -tr\left[L^{-1}(\gamma)\frac{\partial L(\gamma)}{\partial\gamma}\right] + \frac{n}{2}\frac{y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y}{e(\gamma)'e(\gamma)} = 0. \qquad (6.13)$$

6.4 Score of the Marginal Likelihood

The estimating equation from the marginal likelihood (6.6) for model (6.1) is given by Mahmood (2000) as

$$\frac{\partial l_m(y;\gamma)}{\partial\gamma} = -\frac{1}{2}tr\left[\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\right] + \frac{m}{2}\frac{y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y}{e(\gamma)'e(\gamma)} = 0. \qquad (6.14)$$

Substituting $\Delta(\gamma) = \left[\Sigma^{-1}(\gamma) - \Sigma^{-1}(\gamma)X\left(X'\Sigma^{-1}(\gamma)X\right)^{-1}X'\Sigma^{-1}(\gamma)\right]$ in the above equation we have

$$\begin{aligned}\frac{\partial l_m(y;\gamma)}{\partial\gamma} = &-tr\left[L^{-1}(\gamma)\frac{\partial L(\gamma)}{\partial\gamma}\right] - \frac{1}{2}tr\left[\Sigma^{-1}(\gamma)X\left(X'\Sigma^{-1}(\gamma)X\right)^{-1}X'\Sigma^{-1}(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\right] \\ &+ \frac{m}{2}\frac{y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y}{e(\gamma)'e(\gamma)} = 0.\end{aligned} \qquad (6.15)$$

6.5 Least Squares Estimation

In the following subsections, we will discuss two methods of estimating the covariance matrix parameter in the general linear regression model.

6.5.1 Standard method: estimation of parameters by minimising sum of squared errors

In this subsection we are interested in estimating parameters of the model by using the traditional method of minimising the error sum of squares. We return to the general linear regression model given by (6.1), which can be transformed as

$$L^{-1}(\gamma)y = L^{-1}(\gamma)X\beta + L^{-1}(\gamma)u \tag{6.16}$$

which is now a linear regression model with a well behaved error term given that $\mathrm{E}\left(L^{-1}(\gamma)u\right) = 0$ and $Var\left(L^{-1}(\gamma)u\right) = \sigma^2 I$. The method of LS involves finding β and γ which minimise the error sum of squares

$$\begin{aligned} S_1(y;\beta,\gamma) &= \left(L^{-1}(\gamma)y - L^{-1}(\gamma)X\beta\right)'\left(L^{-1}(\gamma)y - L^{-1}(\gamma)X\beta\right) \\ &= (y - X\beta)'\,\Sigma^{-1}(\gamma)(y - X\beta) \\ &= y'\Sigma^{-1}(\gamma)y - 2y'\Sigma^{-1}(\gamma)X\beta + \beta'X'\Sigma^{-1}(\gamma)X\beta. \end{aligned} \tag{6.17}$$

According to the LS method, we can solve for β by using

$$\frac{\partial S_1(y;\beta,\gamma)}{\partial\beta} = 0$$

which results in the generalised least squares (GLS) estimate of β

$$\hat{\beta}(\gamma) = \left(X'\Sigma^{-1}(\gamma)X\right)^{-1} X'\Sigma^{-1}(\gamma)y. \tag{6.18}$$

We also observe that the estimated $\hat{\beta}(\gamma)$ depends on γ which needs to be estimated to get the final estimate of β. To formulate the LS estimation equation for γ we have

$$\frac{\partial S_1(y;\beta,\gamma)}{\partial \gamma} = (y - X\beta)' \frac{\partial \Sigma^{-1}(\gamma)}{\partial \gamma}(y - X\beta) = 0. \tag{6.19}$$

We now substitute the value of β by $\hat{\beta}(\gamma)$ in $\frac{\partial S_1(y;\beta,\gamma)}{\partial \gamma}$ and using $\frac{\partial \Sigma^{-1}(\gamma)}{\partial \gamma} = -\Sigma^{-1}(\gamma)\frac{\partial \Sigma(\gamma)}{\partial \gamma}\Sigma^{-1}(\gamma)$ in (6.19) we have

$$\left.\frac{\partial S_1(y;\beta,\gamma)}{\partial \gamma}\right|_{\beta = \hat{\beta}(\gamma)} = -y'\Delta(\gamma)\frac{\partial \Sigma(\gamma)}{\partial \gamma}\Delta(\gamma)y = 0. \tag{6.20}$$

Mahmood's (2000) work shows that the stochastic part of the estimating equation for the LS method is equivalent to the stochastic part of the estimating equation of the concentrated likelihood but that the two estimating equations are not the same. The estimating equation of the concentrated likelihood contains two components defined in equation (6.13), one is a constant and the other is a stochastic part divided by the error sum of squares. In the next section we will discuss another method which will show that the estimating equation of the LS method is exactly the same as the estimating equation of the concentrated likelihood.

6.5.2 Alternative method: estimation of parameters by modified sum of squared errors

If we pre-multiply model (6.1) by $|L(\gamma)|^{1/n} L^{-1}(\gamma)$, we have

$$|L(\gamma)|^{1/n} L^{-1}(\gamma)y = |L(\gamma)|^{1/n} L^{-1}(\gamma)X\beta + |L(\gamma)|^{1/n} L^{-1}(\gamma)u. \tag{6.21}$$

The modified error sum of squares can be written as

$$S_2(y;\beta,\gamma)=\left(|L(\gamma)|^{1/n}L^{-1}(\gamma)y-|L(\gamma)|^{1/n}L^{-1}(\gamma)X\beta(\gamma)\right)'$$
$$\left(|L(\gamma)|^{1/n}L^{-1}(\gamma)y-|L(\gamma)|^{1/n}L^{-1}(\gamma)X\beta(\gamma)\right)$$
$$=\left(y-X\beta(\gamma)\right)'L^{-1}(\gamma)'L^{-1}(\gamma)\left(y-X\beta(\gamma)\right)|L(\gamma)|^{2/n}. \tag{6.22}$$

In the same way as for the standard method, we can solve for β by using the LS method by setting the first derivative of equation (6.22) equal to zero.

To get the final estimate of β, γ needs to be estimated. To formulate the LS estimation equation for γ, we have the first derivative with respect to γ set equal to zero, which can be written as

$$\left.\frac{\partial S_2(y;\beta,\gamma)}{\partial\gamma}\right|_{\beta=\hat{\beta}(\gamma)}=(y-X\beta)'\frac{\partial\Sigma^{-1}(\gamma)}{\partial\gamma}(y-X\beta)|L(\gamma)|^{\frac{2}{n}}$$
$$+(y-X\beta)'\Sigma^{-1}(\gamma)(y-X\beta)\frac{\partial}{\partial\gamma}|L(\gamma)|^{\frac{2}{n}}=0. \tag{6.23}$$

The first part of the above expression is equivalent to $\left.\frac{\partial S_1(y;\beta,\gamma)}{\partial\gamma}\right|_{\beta=\hat{\beta}(\gamma)}$ except for the term $|L(\gamma)|^{\frac{2}{n}}$. Replacing the parameter β by $\hat{\beta}(\gamma)$, we have

$$\left.\frac{\partial S_1(y;\beta,\gamma)}{\partial\gamma}\right|_{\beta=\hat{\beta}(\gamma)}|L(\gamma)|^{\frac{2}{n}}+\frac{2}{n}\left(y-X\hat{\beta}(\gamma)\right)'\Sigma^{-1}(\gamma)\left(y-X\hat{\beta}(\gamma)\right)$$
$$|L(\gamma)|^{\frac{2}{n}-1}|L(\gamma)|tr\left\{L^{-1}(\gamma)\frac{\partial L(\gamma)}{\partial\gamma}\right\}=0. \tag{6.24}$$

Substituting $\left.\frac{\partial S_1(y;\beta,\gamma)}{\partial\gamma}\right|_{\beta=\hat{\beta}(\gamma)}=-y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y$ in equation (6.24) and after some simple algebra we get

$$-y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y+\frac{2}{n}y'\Delta(\gamma)y\,tr\left\{L^{-1}(\gamma)\frac{\partial L(\gamma)}{\partial\gamma}\right\}=0. \qquad (6.25)$$

Dividing both sides of equation (6.25) by $y'\Delta(\gamma)y$ gives

$$-\frac{-y'\Delta(\gamma)\dfrac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y}{y'\Delta(\gamma)y}+\frac{2}{n}tr\left\{L^{-1}(\gamma)\frac{\partial L(\gamma)}{\partial\gamma}\right\}=0. \qquad (6.26)$$

Multiplying both sides of equation (6.26) by $-\frac{n}{2}$, we have

$$-tr\left[L^{-1}(\gamma)\frac{\partial L(\gamma)}{\partial\gamma}\right]+\frac{n}{2}\frac{y'\Delta(\gamma)\dfrac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y}{e(\gamma)'e(\gamma)}=0. \qquad (6.27)$$

Equation (6.27), which has been derived directly from the least squares estimation equation (6.24), is exactly equation (6.13) which is the estimating equation of the concentrated likelihood. This implies that the ML estimates and the least squares estimates are equivalent and we have a least squares interpretation of the ML estimator without the need for a normality assumption. Thus we conclude that for the general linear regression model, the normality assumption is not crucial for ML estimation.

6.6 Non-stationary AR(1) Error Process

In this section, we discuss estimation of parameters of the model using the alternative method in the case of a nonstationary AR(1) error process. We express the second part of equation (6.27) in terms of a polynomial expression and discuss different avenues for estimating parameters.

6.6.1 Estimation of parameters using method 2

Let us consider the linear model defined by (6.1) with the AR(1) error process

$$u_t = \gamma u_{t-1} + \varepsilon_t \text{ when } t = 2,\cdots,n \qquad (6.28)$$

where $u_1 = \varepsilon_1$ and ε_t is identically and independently distributed $N(0,\sigma^2)$ for $t = 1,\cdots,n$. This is a nonstationary process and differs from a stationary process which requires the variance of each u_t to be constant. This amounts to assuming that $var(\varepsilon_t) = \sigma^2$ for $t \neq 1$ and $\text{var}(\varepsilon_1) = \sigma^2/(1-\gamma^2)$ with $|\gamma| < 1$.

Berenblut and Webb (1973) first considered how to test for non-stationary autocorrelated errors in the linear regression model. In the light of their work, the variance-covariance matrix of the nonstationary process (6.28) can be written as

$$\Sigma(\gamma) = \begin{bmatrix} 1 & \gamma & \gamma^2 & \cdots & \gamma^{n-1} \\ \gamma & 1+\gamma^2 & \gamma^3+\gamma & \cdots & \gamma^n+\gamma^{n-2} \\ \gamma^2 & \gamma^3+\gamma & 1+\gamma^2+\gamma^4 & \cdots & \gamma^{n+1}+\gamma^{n-1}+\gamma^{n-3} \\ \vdots & & & & \vdots \\ \gamma^{n-1} & \gamma^n+\gamma^{n-2} & \gamma^{n+1}+\gamma^{n-1}+\gamma^{n-3} & \cdots & 1+\gamma^2+\cdots+\gamma^{2n-2} \end{bmatrix} \qquad (6.29)$$

in which

$$\Sigma^{-1}(\gamma) = \begin{bmatrix} 1+\gamma^2 & -\gamma & 0 & 0 \cdots & 0 \\ -\gamma & 1+\gamma^2 & -\gamma & 0 \cdots & 0 \\ 0 & -\gamma & 1+\gamma^2 & -\gamma \cdots & 0 \\ & & & & \vdots \\ \vdots & & & & -\gamma \\ 0 & \cdots & \cdots\cdots\cdots & -\gamma & 1 \end{bmatrix} \qquad (6.30)$$

where $\Sigma(\gamma)$ and $\Sigma^{-1}(\gamma)$ are symmetric and positive definite for all γ and we can write $\Sigma(\gamma) = L(\gamma)L(\gamma)'$. Here $L(\gamma)$ is

$$L(\gamma) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \gamma & 1 & 0 & \cdots & 0 \\ \gamma^2 & \gamma & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ \gamma^{n-1} & \gamma^{n-2} & & \cdots & 1 \end{bmatrix} \qquad (6.31)$$

and the inverse of $L(\gamma)$ can be written as

$$L^{-1}(\gamma) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\gamma & 1 & 0 & \cdots & 0 \\ 0 & -\gamma & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & -\gamma & 1 \end{bmatrix} = (I - \gamma B) \qquad (6.32)$$

in which

$$B = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \qquad (6.33)$$

is known as the back shift matrix.

The last term of the first part of equation (6.27) can be expanded in terms of a polynomial as follows

$$tr\left\{L^{-1}(\gamma)\frac{\partial L(\gamma)}{\partial \gamma}\right\}$$

$$= tr[I - \gamma B]\left[\frac{\partial}{\partial \gamma}\left(I + \gamma B + \gamma^2 B^2 + \cdots \gamma^{n-1} B^{n-1}\right)\right]$$

$$= tr[I-\gamma B][B+2\gamma B^2+\cdots+(n-1)\gamma^{n-2}B^{n-1}] \tag{6.34}$$

$$= tr[B+2\gamma B^2+\cdots+(n-1)\gamma^{n-2}B^{n-1}-\gamma B^2-2\gamma^2B^3-\cdots-(n-1)\gamma^{n-1}B^n]$$

$$= [tr(B)+2\gamma tr(B^2)+\cdots-(n-1)\gamma^{n-1}tr(B^n)]=0$$

where

$$L(\gamma)=I+\gamma B+\gamma^2B^2+\cdots\gamma^{n-1}B^{n-1}. \tag{6.35}$$

The first part of the equation (6.27) is equal to zero because of (6.34), which allows us to write equation (6.27) as

$$\frac{n}{2}y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y=0$$

$$\text{or}\quad y'\Delta(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Delta(\gamma)y=0. \tag{6.36}$$

Equation (6.36) can be written as

$$y'\left[\Sigma^{-1}(\gamma)-\Sigma^{-1}(\gamma)X(X\Sigma^{-1}(\gamma)X)^{-1}X\Sigma^{-1}(\gamma)\right]$$
$$\frac{\partial\Sigma(\gamma)}{\partial\gamma}\left[\Sigma^{-1}(\gamma)-\Sigma^{-1}(\gamma)X(X\Sigma^{-1}(\gamma)X)^{-1}X\Sigma^{-1}(\gamma)\right]y=0$$

$$\text{or}\quad y'\left[I-\Sigma^{-1}(\gamma)X(X\Sigma^{-1}(\gamma)X)^{-1}X'\right]\Sigma^{-1}(\gamma)$$
$$\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Sigma^{-1}(\gamma)\left[I-X(X\Sigma^{-1}(\gamma)X)^{-1}X\Sigma^{-1}(\gamma)\right]y=0$$

$$\text{or}\quad \left[y-X(X\Sigma^{-1}(\gamma)X)^{-1}X\Sigma^{-1}(\gamma)y\right]'\Sigma^{-1}(\gamma)$$
$$\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Sigma^{-1}(\gamma)\left[y-X(X\Sigma^{-1}(\gamma)X)^{-1}X\Sigma^{-1}(\gamma)y\right]=0$$

$$\text{or}\quad (y-X\hat\beta(\gamma))'\Sigma^{-1}(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Sigma^{-1}(\gamma)(y-X\hat\beta(\gamma))=0. \tag{6.37}$$

We have

$$\frac{\partial\Sigma(\gamma)}{\partial\gamma}=\frac{\partial}{\partial\gamma}\left\{L(\gamma)L(\gamma)'\right\}=\frac{\partial L(\gamma)}{\partial\gamma}L(\gamma)'+L(\gamma)\frac{\partial L(\gamma)'}{\partial\gamma}$$
$$=\frac{\partial(I-\gamma B)^{-1}}{\partial\gamma}(I-\gamma B)^{-1'}+(I-\gamma B)^{-1}\frac{\partial}{\partial\gamma}(I-\gamma B)^{-1'} \tag{6.38}$$
$$=(I-\gamma B)^{-1}B(I-\gamma B)^{-1}(I-\gamma B)^{-1'}+(I-\gamma B)^{-1}(I-\gamma B)^{-1'}B'(I-\gamma B)^{-1'}.$$

We can write the middle part of equation (6.37) in terms of the B matrix as

$$\Sigma^{-1}(\gamma)\frac{\partial\Sigma(\gamma)}{\partial\gamma}\Sigma^{-1}(\gamma)$$
$$=(I-\gamma B)'(I-\gamma B)(I-\gamma B)^{-1}B(I-\gamma B)^{-1}(I-\gamma B)^{-1'}(I-\gamma B)'(I-\gamma B)$$
$$+(I-\gamma B)'(I-\gamma B)(I-\gamma B)^{-1}(I-\gamma B)^{-1'}B'(I-\gamma B)^{-1'}(I-\gamma B)'(I-\gamma B) \tag{6.39}$$
$$=(I-\gamma B)'B+B'(I-\gamma B)$$
$$=(B+B')-2\gamma B'B.$$

Substituting (6.39) into equation (6.37), we have

$$\left(y-X\hat{\beta}(\gamma)\right)'\left[(B+B')-2\gamma B'B\right]\left(y-X\hat{\beta}(\gamma)\right)=0 \tag{6.40}$$

which is the final form of the estimating equation. Solving equation (6.40), we can obtain the final estimate of the parameter. In the next subsection we show it is possible to express the estimating equation as an infinite or finite degree polynomial in γ.

6.6.2 Expression of the alternative method estimating equation as a polynomial

In this subsection, we express the estimator of β as a polynomial of γ, the autoregressive parameter of the error process. This allows us to substitute this polynomial expression into (6.40) and solve for γ. From equation (6.18) we get

$$\hat{\beta}(\gamma) = \left(X'\Sigma^{-1}(\gamma)X\right)^{-1} X'\Sigma^{-1}(\gamma)y$$

and since $L^{-1}(\gamma) = (I - \gamma B)$, in terms of B, $\hat{\beta}(\gamma)$ is

$$\hat{\beta}(\gamma) = \left\{X'(I-\gamma B)'(I-\gamma B)X\right\}^{-1}\left\{X'(I-\gamma B)'(I-\gamma B)y\right\}$$

$$= \left\{X'\left(I - \gamma(B+B') + \gamma^2 B'B\right)X\right\}^{-1}\left\{X'\left(I-\gamma(B+B')+\gamma^2 B'B\right)y\right\}$$

$$= \left\{X'X - \gamma X'(B+B')X + \gamma^2 X'B'BX\right\}^{-1}\left\{X'y - \gamma X'(B+B')y + \gamma^2 X'B'By\right\}$$

$$= \left\{X'X(X'X)(X'X)^{-1} - \gamma(X'X)(X'X)^{-1}X'(B+B')X + \gamma^2(X'X)(X'X)^{-1}X'B'BX\right\}^{-1}$$
$$\left\{X'y - \gamma X'(B+B')y + \gamma^2 X'B'By\right\}$$
$$= \left\{I - \gamma(X'X)^{-1}X'(B+B')X + \gamma^2(X'X)^{-1}X'B'BX\right\}^{-1}(X'X)^{-1}$$
$$\left\{X'y - \gamma X'(B+B')y + \gamma^2 X'B'By\right\}$$

$$= \left\{I - \left(\gamma(X'X)^{-1}X'(B+B')X - \gamma^2(X'X)^{-1}X'B'BX\right)\right\}^{-1}$$
$$\left\{b - \gamma(X'X)^{-1}X'(B+B')y + \gamma^2(X'X)^{-1}X'B'By\right\}$$
$$= [I + \left(\gamma(X'X)^{-1}X'(B+B')X - \gamma^2(X'X)^{-1}X'B'BX\right)$$
$$+ \left(\gamma(X'X)^{-1}X'(B+B')X - \gamma^2(X'X)^{-1}X'B'BX\right)^2$$
$$+ \left(\gamma(X'X)^{-1}X'(B+B')X - \gamma^2(X'X)^{-1}X'B'BX\right)^3$$
$$+ \left(\gamma(X'X)^{-1}X'(B+B')X - \gamma^2(X'X)^{-1}X'B'BX\right)^4 + \cdots]$$
$$\left\{b - \gamma(X'X)^{-1}X'(B+B')y + \gamma^2(X'X)^{-1}X'B'By\right\}$$

$$
\begin{aligned}
&= [I + \gamma(X'X)^{-1}X'(B'+B)X - \gamma^2(X'X)^{-1}X'B'BX + \gamma^2\left((X'X)^{-1}X'(B'+B)X\right)^2 - \\
&\quad \gamma^3(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX - \gamma^3(X'X)^{-1}X'B'BX(X'X)^{-1} \\
&\quad X'(B'+B)X + \gamma^4\left((X'X)^{-1}X'B'BX\right)^2 + \gamma^3\left((X'X)^{-1}X'B'BX\right)^3 - \\
&\quad \gamma^4\left((X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X\right) - \\
&\quad \gamma^4(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)X - \\
&\quad \gamma^4\left((X'X)^{-1}X'(B'+B)X\right)^2(X'X)^{-1}X'B'BX \\
&\quad + \cdots + \gamma^4\left((X'X)^{-1}X'(B'+B)X\right)^4 + \cdots] \\
&\quad \left[b - \gamma(X'X)^{-1}X'(B'+B)y + \gamma^2(X'X)^{-1}X'B'BX\right] \\
&= b + \gamma(X'X)^{-1}X'(B'+B)Xb - \gamma^2(X'X)^{-1}X'B'BXb + \gamma^2\left((X'X)^{-1}X'(B'+B)X\right)^2 b \\
&\quad - \gamma^3(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)Xb + \gamma^4\left((X'X)^{-1}X'B'BX\right)^2 b \\
&\quad + \gamma^3\left((X'X)^{-1}X'B'BX\right)^3 b - \gamma^4(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX(X'X)^{-1} \\
&\quad X'(B'+B)Xb - \gamma^4(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)Xb - \\
&\quad \gamma^4\left((X'X)^{-1}X'(B'+B)X\right)^2(X'X)^{-1}X'B'BXb + \gamma^4\left((X'X)^{-1}X'(B'+B)X\right)^4 b + \cdots \\
&\quad - \gamma(X'X)^{-1}X'(B'+B)y - \gamma^2(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)y + \\
&\quad \gamma^3(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)y - \gamma^3\left((X'X)^{-1}X'B'BX\right)^2(X'X)^{-1}X'(B'+B)y \\
&\quad + \gamma^4(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)y - \\
&\quad \gamma^4(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)y - \\
&\quad \gamma^4\left((X'X)^{-1}X'(B'+B)X\right)^3(X'X)^{-1}X'(B'+B)y + \cdots + \gamma^2(X'X)^{-1}X'B'By + \\
&\quad \gamma^3(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'By - \gamma^4(X'X)^{-1}X'B'BX(X'X)^{-1}X'B'By \\
&\quad - \gamma^4\left((X'X)^{-1}X'(B'+B)X\right)^2(X'X)^{-1}X'B'By + \cdots
\end{aligned}
$$

$$= b+\gamma[(X'X)^{-1}X'(B'+B)Xb-(X'X)^{-1}X'(B'+B)y]$$

$$+\gamma^2[\left((X'X)^{-1}X'(B'+B)X\right)^2 b-(X'X)^{-1}X'B'BXb$$

$$-(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)y+(X'X)^{-1}X'B'By]$$

$$+\gamma^3[\left((X'X)^{-1}X'(B'+B)X\right)^3 b-(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BXb-$$

$$(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)Xb+(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)y$$

$$-\left((X'X)^{-1}X'(B'+B)X\right)^2(X'X)^{-1}X'(B'+B)y]$$

$$+\gamma^4[\left((X'X)^{-1}X'(B'+B)X\right)^2 b-(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX(X'X)^{-1}$$

$$X'(B'+B)Xb-(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)Xb-$$

$$\left((X'X)^{-1}X'(B'+B)X\right)^2(X'X)^{-1}X'B'BXb+\left((X'X)^{-1}X'(B'+B)X\right)^4 b+$$

$$(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)y$$

$$-(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)y$$

$$-\left((X'X)^{-1}X'(B'+B)X\right)^3(X'X)^{-1}X'(B'+B)y-(X'X)^{-1}X'B'BX(X'X)^{-1}X'B'By$$

$$+\left((X'X)^{-1}X'(B'+B)X\right)^2(X'X)^{-1}X'B'By]+\cdots$$

If we replace the above values in $\hat{\beta}(\gamma)$, we can write

$$\hat{\beta}(\gamma)=b+\gamma A+\gamma^2 C+\gamma^3 D+\gamma^4 E+\cdots \tag{6.41}$$

where

$$A=(X'X)^{-1}X'(B'+B)Xb-(X'X)^{-1}X'(B'+B)y$$

$$b=(X'X)^{-1}X'y$$

$$C=\left((X'X)^{-1}X'(B'+B)X\right)^2 b-(X'X)^{-1}X'B'BXb-(X'X)^{-1}X'(B'+B)X(X'X)^{-1}$$
$$X'(B'+B)y+(X'X)^{-1}X'B'By$$

$$D=\left((X'X)^{-1}X'(B'+B)X\right)^3 b-(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BXb-$$
$$(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)Xb+(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)y$$
$$-\left((X'X)^{-1}X'(B'+B)X\right)^2(X'X)^{-1}X'(B'+B)y$$

$$
\begin{aligned}
E = {} & \left((X'X)^{-1}X'(B'+B)X\right)^2 b - (X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)Xb \\
& -(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)Xb - \\
& \left((X'X)^{-1}X'(B'+B)X\right)^2 (X'X)^{-1}X'B'BXb + \left((X'X)^{-1}X'(B'+B)X\right)^4 b + \\
& (X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)y \\
& -(X'X)^{-1}X'B'BX(X'X)^{-1}X'(B'+B)X(X'X)^{-1}X'(B'+B)y - \left((X'X)^{-1}X'(B'+B)X\right)^3 \\
& (X'X)^{-1}X'(B'+B)y - (X'X)^{-1}X'B'BX(X'X)^{-1}X'B'By + \left((X'X)^{-1}X'(B'+B)X\right)^2 \\
& (X'X)^{-1}X'B'By
\end{aligned}
$$

and B is as in (6.33).

When we replace $\hat{\beta}(\gamma)$ in equation (6.40), we have

$$
\begin{aligned}
& \left(y - X[b+\gamma A+\gamma^2 C+\gamma^3 D+\gamma^4 E+\cdots]\right)'(B+B')\left(y - X[b+\gamma A+\gamma^2 C+\gamma^3 D+\gamma^4 E+\cdots]\right) \\
& -2\gamma\left(y - X[b+\gamma A+\gamma^2 C+\gamma^3 D+\gamma^4 E+\cdots]\right)'(B'B)\left(y - X[b+\gamma A+\gamma^2 C+\gamma^3 D+\gamma^4 E+\cdots]\right) = 0 \\
& \left(y - Xb - \gamma XA - \gamma^2 XC - \gamma^3 XD - \gamma^4 XE - \cdots\right)'(B+B') \\
& \left(y - Xb - \gamma XA - \gamma^2 XC - \gamma^3 XD - \gamma^4 XE - \cdots\right) - 2\gamma\left(y - Xb - \gamma XA - \gamma^2 XC - \gamma^3 XD - \gamma^4 XE - \cdots\right)' \\
& (B'B)\left(y - Xb - \gamma XA - \gamma^2 XC - \gamma^3 XD - \gamma^4 XE - \cdots\right) = 0
\end{aligned}
$$

$$
\begin{aligned}
& (y-Xb)'(B+B')(y-Xb) - \gamma A'X'(B+B')(y-Xb) - \gamma^2 C'X'(B+B')(y-Xb) \\
& \quad - \gamma^3 D'X'(B+B')(y-Xb) - \gamma^4 E'X'(B+B')(y-Xb) - \cdots
\end{aligned}
$$

$$
\begin{aligned}
&(y-Xb)'(B+B')(y-Xb)-\gamma[A'X'(B+B')(y-Xb)+(y-Xb)'(B+B')XA \\
&\qquad -2(y-Xb)'B'B(y-Xb)] \\
&\qquad -\gamma^2[C'X'(B+B')(y-Xb)-A'X'(B+B')XA+(y-Xb)'(B+B')XC \\
&\qquad -A'X'B'B(y-Xb)-(y-Xb)'B'BXA]+\gamma^3[D'X'(B+B')(y-Xb) \\
&\qquad -C'X'(B+B')XA-A'X'(B+B')XC+(y-Xb)'(B+B')XD \\
&\qquad -2C'X'B'B(y-Xb)+2A'X'B'BXA+2(y-Xb)'B'BXC] \\
&\qquad -\gamma^4[E'X'(B+B')(y-Xb)-D'X'(B+B')XA-C'X'(B+B')XC \\
&\qquad -A'X'(B+B')XD+(y-Xb)'(B+B')XE+2D'X'B'B(y-Xb) \\
&\qquad +2C'X'B'BXA+2A'X'B'BXC-2(y-Xb)'B'BXD+\cdots=0.
\end{aligned}
$$

Finally we can write equation (6.40) as a polynomial of infinite order

$$
P+\gamma Q+\gamma^2 R+\gamma^3 S+\gamma^4 T+\cdots=0 \tag{6.42}
$$

where

$$P=(y-Xb)'(B+B')(y-Xb)$$

$$Q=-[A'X'(B+B')(y-Xb)+(y-Xb)'(B+B')XA-2(y-Xb)'B'B(y-Xb)]$$

$$
\begin{aligned}
R=-[&C'X'(B+B')(y-Xb)-A'X'(B+B')XA+(y-Xb)'(B+B')XC-A'X'B'B(y-Xb) \\
&-(y-Xb)'B'BXA]
\end{aligned}
$$

$$
\begin{aligned}
S=-[&D'X'(B+B')(y-Xb)-C'X'(B+B')XA-A'X'(B+B')XC+(y-Xb)'(B+B')XD \\
&-2C'X'B'B(y-Xb)+2A'X'B'BXA+2(y-Xb)'B'BXC]
\end{aligned}
$$

$$
\begin{aligned}
T=-[&E'X'(B+B')(y-Xb)-D'X'(B+B')XA-C'X'(B+B')XC-A'X'(B+B')XD \\
&+(y-Xb)'(B+B')XE+2D'X'B'B(y-Xb)+2C'X'B'BXA+ \\
&2A'X'B'BXC-2(y-Xb)'B'BXD
\end{aligned}
$$

and B is as (6.33).

By using the same approach, we can also express the estimating equation of the marginal likelihood in terms of an infinite degree polynomial. The only difference with that for the concentrated likelihood is that the coefficients are different.

A special case of equation (6.40) is when X is a vector (the order of X is $n \times 1$). In that case, to get rid of the inverse term, we can multiply through by $\left(X'\Sigma(\gamma)^{-1}X\right)$ which is scalar. We then have

$$\left(X'\Sigma(\gamma)^{-1}Xy - XX'\Sigma(\gamma)^{-1}y\right)'[B + B' - 2\gamma B'B]\left(X'\Sigma(\gamma)^{-1}Xy - XX'\Sigma(\gamma)^{-1}y\right) = 0. \tag{6.43}$$

This can be expanded as

$$\left(y\left(X'\left(I-\gamma(B+B')+\gamma^2B'B\right)X\right) - X\left(X'\left(I-\gamma(B+B')+\gamma^2B'B\right)y\right)\right)'(B+B'-2\gamma B'B)$$
$$\left(y\left(X'\left(I-\gamma(B+B')+\gamma^2B'B\right)X\right) - X\left(X'\left(I-\gamma(B+B')+\gamma^2B'B\right)y\right)\right)$$

$$= [(yX'X - XX'y)'(B+B') - \gamma\left((yX'(B+B')X - XX'(B'+B)y)'(B+B') + 2(yX'X - XX'y)'B'B\right)$$
$$+\gamma^2\left((yX'B'BX - XX'B'By)'(B+B') + 2(yX'(B'+B)X - XX'(B'+B)y)'B'B\right)$$
$$-\gamma^3\left(2(yX'B'BX - XX'B'By)'B'B\right)]$$
$$[(yX'X - XX'y) - \gamma(yX'(B'+B)X - XX'(B'+B)y) + \gamma^2(yX'B'BX - XX'B'By)]$$

$$= (yX'X - XX'y)'(B+B')(yX'X - XX'y)$$
$$-\gamma[(yX'(B'+B)X - XX'(B+B')y)'(B+B')(yX'X - XX'y)$$
$$+2(yX'X - XX'y)'B'B(yX'X - XX'y)$$
$$+(yX'X - XX'y)'(B+B')(yX'(B'+B)X - XX'(B+B')y)]$$
$$+\gamma^2[(yX'B'BX - XX'B'By)'(B+B')(yX'X - XX'y)+$$
$$(yX'(B'+B)X - XX'(B+B')y)'(B+B')(yX'(B'+B)X - XX'(B+B')y)$$
$$+2(yX'X - XX'y)'B'B(yX'(B'+B)X - XX'(B+B')y)$$
$$+(yX'B'BX - XX'B'By)(yX'X - XX'y)']$$
$$-\gamma^3[2(yX'B'BX - XX'B'By)'B'B(yX'X - XX'y)+$$
$$(yX'B'BX - XX'B'By)'(B+B')(yX'(B'+B)X - XX'(B+B')y)$$
$$+2(yX'(B'+B)X - XX'(B+B')y)'B'B(yX'(B'+B)X - XX'(B+B')y)$$
$$+(yX'(B'+B)X - XX'(B+B')y)'(B+B')(yX'B'BX - XX'B'By)$$
$$+2(yX'X - XX'y)'B'B(yX'B'BX - XX'B'By)]$$
$$+\gamma^4[(yX'B'BX - XX'B'By)'(B+B')(yX'B'BX - XX'B'By)$$
$$+2(yX'(B'+B)X - XX'(B+B')y)'B'B(yX'B'BX - XX'B'By)$$
$$+2(yX'B'BX - XX'B'By)'B'B(yX'(B'+B)X - XX'(B+B')y)]$$
$$-\gamma^5[2(yX'B'BX - XX'B'By)'B'B(yX'B'BX - XX'B'By)].$$

Thus we can write equation (6.43) as a fifth degree polynomial in γ, namely

$$a_1 + b_1\gamma + c_1\gamma^2 + d_1\gamma^3 + e_1\gamma^4 + f_1\gamma^5 \tag{6.44}$$

where $a_1 = (yX'X - XX'y)'(B+B')(yX'X - XX'y)$,

$$b_1 = -[(yX'(B'+B)X - XX'(B+B')y)'(B+B')(yX'X - XX'y) +$$
$$2(yX'X - XX'y)'B'B(yX'X - XX'y) + (yX'X - XX'y)'(B+B')$$
$$(yX'(B'+B)X - XX'(B+B')y)]$$
$$c_1 = [(yX'B'BX - XX'B'By)'(B+B')(yX'X - XX'y) +$$
$$(yX'(B'+B)X - XX'(B+B')y)'(B+B')(yX'(B'+B)X - XX'(B+B')y)$$
$$+2(yX'X - XX'y)'B'B(yX'(B'+B)X - XX'(B+B')y) +$$
$$(yX'B'BX - XX'B'By)(yX'X - XX'y)']$$
$$d_1 = -[2(yX'B'BX - XX'B'By)'B'B(yX'X - XX'y) +$$
$$(yX'B'BX - XX'B'By)'(B+B')(yX'(B'+B)X - XX'(B+B')y)$$
$$+2(yX'(B'+B)X - XX'(B+B')y)'B'B(yX'(B'+B)X - XX'(B+B')y)$$
$$+(yX'(B'+B)X - XX'(B+B')y)'(B+B')(yX'B'BX - XX'B'By)$$
$$+2(yX'X - XX'y)'B'B(yX'B'BX - XX'B'By)]$$

$$e_1 = [(yX'B'BX - XX'B'By)'(B+B')(yX'B'BX - XX'B'By)$$
$$+2(yX'(B'+B)X - XX'(B+B')y)'B'B(yX'B'BX - XX'B'By)$$
$$+2(yX'B'BX - XX'B'By)'B'B(yX'(B'+B)X - XX'(B+B')y)],$$

and

$$f_1 = -[2(yX'B'BX - XX'B'By)'B'B(yX'B'BX - XX'B'By)].$$

Therefore we finally obtain a special form of the estimating equation that is a fifth degree polynomial when X is a vector. When X is not a vector, we have an infinite order polynomial, which is shown in equation (6.42). Given that it is a polynomial in γ and we are assuming $|\gamma| < 1$, then it can safely be approximated by a finite degree polynomial. The fact that it is a polynomial does suggest strongly the possibility that there may be multiple roots. This translates into multiple local maxima. This issue was discussed in earlier chapters. It helps if we can know where these local maxima might be. In the next section, we will discuss different approaches to finding roots of approximating polynomials with this in mind.

6.7 Solution for Multiple Roots

In the previous section, we expressed the estimating equation as a polynomial of infinite degree for general X and fifth degree when X is a vector for a nonstationary AR(1) process with $|\gamma| < 1$. To approximate an infinite degree polynomial without this restriction is not feasible. That is, an advantage of this assumption is that it allows us to take a finite degree polynomial as an approximation. In this context, there are a few possible approximate solutions that can be calculated by using an approximating polynomial up to degree four. For greater accuracy, an iterative procedure based on (6.40) can be used.

One question we must address is whether the roots of the estimating equation are local or global maxima or minima of the likelihood function. There also may be circumstances in which some of the roots are imaginary. There is always the possibility that all roots of the polynomial are imaginary. It is also possible that the global maxima is not a root but is a boundary point. For example, we would expect the maxima to occur either at the negative boundary or at the positive boundary when all roots are imaginary. In the next subsection, we will discuss a few approximate solutions that exploit the restriction $|\gamma| < 1$.

6.7.1 First approximation

Analytically it is impossible to solve a polynomial of infinite order. So an obvious approach would be to approximate this polynomial with a finite polynomial. In particular we can take account of the fact that $|\gamma| < 1$. This allows us to ignore higher order terms as being zero or near zero.

The simplest polynomial that results from ignoring second order and higher order terms is $P + \gamma Q = 0$. After some simple arithmetic we get the following estimate of the parameter

$$\hat{\gamma} = -P / Q \tag{6.45}$$

where

$P = (y - Xb)'(B + B')(y - Xb)$,

$Q = -[A'X'(B + B')(y - Xb) + (y - Xb)'(B + B')XA - 2(y - Xb)'B'B(y - Xb)]$,

$b = (X'X)^{-1}X'y$ and B is as (6.33).

When the value of γ is close to zero, equation (6.45) provides a good approximate solution and estimation of γ becomes straightforward. On the other hand, if the value of γ is not close to zero, this first order approximation may not provide a good estimate and consideration of a second-degree polynomial approximation may be appropriate. We discuss this in the next subsection.

6.7.2 Second approximation

We now consider the second-degree polynomial approximation to equation (6.42) that takes the form

$$P + \gamma Q + \gamma^2 R = 0. \tag{6.46}$$

Using the quadratic solution we have two roots of the form :

$$\hat{\gamma} = \left(-Q \pm \left\{(-Q)^2 - 4PR\right\}^{1/2}\right)/2P$$

where $P = (y - Xb)'(B + B')(y - Xb)$

$Q = -[A'X'(B + B')(y - Xb) + (y - Xb)'(B + B')XA - 2(y - Xb)'B'B(y - Xb]$,

$R = -[C'X'(B + B')(y - Xb) - A'X'(B + B')XA + (y - Xb)'(B + B')XC$
$\quad - A'X'B'B(y - Xb) - (y - Xb)'B'BXA]$.

The solution of the second-degree polynomial (6.46) may provide different roots, which will either be real roots or imaginary roots. Our concern is with choosing the appropriate root from this quadratic equation that will give an appropriate estimate of the parameters.

When the roots of the quadratic equation (6.46) are both real but different, we can substitute them into the likelihood function in turn, and choose the one that produces the largest likelihood function as our desired estimate. On the other hand, when roots of the quadratic equation are both imaginary, the likelihood function does not have a turning point. In this situation, the likelihood function is maximized either at the positive boundary, $\gamma = 1$, or at the negative boundary, $\gamma = -1$. As we mentioned in previous chapters, boundary points can be local or global maxima. To check whether we have a local or a global maximum, we should always substitute these boundary values in the likelihood function. In the next subsection we consider a third degree approximating polynomial.

6.7.3 Third approximation

For greater accuracy, we can use a third order approximation for finding the approximate root of the polynomial under the assumption that $|\gamma|<1$. In this context, the estimating equation (6.42) is of the form

$$\gamma^3 S+\gamma^2 R+\gamma Q+P=0 \tag{6.47}$$

where $P=(y-Xb)'(B+B')(y-Xb)$,

$$Q=-[A'X'(B+B')(y-Xb)+(y-Xb)'(B+B')XA-2(y-Xb)'B'B(y-Xb)],$$

$$\begin{aligned}R=-[&C'X'(B+B')(y-Xb)-A'X'(B+B')XA+(y-Xb)'(B+B')XC\\&-A'X'B'B(y-Xb)-(y-Xb)'B'BXA],\end{aligned}$$

$$\begin{aligned}S=-[&D'X'(B+B')(y-Xb)-C'X'(B+B')XA-A'X'(B+B')XC\\&+(y-Xb)'(B+B')XD-2C'X'B'B(y-Xb)+2A'X'B'BXA+2(y-Xb)'B'BXC].\end{aligned}$$

We can solve for the roots of equation (6.47) using the Cardan's formula. We follow the same steps as outlined by Conkright (1951). The steps are as follows: First eliminate the second degree from (6.47) by using the transformation

$$\gamma=Z-R/3S. \tag{6.48}$$

Substituting (6.48) in equation (6.47), we get the transformed equation

$$Z^3 + 3HZ + G = 0, \tag{6.49}$$

where $H = (3SQ - R^2)/9S^2$ and $G = (2R^3 - 9SRQ + 27S^2P)/27S^3$. It is convenient to refer to (6.49) as the reduced cubic equation. The roots of equation (6.47) can be obtained from (6.48) when the roots of (6.49) are known. The next calculation follows the different steps for solving the roots of equation (6.49). We wish to choose two numbers in such a way that

$$Z = o + v \tag{6.50}$$

satisfies equation (6.49). Substituting (6.50) in equation (6.49) and after some algebra we have

$$o^3 + v^3 + 3(ov + H)(o + v) + G = 0. \tag{6.51}$$

Observe that equation (6.51) will be satisfied if o and v are chosen such that

$$o^3 + v^3 = -G \tag{6.52}$$

and

$$ov = -H. \tag{6.53}$$

Therefore an appropriate choice of o and v can be found by solving equations (6.52) and (6.53) simultaneously. From (6.53)

$$v = -H / o$$

so that equation (6.52) can be written as

$$o^3 - \frac{H^3}{o^3} + G = 0$$

or

$$o^6 + Go^3 - H^3 = 0. \tag{6.54}$$

Solving equation (6.54) as a quadratic solution in o^3 and v^3, we have

$$o^3 = \frac{-G+\sqrt{G^2+4H^3}}{2} \tag{6.55}$$

and

$$\nu^3 = \frac{-G-\sqrt{G^2+4H^3}}{2}. \tag{6.56}$$

The solution (see Conkright (1951)), known as Cardan's formula is

$$Z = o+\nu = \left[\frac{-G+\sqrt{G^2+4H^3}}{2}\right]^{\frac{1}{3}} + \left[\frac{-G-\sqrt{G^2+4H^3}}{2}\right]^{\frac{1}{3}} \tag{6.57}$$

where the two cube roots must be chosen to satisfy (6.53), i.e., their product must be $-H$. There will be normally three combinations of roots that satisfy this requirement giving equations (6.55) and (6.56). These three roots are $Z_1 = o_1 + \nu_1$, $Z_2 = \omega o_1 + \omega^2 \nu_1$, and $Z_3 = \omega^2 o_1 + \omega \nu_1$ where ω and ω^2 are the imaginary cube roots of unity. Once we know the roots of Z, it is easy to find the roots of γ by solving the relation defined in equation (6.48).

The solution of the third degree polynomial may produce one real root or three real roots. Only one real root can be used directly as our final estimate of the parameter. For three real roots, we follow the same steps as explained in the Subsection 6.7.2. When the third degree polynomial produces real and imaginary roots, we choose the real root as the final root. The question remains unsolved in the case of multiple maxima where the likelihood function has more than one turning points. There is also the possibility that the maximum value of the likelihood is at either the positive or negative boundary. Under these circumstances, in attempting to choose the root of the polynomial, we should also take into account the two boundary points as a part of the maximization. Finally we choose the point among the real root and the two boundary points which provides the largest value of the likelihood function. In the next subsection we will discuss how to find roots of a polynomial of order four for greater accuracy.

6.7.4 Fourth approximation

A fourth degree polynomial approximation of equation (6.42) is

$$\gamma^4 T + \gamma^3 S + \gamma^2 R + \gamma Q + P = 0. \tag{6.58}$$

In order to solve this equation, we introduce r, j and k such that

$$\gamma^4 + \widetilde{S}\gamma^3 + \widetilde{R}\gamma^2 + \widetilde{Q}\gamma + \widetilde{P} + (r\gamma + j)^2 \equiv \left(\gamma^2 + \frac{\widetilde{S}}{2}\gamma + k\right)^2. \tag{6.59}$$

The determination of r, j and k depends on equating the coefficient of the first and the second order of γ or

$$\begin{aligned}\gamma^4 + \widetilde{S}\gamma^3 + (R^2 + r^2)\gamma^2 + (2rj + \widetilde{Q})\gamma + j^2 + \widetilde{P} \\ \equiv \gamma^4 + \widetilde{S}\gamma^3 + \left(\frac{\widetilde{S}^2}{4} + 2k\right)\gamma^2 + \gamma\widetilde{S}k + k^2.\end{aligned} \tag{6.60}$$

where $\widetilde{S} = \frac{S}{T}$, $\widetilde{R} = \frac{S}{T}$, $\widetilde{Q} = \frac{Q}{T}$ and $\widetilde{P} = \frac{P}{T}$. From the above expression we can write

$$\widetilde{R} + r^2 = \frac{\widetilde{S}^2}{4} + 2k \tag{6.61}$$

$$2rj + \widetilde{Q} = \widetilde{S}k \tag{6.62}$$

$$j^2 + \widetilde{P} = k^2. \tag{6.63}$$

From (6.62) we have

$$\left(\widetilde{S}k-\widetilde{Q}\right)^2=4r^2j^2, \tag{6.64}$$

from (6.61) we have

$$r^2=\frac{\widetilde{S}^2}{4}+2k-\widetilde{R} \tag{6.65}$$

and from (6.63) we get

$$j^2=k^2-\widetilde{P}. \tag{6.66}$$

Using (6.65) and (6.66), we can write (6.64) as

$$\left(\widetilde{S}k-\widetilde{Q}\right)^2=4\left(\frac{\widetilde{S}^2}{4}+2k-\widetilde{R}\right)\left(k^2-\widetilde{P}\right)$$

$$\begin{aligned}\widetilde{S}^2k^2-2\widetilde{S}k\widetilde{Q}+\widetilde{Q}^2&=4\left(\frac{\widetilde{S}^2k^2}{4}+2k^3-2k\widetilde{P}-\widetilde{R}k^2+\widetilde{R}\widetilde{P}-\frac{\widetilde{S}^2\widetilde{P}}{4}\right)\\&=\left(\widetilde{S}^2k^2+8k^3-8k\widetilde{P}-4\widetilde{R}k^2+4\widetilde{R}\widetilde{P}-\widetilde{S}^2\widetilde{P}\right)\end{aligned}$$

$$\text{or } 8k^3-4k^2\widetilde{R}-8k\widetilde{P}+4\widetilde{R}\widetilde{P}-\widetilde{S}^2\widetilde{P}+2\widetilde{S}k\widetilde{Q}-\widetilde{Q}^2=0$$

$$\text{or } k^3-\frac{\widetilde{R}}{2}k^2+\frac{1}{4}\left(\widetilde{S}\widetilde{Q}-4\widetilde{P}\right)k+\frac{1}{8}\left(4\widetilde{k}P-\widetilde{S}^2\widetilde{P}-\widetilde{Q}^2\right)=0. \tag{6.67}$$

In (6.67) k is unknown. Once k is known, r and j can be obtained by substituting in (6.61) and (6.62). Adding $(r\gamma+j)^2$ to both members of (6.59), an equation is obtained in which both members are perfect squares. We have in fact,

$$\gamma^2+\frac{\widetilde{S}}{2}\gamma+k=r\gamma+j \tag{6.68}$$

$$\gamma^2+\frac{\widetilde{S}}{2}\gamma+k=-(r\gamma+j) \tag{6.69}$$

and the four roots of (6.59) can be found by solving the quadratic equations (6.68) and (6.69), which are known as Ferrari's solution (see Conkright (1951)). We use the same procedure as discussed above to choose the root from the real roots of the polynomial as the final estimate. That is we choose that root (or boundary value) that has the highest value of the likelihood function. In the case of all four roots being imaginary roots, we seek the global maxima by looking at the values of the likelihood at the two boundary points.

In Subsections 6.7.1 through 6.7.4, we discussed different methods for finding roots of approximating polynomials that provided an estimator of γ with increasing accuracy. We have seen that the order of a good approximating polynomial depends to some degree on the value of $|\gamma|$. For example, for smaller values of $|\gamma|$, all higher terms of the polynomial become zero quickly and equation (6.42) can be approximated by a lower degree polynomial. However, for higher values of $|\gamma|$, we need to consider higher degrees for the approximating polynomial. For greater accuracy, we can use equation (6.40) to derive an iterative procedure as outlined below.

6.7.5 Numerical approximation

From equation (6.40), we can write

$$\tilde{\gamma}_{i+1} = \frac{\left(y - X\beta(\tilde{\gamma}_i)\right)'(B + B')\left(y - X\beta(\tilde{\gamma}_i)\right)}{2\left(y - X\beta(\tilde{\gamma}_i)\right)' B'B\left(y - X\beta(\tilde{\gamma}_i)\right)}. \tag{6.70}$$

We can estimate γ by using an iterative procedure that is based on equation (6.70). For convenience, equation (6.70) can be rewritten as

$$\tilde{\gamma}_{i+1} = f(\tilde{\gamma}_i). \tag{6.71}$$

The iterative procedure begins with a first approximated value of $\tilde{\gamma}$ namely $\tilde{\gamma}_0 = -P/Q$ to start the iterative procedure where P and Q are defined in equation (6.45). Substitute this estimated value in equation (6.71) to obtain the updated value of $\tilde{\gamma}_1$; $\tilde{\gamma}_1$ to obtain $\tilde{\gamma}_2$ and so on. Repeat the process until the difference between successive updated values meets a tolerance limit, i.e., $|\tilde{\gamma}_{i+1} - \tilde{\gamma}_i| \le \delta_c$, where δ_c is a small positive number chosen by the researcher. The final estimate of γ is $\tilde{\gamma}_{i+1}$.

As we know from the previous discussion, the likelihood function may have multiple maxima in the context of MA(1) error process, which raises the issue of whether we have converged to a local or global maximum. In the context of a non–stationary AR(1) process, the iterative process converges to one root which may not be a global maximum. It makes sense to address this issue with multiple starting values to start up the iterative procedure and investigate the convergence of local maxima as we discussed in Chapter 3.

6.8 Conclusion

In this chapter, we investigated the form of the estimating equation for the LS method and that of the concentrated likelihood for estimating the parameters in the covariance matrix of the general linear regression model. In the application of the LS method, we transformed the linear model by pre-multiplying by the n th root of the determinant of the lower triangular Cholesky decomposition of the covariance matrix times the inverse of the same matrix.

The main contribution of this chapter is that the estimating equation for the LS method applied to our modified transformed model is the same as the estimating equation of the concentrated likelihood. This result tells us that instead of using the concentrated likelihood with a normality assumption, we can apply the LS method without the normality assumption when estimating parameters of the covariance matrix. The later results in the same estimates as the former approach but does not need normality to be assumed.

We also showed that the estimating equation can be expressed as a polynomial of fifth degree when X is a vector in the context of the linear model with a non-stationary AR(1) error process. With the same error process, we demonstrated that the estimating equation is an infinite degree polynomial of the parameter when X is not a vector. We conclude that instead of using an infinite degree polynomial, which is analytically difficult or even impossible to solve, we can use an approximating polynomial of finite degree, which is relatively easier to solve.

For smaller values of $|\gamma|$, we discussed different theoretical solutions for finding the roots of the polynomial up to degree four. For higher values of γ, we suggest an iterative formula to obtain the approximate roots of the polynomial with greater accuracy. However, in practice, there is the usual problem concerning whether the process converges to a local maximum or global maximum of the likelihood function.

Usually it is difficult to find the solution of the polynomial of higher order. In this chapter, special attention was paid to the problem of deciding between multiple roots. Our method works well under the restriction of $|\gamma| \leq 1$ for higher degree polynomials. We have seen that the solution of a polynomial contains real roots, imaginary roots or both. In the case of real roots, we suggest choosing the one that gives the largest value of the concentrated likelihood with appropriate consideration also being given to boundary points. In the case of when all roots are imaginary, the possible maximum might be either the positive boundary or negative boundary and we suggest taking the value that makes the concentrated likelihood function largest.

In light of Chapter 3 and Chapter 5, we found that selection of root on the basis of maximising the likelihood function may not always achieve the best estimate of the parameter. It therefore seems a good idea to consider whether the final estimate that comes from the finding the global maximum is actually a sensible estimate. When it does not seem sensible, our results from previous chapters suggest there may be gains in going to a local maximum which gives a more sensible estimate in the case of the concentrated likelihood.

CHAPTER 7

Summary and Concluding Remarks

7.1 Concluding Remarks

Econometrics is concerned with extracting the best possible information from data in order to make inferences using econometric or statistical models. It is imperative to study the properties of estimators with particular emphasis on the accuracy of the ensuing inference. Forecasting plays an important role in many branches of science and it is thought a good forecast depends on the reliable statistical or econometric model.

The work described in this thesis involved the investigation and analysis of four major areas. The first was to study the importance of finding the global maximum of the likelihood function and the consequences of accepting local maxima on the size of the standard test of a regression coefficient when estimation is based on maximum likelihood in the presence of MA(1) errors. The poor results based on maximum likelihood estimates motivated an investigation of the use marginal likelihood based estimates and the consideration of different approaches to counting the degrees of freedom in an attempt to improve the results. The third major area was to reassess the consequences of accepting local maxima of the concentrated and marginal likelihood functions on forecasting performance and make a comparison between OSAF and TSAF. We also investigated whether estimating the first term of the forecasting error improves forecasting performance in comparison to the usual approach of setting the first term to zero.

The fourth area was to examine whether there is equivalence between the LS and the maximum likelihood estimation methods by using the estimating equation approach in the case of the general linear regression model. We also explored the consequences of this estimating equation approach when estimating the covariance matrix parameter in the context of a non-stationary AR(1) error process. In the section that follows, the major contributions of the thesis are summarized and conclusions drawn. We conclude this chapter by giving some suggestions for topics for future research.

7.2 Summary and Conclusions of the Thesis

Chapter 2 reviewed several topics with particular emphasis on the estimation of parameters of the linear regression model in presence of MA(1) errors. The survey began with a brief historical discussion of likelihood-based inference with particular emphasis on estimation of parameters. We also surveyed the literature on the simulated annealing algorithm used for optimization. Our review of numerical estimation difficulties in autocorrelated linear models highlighted that numerical optimisation techniques can have some difficulties in finding the global maxima when estimation is based on maximum likelihood. The literature also suggests that the use of the marginal likelihood in place of the concentrated likelihood can reduce the estimation bias in small samples. We also reviewed the literature regarding the usefulness of estimating the parameters in the covariance matrix of the general linear regression model using estimating equations.

In Chapter 3, we investigated the consequences on the sizes of tests of regression coefficients, of different strategies, including the use of the simulated annealing algorithm, for maximising the likelihood function when a local maximum is distinctly possible. Unfortunately our simulation results found that the traditional likelihood based tests statistics often show very inaccurate and woefully unacceptable sizes particularly in small samples. Our results suggest that always looking for the global maximum is not necessarily the best approach for estimation.

In Chapter 4, we reassessed a number of issues with an eye to improving the accuracy of the estimated sizes. We proposed modified test statistics based on corrected degrees of freedom for both the concentrated likelihood and the marginal likelihood estimates. We investigated whether there is any improvement due to the modified test over the existing test for different optimisation strategies, sample sizes and moving average parameter values, through Monte Carlo simulation. The simulation results indicate that tests based on marginal likelihood perform best irrespective of strategies, samples size, and values of the moving average parameter, in terms of accuracy of sizes. We discovered the not unexpected result that as the sample size increases keeping other factors unchanged, there is a tendency for test sizes to converge to the nominal size. From our results, we recommend the use of the test statistic based on the marginal likelihood combined with the strategy (iv), which involves the use of SA. In other words, finding the global maximum of the marginal likelihood is the best strategy in contrast to that for ML.

In Chapter 5 we looked at the consequences of accepting local maxima on forecasting performance in the context of the linear regression model when the error term follows an MA(1) process. We compared the forecasting performance for two methods of estimation, six different strategies, three different sample sizes, different values of the moving average parameter for one-step-ahead and two-step-ahead forecasts in turn and evaluated the performances of the estimators using AMSFE, AAMFE and AMAFE criteria. We also looked at the issue of whether the first term of the forecasting error recursion should be estimated from the data or just set to zero. The simulation results suggest that forecasting based on the MML estimator is much better than that based on ML estimation. Furthermore, MML provides the best forecasts for different values of MA(1) parameter, different strategies and different sample sizes in the context of stationary and non-stationary design matrices. We found that overall, adaptation of strategy (vi), i.e., always finding the global maximum via grid search, is always the optimal choice when estimation is based on MML. We therefore recommend the use of the MML estimator and strategy (vi) for forecasting with a linear regression model when the errors follow an MA(1) process. An interesting contribution of this chapter is that marginal likelihood based TSAF show an better performance in forecasting for small samples whereas marginal likelihood based OSAF are best only for stationary design matrices in the context of moderate and large samples.

In Chapter 6, we investigated the form of the estimating equation for the LS method and that of the concentrated likelihood for estimating the nuisance parameters of the general linear regression model. We found that the estimating equation for the LS method applied to a transformed model is the same as the estimating equation of the concentrated likelihood. This result tells us that instead of using the concentrated likelihood with a normality assumption we can apply the LS method without the normality assumption when estimating parameters. We found that the estimating equation can be expressed as a fifth degree polynomial when there is a single explanatory variable in the context of the linear model with a non-stationary AR(1) error process and the estimating equation is an infinite degree polynomial when there are multiple explanatory variables. We demonstrated that instead of using an infinite degree polynomial, which is analytically difficult or even impossible to solve, we could use an approximating polynomial of finite degree, which is relatively easier to solve. We discussed different theoretical solutions for finding the roots of an approximating polynomial up to degree four and outlined an iterative formula for finding the roots of the infinite polynomial. Finally special attention was paid to the choice of a root as our desired estimate from the multiple roots.

7.3 Future Work

Given the encouraging results of our research on estimation, testing and forecasting in a linear regression model with MA (1) error process, there is much further work that can be done. A few suggestions for future research are outlined below.

This thesis discusses the issue of finding the global maximum of a likelihood function when estimating the parameters of the general linear regression model with an MA(1) error process using different strategies when a local maximum is distinctly possible. In our case, the error terms are assumed to be normally distributed. A possible extension is to repeat the experiments of Chapter 3 and Chapter 4 for non-normal disturbances and for other error process, such as, MA*(p)* and AR*(q)*.

This thesis discussed the problem of multiple roots for the non-stationary AR(1) processes analytically using an estimating equation. In the literature, there have been various Monte Carlo studies involving likelihood estimators and tests see for example Tanaka (1990), Saikkonen and Luukkonen (1993), Davis and Dunsmuir (1996). There is a question of how much care was taken in some of these studies in choosing the appropriate roots of the estimating equations as estimates. Reworking these studies by taking great care to find the global maximum of the likelihood function may lead to different conclusions. One potential area of considerable interest is the application of the estimating equation to dynamic models.

Another promising result of this thesis is the use of MML methods to estimate the parameters of the model and the use the estimated model for OSAF and TSAF in the context of an MA(1) error process. Further extensions of this work is possible for one-step-ahead forecasts, two-step-ahead forecast and three-step-ahead forecast in the context of linear regression model with MA(2) errors process.

After the simulation experiments in this thesis were largely complete, we came across the work of Small et al. (2000), who considered a number of examples of estimating equation with multiple roots. They also liscussed several methods for choosing a root of an estimating equation when more than one root is present (see the discussion in Section 2.6 of Chapter 2). Further work could be done by looking at their methods in context of the linear regression model with MA(1) errors through simulation in order to see whether their methods improve test sizes and forecasting performance as in Chapters 3-5.

References

Aitchison, J. and S.D. Silvey (1958) Maximum likelihood of parameters subject to restraints, *Annals of Mathematical Statistics*, 29, 813-828.

Aitchison, J. and S.D. Silvey (1960) Maximum likelihood estimation procedures and associated tests of significance, *Journal of the Royal Statistical Society*, B 22, 154-171.

Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle, in B.N. Petrov and F. Csaki eds. *Second International Symposium on Information Theory*, Akademiai Kiado, Budapest, 267-281, Also in S. Kotz and N.L. Johnson eds. *Breakthroughs in Statistics Volume 1: Foundations and Basic Theory*, Springer-Verlag, New York, 601-609.

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

Amemiya, T. (1985) *Advanced Econometrics*, Harvard University Press, Cambridge.

Ansley, C.F. and P. Newbold (1980) Finite sample properties for autoregressive moving average process, *Journal of Econometrics*, 13, 159-83.

Ara, I. (1995) Marginal likelihood based tests of regression disturbances, unpublished Ph.D. thesis, Department of Econometrics and Business Statistics, Monash University, Australia.

Ara, I. and M.L. King (1993) Marginal likelihood based tests of regression disturbances, mimeo, Department of Econometrics, Monash University, Australia.

Ara, I. and M.L. King (1995) Marginal likelihood based tests of a subvector of the parameter vector of regression disturbances, in C.S. Forbes, P. Kofman and T.R.L. Fry eds. *Proceedings of the Econometrics Conference at Monash University*, Monash University, Melbourne, 69-106.

Atukorala, R. (1999) The use of an information criterion for assessing asymptotic approximations in econometrics, unpublished Ph.D. thesis, Department of Econometrics and Business Statistics, Monash University, Australia.

Azam, M.N. (2002) Modelling and forecasting in the presence of structural change in the linear regression model, unpublished Ph.D. thesis, Department of Econometrics and Business Statistics, Monash University, Australia.

Bahadur, R.R. (1958) Examples of inconsistency of maximum likelihood estimates, *Sankhya*, 20, 207-210.

Baltagi, B.H. and Q. Li (1995) Testing AR(1) against MA(1) disturbances in an error component model, *Journal of Econometrics*, 68, 133-151.

Barnett, V.D. (1966) Evaluation of the maximum likelihood estimator where the likelihood estimation has multiple roots, *Biometrika*, 53, 151-165.

Bartlett, M.S. (1953a) Approximate confidence intervals more than one unknown parameter, *Biometrika*, 40, 306-317.

Bartlett, M.S. (1953b) Approximate confidence intervals, *Biometrika*, 40, 12-19.

Basford, K.E. and G.J. McLachlan (1985) Likelihood estimation with normal mixture models, *Applied Statistics*, 34, 282-289.

Baykasoglu, A. and N.N.Z. Gindy (2001) A simulated annealing algorithm for dynamic layout problem, *Computer and Operations Research*, 28, 1403-1426.

Bellhouse, D.R. (1978) Marginal likelihood methods for distributed lag models, *Statistiche Hefte,* 19, 2-14.

Bellhouse, D.R. (1991) Marginal and approximate conditional likelihoods for sampling on successive occasions, *Survey Methodology,* 17, 69-78.

Berenblut, I.I. and G.I. Webb (1973) A new test for autocorrelated errors in the linear regression model, *Journal of the Royal Statistical Society,* B 35, 33-50.

Breusch, T.S. (1978) Testing for autocorrelation in dynamic linear models, *Australian Economic Papers*, 17, 334-355.

Burke, S.P., L.G. Godfrey and A.R. Tremayne (1990) Testing AR(1) against MA(1) disturbances in the linear regression model: An alternative procedure, *Review of Economic Studies,* 57, 135-145.

Carnevali, P., L. Coletti and S. Patarnello (1985) Image processing by simulated annealing, *IBM Journal of Research and Development,* 29, 569-579.

Chaubey, Y.P. and G. Gabor (1981) Another look at Fisher's solution to the problem of weighted mean, *Communications in Statistics,* A 10, 1225-1237.

Chesher, A. and G. Austin (1991) The finite sample distributions of heteroskedasticity robust Wald statistics, *Journal of Econometrics,* 47, 153-173.

Cohen, A.C. (1957) On the solution of estimating equations for truncated and censored samples from normal populations, *Biometrika,* 44, 225-236.

Cohn, H. and M. Fielding (1999) Simulated annealing: Searching for an optimal temperature schedule, *SIAM Journal on Optimization,* 9, 779-802.

Conkright, N.B. (1951) *Introduction to the Theory of Equations*, Ginn and Company, New York.

Cooper, D.M. and R. Thompson (1977) A note on the estimation of the parameters of the autoregressive moving average process, *Biometrika,* 64, 625-628.

Corana, A., M. Marchesi, C. Martini and S. Ridella (1987) Minimizing multinomial functions of continuous variables with the simulated annealing algorithm, *ACM Transactions of Mathematical Software*, 13, 262-280.

Corduas, M. (1986) The use of the marginal likelihood in testing for serial correlation in time series regression, Unpublished M. Phil. thesis, University of Lancaster.

Cox, D.R. (1988) Some aspects of conditional and asymptotic inference: A review, *Sankhya,* 50, 314-337.

Cox, D.R. and D.V. Hinkley (1974) *Theoretical Statistics*, Chapman and Hall, London.

Cramer, H. (1946) *Mathematical Methods of Statistics,* Princeton University Press, Princeton, New Jersey.

Cramer, J.S. (1986) *Econometric Applications of Maximum Likelihood Methods*, Cambridge University Press, Cambridge.

Crowder, M. (1986) On consistency and inconsistency of estimating equations, *Econometric Theory,* 2, 305-330.

Cryer, J.D. and J. Ledolter (1981) Small sample properties of the maximum likelihood estimator in the first-order moving average model, *Biometrika,* 68, 691-694.

Davis, R.A. and W.T.M. Dunsmuir (1996) Maximum likelihood estimation for MA(1) processes with a root on or near the unit circle, *Econometric Theory,* 12, 1-29.

Davis, R.A., M.J. Chen and W.T.M. Dunsmuir (1995) Inference for MA(1) processes with a root or near the unit circle, *Probability and Mathematical Statistics*, 15, 227-242.

Derwent, D. (1988) A better way to control population, *Nature*, 331, 575-578.

Dunsmuir, W.T.M. (1981) Estimation for stationary time series when data are irregularly spaced or missing, in D.F. Findley ed. *Applied Time Series II*, Academic Press, New York, 609-649.

Durbin, J. (1960) Estimation of parameters in time-series regression models, *Journal of the Royal Statistical Society*, B, 22, 139-153.

Fang, H., G. Gong and M. Qian (1997) Annealing of iterative stochastic schemes, *SIAM Journal on Control and Optimization*, 35, 1886-1907.

Ferguson, T.S. (1982) An inconsistent maximum likelihood estimate, *Journal of the American Statistical Association*, 77, 831-834.

Fielding, M. (2000) Simulated annealing with an optimal fixed temperature, *SIAM Journal on Optimization*, 11, 289-307.

Finch, S.J., N.R. Mendell and H.C. Thode (1989) Probability measures of adequacy of a numerical search for a global maximum, *Journal of the American Statistical Association*, 84, 1020-1023.

Fisher, R.A. (1922) On the Mathematical Foundations of Theoretical Statistics, reprinted in *Contributions to Mathematical Statistics (by R.A. Fisher)*, John Wiley and Sons, New York.

Frances, P.H. (1992) A model selection test for an AR(1) versus an MA(1) model, *Statistics and Probability Letters*, 15, 281-284.

Fraser, D.A.S. (1967) Data transformations and the linear model, *Annals of Mathematical Statistics*, 38, 1456-1465.

GAUSS, (1996) *GAUSS Constrained Optimization: Application Module,* Aptech Systems, Inc., Washington.

Gauss, C.F. (1806) II Comet vom Jahr 1805, *Monatliche Correspondenz*, 14,181-186.

Gauss, C.F. (1809) Theoria motus corporum coelestum, *Wereke* 7, (translated into English by C.H. Davis (1963)), Dover, New York.

Gauss, C.F. (1823) Combinationes erroribus minimis obnoxiae, Parts 1, 2 and Supplement, 4, 1-108.

Geman, S. and D. Geman (1984) Stochastic relaxation, Gibbs distribution and the Bayesian restoration in images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721-741.

Ghosh, M. (1991) Estimating functions in survey sampling: A review, in V.P. Godambe, ed. *Estimating Functions*, Clarendon Press, New York, 201-210.

Godambe, V.P and M.E. Thompson (1986) Parameters of superpopulation and survey population, their relationship and estimation, *International Statistical Review*, 54, 127-138.

Godambe, V.P. (1960) An optimum property of regular maximum likelihood estimation, *Annals of Mathematical Statistics*, 31, 1208-1211.

Godambe, V.P. and B.K. Kale (1991) Estimating functions: An overview, in V.P. Godambe, ed. *Estimating Functions*, Clarendon Press, New York, 1-19.

Godambe, V.P. and M.E. Thompson (1974) Estimating equations in the presence of a nuisance parameter, *Annals of Statistics,* 2, 568-571.

Godambe, V.P. and M.E. Thompson (1989) An extension of quasilikelihood estimation (with discussion), *Journal of Statistical Planning and Inference*, 22, 137-172.

Godambe, V.P. and S. Kunte (1993) Optimal estimation under biased allocation of treatments, *Biometrika*, 80, 797-806.

Godfrey, L.G. (1978) Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables, *Econometrica*, 46, 1293-1301.

Godfrey, L.G. (1988) *Misspecification Tests in Econometrics: The Lagrange Multiplier Principle and Other Approaches*, Cambridge University Press, Cambridge.

Godfrey, L.G. and A.R. Tremayne (1988) Checks of model adequacy for univariate time series models and their application to economic relationships, *Econometric Reviews*, 7, 1-42.

Godfrey, L.G. and A.R. Tremayne (1992) Misspecification tests for univariate time series models and their applications in econometrics, in L.G. Godfrey, ed. *The Implementation and Constructive Use of Misspecification Tests in Econometrics*, Manchester University Press, Manchester and New York, 1-66.

Goffe, W.L., M. Heath and J. Rogers (1994) Global optimization of statistical functions with simulated annealing, *Journal of Econometrics*, 60, 65-99.

Gourieroux, C. and A. Monfort (1995) *Statistics and Econometric Models*, Cambridge University Press, Cambridge.

Griffiths, W.E. and K. Surekha (1986) A Monte Carlo evaluation of the power of some tests for heteroskedasticity, *Journal of Econometrics*, 31, 219-273.

Grose, S.D. (1992) Marginal likelihood estimation in the dynamic linear regression model, paper presented at the 1992 Australian Meeting of the Econometric Society, Monash University, Melbourne, Australia.

Hajek, B. (1988) Cooling schedules for optimal annealing, *Mathematics of Operational Research*, 13, 311-329.

Hall, A.D. and M. McAleer (1989) A Monte Carlo study of some tests of model adequacy of time series analysis, *Journal of Business and Economic Statistics*, 7, 95-106.

Hendry, D.F. and F. Srba (1977) The properties of autoregressive instrumental variable estimators in dynamic systems, *Econometrica*, 45, 969-989.

Hendry, D.F. and P.K. Trivedi (1972) Maximum likelihood estimation of difference equations with moving average errors: A simulation study, *The Review of Economic Studies*, 39, 117-145.

Heyde, C.C. and R. Morton (1998) Multiple roots in general estimating equations, *Biometrika*, 85, 954-959.

Hildreth, C. and J.Y. Lu (1960) Demand relations with autocorrelated disturbances, Technical Bulletin 276, Department of Agricultural Economics, Michigan State University.

Hoeschele, I. (1988) A note on local maxima in maximum likelihood, restricted maximum likelihood, and Bayesian estimation of variance components, *Journal of Statistical Computation and Simulation*, 33, 149-160.

Holly, A. (1982) A remark on Hausman's specification test, *Econometrica*, 50, 749-759.

Honda, Y. (1988) A size correction to the Lagrange multiplier test for heteroscedasticity, *Journal of Econometrics*, 38, 375-386.

Huzurbazar, V.S. (1948) The likelihood equation, consistency and the maxima of the likelihood function, *Annals of Eugenics*, 14, 185-200.

Ingber, L. (1989) Very fast simulated re-annealing, *Mathematical and Computer Modelling*, 12, 967-973.

Jeng, F.C. and J.W. Woods (1990) Simulated annealing in compound Gaussian random fields, *IEEE Transactions on Information Theory*, 36, 94-107.

Johnson, D., C. Aragon, L. Mcgedch and C. Schevon (1990) Optimization by simulated annealing, an experimental evaluation, Part I, Graph partitioning, *Operations Research*, 37, 865-892.

Johnson, D., C. Aragon, L. Mcgedch and C. Schevon (1991) Optimization by simulated annealing, an experimental evaluation, Part II, Graph coloring and number partitioning, *Operations Research*, 39, 378-406.

Johnson, D., C. Aragon, L. Mcgedch, and C. Schevon (1992) Optimization by simulated annealing, an experimental evaluation, Part III, The travelling salesman problem, Unpublished manuscript.

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lutkepohl and T.C. Lee (1985) *The Theory and Practice of Econometrics*, 2nd ed., John Wiley, New York.

Kalbfleisch, J.D. and D.A. Sprott (1970) Applications of likelihood methods to models involving large numbers of parameters (with discussion), *Journal of the Royal Statistical Society*, B 32, 175-208.

Kalbfleisch, J.D. and R.L. Prentice (1973) Marginal likelihood methods based on Cox's regression and life model, *Biometrika*, 60, 267-278.

Kang, K.M. (1975) A comparison of estimators for moving average processes, unpublished working paper, Australian Bureau of Statistics.

Kennedy, P. and D. Simons (1991) Fighting the teflon: Comparing classical and Bayesian estimators for autocorrelated errors, *Journal of Econometrics*, 48, 15-27.

Kenward, L.R. (1975) Autocorrelation and dynamic methodology with an application to wage determination models, *Journal of Econometrics*, 3, 179-188.

King, M.L. (1983a) Testing for autoregressive against moving average errors in the linear regression model, *Journal of Econometrics*, 21, 35-51.

King, M.L. (1983b) Testing for moving average disturbances, *Australian Journal of Statistics*, 25, 23-34.

King, M.L. (1987a) An alternative test for regression coefficient stability, *Review of Economics and Statistics*, 69, 379-381.

King, M.L. (1987b) Testing for autocorrelation in the regression models: A survey, in M.L King and D.E.A. Giles, eds. *Specification Analysis in the Linear Model*, Routledge and Kegan Paul, London, 19-74.

King, M.L. (1996) Hypothesis testing in the presence of nuisance parameters, *Journal of Statistical Planning and Inference*, 50, 103-120.

King, M.L. and M.B. Billah (2000) Using simulated annealing to estimate penalty function for time series model selection, *Journal of Statistical Research*, 34, 75-89.

King, M.L. and M.R. Laskar (1998) Estimation and testing of regression disturbances based on modified likelihood functions, *Journal of Statistical Planning and Inference,* 71, 75-92.

King, M.L. and M. McAleer (1987) Further results on testing AR(1) against MA(1) disturbances in the linear regression model, *Review of Economic Studies,* 54, 649-663.

King, M.L. and P.X. Wu (1991) Small-disturbance asymptotics and the Durbin-Watson and related tests in the dynamic regression model, *Journal of Econometrics*, 47, 145-152.

Kirkpatrick, S.C.D., M.P. Gelatt and M.P. Vecchi (1982) Optimization by simulated annealing, *IBM Research Report*, RC 9355.

Kirkpatrick, S.C.D., M.P. Gelatt and M.P. Vecchi (1983) Optimization by simulated annealing, *Science,* 220, 671-680.

Kraft, C.H. and L.M. LeCam (1956) A remark on the roots of the maximum likelihood equation, *Annals of Mathematical Statistics,* 27, 1174-1177.

Laskar, M.R. (1999) Estimation and testing of linear regression disturbances based on modified likelihood and message length functions, unpublished Ph.D. thesis, Department of Econometrics and Business Statistics, Monash University, Australia.

Laskar, M.R. and M.L. King (2001) Modified likelihood and related methods for handling nuisance parameters in the linear regression model, in A.K.M.E. Saleh ed. *Data Analysis from Statistical Foundations*, Nova Science Publisher, Inc. Huntington, New York, 119-142.

Laskar, M.R. and M.L. King (1998) Estimation and testing of regression disturbances based on modified likelihood functions, *Journal of Statistical Planning and Inference,* 71, 75-92.

Laskar, M.R. and M.L. King (1997a) Modified Wald test for regression disturbances, *Economic Letters* 56, 5-11.

Laskar, M.R. and M.L. King (1997b) Testing regression disturbances based on minimum message length, in P. Bardsley and V.L. Martin eds. *Proceedings of the Econometric Society Australian Meeting,* University of Melbourne, 2, 179-202.

Latif, A. and M.L. King (1993) Linear regression forecasting in the presence of AR(1) disturbances, *Journal of Forecasting,* 12, 513-524.

LeCam, L. (1979) Maximum likelihood: An introduction, *Lecture Notes in Statistics,* No. 18, University of Maryland, College Park, Maryland.

LeCam, L. (1990) Maximum likelihood: An introduction, *International Statistical Review,* 58, 153-171.

Lehmann, E.L. (1983) *Theory of Point Estimation.* Wiley, New York.

Levanbach, H. (1972) Estimation of autoregressive parameters from a marginal likelihood function, *Biometrika, 59,* 61-71.

Levanbach, H. (1973) The estimation of heteroscedasticity from a marginal likelihood function, *Journal of the American Statistical Association,* 68, 436-439.

Liang, K.Y. and S.L. Zeger (1986) Longitudinal data analysis using generalized linear models, *Biometrika,* 73, 13-22.

Mahmood, M. (2000) Exploiting estimating equations to improve estimation procedures in the linear model, unpublished Ph.D. thesis, Department of Econometrics and Business Statistics, Monash University, Australia.

Mak, T.K. (1993) Solving non-linear estimation equations, *Journal of the Royal Statistical Society,* B 55, 945-955.

Mardia, K.V. and A.J. Watkins (1989) On multimodality of the likelihood in the spatial linear model, *Biometrika,* 76, 289-295.

Markatou, M., A. Basu and B.G. Lindsay (1998) Weighted likelihood equations with strap root search, *Journal of the American Statistical Association,* 93, 740-750.

McAleer, M., C.R. McKenzie and A.D. Hall (1988) Testing separate time series models, *Journal of Time Series Analysis,* 9, 169-189.

McAleer, M., M.H. Pesaran and A.K. Bera (1990) Alternative approaches to testing non-nested models with autocorrelated disturbances: An application to models of US unemployment, *Communications in Statistics: Theory and Methods,* 19, 3619-3644.

McCullagh, P. and J.A. Nelder (1989) *Generalised Linear Models,* 2nd edition, Chapman & Hall, London.

McKenzie, C.R., M. McAleer and L. Gill (1999) Simple procedures for testing autoregressive versus moving average errors in the linear regression models, *Japanese Economic Review,* 50, 239-252.

Merriman, M. (1877) A list of writings relating to the method of least squares, with historical and critical notes, *Transaction of the Connecticut Academy of Arts and Sciences*, 4, 151-232.

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, 1087-1092.

Moral, P.D. and L. Miclo (1999) On the convergence and applications of generalized simulated annealing, *SIAM Journal on Control and Optimization*, 37, 1222-1250.

Moulton, B.R. and W.C. Randolph (1989) Alternative tests of the error components model, *Econometrica*, 57, 685-693.

Nankervis, J.C. and N.E. Savin (1987) Finite sample distributions of t and F statistics in an AR(1) model with an exogenous variable, *Econometric Theory*, 3, 387-408.

Nelder, J.A. and R. Mead (1965) A simplex method for function minimization, *Computer Journal*, 7, 308-313.

Neyman, J. (1935) On a theorem concerning the concept of sufficient statistic, (in Italian) *Giornale dell'Instituto Italiano degli Atturai*, 6, 320-334.

Neyman, J. and E.L Scott (1948) Consistent estimates based on partially consistent observations, *Econometrica*, 16, 1-32.

Nicholls, D.F., A.R. Pagan and R.D. Terrell (1975) The estimation and the use of models with moving average disturbance terms: a survey, *International Economic Review*, 16, 113-134.

Norden, R.H. (1972) A survey of maximum likelihood estimation, *International Statistical Review*, 40, 329-354.

Norden, R.H. (1973) A survey of maximum likelihood estimation, Part 2, *International Statistical Review*, 41, 39-58.

Olivier, C. (1998) Solving scheduling problems by simulated annealing, *SIAM Journal on Control and Optimization*, 36, 1539-1575.

Oxley, L.T. and C.J. Roberts (1982) Pitfalls in the application of the Cochrane-Orcutt technique, *Oxford Bulletin of Economics and Statistics*, 44, 227-240.

Oxley, L.T. and C.J. Roberts (1986) Multiple minima and the Cochrane-Orcutt technique: Some initial Monte Carlo results, *Economics Letters*, 20, 247-250.

Pesaran, M.H. (1983) A note on the maximum likelihood estimation of regression models with first order moving average errors with roots on the unit circle, *Australian Journal of Statistics*, 25, 442-448.

Press, W.H., B.P. Flannery, S.A. Teukolsky and W.T. Vetterling (1986) *Numerical Recipes, the Art of Scientific Computing*, Cambridge University Press, Cambridge.

Pronzato, L., E. Walter, A. Venot and J.F. Lebruchec (1984) A general-purpose optimizer, implementation and applications, *Mathematics and Computers in Simulation*, 24, 412-422.

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*, University of Melbourne, Melbourne, 2, 613-640.

Rahman, S. and M.L. King (1994) A comparison of marginal likelihood based and approximate point optimal tests for random regression coefficients in the presence of autocorrelation, *Pakistan Journal of Statistics*, 10, 375-394.

Rahman, S. and M.L. King (1998) Marginal likelihood score based tests of regression disturbances in the presence of nuisance parameters, *Journal of Econometrics,* 82, 81-106.

Rao, C.R. (1947) Factorial experiments derivable from combinatorial arrangements of arrays, Supplement to the *Journal of the Royal Statistical Society,* 9 128-139.

Rao, C.R. (1948) Large sample tests of statistical hypotheses concerning several parameters with application to problems of estimation, *Proceedings of the Cambridge Philosophical Society,* 44, 50-57.

Rao, C.R. (1973) *Linear Statistical Inference and Its Applications,* 2nd ed., John Wiley & Sons, New York.

Rosenbrock, H.H. (1960) An automatic method for finding the greatest or least value of a function, *Computer Journal,* 3, 175-184.

Rothenberg, T.J. (1971) Identification in parametric models, *Econometrica,* 39, 577-591.

Rowley, J.C.R and D.A. Wilton (1973) Quarterly models of wage determination: Some new efficient estimates, *American Economic Review,* 63, 380-389.

Saikkonen, P. and R. Luukkonen (1993) Testing for a moving average unit root in autoregressive integrated moving average models, *Journal of the American Statistical Association,* 88, 596-601.

Sargan, J.D. (1964) Wages and prices in the United Kingdom, a study in econometric methodology, in P.T. Hart, G. Mills and J.K. Whitaker, eds. *Econometric Analysis for National Economic Planning,* Butterworths, London.

Sargan, J.D. and A. Bhargava (1983) Maximum likelihood estimation of regression models with first order moving average errors when the root lies on the unit circle, *Econometrica*, 51, 799-820.

Sartori, N., A. Salvan and L. Pace (2003) A note on directed adjusted profile likelihoods, *Journal of Statistical Planning and Inference*, 110, 1-9.

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

Shephard, N. (1993) Maximum likelihood estimation of regression models with stochastic trend components, *Journal of the American Statistical Association*, 88, 590-595.

Silvapulle, P. and M.L. King (1991) Testing moving average against autoregressive disturbances in the linear regression model, *Journal of Business and Economic Statistics*, 9, 329-335.

Silvapulle, P. and M.L. King (1993) Non-nested testing for autocorrelation in the linear regression model, *Journal of Econometrics*, 58, 295-314.

Silvey, S.D. (1959) The Lagrangian multiplier test, *Annals of Mathematical Statistics*, 30, 389-407.

Small, C.G. and D.L. McLeish (1994) *Hilbert Space Methods in Probability and Statistical Inference*, Wiley, New York.

Small, C.G. and Z. Yang (1999) Multiple roots of estimating functions, *Canadian Journal of Statistics*, 27, 585-598.

Small, C.G., J. Wang and Z. Yang (2000) Eliminating multiple root problems in estimation, *Statistical Science*, 15, 313-341.

Smith, M.D. and A.R. Tremayne (1990) Finite sample theory for testing AR(1) against MA(1) time series model, a paper presented at the World Congress of the Econometric Society, Spain.

Stefanski, L.A. and R.J. Carroll (1987) Conditional scores and optimal scores for generalized linear measurement error models, *Biometrika*, 74, 703-716.

Szu, H. and R. Hartley (1987) Fast simulated annealing, *Physics Letters* A, 157-162.

Tanaka, K. (1990) Testing for a moving average unit root, *Econometric Theory*, 6, 433-444.

Theil, H. (1971) *Principles of Econometrics*, John Wiley and Sons, New York.

Trivedi, P.K. (1970) Inventory behavior in U.K. manufacturing, 1956-67, *Review of Economic Studies*, 37, 517-527.

Trouve, A. (1996) Cycle decomposition and simulated annealing, *SIAM Journal on Control and Optimization*, 34, 966-986.

Tucci, M.P. (2002) A note on global optimization in adaptive control, econometrics and macroeconomics, *Journal of Economic Dynamics and Control*, 26, 1739-1764.

Tunnicliffe Wilson, G. (1989) On the use of marginal likelihood in time series model estimation, *Journal of the Royal Statistical Society*, B 51, 15-27.

Tzavelas, G. (1998) A note on the uniqueness of the quasi-likelihood estimator, *Statistical Probability Letters*, 38, 125-130.

Wald, A. (1943) Tests of statistical hypotheses concerning several parameters when the number of observations is large, *Transactions of the American Mathematical Society*, 54, 426-482.

Wald, A. (1947) *Sequential Analysis*, John Wiley and Sons, New York.

Wald, A. (1949) Note on the consistency of the maximum likelihood estimate, *Annals of Mathematical Statistics*, 60, 595-601.

Walker, A.M. (1967) Some tests of separate families of hypothesis in time series analysis, *Biometrika*, 54, 39-68.

Wasserman, P.D. and T. Schwartz (1988) Neural networks, part 2, what are they and why is everybody so interested in them now? *IEEE Expert*, 10-15.

White, H. (1994) Maximum likelihood estimation of misspecified models, *Econometrica*, 50, 1-25.

Wong, D.F., H.W. Leong and C.L. Liu (1988) *Simulated Annealing for VLSI Design*, Kluwer Academic Publishers, Boston.

Yeasmin, M. and M.L. King (2003) Test sizes and the importance of finding global maxima of the likelihood function, accepted for publication, *Journal of Statistical Computation and Simulation*, forthcoming.

Zellner, A. and C. Montmarquette (1971) A study of some aspect of temporal aggregation problems in econometric analysis, *Review of Economics and Statistics*, 53, 335-342.

Zidek, J.V., N.D. Le, H. Wong and R.T. Burnett (1998) Including structural measurement errors in the nonlinear regression analysis of clustered data, *The Canadian Journal of Statistics*, 26, 537-548.