

2475/4071

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

ON.....12 August 2003

Sec. Research Graduate School Committee

Under the Copyright Act 1968, this thesis must be used only under the normal conditions of scholarly fair dealing for the purposes of research, criticism or review. In particular no results or conclusions should be extracted from it, nor should it be copied or closely paraphrased in whole or in part without the written consent of the author. Proper written acknowledgement should be made for any assistance obtained from this thesis.

ERRATA

1. p.14, line 22. Insert after " $y = X\beta + u$," "where y is $n \times 1$, X is an $n \times k$ non-stochastic matrix of rank $k < n$, β is a $k \times 1$ vector of parameters and u is $n \times 1$ disturbance vector".
2. p.15, line 1. Replace "(this is discussed in detail in Section 3.3)" with "where η_0 is a positive scalar, η is a $k \times 1$ vector, z is the ordinary least squares residual vector and P is any $p \times n$ matrix such that $P'P = I_n - X(X'X)^{-1}X'$ and $PP' = I_p$ in which $p = n - k$ ".

3. p.16, line 5. Insert the following paragraph after line 5.
 "Marginal likelihood is one of the likelihood based methods designed to overcome the nuisance parameter problem. It involves transforming the observed data such that the likelihood of the transformed data can be factored into two parts, one being uninformative about the parameter(s) of interest in the absence of knowledge of the value of the nuisance parameter(s) and the other part containing the parameter(s) of interest as the only unknown parameter(s). This second factor is the marginal likelihood as introduced by Fraser (1967) and further developed by Kalbfleisch and Sprott (1970). In the context in which it is used in this thesis, Ara and King (1993) and Ara (1995) have shown it is equivalent to the likelihood function of the standard maximal invariant under transformations of the form $T(y)$. In this form, it therefore has all the properties of a full likelihood. Given how the marginal likelihood is constructed, there is no loss of information in using it in place of the full likelihood."

4. p.33. Replace the last paragraph of Section 2.5 with the following.
 "Andrews and Ploberger (1994) derived asymptotically optimal tests for testing problems when a nuisance parameter is present under the alternative hypothesis but not under the null hypothesis. This is a nonstandard testing problem for which likelihood based asymptotic tests such as the LR, LM and Wald tests are difficult to apply because of problems in determining their distributions. Also, as Andrews and Ploberger point out, the asymptotic optimality properties of these tests no longer apply. They used a weighted average power criterion to generate asymptotically optimal tests similar to such a criterion used by Wald (1943). The weighting function is defined over the space of the nuisance parameter that is present under the alternative hypothesis but not under the null. They constructed general tests that are asymptotically optimal in the sense that they maximize weighted power asymptotically under a sequence of local alternatives. Their new tests are of an average exponential form and are called the exponential Wald, exponential LM and exponential LR tests.

In special cases where through invariance and/or similarity, the testing problem essentially boils down to one in which the null hypothesis is simple and the alternative hypothesis is composite, finite-sample optimal tests can be constructed through the use of the Neyman-Pearson (NP) lemma. This involves applying the lemma to a weighted average of likelihoods under the alternative hypothesis and results in invariant or similar tests that maximize weighted power. Andrews et al. (1996) outlined this approach in the case of

testing for one or more changepoints at unknown times in a multiple linear regression model. Here the timing of the changepoint is an unknown parameter which is present only under the alternative hypothesis. In a further paper, Andrews and Ploberger (1996) applied this approach to testing for white noise against an autoregressive moving average (ARMA(1,1)) model.

In this thesis, we propose a new test procedure for the problem of testing a composite null against a composite alternative hypothesis by applying the generalized NP lemma. This allows us to extend the work of Andrews and Ploberger to the construction of tests that are optimal in small samples in that they maximize weighted power subject to controlling the level of average size over each of a number of subspaces of the null hypothesis parameter space. The test is general and does not require there to be a nuisance parameter that appears only under the alternative hypothesis."

5. p.42, line 10. Insert the following lines after "composite hypotheses."
 "Andrews et al. (1996) and Andrews and Ploberger (1996) constructed tests that maximize weighted power for testing problems in which the null hypothesis is simple after suitable reduction via invariance or similarity. Their results are based on the NP lemma. Our work involves null hypotheses which are still complex after suitable reduction and is based on the generalized NP lemma."
6. p.46, line 17. Replace "prior distribution" with "weighting distribution".
7. p.47, line 9. Replace "prior distribution" with "weighting distribution".
8. p.48. Insert the following paragraph after the equation (3.6).
 "Clearly, the c_i 's are one's preferred level of significance and chosen by the user. An obvious approach is to make them equal $c_1 = c_2 = \dots = c_m$ and then equal to the desired significance level. We believe this is the best approach. However, the generalized NP lemma does allow flexibility in this area when testing composite hypotheses."
8. p.50. Insert the following paragraph before Section 3.3.
 "It may be tempting to use the data to help decide on the weighting function given the data has information regarding the true values of the parameters. This is not recommended because it is likely to result in a loss of power. This is because the generalized NP lemma would no longer apply. Also it is similar to a jump from a point optimal test to a LR test. In the former, the point at which power is to be optimized is chosen in advance with good results. In the latter, parameters under the alternative are replaced with estimates. As pointed out in a survey article on testing for autocorrelation in linear models by King (1987b, p.59), 'It would seem that the LR test is a particularly unreliable test for serial correlation. Dent (1973) found it to have poor power against AR(1) disturbances, while Brandsma and Ketellapper (1979) reported a similar finding for the LR test against first-order spatial autoregression. Maddala and Rao (1973) were so puzzled by the poor power of the LR test against AR(1) disturbances in the lagged dependent variable model, they repeated their calculations using a different algorithm for computing the MLE's'."
9. p.72, line 14. Insert after "control average size." "The use of the uniform distribution means that average sizes, and average powers are unweighted averages."

10. p.74, line 9. Replace "The sizes can be undesirably large in some parts of the null hypothesis parameter space while being smaller than the desired size in other parts of the null hypothesis parameter space" with "The sizes may be undesirably large in some parts of the null hypothesis parameter sub-spaces (Θ_i) while being smaller than the desired size in other parts of the null hypothesis parameter sub-spaces".
11. p.145, line 9. Change " z_{1i} and w_{1i} " to " z_{2i} and w_{2i} ".
12. p.201, line 17. Insert the following paragraph after line 17,
"In this thesis, we have paid very little attention to computational efficiency. We have constructed computer programs to do the task but without worrying about how long the task takes. Clearly, future research is needed on reducing computation time by the use of more efficient algorithms and programming. We conjecture that a greater attention to this issue might reduce computational time by up to 90%."

Additional references

- Brandsma, A.S. and Ketellapper, R.H. (1979), Further evidence on alternative procedures for testing of spatial autocorrelation among regression disturbances, in: C.P.A. Bartels and R.H. Ketellapper, eds., *Exploratory and Explanatory Statistical Analysis of Spatial Data*, Boston: Martinus Nijhoff, 113-136.
- Dent, W.T. (1973), A power study of several tests for autocorrelation, *New Zealand Economic Papers*, 7, 109-120.
- Maddala, G.S. and Rao, A.S. (1973), Tests for serial correlation in regression models with lagged dependent variables and serially correlated errors, *Econometrica*, 41, 761-774.

A New Class of Hypothesis Tests which Maximize Average Power

**A thesis submitted for the degree of
Doctor of Philosophy**

by

Nelufa Begum

B.Sc. (Hons.), M.Sc. (Statistics), Jahangirnagar

Department of Econometrics and Business Statistics

Monash University

Australia

March 2003

Contents

<i>List of Tables</i>	iv
<i>Declaration</i>	viii
<i>Acknowledgments</i>	ix
<i>Abstract</i>	xi

Chapter 1 Introduction

1.1 Background	1
1.2 Motivation and direction of the thesis	3
1.3 Outline of the thesis	6
1.4 Computations	8

Chapter 2 Literature Survey

2.1 Introduction	10
2.2 Nuisance parameters, invariance arguments and the marginal likelihood	13
2.3 Optimal tests	16
2.4 Point optimal tests	24
2.5 Approximate point optimal tests	29
2.6 Brief survey of non-nested testing	34
2.7 Concluding remarks	41

Chapter 3 A New Approach to Testing a Composite Null against a Composite Alternative

3.1 Introduction	43
3.2 Theory and the testing procedure	45

3.3	Application to testing disturbances in the linear regression model	50
3.4	Concluding remarks	65
Chapter 4 Application of the New Approach to a Single Parameter Problem		
4.1	Introduction	67
4.2	The model, testing problem and the test procedure	69
4.3	Monte Carlo experiments	75
4.3.1	Experimental design	75
4.3.2	Size properties of the test with $m = 1$	77
4.3.3	Size properties of the test with $m = 2$	77
4.3.4	Size properties of the test with $m = 3$	81
4.4	Power comparison of the test with $m = 3$	85
4.5	Concluding remarks	87
Chapter 5 An Investigation of the Choice of Boundary Points under the Null Hypothesis		
5.1	Introduction	97
5.2	Methodology	100
5.3	Monte Carlo experiments	101
5.3.1	Experimental design	102
5.3.2	Size results	104
5.3.3	Power results	107
5.4	Concluding remarks	109
Chapter 6 Testing for Heteroscedastic Disturbances in the Linear Regression Model		
6.1	Introduction	136
6.2	The model, testing problem and the test procedure	137
6.3	Monte Carlo experiment	141
6.3.1	Design of experiment	142
6.3.2	Size and power properties of the test with $m = 1$	143
6.3.3	Size and power properties of the test with $m = 2$	145
6.4	Test sizes controlled at $\lambda = 0$ with $m = 1$	149

6.4.1	Properties of the test when sizes controlled at $\lambda = 0$ with $m = 1$	150
6.4.2	Results	150
6.5	Concluding remarks	152
Chapter 7 Testing Joint MA(1)-MA(4) against Joint AR(1)-AR(4) Disturbances in the Linear Regression Model		
7.1	Introduction	161
7.2	The testing problem and the form of the test	164
7.3	Outline of the simulation experiments	173
7.4	Results for $m = 1$	175
7.5	Results for $m = 2$	176
7.6	Results for $m = 2$ with special weights	179
7.7	Results for $m = 3$	181
7.8	Conclusions	183
Chapter 8 Summary and Conclusions		
		194
References		202

List of Tables

Table 4.1	Critical values (c_α) for $X1$, $X2$, $X3$, $X4$ and $X5$ design matrices when $m = 1$ at the five percent level	89
Table 4.2	Calculated sizes of the test when $m = 1$ at the five percent level	89
Table 4.3	Ranges of the two disjoint intervals and the corresponding values of j , where $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$	90
Table 4.4	Final critical values k_1 and k_2 when $m = 2$	91
Table 4.5	Calculated sizes of the test when $m = 2$ at the five percent level	91
Table 4.6	Ranges of the three disjoint intervals and the corresponding values of j , where $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$	92
Table 4.7	Final critical values k_1 , k_2 and k_3 when $m = 3$	93
Table 4.8	Calculated sizes of the test when $m = 3$ at the five percent level	93
Table 4.9	Calculated sizes and powers of the new test and APOI tests for $X1$, $X2$, $X3$, $X4$ and $X5$, when $m = 3$ and $H_0: u_t = e_t + \gamma e_{t-1}$, $0 \leq \gamma \leq 1$, is tested against $H_a: u_t = \rho u_{t-1} + e_t$, $0 < \rho < 1$, at the five percent level	94
Table 5.1	Ranges of three disjoint intervals and the corresponding values of j , where $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$: $X2$ design matrix with $n = 20$	111
Table 5.2	Final critical values k_1 , k_2 and k_3 when $m = 3$: $X2$ design matrix with $n = 20$	120

Table 5.3	Calculated sizes of the test when $m = 3$ at the five percent level: $X2$ design matrix with $n = 20$	122
Table 5.4	Calculated powers of the test when $m = 3$ at the five percent level: $X2$ design matrix with $n = 20$	128
Table 5.5	Average deviations (AD), sum of squared deviations (SSD) of the power and the corresponding rank of accepted test with respect to size condition	134
Table 6.1	Critical values (c_α) of the test for two different sets of non-stochastic regressor values for all design matrices when $m = 1$, at the five percent level	153
Table 6.2	Calculated sizes of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 1$ at the five percent level	153
Table 6.3	Calculated powers of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 1$ at the five percent level	154
Table 6.4	Calculated sizes of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 1$ at the five percent level	154
Table 6.5	Calculated powers of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 1$ at the five percent level	155
Table 6.6	Ranges of the two disjoint intervals and the corresponding values of j , where $\lambda_j = \tau_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$, and $z_{it} = t/n$, $w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$	155
Table 6.7	Final critical values k_1 and k_2 when $z_{it} = t/n$ and $w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, for $m = 2$ at the five percent level	156
Table 6.8	Calculated sizes of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level	156

Table 6.9	Calculated powers of the test when $H_0: z_{1t} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{1t} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level	157
Table 6.10	Ranges of the two disjoint intervals and the corresponding values of j , where $\lambda_j = \tau_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$, and $z_{2t} = t/n$, $w_{2t} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$	157
Table 6.11	Final critical values k_1 and k_2 when $z_{2t} = t/n$ and $w_{2t} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, for $m = 2$ at the five percent level	158
Table 6.12	Calculated sizes of the test when $H_0: z_{2t} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{2t} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level	158
Table 6.13	Calculated powers of the test when $H_0: z_{2t} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{2t} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level	159
Table 6.14	Critical values (c_α) of the test for all design matrices for z_{1t} and w_{1t} when sizes are controlled at $\lambda = 0$, for $m = 1$ at the five percent level	159
Table 6.15	Calculated sizes of the test when $H_0: z_{1t} = t/n$, (sizes are controlled at $\lambda = 0$) is tested against $H_a: w_{1t} = (n+1-t)/n$, $t = 1, 2, \dots, n$, for $m = 1$ at the five percent level	160
Table 6.16	Calculated powers of the test when $H_0: z_{1t} = t/n$, (sizes are controlled at $\lambda = 0$) is tested against $H_a: w_{1t} = (n+1-t)/n$, $t = 1, 2, \dots, n$, for $m = 1$ at the five percent level	160
Table 7.1	Critical values (c_α) of the test with $m = 1$, for testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances for all design matrices at the five percent level	185
Table 7.2a	Selected calculated sizes and powers of the test for the X2 design matrix when $m = 1$, at the five percent level	185

Table 7.2b	Selected calculated sizes and powers of the test for the X3 design matrix when $m = 1$, at the five percent level	186
Table 7.2c	Selected calculated sizes and powers of the test for the X5 design matrix when $m = 1$, at the five percent level	186
Table 7.2d	Selected calculated sizes and powers of the test for the X6 design matrix when $m = 1$, at the five percent level	187
Table 7.3	Ranges of the two disjoint intervals and the corresponding values of j and l , where $\gamma_{1j} = (2j-1)/2q$, $j = 1, 2, \dots, q$ and $\gamma_{4l} = (2l-1)/2q$, $l = 1, 2, \dots, q$, $q = 50$	188
Table 7.4	Critical values k_1 and k_2 when $m = 2$, for testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances at the five percent level	190
Table 7.5a	Selected calculated sizes and powers of the test for the X2 design matrix when $m = 2$, at the five percent level	190
Table 7.5b	Selected calculated sizes and powers of the test for the X3 design matrix when $m = 2$, at the five percent level	191
Table 7.5c	Selected calculated sizes and powers of the test for the X5 design matrix when $m = 2$, at the five percent level	191
Table 7.5d	Selected calculated sizes and powers of the test for the X6 design matrix when $m = 2$, at the five percent level	192
Table 7.6	Selected calculated sizes and powers of the test for the X5 design matrix with $n = 60$ when special weights given for $m = 2$, at the five percent level	192
Table 7.7	Ranges of the three disjoint intervals and the corresponding values of j and l , where $\gamma_{1j} = (2j-1)/2q$, $j = 1, 2, \dots, q$ and $\gamma_{4l} = (2l-1)/2q$, $\gamma_{5l} = (2l-1)/2q$, $l = 1, 2, \dots, q$, $q = 50$ for the X5 design matrix with $n = 60$	193
Table 7.8	Selected calculated sizes and powers of the test for the X5 design matrix with $n = 60$ when $m = 3$, at the five percent level	193

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, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.


Nelufa Begum

Acknowledgments

In the Name of Allah, the Most Gracious, the Most Merciful.

I have to express my deep gratitude to Professor Maxwell L. King, my supervisor for his generous guidance, sagacious suggestions, continuous encouragement, moral support and freedom of work. Without his help, this thesis would not come out into light. He was a source of profound inspiration throughout the course of this work. It would be ungrateful, if I failed to mention his contribution while I had personal difficulties. I am also grateful to him for his editorial advice in the final stages of writing this thesis.

I would like to acknowledge with gratitude to the Department of Immigration and Multicultural Affairs (DIMA) for providing me the International Postgraduate Research Scholarship (IPRS) and Monash University for the award of a Monash Graduate Scholarship (MGS) for this course. I am indebted to Professor Max King for his recommendations to obtain these scholarships. He also gave me a part-time research assistantship which helped me financially. The Department of Econometrics and Business Statistics provided me with a departmental scholarship and a part-time teaching position which gave me financial support and thoughtful experiences. An excellent working environment in the Department of Econometrics and Business Statistics was highly appreciated.

My extended thanks to Associated Professor Paramsothy Silvapule, for giving me permission to use the results of Chapter 3 of her Ph.D. thesis and Mrs Syfun Nahar for

giving me constant company during my stressful life at the beginning of my study, and my present room mate, Mrs Lydia Shenstone, for her cheerful company which has made my study enjoyable.

I would like to thank Ms Inge Meldgaard and Mr. Myles Tooher for their support of my computing needs; Mrs Mary Englefield, Mrs. Philippa Geurens, Mrs. Pauline Froggat, Mrs. Julie Larsen and Ms Deanna Orchard for their administrative support in various ways.

My utmost gratitude goes to my father Md Shawkat Ali and mother Mrs Siddiq Begum who gave impetuous support in all aspects of my work. Without their sacrifice, love and moral support it would be impossible to complete my thesis. I would like to share my feelings with my sisters and brothers for their moral support.

Finally, the sense of my gratefulness would be incomplete if I failed to mention the enormous contribution of my husband, Mohd. Nazrul Islam, to the completion of my thesis. I would like to express my deep debt to my beloved innocent son, Ahmed Tahsin Islam who was deprived from his rightful share of mother's affection during my study.

Abstract

This thesis is concerned with hypothesis testing in econometrics and proposes a new optimal approach for testing composite hypotheses.

Hypothesis testing is a fundamental tool for evaluating econometric models. It can be used to test aspects of econometric theory or to check model specification. It is therefore essential to have reliable test procedures for the small samples econometricians sometimes have to deal with, particularly, when nuisance parameters are present. In the absence of uniformly most powerful (UMP) tests, King (1987a) suggested the use of point optimal (PO) tests which are most powerful at a chosen point under the alternative hypothesis. For situations where PO tests cannot be constructed, King suggested an approximate point optimal (APO) test.

The existing literature shows that King's PO tests are not suitable for all testing problems. For example, Silvapulle and King (1991) could not find a point optimal invariant (POI) test for testing first-order moving average (MA(1)) disturbances against first-order autoregressive (AR(1)) disturbances in the linear regression model. They recommended the use of an approximate point optimal invariant (APOI) test. This motivated us to construct a general test procedure, called the average power test, based on the generalized Neyman Pearson lemma. The test maximizes average power function subject to average size being controlled over subspaces of the null parameter space for testing a composite null hypothesis against a composite alternative.

In this thesis, we apply this new testing approach to the testing problem considered by Silvapulle and King (1991), to testing the form of heteroscedastic disturbances and to testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances in the linear regression model. For the first testing problem, the null and alternative hypothesis parameter spaces are restricted to unit intervals. We compare the small sample size and power properties of the test with those of Silvapulle and King's APOI tests, with encouraging results. The new test is found to have good small sample size and power properties. For the second and third testing problems, both the null and alternative hypotheses parameter spaces are one-sided infinite intervals and two-dimensional squares, respectively. Results of size and power calculations indicate that our new test has fairly good small sample size and power properties. For the third testing problem, we find it extremely time consuming to obtain the critical values for which the average size conditions are simultaneously satisfied. This is because the test procedure is extremely computer intensive when the hypotheses are multidimensional. Our study clearly indicates that the new optimal test procedure performs well at least for the testing problems we considered, especially when the sample size is small.

The overall recommendation of this thesis is that the new test procedure has considerable potential, particularly for problems in which the effective null hypothesis parameter space can be kept small. The new test procedure can be favourably considered in situations where PO and APOI tests cannot be constructed or appear to perform poorly.

Chapter 1

Introduction

1.1 Background

It is well recognised that Ragnar Frisch first introduced the term 'econometrics' in the early 20th century (see for example, Darnell (1984, 1994), Bjerkholt (1995, 1998) and Strom (1998)). This term appeared in his very first paper in economics which was written in French and was published in 1926 (Frisch (1926)).

Econometrics, in its early stages was very much concerned with the application of statistical methods. As Hillier and King (1991, p.1) remarked,

These methods had originally been developed for the natural sciences where careful attention to the experimental design can often justify the confident use of relatively simple stochastic models. ... In particular, there is usually greater uncertainty about the stochastic mechanism that generated the data, making it essential to have available reliable diagnostic procedures that allow one to at least check the adequacy of the model used, and ideally would indicate the direction or directions in which revision of the model might prove fruitful.

Some applied researchers (Klein (1947) and Stone (1954)) quickly appreciated these needs. In this regard, Hillier and King (1991, p.1) further noted:

However, the modern applied econometrician uses more complicated models (nonlinear, dynamic, censored, etc.), and wishes to entertain the possibility of departures from the basic model in a number of directions. Research on diagnostic testing has, according, expanded greatly in its attempt to meet these needs.

Clearly hypothesis testing has an important role to play in econometric practice. It too has a long history. After the publication of Karl Pearson's (1900) goodness of fit test, the systematic development of hypothesis testing began (see Bera and Premaratne (2001, p.38)). Neyman and Pearson (1933) laid the foundation of the theory of hypothesis testing and the concept of an optimal test was introduced through the analysis of the power function. An important building block for test construction is the Neyman-Pearson (NP) lemma which provides a way to find the most powerful (MP) and uniformly most powerful (UMP) tests. Neyman and Pearson (1936) generalized the fundamental NP lemma but compared to the number of applications of the NP lemma, the generalized NP lemma has been very much under utilized.

During the last four decades the likelihood ratio (LR) test, Wald and Lagrange multiplier (LM) tests have emerged as three important classical test principles based on likelihood theory. They have been popular with econometricians but are not totally ideal, particular with respect to power in the case of one-sided testing problems or in applications with large numbers of parameters and a small number of observations. However, econometricians have produced a large number of general test procedures such as those proposed by Hausman (1978), White (1982), Tauchen (1985), Newey (1985) and King (1987a).

A new development in econometric theory during the last two to three decades is the shift of focus from asymptotic theory to finite sample properties. Prior to this, there was little econometricians could do about the exact finite sample distributions of estimators and test statistics (see Chen (2001, p.3)). Therefore most inferential procedures relied on first-order asymptotics. Monte Carlo results provided frequent embarrassing evidence against this approach. Many studies in different settings showed that first-order asymptotic theory provides poor approximations to finite sample distributions and can provide poor inferences in practical applications, see for example Bewley (1986, Section 3.3), Cox and Reid (1987), King (1987b), King and McAleer (1987), Moulton and Randolph (1989), Chesher and Austin (1991) and McAleer (1995).

This shift towards more accurate and more reliable small sample inference is largely facilitated by rapid increases in computing capacity. That is, as in many disciplines, computers have played an essential role in making finite sample econometrics operational. There is software for data management, modelling, estimation, inference, simulation and graphics which make things easier (see for details, Hendry (2001)). The typical building blocks for econometric analysis are model specification, model selection, model estimation, hypothesis testing and forecasting/prediction. The aim of the thesis is to make a contribution in the area of testing a composite null hypothesis against a composite alternative hypothesis.

1.2 Motivation and direction of the thesis

Many applied econometric studies involve testing hypotheses about the parameters of econometric models. The first aim of hypothesis testing is to control the probability of Type I error, that is, the critical value of the test should be chosen so that the test rejects

a correct null hypothesis with a probability α (specified level of significance). The second aim is to make the probability of correctly rejecting the null hypothesis (the power of the test) as high as possible. Most hypotheses in econometrics are composite, the null hypothesis does not completely specify the data generation process in the sense that there are parameters with unknown values involved. Consequently, the sampling distribution of the test statistic is unknown except in special cases. In these circumstances it is highly likely that the sampling distribution of the test statistic depends on one or more unknown population parameters. As a result, the probability of Type I error can vary with different values of the parameters and we cannot fix this error at a desired level. Alternatively, a test statistic is pivotal if its finite-sample distribution under the null does not depend on unknown population parameters. However, pivotal test statistics are not available in many econometric applications unless strong distributional assumptions are made. When the test statistic is not pivotal, its Type I critical value can be very different at different points under the null hypothesis (see for details, Horowitz (2001), and Horowitz and Savin (2000) for a numerical example).

King (1987a) introduced the class of point optimal invariant (POI) tests for testing both simple and composite hypotheses, which are most powerful invariant tests at a chosen point under the alternative hypothesis. POI tests have excellent power properties for a number of testing problems in econometrics. They can have good power properties even at parameter points away from the point of the alternative hypothesis at which power is optimized but are not suitable for all testing problems. There are some situations where the test does not work at all. For example, Silvapulle and King (1991) could not find a POI test for testing MA(1) against AR(1) disturbances in the linear regression model and they recommended the use of an approximately POI (APOI) test. Rahman and King

(1994) considered testing random regression coefficients in the presence of AR(1) errors using APOI tests and they concluded that the extra computation required for the APOI tests hardly seems worthwhile. POI tests are based on the fundamental NP lemma. A starting point of this thesis is that almost no test procedures have been developed to exploit the generalized NP lemma, one exception being the work of Srikanthakumar (2000). She constructed the g test, an APOI test for testing composite hypotheses by using the generalized NP lemma which depends on a choice of representative densities. Unfortunately the particular choice of representative densities determines the form of the test. This motivated us to construct a new test procedure for testing composite hypotheses by using the generalized NP lemma which is introduced in Chapter 3.

Two decades ago, it was not possible to implement some test procedures due to the lack of computer power. Recently there has been a dramatic change in the computing power available to researchers and it is much easier to implement almost all types of test procedure through Monte Carlo simulations. We fully expect these increases in computer power to continue over the next two decades.

The overall objectives of this thesis are:

- i) to design a new test procedure for the problem of testing a composite null against a composite alternative based on the generalized NP lemma;
- ii) to implement the new test procedure, which involves solving a number of practical issues with respect to controlling average size over a number of subregions under the null hypothesis; and

iii) to investigate the small sample performance of the new testing approach with respect to the following individual hypothesis testing problems in the context of the linear regression model:

- a) testing for MA(1) disturbances against AR(1) disturbances;
- b) testing for different forms of heteroscedastic disturbances (based on Hildreth-Houck random coefficient models); and
- c) testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances.

1.3 Outline of the thesis

This thesis is organized into eight chapters. Following this introductory chapter, Chapter 2 reviews the literature related to hypothesis testing. It surveys the literature involving optimal, PO and approximately PO (APO) tests in the context of the linear regression model for composite hypotheses testing problems. Because all the applications considered in this thesis are non-nested, the review also focuses on contributions to tests of non-nested hypotheses and some popular non-nested tests are discussed briefly and their limitations noted. Other topics briefly discussed include nuisance parameters, invariance arguments and marginal likelihood methods.

This literature review reveals the importance of developing optimal tests which have excellent finite-sample properties, such as, PO tests, rather than large-sample based tests. Studies involving APO tests indicate that they are not always suitable for all composite hypotheses testing problems and in some cases they do not work well at all. Almost all existing tests in the context of non-nested testing are large-sample based tests and many of them perform poorly in finite samples. These problems are the key issues of this thesis.

Chapter 3 explores a general solution to testing a composite null against a composite alternative. We discuss the theory behind our new approach and how the test procedure can be applied to the problem of testing the structure of the disturbances in the linear regression model. We refer to this new approach as the 'average power test'. It involves maximizing the average power function of the test subject to average size being controlled over different subregions of the null hypothesis parameter space. We note that controlling average size over the entire null parameter space, namely Θ , may result in undesirably large sizes in some parts of Θ with smaller sizes in other parts of Θ . In order to minimize this possibility, we allow for Θ to be partitioned into m disjoint subregions, namely, $\Theta_1, \Theta_2, \dots, \Theta_m$ so that $\Theta = \Theta_1 \cup \Theta_2 \cup \dots \cup \Theta_m$. This approach involves finding several critical values for which the average size conditions are fulfilled simultaneously. We discuss the standard invariance technique which can be used to simplify the testing problem by eliminating some nuisance parameters. We outline the steps involved in calculating the critical values by using Monte Carlo simulation.

Chapter 4 illustrates the new test procedure by applying it to the problem of testing for MA(1) disturbances against AR(1) disturbances in the linear regression model where Silvapulle and King (1991) could not find a POI test and were forced to use an APOI test. This chapter explores a number practical issues with the aim of finding answers. An important question is whether controlling average size does a good job in controlling size overall. A number of subregions can be chosen to help control size. This chapter compares the performance of our new approach with various versions of Silvapulle and King's (1991) APOI tests.

Chapter 5 investigates the choice of boundary points of the subregions under the null hypothesis parameter space. In Chapter 4, we use what we call the high size/low size technique to determine the boundary points of the subregions under the null hypothesis parameter space. This chapter constructs a pseudo power envelope to trace out the maximum attainable power for a given problem in order to provide a benchmark against which the new test procedure can be evaluated.

Chapter 6 investigates the performance of the new test procedure by applying it to testing for alternative forms of heteroscedastic disturbances in the linear regression model. In this case, we concentrate on the form of heteroscedasticity that results from the Hildreth-Houck random coefficient model. The application investigated in Chapter 4 involved parameter spaces under both hypotheses being restricted to the unit interval. Chapter 6 examines the performance of the test procedure when both the null and alternative hypothesis parameter spaces are extending to an infinite interval (one-sided infinite interval) with suitable choices of subregions under the null parameter space.

Chapter 7 explores the problem of testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances in the linear regression model. This chapter investigates the performance of the test in a situation which involves testing two-dimensional parameter spaces under the null and alternative.

This thesis closes in Chapter 8 with a summary and some concluding remarks.

1.4 Computations

All the calculations reported in this thesis were performed using the GAUSS programming language, Version 3.2.35, Aptech Systems, Inc. (GAUSS (1998)) on IBM-

compatible (Pentium III, 600E with 256 Mb SDRAM) computers. The built-in random number generators (functions RNDS and/or RNDUS) were used to generate all random numbers for the Monte Carlo simulation experiments discussed in this thesis.

Chapter 2

Literature Survey

2.1 Introduction

For econometricians, hypothesis tests are a fundamental tool for evaluating econometric models and a meaningful evaluation of an econometric model depends on the use of accurate statistical test procedures. During last three decades, numerous test statistics have been developed for testing both nested and non-nested models in econometrics. Substantial advances have been made in developing diagnostic tests for the validity of various model specifications. There exist many diagnostic tests such as the DW test for serial correlation, LM tests for various forms of heteroscedasticity and autocorrelation, tests for multiple time series models, tests of distributional shape, tests for functional form, structural stability, structural change, spatial error autocorrelation, non-normality and so on; see for example Poskitt and Tremayne (1981, 1982), Spiegelhalter (1983), Tauchen (1985), Ljung (1986), Andrews (1988), Beggs (1988), Pagan and Vella (1989), Hillier (1991), Ai and Cassou (1993), Anselin et al. (1996), Kang and Inder (1996), Wooldridge (2001) and Mao and Lindsay (2002). Also, there is a special issue of the *Journal of Econometrics* on diagnostic testing edited by Hillier and King (1991). In many econometric applications, there often exist non-nested specifications that can

characterize the same variable of interest. When alternative non-nested models are available, it is particularly important to test which model is correctly specified because correct model specification is essential for inference, forecasting and policy making.

The purpose of this chapter is to survey the literature on optimal tests for testing models in econometrics and other related issues. There is an extensive literature on different approaches to statistical inference from the viewpoints of parametric, non-parametric, semi-parametric, Bayesian and decision theoretic approaches. We will focus our discussion on parametric problems in the context of linear regression models and simple non-linear regression models. The classical tests, namely, the LR, Wald and LM tests can be used both for simple and composite hypotheses testing problems. But problems arise in dealing with likelihood functions when there are nuisance parameters present, particularly with the accuracy of estimates and tests in small samples. Finite sample properties of the test procedures are therefore very important. It is hard to derive the finite sample distributions of test statistics in the presence of nuisance parameters, econometricians often use large sample based distributions as approximations. There are a number of techniques that exist to eliminate the nuisance parameters from the testing problem, an important one being the use of invariance arguments. In the following section we discuss nuisance parameters and invariance arguments in more detail.

The majority of the tests proposed for non-nested problems are large sample based and large sample based tests sometimes perform poorly in terms of size properties in small samples (see McAleer and Pesaran (1986), McAleer (1987, 1995), King and McAleer (1987) and Godfrey (1998)). It is well known that the estimates based on maximum likelihood can be biased in finite samples and the tests based on maximum likelihood

can sometimes perform poorly in finite samples (see McAleer and Pesaran (1986)). Ara (1995), Grose (1998) and Rahman and King (1998) have demonstrated that marginal likelihood based tests can perform better than conventional likelihood based tests in small samples. This indicates that better handling of nuisance parameters may improve finite sample properties of large sample based approaches. King's (1987a) point optimal (PO) and approximately PO (APO) testing approaches are particularly aimed at small sample testing problems and have been applied successfully to a number of testing problems. This chapter also aims to survey the contributions on PO testing of non-nested hypotheses. In this thesis, our interest is in composite hypothesis testing of the regression model in which two forms of disturbance processes i.e., autoregressive and moving average processes, or autocorrelation and heteroscedasticity, or joint autoregressive and moving average processes are possible.

The plan of this chapter is as follows. We briefly discuss nuisance parameters, invariance arguments and the marginal likelihood in Section 2.2. In Section 2.3, the general ideas of optimal test procedures as covered in the literature are outlined. Sections 2.4 and 2.5 introduce PO and APO tests, respectively, and compare PO and APO tests with other existing tests of composite hypotheses. Section 2.6 briefly surveys some non-nested testing problems and discusses the performances of Cox's test, the J , P , JA , F , Cox-type \tilde{N} and adjusted Cox-type \tilde{N} tests. Finally, concluding remarks are made in Section 2.7.

2.2 Nuisance parameters, invariance arguments and the marginal likelihood

In econometrics, almost all hypothesis testing problems of interest involves nuisance parameters. If the testing problem is a parametric one, it may involve a large, perhaps a very large number of unknown parameters. Almost certainly, at least some of these parameters will not be of any particular interest to the researcher. The term 'nuisance parameter' refers to the parameter or set of parameters which do not have any interest to the researcher.

For hypothesis testing problems that involve nuisance parameters, the true size of the test can differ substantially from its nominal size, particularly when many nuisance parameters are present. It is therefore of interest to differentiate the class of tests for a particular testing problem whose size does not depend on nuisance parameters. A critical region whose size is independent of nuisance parameters is known as a similar region (for details, see Hillier (1987)).

Suppose y is an observed data vector which we wish to use to test a hypothesis. Commonly, in a non-experimental discipline such as econometrics, y is assumed to be a variable of interest, which is modelled by a distribution function, assumed known except for its parameter values. Precise and consistent estimation of the model itself is our next concern and it is assumed that at least some of the model parameters will be of interest. Suppose we can write the joint density of $y = (y_1, y_2, \dots, y_n)'$ as $f(y; \mu, \theta)$, where μ is a vector of nuisance parameters and θ is our parameter (vector) of interest.

For the problem of testing $H_0: \theta = \theta_0$, where θ_0 is a specific value of θ , the size of any critical region ω (also known as rejection region, i.e., we reject H_0 if y falls in this region) for testing H_0 , namely $\int_{\omega} f(y; \mu, \theta_0) dy = \alpha(\mu)$, will typically be a function of the nuisance parameter vector μ . A critical region ω is a similar region of size α , if and only if, ω has size α in the conditional distribution of y given t (the test statistic) for almost all t . Then one might seek a test that is best within the class of similar critical regions.

The problem of testing $H_0: \theta = \theta_0$, against the specific value of $H_a: \theta = \theta_1$ is invariant in the sense of Lehmann (1959) under a group G of transformations acting on the sample y if for any transformation $T(y) \in G$, the probability distribution of $T(y)$ belongs to the same set (H_0 or H_a) as y . That is, a test with critical region ω is an invariant test if $y \in \omega$ implies $T(y) \in \omega$ for all $T(y) \in G$ and $y \notin \omega$ implies $T(y) \notin \omega$. With respect to these transformations, the vector $v = g(y)$ is a maximal invariant statistic (see Lehmann (1959)) if it is a statistic which takes the same value for y vectors that are connected by transformations and different values for y 's not connected by transformations. Now if we wish to restrict our attention to tests that are invariant under the transformations $T(y)$, then we can treat the maximal invariant as though it is the observed data. This is because all invariant test statistics can be written as a function of the maximal invariant. A test is invariant under a group of transformations on y , if and only if, the critical region for the test is defined in terms of the maximal invariant. For example, in the case of testing the disturbance structure in the linear regression model, $y = X\beta + u$, the testing problem is invariant to transformations of the form

$T(y) = \eta_0 y + X\eta$ and $v = Pz/(z'z)^{1/2}$ is a maximal invariant (this is discussed in detail in Section 3.3).

A number of researchers have used invariance arguments to eliminate nuisance parameters from their testing problems (see for example, Durbin and Watson (1971), Kariya and Eaton (1977), King (1980, 1981a, 1981b, 1982a, 1983a), Franzini and Harvey (1983), Sargan and Bhargava (1983), Lehmann (1986), King and Inder (1986), King and Smith (1986), Honda (1988), Ara and King (1993, 1995) and Ara (1995)). In addition King and Hillier (1985) constructed a locally best invariant (LBI) test for the case of testing a single parameter when the alternative is one-sided, i.e., for the wider problem of testing the null hypothesis $H_0: \theta = 0$ against $H_a: \theta > 0$, when the covariance matrix of the regression disturbance vector u is of the form $\sigma^2 \Omega(\theta)$. Also, they constructed a locally best unbiased invariant (LBUI) test when the alternative is two-sided, that is, for testing $H_0: \theta = 0$ against $H_a: \theta \neq 0$.

The most popular basis for inference in econometrics is that based on the likelihood function. It is sometimes described as the likelihood principle. Its strength is that it is easily applied to problems of inference in multiparameter models. A large part of the literature has generally concentrated on the idea of removing, in some sense, at least some of the parameters from the likelihood in order to improve inference on the others. An extended list of 'modified' likelihoods has resulted; ranging from the now conventional concentrated, or profile, likelihood, in which some parameters are replaced by "estimators" that solve the first-order equations but are a function of other parameters, through to conditional, marginal and integrated likelihoods (Kalbfleisch and Sprott (1970)), various conditional, adjusted and modified profile likelihoods (Cox and

Reid (1987, 1993), Barndorff-Nielsen (1983), McCullagh and Tibshirani (1990), Laskar (1998), Laskar and King (1998)), to canonical and expected likelihoods (Hinde and Aitkin (1987), Conniffe (1987)). All of these are intended to allow inference on the parameter or parameters of interest, along with minimising interference from nuisance parameters.

The concept of marginal likelihood was first introduced by Fraser (1967) in the context of structural inference. Since then this approach has been applied to the problem of testing the parameters of the linear regression model, their marginal likelihood is easily obtained and well known. The properties of marginal likelihood based tests have been considered in the context of testing for autocorrelation (Corduas (1986), Ara (1995), Ara and King (1993)); testing for heteroscedasticity (Levenbach (1973)) and testing for random coefficients (Ara and King (1993)). In all cases use of the marginal likelihood has resulted in an improvement in the finite sample properties of inferential procedures. Laskar and King (1995) considered LR, LM and Wald tests of the MA(1) error model and reported improved size and power for marginal likelihood based tests. Rahman and King (1998) observed that the LM test based on the marginal likelihood tends to be more powerful than its conventional counterparts. The general theme of this literature is that procedures based on the marginal likelihood show a distinct improvement over equivalent procedures based on the conventional concentrated likelihood.

2.3 Optimal tests

In this section we survey important ideas related to statistical inference, with respect to hypothesis testing in econometrics. We start the discussion of particular methods of inference by looking at problems with a very simple structure. Suppose that we have the

data vector y and a hypothesis H_0 concerning its density $f_0(y)$. The purpose of a hypothesis test is to examine the consistency of the data with H_0 . In order to do this we need to have some idea of the type of departure from the null hypothesis that we wish to be looking out for. We now assume that, in addition to H_0 , we have one or more alternative hypotheses representing the directions of the departures from H_0 . Let H_a ($\theta = \theta_1$) be a particular simple (a hypothesis is simple when it completely specifies the distribution of y) alternative hypothesis. The problem is basically one of deciding whether H_0 or H_a is more appropriate.

The main tool we have for choosing a test statistic in order to maximize power is the NP lemma which suggests the use of the LR

$$t(y) = \frac{f_a(y)}{f_0(y)}, \quad (2.1)$$

when just those two distributions are under consideration. Clearly, the larger the value of $t(y)$, the worse the fit under H_0 , so we reject H_0 for large values of $t(y)$.

We suppose that $t(y)$, under H_0 , is a continuous random variable such that for all α , $0 < \alpha < 1$, there exists a unique critical value, c_α , which is defined by

$$\Pr(t(y) \geq c_\alpha | H_0) = \alpha, \quad (2.2)$$

for c_α when a significance level α is chosen. The NP lemma states that for any size α , the LR critical region,

$$t(y) \geq c_\alpha, \quad (2.3)$$

is the most powerful (MP) test. The NP lemma is the foundation of tests with good power properties.

Suppose now that under H_0 we have just one distribution but under H_a we have a collection of distributions, $f(y, \theta)$, one for each different value of θ such that $\theta > 0$.

Cox and Hinkley (1974, p.101) state that two cases may now arise:

- i) we get the same size α best (or MP) critical region (via the NP lemma) for all distributions of y under H_a ;
- ii) the best (or MP) critical region depends on the particular distribution of y under H_a .

In order to find out which of these two possibilities is true, we would first select a particular distribution of y under H_a (say, $\theta = \theta_1 > 0$) and then treat this choice as a simple alternative hypothesis. With respect to testing a simple null hypothesis against this simple alternative, the NP lemma gives us the MP critical region. If after taking various monotonic transformations (quite often useful in numerical work) to simplify the form of the test statistic, we find it does not depend on the choice of θ_1 under H_a , then we have case (i). Otherwise, the test depends on the choice of θ_1 and will change as θ_1 changes, so case (ii) is the outcome.

The performance of a test is typically assessed by its size and power properties. Econometricians are always concerned with optimality of power. Neyman and Pearson (1933) laid down the foundation for the theory of optimal test procedures by introducing the uniformly most powerful (UMP) test. A test is said to be UMP when it is MP for

every alternative distribution. For a parametric hypothesis testing problem, we would always prefer to use an UMP test, which maximizes the power curve over the entire parameter space.

King (1996) pointed out that statistical theory is good at suggesting optimal test procedures for situations involving only a few parameters, but beyond this is rather less helpful. Later, Grose (1998, p.11) commented that this point might be extended by noting that only for a single parameter system can the estimation/inference problem be said to be solved in any general sense and cases involving more than one parameter must be handled according to their own nature.

In our case (i) above, where the test does not depend on θ_1 , we have a test that is MP for every value of θ under H_a . This is an ideal test and is UMP. Its power functions always dominate the power curves of any other test of the same significance level.

A UMP test may not always exist. Unfortunately, it is rarely possible to find a UMP test when the alternative hypothesis is composite (a hypothesis is composite when it does not completely specify the distribution of the data vector) and/or in the presence of nuisance parameters. A more realistic situation is case (ii), in which case we need to decide how to make a choice of test.

There are a number of possible approaches. We consider here the situation when the distributions under H_a are indexed by a single parameter θ . Cox and Hinkley (1974, p.102) consider various techniques for constructing tests of a simple hypothesis against a composite alternative when no UMP test exists. They suggest three different approaches to choose the point at which the power optimised:

- i) to pick, somewhat arbitrarily, a "typical" point θ_1 in the alternative parameter space and use it in the test in order to find a test that is MP for that particular chosen point under the alternative;
- ii) to take θ_1 , a chosen point from the alternative parameter space which is very close to its H_0 value, i.e., to maximize the power locally near the null hypothesis;
- iii) to maximize some weighted average of power over the alternative parameter space.

Thus if uniform optimization is impossible, other approaches of test construction such as the principles of unbiased tests, consistent tests, locally most powerful (LMP) tests, invariant tests, similar tests that may ensure certain known but lesser optimal power properties, can be claimed. Davies (1969) defined a PO test to be ' β optimal' if its power function is always a monotonic decreasing function of the parameter under test and its power reaches a predetermined value most quickly. Bhatti and King (1990) have given an example of a β optimal test.

The locally best (LB) test is optimal in the sense that its power curve has the steepest slope of all power curves with the same size under the null hypothesis. For a single parameter, Neyman and Pearson (1936) proposed a test which they called a type A test and which is a locally best unbiased (LBU) or locally most powerful unbiased (LMPU) test. They also introduced the A_1 type test which is a uniformly most powerful unbiased (UMPU) test. In the case of a composite hypothesis, Neyman (1935) constructed type B and type B_1 tests in the presence of a nuisance parameter, that are LBU and UPMU,

respectively. For two unknown parameters, Neyman and Pearson (1938) constructed a type C test. Isaacson (1951) introduced a type D test, but unfortunately in practice, type D tests are very difficult to apply. Later, Lehmann (1986, p.529) proposed a modified version of the type D test and named it a type E test which admits a nuisance parameter.

Following Neyman and Pearson (1936); Ferguson (1967), Efron (1975), King and Hiller (1985), Sen Gupta (1987), and Wu and King (1994) reviewed the LB test. For testing a single parameter against a one-sided alternative, the LB test is a well defined and accepted test. King and Hiller (1985) and Wu and King (1994) noted that the LB test is equivalent to the LM test based on the square root of the standard LM test statistic. King and Wu (1997) recommended a one-sided LM-type test which is locally mean most powerful (LMMP), a generalisation of a single parameter LB test. Majumder (1999) proposed a distance-based approach to a single parameter or scalar case to one-sided or partially one-sided testing problems. He developed a modification of King and Wu's LMMP test and other LM-type tests. As this LMMP test is not suitable for testing partially one-sided alternatives, he suggested that the generalization would solve the problem.

Neyman and Pearson (1936) first introduced and proved the generalized Neyman-Pearson lemma is basic in the theory of statistical hypotheses testing. Here we present the lemma only, for full mathematical details see Dantzig and Wald (1951). We quote the generalization of the fundamental NP lemma from Lehmann (1986, p.96):

Theorem: Let f_1, \dots, f_{m+1} be real-valued functions defined on a Euclidean space A and integrable μ , and suppose that for given constants c_1, \dots, c_m there exists a critical function ϕ satisfying

$$\int \phi f_i d\mu = c_i, \quad i = 1, \dots, m. \quad (2.4)$$

Let ζ be the class of critical functions ϕ for which (2.4) holds.

(i) Among all members of ζ there exists one that maximizes

$$\int \phi f_{m+1} d\mu. \quad (2.5)$$

(ii) A sufficient condition for a member ζ to maximize (2.5) is the existence of constants k_1, \dots, k_m such that

$$\begin{aligned} \phi(y) &= 1 \text{ when } f_{m+1}(y) > \sum_{i=1}^m k_i f_i(y), \\ \phi(y) &= 0 \text{ when } f_{m+1}(y) < \sum_{i=1}^m k_i f_i(y). \end{aligned} \quad (2.6)$$

(iii) If a member of ζ satisfies (2.6) with $k_1, \dots, k_m \geq 0$, then it maximizes (2.5) among all critical functions satisfying

$$\int \phi f_i d\mu \leq c_i, \quad i = 1, \dots, m. \quad (2.7)$$

(iv) The set M of points in m -dimensional space whose coordinates are $(\int \phi f_1 d\mu, \dots, \int \phi f_m d\mu)$ for some critical function ϕ is convex and closed. If (c_1, \dots, c_m) is an inner point of M , then there exist constants k_1, \dots, k_m and a test ϕ satisfying (2.4) and (2.6), and a necessary condition for a member of ζ to maximize (2.5) is that (2.6) holds.

Suppose $f_1(y), f_2(y), \dots, f_m(y)$ are m separate density functions under H_0 and $f_{m+1}(y)$ is the density under H_a . Also suppose for given probability constants c_1, c_2, \dots, c_m there exists a critical region ω of the form that H_0 is rejected if $y \in \omega$ satisfying

$$\int_{\omega} f_i(y) dy = c_i, \quad i = 1, 2, \dots, m, \quad (2.8)$$

then the generalized NP lemma tells us that the critical region to maximize power, i.e.,

$$\int_{\omega} f_{m+1}(y) dy. \quad (2.9)$$

evolves rejecting H_0 when

$$f_{m+1}(y) > \sum_{i=1}^m k_i f_i(y), \quad (2.10)$$

where k_1, k_2, \dots, k_m are "critical values" chosen to satisfy (2.8).

If any member of the class of critical regions, ω , satisfies (2.10) with $k_1, k_2, \dots, k_m \geq 0$, then it maximizes (2.9) among all critical functions satisfying

$$\int_{\omega} f_i(y) dy \leq c_i, \quad i = 1, 2, \dots, m. \quad (2.11)$$

This is one example of an application of the generalized NP lemma. Note also that $f_1(y), f_2(y), \dots, f_{m+1}(y)$ do not need to be likelihood functions or density functions. In fact, in our application in Chapter 3, we will be using integrated likelihood functions.

The generalized NP lemma states that if one finds appropriate critical values k_1, k_2, \dots, k_m such that the m size conditions (2.8) hold simultaneously, for those critical values the test will be the most powerful. We note that there are not a lot of examples of applications of the generalized NP lemma, although Sriananthakumar (2000) used this lemma to construct an APOI test.

There seems to be a range of tests, each optimal in one sense but no obvious choice of which is best. If this is the case, according to Lehmann (1947) the choice has to be based on information not contained in the general formulation of the testing problem. If no

such additional information is available, the choice must of necessity be somewhat arbitrary.

2.4 Point optimal tests

The origin of the key idea behind the PO test goes back to Lehmann and Stein (1948). Other references are Lehmann (1959), Davies (1969), Kadiyala (1970) and Berenblut and Webb (1973). It was King (1985a, 1985b) who used the term "point optimal" and sparked a revival of interest in the tradition of optimal test procedures. King (1983a) provides a discussion of the theoretical foundations of PO tests and King (1987a) presented a clear and thorough explanation of PO testing by simplifying the method of test construction and by introducing computational checks for accuracy.

We start by describing a PO test in the context of testing

- i) a simple null against a simple alternative hypothesis;
- ii) a simple null against a composite alternative hypothesis; and
- iii) a composite null against a composite alternative.

When a test achieves optimum power at a particular parameter point it is then referred to as a PO test. For a testing problem that involves a simple null against a simple alternative, it is well known that the NP lemma provides us with an MP test. That is, the LR test (see equations (2.1) and (2.3)) is the MP test and hence it is the PO test.

For a testing problem that involves a simple null against a composite alternative, then for a chosen point in the alternative parameter space, the LR test of this fixed point under the alternative is a test that is MP in the neighbourhood of the chosen point; hence

it is a PO test. When the testing problem involves both composite null and alternative hypotheses, one might use the LR test that corresponds to a fixed point under the null and a fixed point under the alternative. This does not necessarily result in a test which is MP in the neighbourhood of the chosen point in the alternative parameter space. The reason for this is that since the null hypothesis is composite, the probability of a Type I error, that is, the size of the test, may be a function of the null hypothesis parameters; consequently the maximum size of the test may exceed the nominal level. The standard approach in this case (see, for example, Lehmann and Stein (1948)), is to control the maximum size to be less than or equal to some desired level of significance. Unfortunately, there are many circumstances in which size will vary across the null hypothesis parameter space. If such a critical value is chosen so that the global maximum of the size function of the LR test is attained at some point in the null parameter space, then King (1987a) called the test a Pseudo PO (PPO) test. If such a critical value is chosen so that the maximum size is attained precisely at the chosen point in the null hypothesis parameter space, then it is a MP test in the neighbourhood of the chosen point in the alternative parameter space and therefore a PO test.

In the absence of UMP tests, King (1987a) constructed the PO test for a very general framework and applied this to the problem of testing AR(1) errors against MA(1) errors in the linear regression model. King proposed the PO test for the general problem of testing (not in all cases) which includes both nested and non-nested problems as special cases and involves testing

$$H_0: y \text{ has density } f(y, \theta) \quad (2.12)$$

against

$$H_a: y \text{ has density } g(y, \phi), \quad (2.13)$$

where y is the observed sample, θ is a $p \times 1$ vector of parameters restricted to the set Ω and ϕ is a $q \times 1$ vector of parameters restricted to the set Φ . It is assumed that the possible range of parameter sets, Ω and Φ , are as small as possible.

As mentioned in Section 2.2, for the simpler problem of testing

$$H'_0: y \text{ has density } f(y, \theta_1) \quad (2.14)$$

against

$$H'_a: y \text{ has density } g(y, \phi_1), \quad (2.15)$$

where $\theta_1 \in \Omega$ and $\phi_1 \in \Phi$ are fixed and known, we have simple null and simple alternative hypotheses. Therefore the NP lemma implies that rejecting H'_0 for large values of

$$t(\theta_1, \phi_1) = \frac{g(y, \phi_1)}{f(y, \theta_1)}$$

is a MP test. If $t(\theta_1, \phi_1)$ is used as a test statistic for the wider problem of testing the simple null hypothesis, H'_0 , against the composite alternative, H_a , then this test serves as the MP test in the neighbourhood of $\phi = \phi_1$. However, as already noted the same test does not necessarily result in a MP test in the neighbourhood of $\phi = \phi_1$, when testing the composite null, H_0 , against the composite alternative, H_a . For the test, when H_0 is composite, the distribution of t under H_0 and hence the probability of Type I error is

likely to be a function of θ . Hence, for this case, the critical value, c_α , is found by solving

$$\Pr[t(\theta_1, \phi_1) > c_\alpha | y \text{ has density } f(y, \theta)] \leq \alpha, \text{ for all } \theta \in \Omega, \text{ for } c_\alpha, \quad (2.16)$$

with equality for at least one value of θ . That is, we have to choose θ_1 such that the maximum size of the test occurs at this point. Then testing H_0 against H_a , the test is MP in the neighbourhood of $\phi = \phi_1$, and we have the PO test.

Since King (1980), PO testing has been an active field of research. The most successful applications of PO testing have been in the context of the linear regression model. Some specific examples of PO test are: testing for AR(1) disturbances (Berenblut and Webb (1973), King (1985a), Dufour and King (1991)); testing for fourth-order autocorrelation in presence of first-order autocorrelation (King (1989)); testing for MA(1) disturbances (King (1985b)); testing for AR(1) against MA(1) disturbances (King and McAleer (1987)); testing for heteroscedasticity (Evans and King (1985b, 1988)); testing for serial correlation and heteroscedasticity (King (1986), King and Evans (1984)); testing for random walk disturbances (Sargan and Bhargava (1983)); testing for random coefficients (Franzini and Harvey (1983), Shively (1986, 1988a, 1988b), Brooks (1993, 1995), Brooks and King (1994a, 1994b)); testing for unit roots in observed time series (Bhargava (1986), Jaeyoun and Schmidt (1996)); testing for moving average unit roots in ARIMA models (Saikkonen and Luukkonen (1993)); testing for autoregressive disturbances in a time series regression with missing observations (Shively (1993)); testing for block effects in regression disturbances (King and Evans (1986), Bhatti (1992a), Bhatti and King (1993, 1994)); testing for serial correlation in a large number

of small samples (Bhatti (1992b)); tests of linear regression disturbances (Evans and King (1985a)); testing for trend stationarity (Hwang and Schmidt (1993)); tests of the error covariance matrix (Honda (1989)); point-optimal Cox tests (Dastoor and Fisher (1988a)); tests of non-nested error processes (King (1983a, 1987a), Silvapulle (1991, 1994a, 1994b), Silvapulle and King (1991, 1993)), inference in integrated autoregressive models (Rothenberg and Stock (1997)) and inference in nearly integrated non Gaussian models (Thompson (2002)).

Regardless of its excellent performance in many instances, King's (1987a) PO test has had its opponents (see, Dastoor and Fisher (1988b), Bierens (1988) and Pötscher (1988)). The main criticisms King's PO test have received are mentioned briefly in Srikanthakumar (2000, p.13). One important criticism is that if the null hypothesis is simple (possibly after reduction by invariance), PO tests are easily constructed using the NP lemma. But in practice, there are many existing testing problems that cannot be simplified by using invariance arguments or by any other means. In other words, if the null hypothesis is not simple (or cannot be reduced to a simple hypothesis by invariance arguments), the construction of the PO test poses unresolved problems. Thus, in the case where the null hypothesis is composite, under what conditions do PO tests, exist? How can PO tests be found in this case?

As the complete knowledge of the distribution is a prerequisite for the application of PO tests to regression problems, some researchers have ignored the PO test (even for situations where it can be easily constructed). They have developed or used tests, which have known asymptotic distributions (see, Burke et al. (1990) and Smith and Tremayne (1990)). Some researchers find it very straightforward to apply asymptotic tests because

of their known asymptotic null distribution. However, computational simplicity is not the only criterion for making a choice of testing procedure. In the following section we review approximate PO (APO) tests.

2.5 Approximate point optimal tests

In the previous section, we mentioned that King (1987a) reviewed the theory of PO test construction for a very general framework and applied it to the problem of testing AR(1) disturbances against MA(1) disturbances in the linear regression model. In cases where the null hypothesis cannot be reduced to a simple hypothesis, there is no clear approach for constructing a test with optimal power properties. King observed that his approach does not suit all testing problems and does not work in a number of cases. One example of a situation in which King's test does not work involves testing MA(1) disturbances against AR(1) disturbances in the linear regression model. Silvapulle and King (1991) could not find a PO invariant (POI) test (if a test achieves optimum power at a particular point within the class of invariant tests then it is called a POI test) for this testing problem. Unfortunately, sometimes the presence of nuisance parameters that cannot be eliminated through invariance arguments considerably complicates the construction of PO tests. In fact, there is no guarantee that the method of construction outlined by King (1987a) will work in all situations.

King (1989) applied the APOI test to the problem of testing for fourth-order autocorrelation disturbances in the presence of first-order autocorrelation in the linear regression model. He found that it works well. For the problem of testing MA(1) disturbances against AR(1) disturbances in the linear regression model, Silvapulle and King (1991) could not find a POI test, they recommended the use of an APOI test. For

this testing problem following King (1987a), the critical regions of a POI test with optimal power at $\rho = \rho_0 > 0$ can be written as

$$s(\rho_0, \gamma_0) = \hat{u}' \Sigma(\rho_0)^{-1} \hat{u} / \tilde{u}' \Omega(\gamma_0)^{-1} \tilde{u} < c, \quad (2.17)$$

where \hat{u} and \tilde{u} are the generalized least squares (GLS) residual vectors of the general linear regression model $y = X\beta + u$, assuming the covariance matrices $\Sigma(\rho_0)$ and $\Omega(\gamma_0)$, respectively. We assume that u is normally distributed, although this can be extended to elliptical symmetry without changing any of what follows.

Now the test statistic given in (2.17), requires the critical value c and the parameter γ_0 to be chosen in such a way that

$$\Pr[s(\rho_0, \gamma_0) < c \mid u \sim N(0, \Omega(\gamma_0))] = \alpha, \quad (2.18)$$

and

$$\Pr[s(\rho_0, \gamma_0) < c \mid u \sim N(0, \Omega(\gamma))] \leq \alpha, \text{ for all } \gamma, \quad (2.19)$$

where α is the desired level of significance. King mentioned that Imhof's (1961) algorithm could be used to solve (2.19) and this may be achieved using Koerts and Abrahamse's (1969) FQUAD subroutine or Imhof's algorithm coded by Davies (1980). For more detail see King (1987a).

To obtain appropriate γ_0 and c values to construct an APOI test, King (1987a) suggested an iterative procedure (see for details Silvapulle and King (1991)). He also suggested that the APOI test might work well if it is nearly a PO test, otherwise it may not always be reliable.

Sriananthakumar (2000) constructed the g test: an APOI test based on the generalized NP lemma. For the g test, if the observed sample is generated from one of a finite number of densities under the null, then the generalized NP lemma provides a PO test for the problem of testing for a finite number of observable density functions against a single alternative density function. Suppose, one wishes to test composite hypotheses of the form given by equations (2.12) and (2.13) and $\phi^* \in \Phi$ is the point under the alternative hypothesis at which the power is to be optimized. Then the testing problem can be written as

$$H_0: y \text{ has density } f(y, \theta) \quad (2.20)$$

against

$$H_a: y \text{ has density } g(y, \phi^*). \quad (2.21)$$

The g test involves approximating $f(y, \theta)$, $\theta \in \Omega$, by a finite number (say, r) of densities by the selection of r separate θ points in Ω , namely, $\theta_1, \theta_2, \dots, \theta_r$. The corresponding densities,

$$f_i = f(y, \theta_i), \quad i = 1, 2, \dots, r, \quad (2.22)$$

are regarded as representative densities of $f(y, \theta)$. The generalized NP lemma is used to obtain a PO solution for this type of approximation.

In construction of the g test, for the selected chosen points under the null, a corresponding number of representative densities are needed for which the size

conditions hold simultaneously. That is, for the existence of the g test, the following r size conditions,

$$\Pr[f_{r+1}(y) \geq \sum_{i=1}^r k_i f_i(y) | y \sim f_j(y)] = \alpha_j, \quad j = 1, 2, \dots, r, \quad (2.23)$$

where α_j is the desired level of significance, need to be solved simultaneously by appropriate choices of values for k_1, k_2, \dots, k_r .

If this does not work in terms of giving acceptable size properties across the entire null hypothesis parameter space, Srikanthakumar suggested adding another representative density under the null and solving for an extra size condition. The process of increasing the number of representative densities is followed until the desired outcome is achieved.

The g test is applied with a minimum number of representative densities under the null, because as the number of representative densities increases, the required computing time increases exponentially. She applied the g test to two testing problems, namely, testing for MA(1) disturbances against AR(1) disturbances and testing for AR(1) disturbances against IMA(1,1) disturbances in the linear regression model. She compared its size and power properties in small samples with those of King's APOI tests and the asymptotic tests of Silvapulle and King (1991) and Silvapulle (1994a). She found that the g test performed well for both testing problems, while King's APOI tests performed extremely well for the former problem but performed poorly for the latter, in terms of size and power properties.

The existing literature shows that the traditional asymptotic tests based on the marginal likelihood sometimes performs better than those based on the standard likelihood in

finite samples. Also, POI tests can have excellent finite sample properties and APOI tests perform well in a few cases. Rahman and King (1994) compared testing for random regression coefficients in the presence of autocorrelation, using marginal likelihood based LM tests, King and Wu's (1997) asymptotic LMMP (ALMMP) tests and APOI tests. They calculated small sample sizes and powers and found that marginal likelihood based tests are very competitive in terms of power with APOI tests. ALMMP and marginal likelihood based LM tests also appeared to have more desirable sizes than APOI tests. Rahman and King further concluded that the extra computation required to apply APOI tests hardly seems worthwhile. Particularly for large sample sizes, APOI tests do not seem to result in clear-cut improvements in either size or power. On the other hand, the marginal likelihood based LM test is reasonably reliable for large sample sizes. Rahman and King also noted that the marginal likelihood based ALMMP test is best in terms of average power but for data sets in which the component scores used in the test statistic are negatively correlated, the test performs poorly. In their test, it appeared that the lower than nominal sizes for middle values of the nuisance parameter, is a cause of the APOI tests not performing better than the other tests. This means that the APOI test may not be optimizing power at the chosen point under the alternative hypothesis. Thus, a more appropriate choice of APOI test could have changed Rahman and King's conclusion.

Andrews and Ploberger (1994) derived asymptotically optimal tests for testing problem when a nuisance parameter is present under the alternative hypothesis but not under the null hypothesis. They considered a nonstandard testing problem which was of interest because the classical asymptotic optimality properties of LM, Wald, and LR tests do not hold for nonstandard testing problems. They used the weighted average power criterion

(prior) to generate the optimal tests similar to that used by Wald (1943). Their study indicates that the LR test is not an optimal test using the weighting functions. However, it is asymptotically admissible (see Andrews and Ploberger (1995)). Andrews and Ploberger (1994) also introduced the average exponential tests. These tests are asymptotically optimal in the sense that they minimize the weighted average power for specific weight functions. Subsequently, Andrews et al. (1996) derived a class of finite sample optimal tests for one or more changepoints at unknown times in a multiple linear regression model. They considered a weighted average power by replacing a weight function over the number of changepoints. Andrews and Ploberger (1996) commented that the standard LR and sup LM tests beat any given test in terms of weighted average power against alternatives that are local to, but sufficiently distant from, the null. In this thesis, we propose a new test procedure for the problem of testing a composite null against a composite alternative hypothesis where we optimize a weighted average power function. This extends the of Andrews, Ploberger and others along with the third option (see p.19) suggested by Cox and Hinkley (1974, p.102).

2.6 Brief survey of non-nested testing

As Pesaran and Weeks (2001, p.289) noted,

From a statistical view point the main difference between the nested and nonnested hypothesis testing lies in the fact that the usual loglikelihood ratio or Wald statistics used in the conventional hypothesis testing are automatically centred at zero under the null when the hypotheses under consideration are nested while this is not true in the case of nonnested hypotheses. However, once the conventional test statistics are approximately centred (at least asymptotically) the same classical techniques can be applied to testing of non-nested hypotheses.

In this section we briefly survey the literature on non-nested hypotheses testing in econometrics and statistics. The seminal papers of Cox (1961, 1962) were the pioneering contribution to non-nested hypotheses testing in statistics. In these papers, three general approaches to non-nested testing problems are discussed. The modified (centred) log-likelihood ratio principle, hereafter known as the Cox-test was subsequently applied in linear econometric models by Pesaran (1974) and in simple non-linear equations by Pesaran and Deaton (1978), Evans and Deaton (1980) and Fisher and McAleer (1981). Unfortunately, some confusion has arisen in respect to the interpretation of these tests. There would appear to be several reasons for this, the distinction to be drawn between discrimination and significance testing. It is not clear which of these is being considered because the Cox-test may be used for either. Therefore, it is essential to have a clear understanding of the inferences that might be drawn from the Cox-test (for details, see Fisher and McAleer (1979)). Dastoor (1981) in commenting on Fisher and McAleer's (1979) paper, observed that the Cox procedure in general, certainly yields four possible conclusions and to suggest that nine conclusions might be possible could mislead the researcher.

The Cox-test has mainly been applied to linear and simple non-linear regression models. The reasons may be due to the complex computation of the numerator of the Cox statistic in non-regression circumstances. To overcome this problem, Pesaran and Pesaran (1993) proposed a new procedure for computing the Cox statistic for tests of non-nested hypotheses based on stochastic simulation which have been found works reasonably well even for a moderate number of replications.

Atkinson (1970) advocated the comprehensive model approach, whereby non-nested models are tested against an artificially constructed general model that includes the non-nested models as special cases. This approach is based on exponential weighting of the probability density functions underlying the respective hypotheses, see also Quandt (1974) and Davidson and MacKinnon (1981). But this approach has some important limitations because the testing framework is nested, the test under the null is non-standard due to the fact that under the null, the parameters of the alternative hypothesis disappear (see Pesaran and Weeks (2001)). This problem is known as Davies's problem (see Davies (1977)).

From the existing literature on non-nested hypothesis testing, it is evident that there are many other tests that under certain conditions are asymptotically equivalent to the Cox test. These asymptotic tests have been proposed and investigated by several authors in the context of non-nested, non-linear and multivariate regression models. Among these, the J and P tests proposed by Davidson and MacKinnon (1981, 1982) seem to be very popular because they are conceptually simpler and are very easy to implement using existing computer software. Moreover, the J test can be generalized relatively easily to cases where the regression function under the alternative hypothesis is non-parametric, see Delgado and Stengos (1994). Also, the J and P tests can be regarded as a way of implementing the artificial nesting (AN) procedure of Atkinson (1970) which solved the identification problem normally associated with such procedures. The JA test proposed by Fisher and McAleer (1981), the Cox-type N test proposed by Pesaran (1974) and adjusted Cox-type tests derived by Godfrey and Pesaran (1983) also seem to be popular in contrast to Deaton's (1982) F test which is less frequently used than either the J or

JA tests. This may be due to its poor power performances (see McAleer and Pesaran (1986)).

Michelis (1999) investigated the asymptotic null distributions of the J and Cox non-nested tests under a specific assumption of model orthogonality. He concluded that the J and Cox tests are no longer asymptotically equivalent as they are in the standard case of non-orthogonal models. The J and Cox tests tend to over-reject more severely when the correlation among the non-nested regressors is weakened (Godfrey and Pesaran, 1983). Also, Michelis's simulation results indicate that, with nearly orthogonal models, the J and Cox tests have excessive size distortion that increases with the number of non-nested regressors in the alternative model, and the size of the Cox test becomes substantially larger than the size of the J test. To overcome the poor performance of the Cox test, Godfrey and Pesaran proposed some small sample corrections and in particular the adjusted Cox-type \tilde{N} test which results in substantial improvement in its finite sample performance. Then the estimated sizes of the adjusted Cox-type \tilde{N} test are very close to the nominal size and powers are typically higher than those of the F test.

From the literature, it is evident that the Cox and Cox-type \tilde{N} tests of non-nested regression models have significantly higher than nominal sizes, particularly in finite samples. The JA and F tests have accurate sizes, however the JA test may be less powerful than the J test in certain situations, especially when two sets of regressors are nearly orthogonal (Davidson and MacKinnon, 1982, p. 563). There is a need for a better approximation to the finite sample distribution of the J test statistic to overcome the size distortion due to near orthogonality and to improve on the normal approximation based on asymptotic theory in small samples. Fan and Li (1995) proposed a bootstrap

version of the J and JA tests to improve on the normal approximation to the finite sample distributions. It is evident that the empirical sizes for both the bootstrap J and JA tests are very close to the nominal size and their power properties are good.

McAleer (1995) pointed out that there is a considerable body of evidence that the J and P tests and Cox-type \tilde{N} test can perform badly in small samples, particularly in the context of testing two non-nested linear regression models. Godfrey (1998, p. 60) noted that the J test has several useful features: it is easily generalized to allow for several non-nested alternative regression models; and it has considerable intuitive appeal. He argued that after adjustment of critical values, the J test might be more powerful than other procedures. He applied the bootstrapping method considered by Horowitz (1994) to the J test to reduce the problem of over rejection of true models. He also applied the same technique to the Cox-type test, JA test, F test and adjusted Cox-type \tilde{N} test of Godfrey and Pesaran and observed a substantial improvement in finite sample sizes when the errors are normally and independently distributed. His power results indicate that the J and adjusted Cox-type \tilde{N} tests are equally powerful, whereas the JA and F tests are less powerful. He considered applying the bootstrap method to non-nested multiple alternatives based upon an appropriate joint J test and recommended that bootstrap samples could be used to make more reliable inferences from diagnostic checks as well. Davidson and MacKinnon (1996) also successfully applied the bootstrap J -type test and Davidson and MacKinnon (2002) investigated its properties in the context of non-nested linear regression models. Their analysis implied that the J test performed very well when bootstrapped, except in certain extreme cases.

The encompassing principle of Hendry and Richard (1982), Mizon (1984), and Mizon and Richard (1986) is a leading approach to a non-nested testing framework; see also Gouriéroux et al. (1983). Mizon and Richard (1986) proposed two encompassing testing procedures: the Wald encompassing test and the score encompassing test; for a definition of encompassing, see Gouriéroux and Monfort (1995). Their analysis indicated that the conventional F test, as well as the one degree of freedom non-nested tests, has an encompassing interpretation and that the F test is a complete parametric encompassing test. They introduced the concept of implicit null hypotheses which enables alternative test statistics to be compared. They also mentioned that a complete parametric encompassing test is powerful for a wider range of alternatives than incomplete parametric encompassing tests. McAleer and Pesaran (1986) noted that encompassing tests are simple applications of Cox's principle. Encompassing tests are harder to implement because the test statistics are not always easy to evaluate. Chen and Kuan (2002) proposed a more operational encompassing test that is asymptotically equivalent to the Wald and score encompassing tests. Their proposed pseudo-true score encompassing test is relatively easy to evaluate. Their simulations indicated that this test has better finite sample performance than the J , JA and Cox tests. This test is not restricted to testing conditional mean specifications and hence extends the applicability of the conditional mean encompassing test of Wooldridge (1990) and the proposed test serves as a useful complement to existing non-nested tests. Recently, Ramalho and Smith (2002) investigated non-nested tests for competing moment condition models using a semi-parametric generalized empirical likelihood framework. They proposed Cox-type, moment encompassing and parametric encompassing non-nested tests. It is evident that some tests are most powerful, their empirical size characteristics oppose

their practical use. Also, their parametric encompassing statistics represent an important method for the assessment of models against specific non-nested alternatives.

Based on Cox's principle, Walker (1967) first introduced a test for testing non-nested time series models, specifically $AR(p)$ against $MA(q)$ processes, but his test was not popular because of its high computational cost particularly, when both p and q exceed one. Godfrey and Tremayne (1988), Smith and Tremayne (1990) and Burke et al. (1990) proposed and implemented pure significance (PS) tests for testing $AR(1)$ against $MA(1)$ errors in the linear regression model. King (1983a, 1987b), King and McAleer (1987) and Franses (1992) considered the problem of testing between AR and MA disturbance models for the linear regression model. McAleer et al. (1988), Hall and McAleer (1989), Godfrey and Tremayne (1992) studied the problem of testing between AR and MA time series models. McAleer et al. (1990) presented several straightforward procedures regarding the disturbances in a linear regression model for testing non-nested models of first or higher order autocorrelation processes which are alternative approaches to testing non-nested models with autocorrelated disturbances. They illustrated the test procedures with an application to US unemployment using annual time series data.

Silvapulle and King (1993) constructed a non-nested test of joint $AR(1)$ - $AR(4)$ disturbances against joint $MA(1)$ - $MA(4)$ disturbances in the linear regression model and illustrated the test with an application to quarterly time series data. They concluded with some discussion of the problem of testing general $AR(p)$ disturbances against $MA(q)$ disturbances. Baltagi and Li (1995) derived two extensions of the PS test proposed by Burke et al. (1990) and applied the first test to the problem of testing for $AR(1)$

disturbances against $MA(1)$ disturbances, and the second test for $MA(1)$ disturbances against $AR(1)$ disturbances in an error component model. It is evident from their Monte Carlo results that for testing $AR(1)$ disturbances against $MA(1)$ disturbances, the PS test performs well when the sample size, $n \geq 60$ and they do not recommend the test when the sample size is small or moderate and the number of individuals is large. However, the PS test performs well for testing $MA(1)$ disturbances against $AR(1)$ disturbances when the number of individuals is large and does not rely on sample size (n) to achieve its asymptotic distribution. McKenzie et al. (1999) developed some simple prediction tests for testing $AR(p)$ against $MA(q)$ errors and vice versa in the linear regression model. For the simpler case, testing for $AR(1)$ against $MA(1)$ errors, they compared the finite sample properties of their tests with those of Burke et al.'s PS test and the LM tests of $AR(1)$ against $ARMA(1,1)$ errors. The Monte Carlo results supported their prediction tests, the LM and PS tests perform worst.

2.7 Concluding remarks

This chapter reviewed the literature on optimal, PO and APO tests for composite hypothesis testing problems in the context of linear regression model and several potential problems were highlighted. We also focused on contributions to tests of non-nested hypotheses and discussed some popular non-nested tests. We briefly discussed nuisance parameters, invariance arguments and the marginal likelihood. In the literature, the solution for testing simple hypotheses is well established, therefore we focus our attention on composite hypothesis testing problems.

The literature surveyed in this chapter indicates that there are many asymptotic tests for composite non-nested testing but all of them are not equally efficient for diagnostic

testing in econometrics. Some limitations of these tests were also surveyed. Among the optimal tests, King's (1987a) PO tests seem to be popular for composite hypotheses testing problems involving regression disturbances. His PO tests can have excellent finite sample properties compared to existing popular non-nested tests. Unfortunately, the PO test cannot be constructed for all testing problems. For example, in one case Silvapulle and King (1991) could not find a PO test, so they recommended the use of an APO test. The APO test does not always seem to be reliable. Thus, it is important to find a general solution for testing a composite null against a composite alternative.

In this thesis, we propose a new approach, called the average power test which is based on the generalized NP lemma for testing composite hypotheses. Srikanthakumar (2000) proposed the g test: an APOI test for testing a composite null hypothesis. The generalized NP lemma provides a PO test for the problem of testing a finite number of representative densities against a single alternative. For the g test, if the number of representative densities is high then it is hard to find the critical values for which the size condition holds simultaneously. The new approach that we propose does not rely on the choice of representative densities but rather constructs the test with maximum average power while controlling average size over different subsets of the null hypothesis parameter space.

Chapter 3

A New Approach to Testing a Composite Null against a Composite Alternative

3.1 Introduction

As noted in Chapter 2, the theory of hypothesis testing is well developed in the case of testing a simple null hypothesis against a simple alternative hypothesis and a simple null hypothesis against a composite alternative hypothesis. We know in the former case that the NP lemma provides the most powerful test while in the latter case, the same lemma can be used to construct a test that is optimal (most powerful) at a chosen point under the alternative hypothesis parameter space. This test is known as a point optimal (PO) test. In hypothesis testing, much less developed is the problem of testing a composite null hypothesis against a composite alternative. There are a number of procedures that involve techniques for reducing the dimension of a composite null hypothesis, which in some cases can reduce a composite null hypothesis to a simple hypothesis. These include invariance arguments, which can be used to simplify the testing problem and eliminate some or all the nuisance parameters in the model under the null hypothesis. For example, see King and Hillier (1985). The theory of sufficient statistics and similar critical regions can also be used to reduce the complexity of the testing problem, see for example Cox and Hinkley (1974, p.135).

However, in cases where the null hypothesis cannot be reduced to a simple hypothesis, there is no clear approach for constructing a test with optimal small-sample power properties. King (1987a) reviewed the theory of point optimal invariant (POI) tests, which have been found to work well for some composite hypothesis testing problems in finite samples but he observed that his approach does not suit all testing problems and does not work in a number of cases. One example of a situation in which it does not work involves testing for first-order moving average (MA(1)) disturbances against first-order autoregressive (AR(1)) disturbances. Silvapulle and King (1991) could not find a POI test for this testing problem so they recommended the use of an approximately POI (APOI) test. In this regard, we can say that when testing a composite null against a composite alternative, the NP lemma does not provide a test which is PO. However, it is not clear that the APOI test is necessarily the best test in this case.

POI and APOI tests are based purely on the NP lemma, which requires a simple null hypothesis and a simple alternative. The generalized NP lemma allows for a composite null hypothesis that has a finite number of possible distributions under the null hypothesis. This does not suit composite hypotheses which involve nuisance parameters, because such parameters typically can take an infinite (uncountable) number of values.

The aim of this chapter is to explore a general solution to testing a composite null hypothesis against a composite alternative. It involves maximizing the average power of the test subject to average size being controlled over different subsets of the null hypothesis parameter space. The choice of maximizing average power comes from Cox and Hinkley's (1974, p.102) suggestion discussed in Chapter 2 and the recent work by Andrews and Ploberger (1994, 1995, 1996) in a different context. Working with average

sizes over a countable number of subsets of the null hypothesis parameter space allows us to use the generalized NP lemma to construct the optimal test. A composite null hypothesis brings with it the possibility that a test's size will vary with different parameter values under the null hypothesis. In such cases, the standard approach to test construction (see, for example, Lehmann and Stein (1948)) is to control the maximum size to be less than or equal to some desired level of significance. Typically this is extremely difficult and time consuming. One may not always be certain that a critical value, which controls size locally, also controls it globally. Controlling average size over subregions selected to reduce variability in size over the entire parameter space does seem to be an alternative worthy of investigation.

In this chapter we illustrate the new test procedure by showing how it can be applied to the problem of testing disturbances in the linear regression model.

The plan of this chapter is as follows. The theory and general testing procedure is discussed and outlined in Section 3.2. Section 3.3 covers the application of the test procedure to the problem of testing the disturbances in the linear regression model and provides solutions to a number of practical problems that arise when applying the test. Some concluding remarks are made in the final section.

3.2 Theory and the testing procedure

Let y be an observable $n \times 1$ vector and suppose we wish to test

$$H_0: y \text{ has density } f(y, \theta)$$

against

$H_a: y$ has density $g(y, \phi)$

where θ is a $j \times 1$ vector of unknown parameters restricted to the set Θ and ϕ is a $k \times 1$ vector of parameters restricted to the set Φ . This is a very general form of testing problem and it is assumed that any knowledge about the possible range of parameter values has been used to keep the parameter sets, Θ and Φ , as small as possible.

Our aim is to find the critical region $\omega \subset R^n$ for which we reject H_0 when $y \in \omega$ and which maximizes average power subject to controlling the average size (probability of a Type I error) of the test. We note that controlling average size over the entire parameter space under H_0 , namely Θ , may result in undesirably large sizes in some parts of Θ while smaller in other parts of Θ . In order to minimize this possibility, we allow for Θ to be partitioned into m disjoint subregions, namely $\Theta_1, \Theta_2, \dots, \Theta_m$ so that

$$\Theta = \Theta_1 \cup \Theta_2 \cup \dots \cup \Theta_m,$$

and we aim to control the average size over each of these subregions under H_0 .

Because H_0 is rejected when $y \in \omega$, $\int_{\omega} f(y, \theta) dy$ is the size of the test for a given θ value under H_0 . In order to define the average size over the subregion Θ_i , we need a weighting function, $p_{\theta_i}(\theta)$, defined for $\theta \in \Theta_i$, $i = 1, 2, \dots, m$. This function is rather like the density function of a prior distribution for θ defined over Θ_i . Then the average size of the test over the subregion Θ_i under H_0 is

$$\int_{\Theta_i} \int_{\omega} f(y, \theta) dy p_{\theta_i}(\theta) d\theta$$

$$= \int_{\omega} \int_{\Theta_i} f(y, \theta) p_{\theta_i}(\theta) d\theta dy$$

$$= \int_{\omega} f_i(y) dy, \quad i = 1, 2, \dots, m,$$

where

$$f_i(y) = \int_{\Theta_i} f(y, \theta) p_{\theta_i}(\theta) d\theta \quad (3.1)$$

is the weighted average over Θ_i of the density functions of y . It can also be thought of as an integrated likelihood function for $\theta \in \Theta_i$.

The power of the test for a given parameter vector ϕ under H_a is $\int_{\omega} g(y, \phi) dy$. In order to define the average power over Φ , we need a weighting function $p_{\phi}(\phi)$ defined over $\phi \in \Phi$. Again this can be the density function of a prior distribution for ϕ , although given its role in weighting powers, it is much better thought of as the density function of a weighting distribution. The average power function of the test is

$$\int_{\Phi} \int_{\omega} g(y, \phi) dy p_{\phi}(\phi) d\phi$$

$$= \int_{\omega} \int_{\Phi} g(y, \phi) p_{\phi}(\phi) d\phi dy$$

$$= \int_{\omega} f_{m+1}(y) dy,$$

where

$$f_{m+1}(y) = \int_{\Phi} g(y, \phi) p_{\phi}(\phi) d\phi \quad (3.2)$$

is the weighted average over Φ of the density functions of y under H_0 . It can also be thought of as an integrated likelihood function for H_0 .

Given $f_1(y), \dots, f_m(y), f_{m+1}(y)$ defined by (3.1) and (3.2), our aim is to find the critical region ω which maximizes average power, namely

$$\int_{\omega} f_{m+1}(y) dy \quad (3.3)$$

subject to controlling average size over the subregions $\Theta_i, i = 1, 2, \dots, m$, which can be written mathematically as

$$\int_{\omega} f_i(y) dy \leq c_i, \quad i = 1, 2, \dots, m, \quad (3.4)$$

where c_i is the desired maximum average size over the subregion Θ_i under H_0 . The solution to this problem is given by the generalized NP lemma (introduced in Section 2.3, also see Lehmann, 1986, p.96). The critical region defined by

$$\omega = \left\{ y: f_{m+1}(y) \geq \sum_{i=1}^m k_i f_i(y) \right\}, \quad (3.5)$$

provided that the constants k_1, k_2, \dots, k_m exist such that the inequalities (3.4) hold, maximizes average power. Therefore, under H_0 , we have m size conditions in order to solve for the m constants k_1, k_2, \dots, k_m . Note that the generalized NP lemma requires the constants, k_1, k_2, \dots, k_m , to all be positive real numbers. If one or more of the k_i is negative, then the test is optimal in the class of tests such that

$$\int_{\omega} f_i(y) dy = c_i, \quad i = 1, 2, \dots, m. \quad (3.6)$$

In summary, the maximum average power test proposed in this chapter for testing a composite null hypothesis against a composite alternative is based on rejecting H_0 for observed y such that

$$f_{m+1}(y) \geq \sum_{i=1}^m k_i f_i(y) \quad (3.7)$$

where k_1, k_2, \dots, k_m are constants (critical values) chosen such that the size conditions given by (3.4) or (3.6) are satisfied.

There are two special cases worthy of mention. Observe that when $m=1$, then (3.7) reduces to

$$f_2(y) \geq k_1 f_1(y)$$

which is equivalent to

$$\frac{f_2(y)}{f_1(y)} \geq k_1$$

in which the left-hand expression is a ratio of integrated likelihoods and k_1 is the critical value calculated to control the average size of the test. In the case of the null hypothesis being the simple hypothesis that y has density $f(y)$, then $m=1$, $f_1(y) = f(y)$, $f_{m+1}(y) = f_2(y)$ which is still defined by (3.2), and (3.4) becomes

$$\int_{\omega} f(y) dy \leq c_1. \quad (3.8)$$

In addition, (3.7) becomes

$$\frac{\int_{\phi} g(y, \phi) p_{\theta}(\phi) d\phi}{f(y)} \geq k_1 \quad (3.9)$$

where k_1 is such that (3.8) holds. The test is defined by (3.9) and can be thought of as being based on a weighted average of the likelihood ratios $g(y, \phi)/f(y, \theta)$.

3.3 Application to testing disturbances in the linear regression model

In order to apply the test given by (3.7), there are a number of practical issues that have to be considered such as the choice of m , construction of the integrated likelihoods $f_1(y), \dots, f_m(y), f_{m+1}(y)$ and the calculation of the critical values k_1, k_2, \dots, k_m . Some of these have specific solutions, which will best be left to the next chapter. In this section, we discuss some solutions in the context of testing disturbances in the linear regression model.

Consider the linear regression model

$$y = X\beta + u \quad (3.10)$$

where y is $n \times 1$, X is an $n \times k$ non-stochastic matrix of rank $k < n$ and β is a $k \times 1$ vector of parameters. Let the $n \times 1$ disturbance vector u be distributed as $N(0, \sigma^2 \Omega(\theta))$ under H_0 and $N(0, \sigma^2 \Sigma(\phi))$ under H_a , where $\Omega(\theta)$ and $\Sigma(\phi)$ are $n \times n$ matrix functions of the scalars θ and ϕ , respectively.

Our interest is in testing

$$H_0: u \sim N(0, \sigma^2 \Omega(\theta)), \quad l_1 \leq \theta \leq l_2, \quad (3.11)$$

against

$$H_a: u \sim N(0, \sigma^2 \Sigma(\phi)), \quad l_3 \leq \phi \leq l_4, \quad (3.12)$$

in the context of the linear regression model (3.10), where l_1 could be $-\infty$, and l_2 could be $+\infty$. Similarly, l_3 could be $-\infty$, and l_4 could be $+\infty$.

For this testing problem, β and σ^2 are nuisance parameters. Their influence can be removed through standard invariance arguments because the testing problem defined by (3.10), (3.11) and (3.12) is invariant to transformations of the form

$$y^* = \eta_0 y + X\eta, \quad (3.13)$$

where η_0 is a positive scalar and η is a $k \times 1$ vector. Clearly we wish to consider only tests which are invariant to transformations of y of the form of (3.13). With respect to these transformations, the $p \times 1$ vector

$$v = \frac{Pz}{(z'z)^{1/2}}$$

is a maximal invariant, where $p = n - k$, $z = My$ is the OLS residual vector from (3.10),

$M = I_n - X(X'X)^{-1}X'$ and P is any $p \times n$ matrix such that $P'P = M$ and $PP' = I_p$.

Because all statistics invariant to transformations of the form of (3.13) can be written as functions of v , for the purpose of our testing problem, we can treat v as the observed data.

The joint density function of v when $u \sim N(0, \sigma^2 \Lambda(\lambda))$ where $\Lambda(\lambda)$ is an $n \times n$ matrix function of λ , has been found by King (1979) to be

$$h(v) = \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |P\Lambda(\lambda)P'|^{-1/2} \left(v'(P\Lambda(\lambda)P')^{-1} v \right)^{-p/2} \quad (3.14)$$

with respect to the uniform measure on $\{v: v \in R^n, v'v = 1\}$. It is possible to show, as pointed out by Verbyla (1990), that

$$|P\Lambda(\lambda)P'| = |X'X|^{-1} |\Lambda(\lambda)| |X'\Lambda(\lambda)^{-1} X|$$

and from lemma 2 in King (1980) that

$$v'(P\Lambda(\lambda)P')^{-1} v = \frac{\tilde{z}(\lambda)' \tilde{z}(\lambda)}{z'z}$$

where $\tilde{z}(\lambda)$ is the OLS residual vector from the regression

$$\Lambda(\lambda)^{-1/2} y = \Lambda(\lambda)^{-1/2} X\beta + \Lambda(\lambda)^{-1/2} u \quad (3.15)$$

in which $\Lambda(\lambda)^{-1/2}$ is the inverse of $\Lambda(\lambda)^{1/2}$ which is defined by

$$\Lambda(\lambda)^{1/2} (\Lambda(\lambda)^{1/2})' = \Lambda(\lambda).$$

Using these results, our problem of testing (3.11) against (3.12) in the context of (3.10),

becomes one of testing

$$H_0: h_0(v, \theta) = \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{1/2} |\Omega(\theta)|^{-1/2} |X'\Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2},$$

$$l_1 \leq \theta \leq l_2, \quad (3.16)$$

against

$$H_a: h_a(v, \phi) = \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{1/2} |\Sigma(\phi)|^{-1/2} |X'\Sigma(\phi)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi)' \tilde{z}(\phi)}{z'z} \right)^{-p/2},$$

$$l_3 \leq \phi \leq l_4, \quad (3.17)$$

where $\tilde{z}(\theta)$ denotes $\tilde{z}(\lambda)$ with $\Lambda(\lambda) = \Omega(\theta)$ and $\tilde{z}(\phi)$ denotes $\tilde{z}(\lambda)$ with $\Lambda(\lambda) = \Sigma(\phi)$.

Observe that through invariance arguments, we have the respective parameter spaces as

$$\Theta = \{\theta: l_1 \leq \theta \leq l_2\} \text{ and } \Phi = \{\phi: l_3 \leq \phi \leq l_4\}.$$

First let us consider the test for which we control average size over the whole of Θ , i.e.,

$m = 1$. Denote the weighting functions for each hypothesis under consideration as

$p_{01}(\theta)$ and $p_a(\phi)$, respectively. The test which maximizes average power subject to

controlling average size over Θ , involves rejecting H_0 for

$$\begin{aligned} & \int_{l_1}^{l_2} \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{1/2} |\Sigma(\phi)|^{-1/2} |X'\Sigma(\phi)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi)' \tilde{z}(\phi)}{z'z} \right)^{-p/2} p_a(\phi) d\phi \\ & \geq c_a \int_{l_1}^{l_2} \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{1/2} |\Omega(\theta)|^{-1/2} |X'\Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta. \end{aligned}$$

After cancellation of some constants, this is equivalent to

$$\frac{\int_{l_3}^{l_4} |\Sigma(\phi)|^{-1/2} |X' \Sigma(\phi)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi)' \tilde{z}(\phi)}{z'z} \right)^{-p/2} p_a(\phi) d\phi}{\int_{l_1}^{l_2} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta} \geq c_a \quad (3.18)$$

where l_1, l_2, l_3 and l_4 all are appropriate boundaries of the parameter space which may be infinite and c_a is the critical value calculated to control average size.

If we denote the critical region given by (3.18) as $\omega(c_a)$, then finding c_a involves solving

$$\int_{\omega(c_a)} \int_{l_1}^{l_2} h_0(v, \theta) d\theta dv = \int_{l_1}^{l_2} \int_{\omega(c_a)} h_0(v, \theta) dv d\theta = \alpha \quad (3.19)$$

for c_a , where α is the desired average size or "significance level".

If one or more of l_i ($i = 1, 2, 3$ and 4) are infinite, there is an issue of truncation error to find the solution of (3.18) and (3.19). First consider the denominator of (3.18). Four cases then may arise. These are as follows.

Case I. If l_i ($i = 1, 2$) are all finite, then there is no issue of truncation error.

Case II. If $l_1 = -\infty$ and l_2 is finite, then the denominator of (3.18) is

$$\int_{-\infty}^{l_2} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta, \quad (3.20)$$

which is equivalent to

$$\begin{aligned} & \int_{-\infty}^{M_1} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta \\ & + \int_{M_1}^{l_2} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta, \end{aligned} \quad (3.21)$$

for which M_1 , the truncation point, is chosen so that

$$\int_{-\infty}^{M_1} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta < \varepsilon,$$

where ε is a small value known as the truncation error. Therefore (3.20) can be approximated by the second term of (3.21) and this approximation will result in a truncation error of less than ε .

Case III. If l_1 is finite and $l_2 = \infty$, then the denominator of (3.18) is

$$\int_{l_1}^{\infty} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta, \quad (3.22)$$

which is equivalent to

$$\begin{aligned} & \int_{l_1}^{M_2} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta \\ & + \int_{M_2}^{\infty} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z'z} \right)^{-p/2} p_{01}(\theta) d\theta, \end{aligned} \quad (3.23)$$

for which M_2 , the truncation point, is chosen so that

$$\int_{M_2}^{\infty} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z' z} \right)^{-p/2} p_{01}(\theta) d\theta < \varepsilon,$$

where ε is a small value being the maximum allowable truncation error.

Therefore (3.22) can be approximated by the first term of (3.23) and this approximation will result in a truncation error of less than ε .

Case IV. If $l_1 = -\infty$ and $l_2 = \infty$, then the denominator of (3.18) is

$$\int_{-\infty}^{\infty} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z' z} \right)^{-p/2} p_{01}(\theta) d\theta, \quad (3.24)$$

which is equivalent to

$$\begin{aligned} & \int_{-\infty}^{M_1} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z' z} \right)^{-p/2} p_{01}(\theta) d\theta \\ & + \int_{M_1}^{M_2} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z' z} \right)^{-p/2} p_{01}(\theta) d\theta \\ & + \int_{M_2}^{\infty} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z' z} \right)^{-p/2} p_{01}(\theta) d\theta, \end{aligned} \quad (3.25)$$

for which M_1 and M_2 are the truncation points chosen so that

$$\begin{aligned} & \left[\int_{-\infty}^{M_1} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z' z} \right)^{-p/2} p_{01}(\theta) d\theta \right. \\ & \left. + \int_{M_2}^{\infty} |\Omega(\theta)|^{-1/2} |X' \Omega(\theta)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta)' \tilde{z}(\theta)}{z' z} \right)^{-p/2} p_{01}(\theta) d\theta \right] < \varepsilon, \end{aligned}$$

where ε is a small value being the maximum allowable truncation error.

Therefore (3.24) can be approximated by the middle term of (3.25) and this approximation will result in a truncation error of less than ε .

Now if we consider the numerator of (3.18), similar cases will happen for l_3 and l_4 .

The consequence of this is that for the purpose of evaluating (3.18) and (3.19), l_1 , l_2 , l_3 and l_4 can be regarded as finite without any loss of generality.

Unfortunately, (3.18), and (3.19) both involve integrals, which do not have any analytical solutions. We therefore have to resort to numerical approximation methods. We can approximate the integrals in (3.18) by summing the areas under a large number, say q , of rectangles of width $(l_4 - l_3)/q$ for the numerator and $(l_2 - l_1)/q$ for the denominator whose height is equal to the value of the function being integrated at the midpoint of the base of the rectangle (for both integration with respect to θ and ϕ). The bases of the rectangles combined cover the interval of integration. This involves replacing θ and ϕ in the integrands by $l_1 + (2j-1)(l_2 - l_1)/(2q)$ and $l_3 + (2j-1)(l_4 - l_3)/(2q)$, where $j = 1, 2, \dots, q$, multiplying by $(l_2 - l_1)/q$ and $(l_4 - l_3)/q$ respectively and summing the areas of the rectangles.

Thus (3.18) can be approximated by (see Conte, 1965, p.120)

$$\frac{\sum_{j=1}^q |\Sigma(\phi_j)|^{-1/2} |X' \Sigma(\phi_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi_j)' \tilde{z}(\phi_j)}{z'z} \right)^{-p/2} (l_4 - l_3) p_a(\phi_j)}{\sum_{j=1}^q |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} (l_2 - l_1) p_{01}(\theta_j)} \geq c_\alpha \quad (3.26)$$

where $\theta_j = l_1 + (2j-1)(l_2 - l_1)/(2q)$ and $\phi_j = l_3 + (2j-1)(l_4 - l_3)/(2q)$, $j = 1, 2, \dots, q$. Observe that we can cancel $z'z$ from the numerator and denominator of the left-hand side (LHS) of (3.26). We retain these $z'z$ terms because they help with the numerical stability of the functions we are dealing with, particularly since they involve powers of the order $p/2$.

The LHS of (3.19) can be approximated using Monte Carlo integration in the following way. First draw a θ value from $p_{01}(\theta)$. Then generate an $n \times 1$ y vector via (3.10) and (3.11). Because of invariance, we can set $\beta = 0$ and $\sigma^2 = 1$. For the given value of c_α (the LHS of (3.19) is a function of c_α), conduct the test using (3.26). Repeat this a large number of times. The proportion of times H_0 is rejected in these simulations provides an estimate of the LHS of (3.19). This allows us to solve (3.19) for c_α via trial and error.

As discussed in Section 3.2, a potential problem with controlling the average size of our test is that sizes can be undesirably large in some parts of Θ while being smaller than desired in other parts of Θ . For example, if we divided Θ into three subregions ($m = 3$), the critical region (3.26) now becomes

$$\begin{aligned} & \sum_{j=1}^q |\Sigma(\phi_j)|^{-1/2} |X' \Sigma(\phi_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi_j)' \tilde{z}(\phi_j)}{z'z} \right)^{-p/2} p_a(\phi_j) \\ & \geq \left[k_1 \sum_{j=1}^{q_1} |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{01}(\theta_j) \right. \\ & \quad + k_2 \sum_{j=q_1+1}^{q_2} |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{02}(\theta_j) \\ & \quad \left. + k_3 \sum_{j=q_2+1}^q |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{03}(\theta_j) \right] \quad (3.27) \end{aligned}$$

where k_1 , k_2 and k_3 are critical values calculated to control average size over the three subregions of Θ and the values of q_1 and q_2 are determined by the boundaries of these subregions. The terms $(l_4 - l_3)/q$ and $(l_2 - l_1)/q$ are absorbed into the constants k_1 , k_2 and k_3 .

Similarly, if we divided Θ into m subregions with inner boundary points $\theta_1, \theta_2, \dots, \theta_{m-1}$, the boundary points for the purpose of summation, q_1, q_2, \dots, q_{m-1} , are determined by rounding $((\theta_1 - l_1)q/(l_2 - l_1) + (1/2))$, $((\theta_2 - l_1)q/(l_2 - l_1) + (1/2))$, \dots , $((\theta_{m-1} - l_1)q/(l_2 - l_1) + (1/2))$, respectively, down to the nearest integer, then the critical region becomes

$$\begin{aligned}
& \sum_{j=1}^q |\Sigma(\phi_j)|^{-1/2} |X' \Sigma(\phi_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi_j)' \tilde{z}(\phi_j)}{z'z} \right)^{-p/2} p_a(\phi_j) \\
& \geq \left[k_1 \sum_{j=1}^{q_1} |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{01}(\theta_j) \right. \\
& \quad + k_2 \sum_{j=q_1+1}^{q_2} |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{02}(\theta_j) \\
& \quad \left. \dots + k_m \sum_{j=q_{m-1}+1}^q |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{0m}(\theta_j) \right] \quad (3.28)
\end{aligned}$$

where k_1, k_2, \dots, k_m are critical values calculated to control average size over the m subregions of Θ .

Returning to the case of $m=3$ for illustration, if we denote the critical region given by (3.27) as $\omega(k_1, k_2, k_3)$ and if $\Theta_1 = \{\theta: \theta_0 \leq \theta \leq \theta_1\}$, $\Theta_2 = \{\theta: \theta_1 < \theta \leq \theta_2\}$ and $\Theta_3 = \{\theta: \theta_2 < \theta \leq \theta_3\}$ where $\theta_0 = l_1$, $\theta_3 = l_2$ and θ_1 and θ_2 are the chosen boundaries of the subregions of Θ , then finding k_1 , k_2 and k_3 involves solving

$$\int_{\Theta_i} \int_{\omega(k_1, k_2, k_3)} h(v, \theta) dv d\theta = \alpha, \quad i = 1, 2, 3, \quad (3.29)$$

jointly for k_1 , k_2 and k_3 . Note that q_1 and q_2 in (3.27) are determined by rounding $((\theta_1 - l_1)q/(l_2 - l_1) + (1/2))$ and $((\theta_2 - l_1)q/(l_2 - l_1) + (1/2))$, respectively, down to the

nearest integer. Again the LHS of (3.29), for $i=1, 2, 3$, can be approximated using Monte Carlo integration as follows. For the given value of i , first draw a θ value from $p_{0i}(\theta)$, then generate y via (3.10) and (3.11) with $\beta=0$ and $\sigma^2=1$. For the given values of k_1 , k_2 and k_3 , conduct the test for this y vector using (3.27). Repeat this a large number of times. The proportion of times H_0 is rejected provides an estimate of the LHS of (3.29).

Solving the three equations given by (3.29) for k_1 , k_2 , k_3 can be difficult and time consuming. In the experiments reported in the next chapter, we used the following steps to solve for k_1 , k_2 and k_3 .

Step 1. Fix the values of k_2 and k_3 to some reasonable initial values. Solve (3.29)

with $i=1$ for k_1 . This can be done by rewriting (3.27) as

$$\begin{aligned}
& \left[\sum_{j=1}^q |\Sigma(\phi_j)|^{-1/2} |X' \Sigma(\phi_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi_j)' \tilde{z}(\phi_j)}{z'z} \right)^{-p/2} p_a(\phi_j) \right. \\
& \quad - k_2 \sum_{j=q_1+1}^{q_2} |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{02}(\theta_j) \\
& \quad \left. - k_3 \sum_{j=q_2+1}^q |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{03}(\theta_j) \right]
\end{aligned}$$

$$\left/ \sum_{j=1}^{q_1} |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{01}(\theta_j) \geq k_1 \right. \quad (3.30)$$

repeatedly drawing θ values to generate y from $p_{01}(\theta)$, calculating the LHS of (3.30) and then sorting these values in ascending order so as to find the $100(1-\alpha)$ percentile value which is k_1 .

Step 2. Taking these values of k_1, k_2 and k_3 , calculate the LHS of (3.29) for $i=2$.

Let this value be denoted by s_2 . If $s_2 = \alpha$, go to step 3. If $s_2 > \alpha$, make k_2 bigger otherwise make k_2 smaller with the aim of getting s_2 close to α and repeat steps 1 and 2.

Step 3. Taking the values of k_1, k_2 and k_3 used in step 2, calculate the LHS of (3.29)

for $i=3$. Denote this value by s_3 . If $s_3 = \alpha$, we have found the required values of k_1, k_2 and k_3 . If $s_3 > \alpha$, make k_3 bigger otherwise make k_3 smaller with the aim of getting s_3 close to α and repeat steps 1, 2 and 3.

The above algorithm applies for the case of $m=3$, it can easily be generalized for general m . If we denote the critical region given by (3.28) as $\omega(k_1, k_2, \dots, k_m)$ and if $\Theta_i = \{\theta: \theta_{i-1} \leq \theta \leq \theta_i\}$, $i=1, 2, \dots, m$, where $\theta_0 = l_1$, $\theta_m = l_2$ and $\theta_1, \theta_2, \dots, \theta_{m-1}$ are our chosen boundaries of our subregions of Θ , then finding k_1, k_2, \dots, k_m involves solving

$$\int_{\Theta_i} \int_{\omega(k_1, k_2, \dots, k_m)} h(v, \theta) dv d\theta = \alpha, \quad i=1, 2, \dots, m, \quad (3.31)$$

jointly for k_1, k_2, \dots, k_m . The integers q_1, q_2, \dots, q_{m-1} given in (3.28) are defined above. Again the LHS of (3.31), for $i=1, 2, \dots, m$, can be approximated using Monte Carlo integration as follows. For the given value of i , first draw a θ value from $p_{0i}(\theta)$, generate y via (3.10) and (3.11) with $\beta=0$ and $\sigma^2=1$. For the given values of k_1, k_2, \dots, k_m , conduct the test for this y vector using (3.28). Repeat this a large number of times. The proportion of times H_0 is rejected provides an estimate of the LHS of (3.31).

Solving the m equations given by (3.31) for k_1, k_2, \dots, k_m is really difficult and time consuming. We can use the following m steps to solve for k_1, k_2, \dots, k_m .

Step 1. Fix the values of k_2, k_3, \dots, k_m to some reasonable initial values. Solve (3.31) with $i=1$ for k_1 . This can be done by rewriting (3.28) as

$$\left[\sum_{j=1}^q |\Sigma(\phi_j)|^{-1/2} |X' \Sigma(\phi_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\phi_j)' \tilde{z}(\phi_j)}{z'z} \right)^{-p/2} p_a(\phi_j) \right. \\ \left. - k_2 \sum_{j=q_1+1}^{q_2} |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{02}(\theta_j) \right. \\ \left. \dots - k_m \sum_{j=q_{m-1}+1}^q |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\theta_j)' \tilde{z}(\theta_j)}{z'z} \right)^{-p/2} p_{0m}(\theta_j) \right]$$

$$\left/ \sum_{j=1}^n |\Omega(\theta_j)|^{-1/2} |X' \Omega(\theta_j)^{-1} X|^{-1/2} \left(\frac{\bar{z}(\theta_j)' \bar{z}(\theta_j)}{z' z} \right)^{-p/2} \right. p_{01}(\theta_j) \geq k_1 \quad (3.32)$$

repeatedly drawing θ values to generate y vectors from $p_{01}(\theta)$, calculating the LHS of (3.32) and then sorting these values in ascending order so as to find the $100(1-\alpha)$ percentile value which is k_1 .

Step 2. Taking these values of k_1, k_2, \dots, k_m , calculate the LHS of (3.31) for $i=2$.

Let this value be denoted by s_2 . If $s_2 = \alpha$, go to step 3. If $s_2 > \alpha$, make k_2 bigger otherwise make k_2 smaller in an attempt to get s_2 close to α and repeat steps 1 and 2.

Step 3. Taking the values of k_1, k_2, \dots, k_m used in step 2, calculate the LHS of (3.31)

for $i=3$. Denote this value by s_3 . If $s_3 = \alpha$, go to step 4. If $s_3 > \alpha$, make k_3 bigger otherwise make k_3 smaller with the aim of getting s_3 close to α and repeat steps 1, 2 and 3.

Step 4. Taking the values of k_1, k_2, \dots, k_m used in step 3, calculate the LHS of (3.31)

for $i=4$. Denote this value by s_4 . If $s_4 = \alpha$, go to step 5. If $s_4 > \alpha$, make k_4 bigger otherwise make k_4 smaller and repeat steps 1, 2, 3 and 4.

⋮

Step j . Taking the values of k_1, k_2, \dots, k_m used in step $(j-1)$, calculate the LHS of

(3.31) for $i=j$. Denote this value by s_j . If $s_j = \alpha$, go to step $(j+1)$. If

$s_j > \alpha$, make k_j bigger otherwise make k_j smaller and repeat steps 1, 2, ..., j .

⋮

Step m . Taking the values of k_1, k_2, \dots, k_m used in step $(m-1)$, calculate the LHS of

(3.31) for $i=m$. Denote this value by s_m . If $s_m = \alpha$, we have found the

required values of k_1, k_2, \dots, k_m . If $s_m > \alpha$, make k_m bigger otherwise make

k_m smaller and repeat steps 1, 2, ..., m .

Here we see that the amount of computation involved does increase exponentially with m . It is best, therefore, that m be kept as small as possible.

3.4 Concluding remarks

In this chapter we have presented a new approach to test construction when testing a composite null hypothesis against a composite alternative based on using the generalized NP lemma to maximize average power subject to controlling average size over different subsets of the null hypothesis parameter space. We discussed the new test procedure by applying it to the problem of testing disturbance errors in the linear regression model. This problem was deliberately chosen because, through invariance arguments, it reduces to composite null and alternative hypotheses involving scalar parameters. It is also a testing problem for which in some cases it has been difficult to find a satisfactory test. The new test procedure has considerable potential, particularly for problems in which the effective null hypothesis parameter space can be kept small.

In the next chapter we illustrate the new test procedure by applying it to the problem Silvapulle and King (1991) could not find a POI test for and were forced to recommend the use of an APOI test for. This problem is testing for MA(1) disturbances against AR(1) disturbances in the linear regression model. A small sample power comparison reveals that the new test can be substantially more powerful than various versions of the APOI test recommended by Silvapulle and King. Following the next two chapters, we will investigate how well the procedure works for other types of parameter spaces under the null hypothesis.

Chapter 4

Application of the New Approach to a Single Parameter Problem

4.1 Introduction

In the previous chapter, we introduced a new approach to testing a composite null hypothesis against a composite alternative. We outlined the theory behind this approach with particular emphasis on testing the disturbances in the linear regression model. In this testing problem, we are interested in a particular parameter or a set of parameters. Clearly the remaining parameters are nuisance parameters which it would be nice to be able to eliminate. We discussed in Section 3.3 the standard invariance technique that can be used to simplify the testing problem and eliminate the nuisance parameters involved in the regression model under consideration.

There are a number of practical issues that were not explored in Chapter 3. These include: (a) Does controlling average size do a good job in controlling size overall? (b) How should m be chosen? (c) How easy is it to apply this test procedure? (d) Does this test procedure perform well? (e) If so, how well will it perform in comparison with other tests? These are some of the questions this chapter seeks to answer.

We will address these questions by considering the application of the new test procedure to testing for MA(1) disturbances against AR(1) disturbances in the linear regression model.

Nicholls, Pagan, and Terrel (1975) ascribed the general lack of interest in models with MA disturbances to the computational difficulty involved in estimating its regression parameters. This problem has now been overcome by the arrival of powerful personal computers. King (1983b) considered a point optimal procedure for testing MA(1) disturbances in the linear regression model. The reverse problem of testing AR(1) against MA(1) disturbances has been extensively studied in the time series framework, see for example, Walker (1967), King (1983a), King (1987a), King and McAleer (1987), Burke et al. (1990) and Baltagi and Li (1995).

King (1983a) constructed and investigated the properties of a test for testing AR(1) against MA(1) disturbances, called a pseudo POI (PPOI) test, by applying the NP lemma to a maximal invariant statistic. In his paper, King mentioned that wrongly correcting for AR(1) disturbances in a model with MA(1) errors can lead to inefficient parameter estimates and, more importantly, misleading inferences. Also, he pointed out that the confirmation or rejection of a particular economic theory could plausibly depend upon whether a certain set of regression disturbances are AR(1) or MA(1), see for example, Rowley and Wilton (1973). King and McAleer (1987) further compared the small sample properties of the Cox test, some linearized Cox tests, and an APOI test, as well as the LM test of AR(1) against ARMA(1,1) disturbances in the linear regression model. Their comparison showed King's test to have the best power properties. King (1987a)

discussed an application to the non-nested problem of testing for AR(1) disturbances against MA(1) disturbances in the linear regression model and calculated powers for a POI test procedure. Unfortunately the POI test cannot be constructed for all testing problems. For example, Silvapulle and King (1991) could not find a POI test for testing MA(1) against AR(1) disturbances in the linear regression model, so they recommended the use of an APOI test. In this chapter we will evaluate the small sample properties of our new test procedure in the context of the problem of testing MA(1) against AR(1) disturbances in the linear regression model, and compare these properties with those of Silvapulle and King's test.

The plan of this chapter is as follows. The model, testing problem and the test procedure are discussed and outlined in Section 4.2 which covers the application of the test to the problem of testing for MA(1) disturbances against AR(1) disturbances in the linear regression model. A Monte Carlo experiment for exploring the properties of the new test procedure is outlined in Section 4.3. Sections 4.3.1, 4.3.2, 4.3.3 and 4.3.4 cover the experimental design and the size properties of the new test when $m=1$, $m=2$ and $m=3$, respectively. A comparison of the small sample size and power properties of the new test and various versions of the APOI test recommended by Silvapulle and King (1991) is reported in Section 4.4. It demonstrates the potential of the new test procedure. Some concluding remarks are made in the final section.

4.2 The model, testing problem and the test procedure

Consider the linear regression model (3.10)

$$y = X\beta + u.$$

If the components of the $n \times 1$ disturbance vector u are generated by the MA(1) process

$$u_t = e_t + \gamma e_{t-1}, \quad t = 1, 2, \dots, n, \quad (4.1)$$

where $e^* = (e_0, e_1, \dots, e_n)' \sim N(0, \sigma^2 I_{n+1})$, then $u \sim N(0, \sigma^2 \Sigma(\gamma))$, where $\Sigma(\gamma)$ is the tridiagonal matrix with $1 + \gamma^2$ as the main diagonal elements and γ as the non-zero off-diagonal elements. If the components of u are generated by the stationary AR(1) process

$$u_t = \rho u_{t-1} + e_t, \quad |\rho| < 1, \quad t = 1, 2, \dots, n, \quad (4.2)$$

where $u_0 \sim N(0, \sigma^2 / (1 - \rho^2))$ and $e = (e_1, e_2, \dots, e_n)' \sim N(0, \sigma^2 I_n)$, then

$u \sim N(0, \sigma^2 \Omega(\rho))$ in which $\Omega(\rho)$ is an $n \times n$ matrix whose $(i, j)^{\text{th}}$ element is $\rho^{|j-i|} / (1 - \rho^2)$.

Our interest is in testing

$$H_0: u \sim N(0, \sigma^2 \Sigma(\gamma)), \quad 0 \leq \gamma \leq 1, \quad (4.3)$$

against

$$H_a: u \sim N(0, \sigma^2 \Omega(\rho)), \quad 0 < \rho < 1, \quad (4.4)$$

in the context of the linear regression model (3.10). Observe that we are assuming non-negative autocorrelation in the disturbances. The discussion below can easily be generalized to the case of negative autocorrelation. Typically when one is considering

disturbance models of the form of (4.1) and (4.2), one is also willing to make an assumption about whether consecutive disturbances are positively or negatively correlated. As noted by Silvapulle and King (1991), if available, this knowledge should be incorporated into the test in order to improve power.

As we mentioned in the previous chapter, the parameters β and σ^2 are nuisance parameters for our testing problem. Their influence can be removed through standard invariance arguments because the testing problem defined by (3.10), (4.3) and (4.4) is invariant to transformations of the form (3.13).

Now using the results of (3.16) and (3.17), our problem of testing (4.3) against (4.4) in the context of (3.10), becomes one of testing

$$H_0: h_0(v, \gamma) = \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{1/2} |\Sigma(\gamma)|^{-1/2} |X'\Sigma(\gamma)^{-1}X|^{-1/2} \left(\frac{\tilde{z}(\gamma)' \tilde{z}(\gamma)}{z'z} \right)^{-p/2},$$

$$0 \leq \gamma \leq 1, \quad (4.5)$$

against

$$H_a: h_a(v, \rho) = \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{1/2} |\Omega(\rho)|^{-1/2} |X'\Omega(\rho)^{-1}X|^{-1/2} \left(\frac{\tilde{z}(\rho)' \tilde{z}(\rho)}{z'z} \right)^{-p/2},$$

$$0 < \rho < 1, \quad (4.6)$$

where $\tilde{z}(\gamma)$ is the OLS residual vector from the regression defined in (3.15), $\tilde{z}(\gamma)$

denotes $\tilde{z}(\lambda)$ with $\Lambda(\lambda) = \Sigma(\gamma)$ and also, $\tilde{z}(\rho)$ denotes $\tilde{z}(\lambda)$ with $\Lambda(\lambda) = \Omega(\rho)$. Observe that through invariance arguments and our assumption of positive autocorrelation, we have restricted the respective parameter spaces to $\Theta = \{\gamma: 0 \leq \gamma \leq 1\}$ and $\Phi = \{\rho: 0 < \rho < 1\}$.

First let us consider the test for which we control average size over the entire parameter space of Θ , i.e., $m=1$. Now the fundamental NP lemma implies that the test which maximizes average power subject to controlling average size over the entire null hypothesis parameter space, can be based on the critical region of the form (after cancellation of some constants)

$$\begin{aligned} & \int_0^1 |\Omega(\rho)|^{-1/2} |X' \Omega(\rho)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho)' \tilde{z}(\rho)}{z' z} \right)^{-p/2} p(\rho) d\rho \\ & \geq c_\alpha \int_0^1 |\Sigma(\gamma)|^{-1/2} |X' \Sigma(\gamma)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma)' \tilde{z}(\gamma)}{z' z} \right)^{-p/2} p(\gamma) d\gamma \\ \text{or, } & \frac{\int_0^1 |\Omega(\rho)|^{-1/2} |X' \Omega(\rho)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho)' \tilde{z}(\rho)}{z' z} \right)^{-p/2} d\rho}{\int_0^1 |\Sigma(\gamma)|^{-1/2} |X' \Sigma(\gamma)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma)' \tilde{z}(\gamma)}{z' z} \right)^{-p/2} d\gamma} \geq c_\alpha \end{aligned} \quad (4.7)$$

where $p(\rho)$ and $p(\gamma)$ are the weighting functions of ρ and γ respectively, and c_α is the critical value calculated to control average size. The uniform distribution would appear to be a good choice of weighting function for ρ and γ because both ρ and γ

are restricted to lie in an interval. In equation (4.7), $p(\rho)$ and $p(\gamma)$ have been replaced by uniform density functions which cancel.

If we denote the critical region given by (4.7) as $\omega(c_\alpha)$, then using equation (3.19), finding c_α involves solving

$$\int_{\omega(c_\alpha)} \int_0^1 h_0(v, \gamma) d\gamma dv = \int_0^1 \int_{\omega(c_\alpha)} h_0(v, \gamma) v dv d\gamma = \alpha \quad (4.8)$$

for c_α , where α is the desired average size of the test.

Equations (4.7) and (4.8) both involve integrals, which do not have analytical solutions. Therefore, as discussed in Chapter 3, we can approximate the integrals in (4.7) by summing the areas under a large number, say q , of rectangles of width $1/q$. The bases of the rectangles combined cover the unit interval. This involves replacing ρ and γ in the integrands by $(2j-1)/(2q)$, where j is a positive integer, namely $j=1, 2, \dots, q$, multiplying by $1/q$ and summing the areas of the rectangles. Thus (4.7) can be approximated by (see Conte, 1965, p.120)

$$\frac{\sum_{j=1}^q |\Omega(\rho_j)|^{-1/2} |X' \Omega(\rho_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_j)' \tilde{z}(\rho_j)}{z' z} \right)^{-p/2}}{\sum_{j=1}^q |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z' z} \right)^{-p/2}} \geq c_\alpha \quad (4.9)$$

where $\rho_j = \gamma_j = (2j-1)/(2q)$, $j=1, 2, \dots, q$.

The LHS of (4.8) can be approximated using Monte Carlo integration by repeatedly drawing a γ value from the uniform distribution on $[0,1]$, then generating an $n \times 1$ y vector via (3.10) and (4.1). Because of invariance, we set $\beta = 0$ and $\sigma^2 = 1$. To calculate the critical value c_α , again we repeatedly draw γ values from the uniform distribution on $[0,1]$ to generate y vectors, each time calculating the LHS of (4.9) and then sorting these values in ascending order so as to find the $(1-\alpha)100$ percentile value at the α percent level. This is our estimated value of c_α .

As discussed in Section 3.2 of Chapter 3, a potential problem is with controlling the average size of our test. The sizes can be undesirably large in some parts of the null hypothesis parameter space while being smaller than the desired size in other parts of the null hypothesis parameter space. We conducted some simulation experiments, for $m = 1$, $m = 2$ and $m = 3$ for testing (4.3) against (4.4) in the context of (3.10). In this testing problem when $m = 1$ and $m = 2$, we did indeed find regions of larger sizes than α and regions of lower sizes than α under the null hypothesis. For this testing problem, finally, we divided the null hypothesis parameter space Θ into three disjoint intervals ($m = 3$), determined by boundaries where the size in the case of $m = 1$ as a function of γ crosses $\alpha = 0.05$ for the test given by (4.9). When this was done we saw that sizes of the resultant test are reasonably acceptable.

A detailed discussion of the test procedures and the size properties of the test with $m = 1$, $m = 2$ and $m = 3$ are given below.

4.3 Monte Carlo experiments

In order to investigate the small sample size and power properties of our test in the context of testing MA(1) against AR(1) disturbances in the linear regression model, we conducted a Monte Carlo experiment. The experiment was also carried out to compare the small sample size and power properties of our new test and three versions of the APOI test recommended by Silvapulle and King (1991).

4.3.1 Experimental design

Two thousand replications were used to investigate the size and power properties of our test. The critical values were calculated at the five percent level. For two thousand replications, estimated rejection probabilities outside the range $[0.037, 0.063]$ are significantly different from the five percent level at the 0.01 level.

In order to compare the size and power properties of our test with different versions of Silvapulle and King's (1991) APOI tests, we used their experimental design matrices (also see Silvapulle, 1991) which are as follows:

- X1: ($n \times 3$; $n = 20, 60$). The first n observations of Durbin and Watson's (1951, p.160) consumption of spirits example; that is, a constant, plus annual data on the price of spirits and household income in the UK.
- X2: ($n \times 3$; $n = 20, 60$). A constant dummy, the quarterly Australian Consumer Price Index (CPI) commencing 1959(1) and the same index lagged one quarter.

- $X3$: ($n \times 5$; $n = 20, 60$). $X2$ augmented by adding the CPI lagged two quarters and three quarters as additional regressors.
- $X4$: ($n \times 3$; $n = 20, 30$). The regressors are a constant, logarithms of Chow's (1957, Table 1, p.32) automobile stock per capita and logarithm of personal money stock per capita variables for the United States 1921-1950.
- $X5$: ($n \times 3$; $n = 20, 60$). The regressors are the eigenvectors corresponding to the three smallest eigenvalues of the $n \times n$ Durbin-Watson (DW) differencing matrix A_1 , which is the tridiagonal matrix, whose main diagonal is $(1, 2, 2, \dots, 2, 1)$ and whose leading diagonals are $(-1, -1, \dots, -1)$. We note that the first regressor is a constant.

All the design matrices cover a variety of economic phenomena. The design matrix $X1$ which includes a constant dummy, has smoothly evolving regressors with no hint of seasonality, $X2$ and $X3$ exhibit practical degrees of multicollinearity. The matrix $X5$ was included because the DW test is approximately uniformly most powerful invariant for this matrix. The sample sizes for all the design matrices are 20 and 60 except for $X4$ where the sample sizes are 20 and 30. The $X3$ matrix was included to see if the number of regressors affects results.

For the problem under consideration, the proposed test under H_0 and H_a is invariant with respect to β and σ^2 , and we have chosen, $\beta = 0$ and $\sigma^2 = 1$ for the simulation experiment.

The first step involved deciding on the value of m for which the size of our test is approximately controlled. We mentioned earlier that we conducted some simulation experiments to choose the value of m and we tried $m = 1$, $m = 2$ and $m = 3$. Finally we saw that the test performs well for $m = 3$. Detailed discussions follow.

4.3.2 Size properties of the test with $m = 1$

Using 2000 replications with $m = 1$, we calculated the critical values c_α for all the design matrices based on (4.7). These values are presented in Table 4.1. For these critical values, using (4.9) we calculated the level (sizes) of the test for different values of γ , namely, $\gamma = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9 under H_0 . These estimated sizes are presented in Table 4.2. From this table we see that the estimated sizes are far from the average size ($\alpha = 0.05$), some sizes are undesirably larger than the average size while others are much smaller than the average size. We do not view this as successfully controlling the probabilities of Type I errors across the null hypothesis parameter space because most of the estimated sizes are outside the rejection probabilities range $[0.037, 0.063]$ and are therefore significantly different from five percent at the 0.01 level. This led us to consider $m = 2$. In the following section we discuss the size properties of the test for $m = 2$.

4.3.3 Size properties of the test with $m = 2$

We wish to control the average estimated sizes of our test over suitable regions under H_0 . When we controlled the average estimated size of our test over the whole parameter space ($m = 1$) under H_0 , naturally we got some sizes that are above average and some

that are below average. Our suggestion is to split the sample space under H_0 into two regions and control average size over each of these regions. An obvious approach is to split into those regions for which size was above average when $m = 1$ and those regions for which size was below average. In the former region, this should have the desirable effect of lowering sizes when they are previously above average and for the latter, this should raise estimated sizes that have previously been below average. With this in mind and using the calculated sizes for $m = 1$ presented in Table 4.2, we divided the null hypothesis parameter space ($0 \leq \gamma \leq 1$) into two disjoint intervals, with the boundary points calculated through simple interpolation. For example, in the case of the $X1$ design matrix with $n = 20$, through interpolation we estimated that smaller than average sizes lie in the ranges $0 \leq \gamma \leq 0.24$ and $0.64 < \gamma \leq 1$, and therefore larger than average sizes lie in the interval $0.24 < \gamma \leq 0.64$; see Table 4.2. In this case, our aim is to control average size over these two regions in the hope that this gives a test with better size properties. Ranges of the two disjoint intervals and the corresponding values of j , where $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, 50$, for all design matrices are presented in Table 4.3.

Thus for $m = 2$, equation (3.7) gives the test that maximizes average power subject to two size conditions and involves a critical region of the form

$$\sum_{j=1}^q |\Omega(\rho_j)|^{-1/2} |X' \Omega(\rho_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_j)' \tilde{z}(\rho_j)}{z'z} \right)^{-p/2}$$

$$\begin{aligned} &\geq k_1 \left[\sum_{j=1}^{q_1} |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right. \\ &\quad \left. + \sum_{j=q_2+1}^q |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right] \\ &\quad + k_2 \sum_{j=q_1+1}^{q_2} |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2}, \end{aligned} \quad (4.10)$$

where k_1 and k_2 are critical values calculated to control average size over the two subregions of Θ and the values of q_1 and q_2 are determined by the boundaries of the subregions.

If we denote the critical region given by (4.10) as $\omega(k_1, k_2)$ and if $\Theta_1 = \{\gamma: \gamma_0 \leq \gamma \leq \gamma_1\} \cup \{\gamma: \gamma_2 < \gamma \leq \gamma_3\}$ and $\Theta_2 = \{\gamma: \gamma_1 < \gamma \leq \gamma_2\}$, where $\gamma_0 = 0$, $\gamma_3 = 1$, γ_1 and γ_2 are the chosen boundaries of the subregions of Θ , then finding k_1 and k_2 involves solving

$$\int_{\Theta_i} \int_{\omega(k_1, k_2)} h_0(v, \gamma) dv d\gamma = \alpha, \quad i = 1, 2, \quad (4.11)$$

jointly for k_1 and k_2 . Here q_1 and q_2 in (4.10) are determined by rounding $(\gamma_1 q + (1/2))$ and $(\gamma_2 q + (1/2))$ respectively, down to the nearest integer.

The LHS of (4.11) can be approximated using Monte Carlo integration in the following

way. First, fix the value of k_2 to some reasonable initial value. Solve (4.11) with $i = 1$ for k_1 . Rewriting (4.10) as

$$\left[\sum_{j=1}^{q_1} |\Omega(\rho_j)|^{-1/2} |X' \Omega(\rho_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_j)' \tilde{z}(\rho_j)}{z'z} \right)^{-p/2} \right. \\ \left. - k_2 \sum_{j=q_1+1}^{q_2} |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right] \\ \sqrt{\left[\sum_{j=1}^{q_1} |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right.} \\ \left. + \sum_{j=q_2+1}^q |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right]} \geq k_1 \quad (4.12)$$

can help do this. It involves repeatedly drawing γ values from the uniform distribution over $[0, 1]$. If such a γ value belongs to Θ_1 , then generate the y vectors via (3.10) and (4.1) with $\beta = 0$ and $\sigma^2 = 1$. Otherwise, again we repeat this step by drawing another γ value from the uniform distribution. The next step involves calculating the LHS of (4.12), and then sorting these values in ascending order so as to find the ninety-fifth percentile value at the five percent level, which is k_1 . For these values of k_1 and k_2 , calculate the LHS of (4.11) for $i = 2$. Let this value be denoted by s_2 . If $s_2 = \alpha$, we have found the required values of k_1 and k_2 . If $s_2 > \alpha$, make k_2 bigger, otherwise make

k_2 smaller with the aim of getting s_2 closer to α and repeat these steps again. This allows us to solve (4.11) for k_1 and k_2 via trial and error.

Final critical values of k_1 and k_2 for all design matrices are given in Table 4.4. From this table we observe that k_1 and k_2 are always positive numbers and the k_2 values are always far larger than k_1 values for all design matrices. From the critical values, we also observe that there is a decreasing trend for the k_1 and k_2 values as the sample size increases, the one exception for k_1 being for $X3$ and for k_2 being for $X4$.

Using the critical values k_1 and k_2 , we calculated the sizes of the test using (4.10) for different values of γ , namely, $\gamma = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9 under H_0 . The estimated sizes are presented in Table 4.5. Again we see that the sizes of the test obtained were not satisfactory due to the fact that we are controlling average size. Here the sizes at the end points of the null hypothesis parameter space are the worst. These sizes are significantly different from our nominal size. This caused us to decide to divide the null hypothesis parameter space into three disjoint intervals and control average sizes of these three intervals. The detailed discussions of the size properties of the test for $m = 3$ are given below.

4.3.4 Size properties of the test with $m = 3$

In the previous subsection, we discussed the option of controlling the average estimated size of our test over two regions of the null hypothesis parameter space ($m = 2$). The choice of regions was determined by sizes for the $m = 1$ case. Unfortunately the resultant

test sizes were judged to be less than acceptable. Part of the problem seems to be that one of the regions is made up of two intervals, with one interval having typically higher sizes balanced on average by lower sizes for the other interval. This led us to consider three intervals based on the $m = 1$ sizes. These are the LHS interval for which sizes are below average, the middle interval for which sizes are above average and the RHS interval for which sizes are below average. For example, in the case of the $X1$ design matrix with $n = 20$, these intervals are $0 \leq \gamma \leq 0.24$ (Θ_1), $0.24 < \gamma \leq 0.64$ (Θ_2) and $0.64 < \gamma \leq 1$ (Θ_3), respectively, see Table 4.6. In this case, ranges of the three disjoint intervals and the corresponding values of j , where $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, 50$, for all design matrices are presented in Table 4.6.

Thus for $m = 3$, equation (3.7) gives the resultant test that maximizes average power subject to three size conditions and involves a critical region of the form:

$$\begin{aligned} & \sum_{j=1}^q |\Omega(\rho_j)|^{-1/2} |X' \Omega(\rho_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_j)' \tilde{z}(\rho_j)}{z'z} \right)^{-p/2} \\ & \geq \left[k_1 \sum_{j=1}^{q_1} |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right. \\ & \quad \left. + k_2 \sum_{j=q_1+1}^{q_2} |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right. \\ & \quad \left. + k_3 \sum_{j=q_2+1}^q |\Sigma(\gamma_j)|^{-1/2} |X' \Sigma(\gamma_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_j)' \tilde{z}(\gamma_j)}{z'z} \right)^{-p/2} \right] \end{aligned} \quad (4.13)$$

where k_1 , k_2 and k_3 are critical values calculated to control average size over the three subregions of Θ and the values of q_1 and q_2 are determined by the boundaries of these subregions.

If we denote the critical region given by (4.13) as $\omega(k_1, k_2, k_3)$ and if $\Theta_1 = \{\gamma: \gamma_0 \leq \gamma \leq \gamma_1\}$, $\Theta_2 = \{\gamma: \gamma_1 < \gamma \leq \gamma_2\}$ and $\Theta_3 = \{\gamma: \gamma_2 < \gamma \leq \gamma_3\}$ where $\gamma_0 = 0$, $\gamma_3 = 1$, and γ_1 and γ_2 are our chosen boundaries of our subregions of Θ , then finding k_1 , k_2 and k_3 involves solving

$$\int_{\Theta_i} \int_{\omega(k_1, k_2, k_3)} h(v, \gamma) dv d\gamma = \alpha, \quad i = 1, 2, 3, \quad (4.14)$$

jointly for k_1 , k_2 and k_3 . Here q_1 and q_2 in (4.13) are determined by rounding $(\gamma_1 q + (1/2))$ and $(\gamma_2 q + (1/2))$, respectively, down to the nearest integer.

Again the LHS of (4.14), for $i = 1, 2, 3$, can be approximated using Monte Carlo integration. Solving the three equations given by (4.14) for k_1, k_2, k_3 can be difficult and time consuming, but it can be solved via trial and error. When solving for the critical values k_i ($i = 1, 2, 3$) using (4.14), we used Steps 1 to 3, discussed in Section 3.3 for $m = 3$, so the average size conditions are simultaneously satisfied.

Calculated critical values of k_1 , k_2 and k_3 for different design matrices are given in

Table 4.7. From this table we observe that the k_2 values are always far larger than the k_1 and k_3 values for all design matrices. We also observe that there is a decreasing trend for the k_1 , k_2 and k_3 values when the sample size is increasing, the only exceptions being for k_1 for $X3$ and k_2 for $X4$. From the values presented in Table 4.7, we observe that k_1 and k_2 are always positive numbers. Moreover, for some cases, the k_3 values are close to zero. In addition to that, k_3 takes a very small negative value when the sample size is increased. Note that Lehmann (1986) required the constants k_1, k_2, \dots, k_m to all be positive numbers. Arthanari and Dodge (1980, *Result 4.5.1*, p.196) have been able to show that the NP lemma and the generalized version of the lemma only require the existence of real numbers k_1, k_2, \dots, k_m . From the two experiments with $m=2$ and $m=3$ we observed that the behaviour and the pattern of the critical values k_1 and k_2 are very similar.

Once we obtained the critical values k_i ($i=1, 2, 3$), the sizes and powers were calculated using (4.13) for $\gamma=0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9 under H_0 and for $\rho=0.1, 0.3, 0.5, 0.7$ and 0.9 under H_a . The calculated sizes are presented in Table 4.8.

Our interest in this table is to see the behaviour of the sizes of our new test given that we are controlling average size. It is pleasing to note that the calculated sizes for all design matrices with $n=20$ fall within the range $[0.037, 0.063]$ making them not significantly different from five percent at the 0.01 level. Note that we are more anxious when the estimated sizes are greater than 0.063 rather than below 0.037. With this in mind, there are a few points (most of them are near to the boundary points, $\gamma=0.0$ and 0.9) which

are significantly different from five percent at the 0.01 level for the larger sample sizes. Overall we find the sizes of our test to be very acceptable for small sample sizes, when $m=3$, at least for the ten γ values we considered.

4.4 Power comparison of the test with $m=3$

Table 4.9 presents the calculated sizes and powers of our average power test and the corresponding sizes and powers of three versions of Silvapulle and King's (1991) APOI test, namely the $s(0.3, \gamma_0^*)$, $s(0.5, \gamma_0^*)$ and $s(0.75, \gamma_0^*)$ tests. Here we note that the size and power results for the three different versions of APOI tests for $X2$ design matrix come from Silvapulle and King (1991), and for all other design matrices, the results come from Silvapulle¹ (1991). The sizes of our test have been estimated via simulation methods using 2000 replications, sizes within the range $[0.037, 0.063]$ are not significantly different from 0.05 at the 0.01 level. On the other hand, the sizes and powers for the APOI tests have been calculated to a much higher level of accuracy using standard numerical methods based on Imhof's (1961) algorithm as discussed by King (1987a).

The first aspect that we have discussed in the previous section is the behaviour of the sizes of our new test. Now we want to compare sizes of our test with three different versions of Silvapulle and King's (1991) APOI test. We observe from Table 4.9 that sizes for all design matrices with $n=20$ fall within the rejection probability interval $[0.037, 0.063]$. Note that when the sample size is small there are a few sizes that are

¹ The author wishes thank Associate Professor Silvapulle for giving her permission to use these results from her Ph.D. thesis.

higher than 0.05 while all sizes of three different versions of the APOI test are 0.05 or below. One reason for this difference may be the fact that because the new test seeks to control average size at 0.05, it should have slightly higher sizes than the APOI tests. There are a few sizes of our test that are significantly different from five percent at the 0.01 level when the sample size is increased. Overall we find that the sizes of the new test perform very well for small sample sizes, when $m = 3$. With respect to the size condition, we can say that our test compares favourably with each of the three different versions of Silvapulle and King's (1991) APOI tests when the sample size is small.

The second point of Table 4.9 is to compare the power of our test with that of the APOI tests. We observe that for all design matrices, when the sample size is small except when $\rho = 0.3$ for $X1$ and $\rho = 0.1, 0.3$ for $X3$ and $X5$ the powers of the new test are typically higher than those of the APOI tests. On the other hand, when the sample size is larger except when $\rho = 0.1$ for all design matrices, $\rho = 0.9$ for $X3$, $\rho = 0.3, 0.5, 0.7$ for $X4$ and $\rho = 0.3, 0.9$ for $X5$ the powers of the new test are even higher than those of the APOI tests. The new test still has a very clear power advantage, that is, when the sample size is small, on an average the new test gained 0.7 to 5.7 percentage points for all design matrices and when the sample size is large, the new test gained 0.8 to 1.9 percentage points in terms power improvement except for $X4$. Also, we observe that when the sample size is small for all design matrices when $\rho = 0.7, 0.9$ the powers of the new test are much higher than those of the APOI tests. The better powers could be due to the slightly higher sizes of the new test. Considering all these cases, the new test continues to show its very clear superiority in terms of power when the sample size is small. This leads us to conclude that, at least for the problem of testing for MA(1) disturbances

against AR(1) disturbances in the linear regression model, the new test is a better test than the APOI test which Silvapulle and King (1991) found out-performed the LM test and an analogous test to Burke et al.'s (1990) asymptotic test.

4.5 Concluding remarks

In this chapter we have considered the application of the new approach which was constructed in Chapter 3 for testing a composite null hypothesis against a composite alternative. We illustrated the new test procedure by applying it to the problem of testing for MA(1) disturbances against AR(1) disturbances in the linear regression model as a single parameter testing problem. The problem was deliberately chosen because through invariance arguments, it reduces to composite null and alternative hypotheses involving a scalar parameter defined over the unit interval. It is also a testing problem for which it has been difficult to find a satisfactory test. The main practical difficulty is with the finding of critical values for which the average size conditions of the test are simultaneously satisfied. In this approach, we only used the uniform distribution as a weighting function. It can easily be applied using a non-uniform distribution. Using Silvapulle and King's (1991) APOI tests as benchmarks, the new test was found to have good small sample size and power properties. This suggests the new procedure has considerable potential, particularly for problems in which the effective null hypothesis parameter space can be kept small. Further research is needed in order to see how well the test procedure works when the boundary points of the subregions of the null hypothesis parameter space are chosen without any reference to lower than average size and higher than average size regions in the $m = 1$ case, we will call this the high

size/low size technique in the remainder of the thesis. Also, we need to see how well the procedure works for other types of parameter spaces under the null hypothesis. In the following chapters we will look at these issues.

Table 4.1

Critical values (c_α) for $X1$, $X2$, $X3$, $X4$ and $X5$ design matrices when $m = 1$ at the five percent level.

Design matrix	Sample Size (n)	Critical values (c_α)
$X1(n \times 3)$	20	2.5521
	60	1.9359
$X2(n \times 3)$	20	2.5748
	60	2.0125
$X3(n \times 5)$	20	2.2807
	60	1.9539
$X4(n \times 3)$	20	2.6958
	30	2.2886
$X5(n \times 3)$	20	2.8923
	60	2.0428

Table 4.2

Calculated sizes of the test when $m = 1$ at the five percent level.

Parameter values	Sizes									
	$X1(n \times 3)$		$X2(n \times 3)$		$X3(n \times 5)$		$X4(n \times 3)$		$X5(n \times 3)$	
γ	$n = 20$	60	20	60	20	60	20	30	20	60
0.0	0.013*	0.009*	0.017*	0.009*	0.030*	0.010*	0.014*	0.016*	0.013*	0.008*
0.1	0.025*	0.041	0.027*	0.035*	0.041	0.036*	0.025*	0.036*	0.024*	0.034*
0.2	0.043	0.081*	0.046	0.074*	0.055	0.080*	0.045	0.064*	0.040	0.068*
0.3	0.060	0.114*	0.061	0.109*	0.068*	0.114*	0.061	0.088*	0.058	0.109*
0.4	0.068*	0.123*	0.069*	0.113*	0.073*	0.119*	0.069*	0.097*	0.068*	0.117*
0.5	0.068*	0.101*	0.062	0.092*	0.069*	0.096*	0.061	0.089*	0.064*	0.091*
0.6	0.055	0.053	0.053	0.050	0.062	0.055	0.052	0.069*	0.048	0.047
0.7	0.041	0.019*	0.044	0.023*	0.047	0.024*	0.035*	0.042	0.029*	0.017*
0.8	0.025*	0.006*	0.029*	0.007*	0.033*	0.007*	0.020*	0.022*	0.017*	0.006*
0.9	0.017*	0.000*	0.021*	0.001*	0.027*	0.002*	0.015*	0.010*	0.011*	0.000*

* denotes value significantly different from five percent at the 0.01 level.

Table 4.3

Ranges of the two disjoint intervals and the corresponding values of j , where
 $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$.

Sample Size (n)	Subregions	
	Θ_1 (smaller than average size for $m = 1$)	Θ_2 (larger than average size for $m = 1$)
For design matrix $X1(n \times 3)$		
20	$0 \leq \gamma \leq 0.24 \cup 0.64 < \gamma \leq 1$ ($j = 1, 2, \dots, 12 \cup j = 33, 34, \dots, 50$)	$0.24 < \gamma \leq 0.64$ ($j = 13, 14, \dots, 32$)
60	$0 \leq \gamma \leq 0.12 \cup 0.61 < \gamma \leq 1$ ($j = 1, 2, \dots, 6 \cup j = 32, 33, \dots, 50$)	$0.12 < \gamma \leq 0.61$ ($j = 7, 8, \dots, 31$)
For design matrix $X2(n \times 3)$		
20	$0 \leq \gamma \leq 0.23 \cup 0.63 < \gamma \leq 1$ ($j = 1, 2, \dots, 12 \cup j = 33, 34, \dots, 50$)	$0.23 < \gamma \leq 0.63$ ($j = 13, 14, \dots, 32$)
60	$0 \leq \gamma \leq 0.14 \cup 0.60 < \gamma \leq 1$ ($j = 1, 2, \dots, 7 \cup j = 31, 32, \dots, 50$)	$0.14 < \gamma \leq 0.60$ ($j = 8, 9, \dots, 30$)
For design matrix $X3(n \times 5)$		
20	$0 \leq \gamma \leq 0.16 \cup 0.68 < \gamma \leq 1$ ($j = 1, 2, \dots, 8 \cup j = 35, 36, \dots, 50$)	$0.16 < \gamma \leq 0.68$ ($j = 9, 10, \dots, 34$)
60	$0 \leq \gamma \leq 0.13 \cup 0.61 < \gamma \leq 1$ ($j = 1, 2, \dots, 7 \cup j = 32, 33, \dots, 50$)	$0.13 < \gamma \leq 0.61$ ($j = 8, 9, \dots, 31$)
For design matrix $X4(n \times 3)$		
20	$0 \leq \gamma \leq 0.23 \cup 0.61 < \gamma \leq 1$ ($j = 1, 2, \dots, 12 \cup j = 32, 33, \dots, 50$)	$0.23 < \gamma \leq 0.61$ ($j = 13, 14, \dots, 31$)
30	$0 \leq \gamma \leq 0.15 \cup 0.67 < \gamma \leq 1$ ($j = 1, 2, \dots, 8 \cup j = 35, 36, \dots, 50$)	$0.15 < \gamma \leq 0.67$ ($j = 9, 10, \dots, 34$)
For design matrix $X5(n \times 3)$		
20	$0 \leq \gamma \leq 0.25 \cup 0.58 < \gamma \leq 1$ ($j = 1, 2, \dots, 13 \cup j = 30, 31, \dots, 50$)	$0.25 < \gamma \leq 0.58$ ($j = 14, 15, \dots, 29$)
60	$0 \leq \gamma \leq 0.15 \cup 0.59 < \gamma \leq 1$ ($j = 1, 2, \dots, 8 \cup j = 31, 32, \dots, 50$)	$0.15 < \gamma \leq 0.59$ ($j = 9, 10, \dots, 30$)

Table 4.4

Final critical values k_1 and k_2 when $m = 2$.

Sample size (n)	Critical values	
	k_1	k_2
For design matrix $X1(n \times 3)$		
20	0.3938	1.980
60	0.2000	1.710
For design matrix $X2(n \times 3)$		
20	0.2969	2.100
60	0.2918	1.600
For design matrix $X3(n \times 5)$		
20	0.0420	2.100
60	0.3271	1.6085
For design matrix $X4(n \times 3)$		
20	0.5423	1.690
30	0.0911	2.320
For design matrix $X5(n \times 3)$		
20	0.6002	1.890
60	0.3494	1.800

Table 4.5

Calculated sizes of the test when $m = 2$ at the five percent level.

Parameter values	Sizes									
	$X1(n \times 3)$		$X2(n \times 3)$		$X3(n \times 5)$		$X4(n \times 3)$		$X5(n \times 3)$	
γ	$n = 20$	60	20	60	20	60	20	30	20	60
0.0	0.136*	0.224*	0.164*	0.269*	0.074*	0.207*	0.097*	0.165*	0.096*	0.263*
0.1	0.092*	0.082*	0.112*	0.113*	0.056	0.085*	0.081*	0.089*	0.071*	0.108*
0.2	0.063	0.047	0.075*	0.054	0.046	0.048	0.072*	0.052	0.059	0.045
0.3	0.052	0.054	0.055	0.057	0.048	0.058	0.068*	0.043	0.052	0.046
0.4	0.043	0.065*	0.046	0.065*	0.048	0.062	0.060	0.050	0.045	0.053
0.5	0.044	0.057	0.044	0.056	0.044	0.057	0.057	0.049	0.042	0.048
0.6	0.043	0.039	0.041	0.042	0.046	0.043	0.051	0.046	0.039	0.032*
0.7	0.039	0.022*	0.038	0.023*	0.047	0.027*	0.042	0.038	0.034*	0.019*
0.8	0.037	0.012*	0.035*	0.014*	0.043	0.011*	0.029*	0.032*	0.027*	0.008*
0.9	0.032*	0.004*	0.033*	0.006*	0.040	0.006*	0.024*	0.025*	0.022*	0.004*

* denotes value significantly different from five percent at the 0.01 level.

Table 4.6

Ranges of the three disjoint intervals and the corresponding values of j , where
 $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$.

Sample size (n)	Subregions		
	Θ_1 (smaller than average size)	Θ_2 (larger than average size)	Θ_3 (smaller than * average size)
For design matrix $X1(n \times 3)$			
20	$0 \leq \gamma \leq 0.24$ ($j=1, 2, \dots, 12$)	$0.24 < \gamma \leq 0.64$ ($j=13, 14, \dots, 32$)	$0.64 < \gamma \leq 1$ ($j=33, 34, \dots, 50$)
60	$0 \leq \gamma \leq 0.12$ ($j=1, 2, \dots, 6$)	$0.12 < \gamma \leq 0.61$ ($j=7, 8, \dots, 31$)	$0.61 < \gamma \leq 1$ ($j=32, 33, \dots, 50$)
For design matrix $X2(n \times 3)$			
20	$0 \leq \gamma \leq 0.23$ ($j=1, 2, \dots, 12$)	$0.23 < \gamma \leq 0.63$ ($j=13, 14, \dots, 32$)	$0.63 < \gamma \leq 1$ ($j=33, 34, \dots, 50$)
60	$0 \leq \gamma \leq 0.14$ ($j=1, 2, \dots, 7$)	$0.14 < \gamma \leq 0.60$ ($j=8, 9, \dots, 30$)	$0.60 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
For design matrix $X3(n \times 5)$			
20	$0 \leq \gamma \leq 0.16$ ($j=1, 2, \dots, 8$)	$0.16 < \gamma \leq 0.68$ ($j=9, 10, \dots, 34$)	$0.68 < \gamma \leq 1$ ($j=35, 36, \dots, 50$)
60	$0 \leq \gamma \leq 0.13$ ($j=1, 2, \dots, 7$)	$0.13 < \gamma \leq 0.61$ ($j=8, 9, \dots, 31$)	$0.61 < \gamma \leq 1$ ($j=32, 33, \dots, 50$)
For design matrix $X4(n \times 3)$			
20	$0 \leq \gamma \leq 0.23$ ($j=1, 2, \dots, 12$)	$0.23 < \gamma \leq 0.61$ ($j=13, 14, \dots, 31$)	$0.61 < \gamma \leq 1$ ($j=32, 33, \dots, 50$)
30	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.67$ ($j=9, 10, \dots, 34$)	$0.67 < \gamma \leq 1$ ($j=35, 36, \dots, 50$)
For design matrix $X5(n \times 3)$			
20	$0 \leq \gamma \leq 0.25$ ($j=1, 2, \dots, 13$)	$0.25 < \gamma \leq 0.58$ ($j=14, 15, \dots, 29$)	$0.58 < \gamma \leq 1$ ($j=30, 31, \dots, 50$)
60	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.59$ ($j=9, 10, \dots, 30$)	$0.59 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)

Table 4.7
Final critical values k_1, k_2 and k_3 when $m = 3$.

Sample size (n)	Critical values		
	k_1	k_2	k_3
For design matrix $X1(n \times 3)$			
20	0.2072	1.950	0.0350
60	0.0861	1.610	-0.0011
For design matrix $X2(n \times 3)$			
20	0.2098	1.840	0.0550
60	0.1202	1.520	-0.0012
For design matrix $X3(n \times 5)$			
20	0.0448	2.014	0.0300
60	0.1083	1.590	-0.0025
For design matrix $X4(n \times 3)$			
20	0.2242	1.720	0.1000
30	0.0689	2.350	-0.0200
For design matrix $X5(n \times 3)$			
20	0.2423	1.97	0.0200
60	0.1335	1.683	-0.0006

Table 4.8
Calculated sizes of the test when $m = 3$ at the five percent level.

Parameter values	Sizes									
	$X1(n \times 3)$		$X2(n \times 3)$		$X3(n \times 5)$		$X4(n \times 3)$		$X5(n \times 3)$	
γ	$n=20$	60	20	60	20	50	20	30	20	60
0.0	0.061	0.080*	0.058	0.076*	0.054	0.075*	0.056	0.075*	0.055	0.091*
0.1	0.048	0.030*	0.055	0.031*	0.046	0.030*	0.053	0.043	0.044	0.036*
0.2	0.042	0.038	0.049	0.041	0.045	0.035*	0.060	0.030*	0.042	0.032*
0.3	0.040	0.060	0.049	0.058	0.048	0.058	0.063	0.038	0.045	0.047
0.4	0.046	0.070*	0.053	0.069*	0.049	0.064*	0.060	0.048	0.048	0.059
0.5	0.049	0.064*	0.052	0.063	0.046	0.061	0.063	0.051	0.049	0.055
0.6	0.047	0.048	0.055	0.050	0.049	0.052	0.061	0.050	0.050	0.038
0.7	0.050	0.031*	0.053	0.045	0.049	0.043	0.060	0.045	0.048	0.031*
0.8	0.050	0.034*	0.050	0.046	0.045	0.045	0.051	0.042	0.048	0.033*
0.9	0.048	0.076*	0.046	0.072*	0.042	0.076*	0.043	0.042	0.047	0.058

* denotes value significantly different from five percent at the 0.01 level.

Table 4.9

Calculated sizes and powers of the new test and APOI tests for $X1$, $X2$, $X3$, $X4$ and $X5$, when $m=3$ and $H_0: u_t = e_t + \gamma e_{t-1}$, $0 \leq \gamma \leq 1$, is tested against $H_a: u_t = \rho u_{t-1} + e_t$, $0 < \rho < 1$, at the five percent level.

Hypothesis/ parameter values	New test	APOI tests			New test	APOI tests		
		$s(0.3, \gamma_a^*)$	$s(0.5, \gamma_a^*)$	$s(0.75, \gamma_a^*)$		$s(0.3, \gamma_a^*)$	$s(0.5, \gamma_a^*)$	$s(0.75, \gamma_a^*)$
X1: $n = 20$					X1: $n = 60$			
H_0								
$\gamma = 0.9$	0.048	0.048	0.049	0.048	0.076	0.048	0.047	0.047
0.7	0.050	0.047	0.048	0.046	0.031	0.046	0.044	0.044
0.5	0.049	0.046	0.045	0.044	0.064	0.044	0.042	0.042
0.3	0.040	0.045	0.043	0.043	0.060	0.042	0.039	0.039
0.1	0.048	0.046	0.044	0.044	0.030	0.047	0.044	0.046
0.0	0.061	0.050	0.050	0.050	0.080	0.050	0.050	0.050
H_a								
$\rho = 0.1$	0.055	0.052	0.051	0.051	0.031	0.054	0.052	0.051
0.3	0.073	0.079	0.076	0.075	0.138	0.131	0.122	0.119
0.5	0.177	0.116	0.163	0.159	0.492	0.434	0.426	0.419
0.7	0.363	0.335	0.337	0.336	0.868	0.839	0.842	0.845
0.9	0.557	0.515	0.526	0.536	0.984	0.975	0.979	0.984
X2: $n = 20$					X2: $n = 60$			
H_0								
$\gamma = 0.9$	0.046	0.049	0.049	0.049	0.072	0.049	0.049	0.049
0.7	0.053	0.049	0.048	0.048	0.045	0.048	0.047	0.046
0.5	0.052	0.046	0.046	0.045	0.063	0.045	0.043	0.042
0.3	0.049	0.044	0.045	0.044	0.058	0.043	0.042	0.041
0.1	0.055	0.048	0.047	0.046	0.031	0.046	0.044	0.043
0.0	0.058	0.050	0.050	0.050	0.076	0.050	0.050	0.050
H_a								
$\rho = 0.1$	0.057	0.052	0.051	0.050	0.036	0.053	0.052	0.051
0.3	0.081	0.076	0.077	0.073	0.149	0.136	0.132	0.126
0.5	0.171	0.156	0.156	0.154	0.528	0.455	0.453	0.442
0.7	0.344	0.311	0.324	0.324	0.896	0.870	0.875	0.872
0.9	0.566	0.504	0.529	0.540	0.990	0.990	0.991	0.992

Table 4.9 (contn'd)

Hypothesis/ parameter values	New test	APOI tests			New test	APOI tests			
		$s(0.3, \gamma_a^*)$	$s(0.5, \gamma_a^*)$	$s(0.75, \gamma_a^*)$		$s(0.3, \gamma_a^*)$	$s(0.5, \gamma_a^*)$	$s(0.75, \gamma_a^*)$	
$X3: n = 20$					$X3: n = 60$				
H_0									
$\gamma = 0.9$	0.042	0.048	0.049	0.050	0.076	0.049	0.049	0.050	
0.7	0.049	0.046	0.047	0.048	0.043	0.048	0.047	0.047	
0.5	0.046	0.044	0.045	0.045	0.061	0.045	0.043	0.042	
0.3	0.048	0.044	0.045	0.045	0.058	0.043	0.041	0.041	
0.1	0.046	0.045	0.046	0.047	0.030	0.045	0.044	0.044	
0.0	0.054	0.050	0.050	0.050	0.075	0.050	0.050	0.050	
H_a									
$\rho = 0.1$	0.048	0.052	0.053	0.053	0.033	0.058	0.056	0.057	
0.3	0.064	0.070	0.071	0.071	0.139	0.132	0.127	0.123	
0.5	0.138	0.130	0.133	0.132	0.504	0.441	0.436	0.429	
0.7	0.283	0.252	0.261	0.264	0.880	0.859	0.862	0.862	
0.9	0.455	0.418	0.435	0.447	0.985	0.989	0.989	0.990	
$X4: n = 20$					$X4: n = 30$				
H_0									
$\gamma = 0.9$	0.043	0.048	0.049	0.048	0.042	0.049	0.049	0.050	
0.7	0.060	0.046	0.048	0.047	0.045	0.048	0.048	0.048	
0.5	0.063	0.044	0.045	0.045	0.051	0.045	0.044	0.043	
0.3	0.063	0.044	0.044	0.045	0.038	0.043	0.043	0.042	
0.1	0.053	0.046	0.045	0.046	0.043	0.044	0.045	0.044	
0.0	0.056	0.050	0.050	0.050	0.075	0.050	0.050	0.050	
H_a									
$\rho = 0.1$	0.058	0.054	0.053	0.053	0.049	0.052	0.054	0.054	
0.3	0.094	0.076	0.076	0.074	0.074	0.094	0.092	0.090	
0.5	0.185	0.152	0.155	0.150	0.235	0.243	0.244	0.238	
0.7	0.371	0.293	0.310	0.317	0.531	0.543	0.538	0.538	
0.9	0.588	0.438	0.484	0.539	0.799	0.778	0.792	0.803	

Table 4.9 (contn'd)

Hypothesis/ parameter values	New test	APOI tests			New test	APOI tests			
		$s(0.3, \gamma'_n)$	$s(0.5, \gamma'_n)$	$s(0.75, \gamma'_n)$		$s(0.3, \gamma'_n)$	$s(0.5, \gamma'_n)$	$s(0.75, \gamma'_n)$	
$X5: n = 20$					$X5: n = 60$				
H_0									
$\gamma = 0.9$	0.047	0.050	0.050	0.049	0.058	0.049	0.050	0.050	
0.7	0.048	0.049	0.048	0.047	0.031	0.047	0.047	0.047	
0.5	0.049	0.046	0.045	0.044	0.055	0.044	0.043	0.042	
0.3	0.045	0.045	0.043	0.043	0.047	0.043	0.042	0.040	
0.1	0.044	0.046	0.045	0.045	0.036	0.045	0.043	0.042	
0.0	0.055	0.050	0.050	0.050	0.091	0.050	0.050	0.050	
H_a									
$\rho = 0.1$	0.047	0.053	0.052	0.053	0.042	0.058	0.056	0.054	
0.3	0.073	0.075	0.073	0.072	0.128	0.133	0.130	0.122	
0.5	0.146	0.141	0.139	0.137	0.483	0.435	0.434	0.423	
0.7	0.265	0.250	0.251	0.249	0.857	0.840	0.844	0.842	
0.9	0.385	0.343	0.348	0.350	0.975	0.977	0.979	0.979	

Chapter 5

An Investigation of the Choice of Boundary Points under the Null Hypothesis

5.1 Introduction

A new approach to testing a composite null hypothesis against a composite alternative using the generalized NP lemma was introduced in Chapter 3. In that chapter we outlined the theory of this new approach with particular emphasis on testing the disturbances in the general linear regression model. These testing problems involve nuisance parameters which in this case can be eliminated through invariance arguments. In Chapter 4, we applied the new test procedure to the problem of testing for MA(1) disturbances against AR(1) disturbances in the linear regression model with the aim of exploring a number of practical issues.

Our new test involves maximizing the average power subject to controlling the average size over different regions under the null hypothesis parameter space. A major question is how many regions should we use. With respect to testing for MA(1) disturbances against AR(1) disturbances in the linear regression model, we conducted some simulation experiments for tests constructed based on one region ($m = 1$), the entire null hypothesis parameter space ($\Theta = \{\gamma: 0 \leq \gamma \leq 1\}$), and for $m = 2$, two subsets of the null

hypothesis parameter space. After completing these simulation experiments, we decided we needed to control the average size of the test over the three subsets ($m = 3$) of the null hypothesis parameter space. When conducting the simulation experiments for $m = 2$ and $m = 3$, the boundary points of the subregions of the null hypothesis parameter space were chosen through the high size/low size technique. We saw that our test for $m = 3$, appeared to do a good job of controlling sizes overall, especially when the sample size is small. In addition, the power of the test appeared to be satisfactory. Overall our test performed well with respect to sizes and powers for $m = 3$, when the boundary points were selected by the high size/low size method. Unfortunately we are not sure that this is a good method for deciding on the boundary points of these subregions. There may be another way to choose the boundary points such that sizes are controlled adequately and the resultant test has better power.

How can we check whether our high size/low size technique is a good one for choosing the boundary points? One approach we could take to answering this question is to devise a power envelope. The power envelope is the maximum power that can be obtained at each particular point under the alternative hypothesis over all tests of a designated size. We will investigate this by conducting an experiment with many different choices of boundary points. That is, we systematically will choose different boundary points of the three subregions ($m = 3$) of the null hypothesis parameter space for separate tests and using the power results of the resultant tests with acceptable sizes, we will trace out a power envelope of sorts. This is not a true power envelope because we are not considering all possible tests. However it will provide us with a guide. We will call it a pseudo power envelope. It will be interesting to see if the new approach introduced in

Chapter 3 and as applied in Chapter 4, leads to useful results compared to this pseudo power envelope.

In this chapter, our aim is to investigate the performance of our test for different boundary points of the subregions of the test when $m = 3$. We realise that changing the middle two boundary points of the three subsets of the null hypothesis parameter space Θ also changes the test, its critical values (k_i , $i = 1, 2, 3$), and hence the calculated sizes and powers of the test. With this in mind, we chose different boundary points of the three subregions and then calculated the critical values, sizes and powers of the resultant test.

The calculation of the pseudo power envelope provides a benchmark against which the power performance of different tests can be measured. For example we can calculate the average deviation (AD) of the test's power from the pseudo power envelope. Alternatively we could calculate the sum of squared deviations (SSD) of the actual power from the pseudo power envelope. Clearly the best test would be the test with smallest AD or SSD, so our approach will lead to the construction of the best test within our class of tests.

We can also use the pseudo power envelope to judge the relative performance of our test proposed in Chapter 4 that is based on the high size/low size technique. If our proposed test has power close to the pseudo power envelope, our proposed approach can be judged to be a reasonable one. A small improvement in AD of the power of the best test over the proposed test would mean that the huge amount of extra computation in finding the best test is not well justified.

The plan of this chapter is as follows. The methodology of the experiment is discussed in Section 5.2. Details of the Monte Carlo experiment are outlined in Section 5.3. We report the experimental design, size results and power results for all experiments for the test with $m=3$ in subsections 5.3.1, 5.3.2 and 5.3.3, respectively. The best test is identified in subsection 5.3.3. The final section contains some concluding remarks.

5.2 Methodology

As stated in the above section, our aim is to investigate different splits of the null hypothesis parameter space into three subregions with the aim of maximising the power of our test. This requires a benchmark for power, an obvious one being the power envelope as discussed above.

We constructed a pseudo power envelope as follows. For each test, the boundary points of the three subregions of $(\Theta = \{\gamma: 0 \leq \gamma \leq 1\})$ are chosen systematically without the use of a numerical technique. We have to choose only the middle two boundary points for the three subregions of the null hypothesis parameter space, because the starting boundary point and the end boundary point are known to us, being zero and one, respectively. The middle two boundary points were systematically set as a pair from $0.05i$, $i = 1, 2, \dots, 19$, resulting in ${}^{19}C_2$ combinations of choice for these two boundary points, i.e. the total number of choices for the middle two boundary points was one hundred and seventy-one. For example, for the first set of systematic choices, we made eighteen choices of the middle two boundary points in which case the three subregions under the null hypothesis parameter space were $\Theta_1 = \{\gamma: 0 \leq \gamma \leq 0.05\}$, $\Theta_2 = \{\gamma: 0.05 < \gamma \leq 0.05j\}$ and $\Theta_3 = \{\gamma: 0.05j < \gamma \leq 1\}$, for $j = 2, 3, \dots, 19$,

consecutively. For the second set of systematic choices, we made seventeen choices of the middle two boundary points and in this case the three disjoint intervals under the null hypothesis parameter space were $\Theta_1 = \{\gamma: 0 \leq \gamma \leq 0.1\}$, $\Theta_2 = \{\gamma: 0.1 < \gamma \leq 0.05j\}$ and $\Theta_3 = \{\gamma: 0.05j < \gamma \leq 1\}$, for $j = 3, 4, \dots, 19$, consecutively. Similarly, for the third set of systematic choices, we made sixteen choices of the middle two boundary points and the three disjoint intervals under the null hypothesis parameter space were $\Theta_1 = \{\gamma: 0 \leq \gamma \leq 0.15\}$, $\Theta_2 = \{\gamma: 0.15 < \gamma \leq 0.05j\}$ and $\Theta_3 = \{\gamma: 0.05j < \gamma \leq 1\}$, for $j = 4, 5, \dots, 19$, consecutively and so on. For all combinations of the choice of the two boundary points of the three subregions of the null hypothesis parameter space, we conducted a total of one hundred and seventy-one separate tests. Using the size results of one hundred and seventy-one tests, we calculated the powers of those tests with acceptable sizes in order to calculate the pseudo power envelope. The pseudo power envelope was then calculated as the maximum power of all these remaining tests for each value of ρ . Detailed descriptions are given in the following section.

5.3 Monte Carlo experiments

In order to investigate the performance of our test for different boundary points of the three subsets of the null hypothesis parameter space in the context of testing MA(1) disturbances against AR(1) disturbances in the linear regression model, we conducted a Monte Carlo experiment. The experiment involved constructing the pseudo power envelope as outlined above. We calculated the AD and SSD of the power of each of the tests under the pseudo power envelope. From the calculated AD we can find the overall best test where AD is least. This will allow us to find the appropriate boundary points of the three subregions of Θ of the test, which maximizes average power. Also we

compared the AD of the power of the test proposed in previous chapter with the AD of the power of the best test. From this comparison, if we find that the power of the test proposed in Chapter 4 is close to the power of the best test as identified against the pseudo power envelope, then we can say that our proposed test is realistically a good test. It would indicate that there is little to be gained from the vast amount of extra computation required to find the best test using the pseudo power envelope.

5.3.1 Experimental design

For the above-mentioned testing problem, two thousand replications were used to calculate the sizes and powers of our test for different boundary points of the subregions under the null hypothesis. The critical values were calculated at the five percent level. To construct a pseudo power envelope of the test, we chose only the $X2$ (used in Chapter 4) design matrix with $n = 20$, because this matrix exhibits practical degrees of multicollinearity. For this choice of design matrix we have constructed a total of one hundred and seventy-one different tests. We confined our choice to a single design matrix, because for each test we have to calculate the critical values k_i , $i = 1, 2, 3$, so that the three size conditions are simultaneously fulfilled, via the trial and error. The calculation for one trial takes about four minutes. Typically, calculating of a set of critical values requires 10-50 trials, i.e., 40-200 minutes. Our single design matrix for sample size $n = 20$ required approximately 342 hours to calculate the one hundred and seventy-one sets of critical values. Then follows the calculation of sizes and powers of the test. Therefore, calculating critical values and the required sizes and powers is laborious and time consuming. In other words for a single design matrix, there is a huge amount of calculation required to calculate the pseudo power envelope. So our choice

had to be limited to one design matrix. This is a clear drawback of our simulation experiment.

For two thousand replications, if the calculated size for a test falls outside the estimated rejection probability range $[0.037, 0.063]$, then it is judged to be significantly different from five percent at the 0.01 level.

We mentioned in Chapter 4 that for the problem under consideration, our proposed test under the null hypothesis and alternative hypothesis is invariant with respect to the nuisance parameters, β and σ^2 , and we have chosen $\beta = 0$ and $\sigma^2 = 1$ for the simulation experiment. We already decided that the value of m is three, and we have seen that the size of our test is approximately controlled for this value of m .

Thus for $m = 3$, equation (3.7) gives the resultant test that involves a critical region of the form (4.9) given in (4.13). In equation (4.13), k_i , $i = 1, 2, 3$, are the critical values calculated to control average size over the three subregions of the null hypothesis parameter space, and q_1 and q_2 are the outer limits of summation over these subregions. These two limit integers are determined by rounding $(\gamma_1 q + 1/2)$ and $(\gamma_2 q + 1/2)$ respectively, down to the nearest integer, where γ_1 and γ_2 are the systematically chosen middle two boundary points of the subregions of Θ . Ranges of the three disjoint intervals and the corresponding values of j together with the test number for a total of one hundred and seventy-one tests for the design matrix $X2$ with $n = 20$ are given in Table 5.1.

In order to solve for the critical values k_i , $i = 1, 2, 3$, using equation (4.14), we used Steps 1 to 3, discussed in Section 3.3 for $m = 3$, so that the average size conditions are

simultaneously satisfied. The calculated critical values k_i , $i = 1, 2, 3$, for different tests are presented in Table 5.2. From this table we observe that all k_2 values are positive except six values (where q_2 is greater than or equal 40), while k_1 and k_3 take both positive as well as negative values. We observe that k_1 takes large positive values when the interval of the first subregion (Θ_1) is large and k_3 takes large positive values when the third subregion (Θ_3) interval is large. We also observe that when q_1 is less or equal to 10, then k_1 takes a negative value and when q_2 is greater than 28, then k_3 takes a negative value.

Once we obtain the critical values k_i , $i = 1, 2, 3$, we can calculate the sizes and powers of our test. In the following sections, we report the size and power results of our tests.

5.3.2 Size results

Using the critical values k_i , $i = 1, 2, 3$, presented in Table 5.2, we calculated the sizes of our test statistic through equation (4.13) for $\gamma = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9 under H_0 . The calculated sizes are presented in Table 5.3.

From Table 5.3, we observe that out of a total of one hundred and seventy-one tests, there are sixty-two (36%) tests where the calculated sizes first decrease, then increase and again decrease as γ increases, and for sixty (35%) tests, the sizes increase then decrease and again increase. For twenty-eight (16%) tests, the sizes first increase and then decrease and for one test, the sizes are first decreasing and then increasing. There is also only one test where all sizes decrease and for the other remaining tests, sizes do not follow any pattern. There are two tests (test numbers 74 and 87) where all calculated

sizes are lower than or equal to 0.05. There are another two tests (test numbers 98 and 109) where all calculated sizes are greater than 0.05. The worst calculated size is 0.096 which occurs for test number 60.

The calculated sizes that fall outside the rejection probability range [0.037, 0.063] are significantly different from five percent at the 0.01 level, for two thousand replications. A "star" in Table 5.3 indicates a significantly different calculated size of a test. We observe that some sizes are significantly different from five percent at the 0.01 level when the interval of the first subregion (Θ_1) is very small (i.e., close to the starting boundary point) and the third subregion (Θ_3) interval is large under the null hypothesis parameter space. These significantly different values belong to Θ_3 and most of these sizes occur at the middle and end boundary points under H_0 . On the other hand, when both intervals Θ_1 and Θ_3 are large, some calculated sizes which belong to Θ_1 are significantly different from five percent. Also we observe that when the interval Θ_1 is small, calculated sizes are must likely to be small at the end boundary point under H_0 . Similarly, when the interval Θ_1 is large, calculated sizes are likely to be small at the starting boundary point under H_0 .

In our testing approach, we are controlling the average size over different subregions under the null hypothesis parameter space at the five percent level of significance. In this regard, if the calculated size of the test is less than 0.05 we are not concerned about it but if it is greater than 0.05 especially when it exceeds 0.063, the upper limit of the rejection probability range, we are much more concerned about it. From this point of view, we also accepted the test when the only calculated sizes significantly different from 0.05 are less than 0.037, the lower limit of the rejection probability range. Otherwise, we rejected

the test in the sense that we did not use it in the construction of the pseudo power envelope. With respect to these size conditions, we discarded some tests. A total of one hundred and fifteen (67%) tests were accepted out of one hundred and seventy-one tests, and these are indicated by a "double star" in Table 5.3. The powers of these one hundred and fifteen accepted tests were used to calculate a pseudo power envelope.

From the size analysis we observe that in most cases, the size performance of our accepted test is very good when the intervals Θ_1 , Θ_2 and Θ_3 are not large. This indicates that when the intervals of the three subregions are comparatively small, our test successfully controls the average size.

On the other hand, according to the size conditions discussed above, a total of fifty-six (33%) tests were rejected out of one hundred and seventy-one tests. From the length analysis of the three subsets of Θ of the rejected tests, we observe that for thirty-three (59%) tests, the largest interval of one subset is greater or equal 0.65 and for other fifteen (41%) tests, the largest interval of one subset is in between 0.45 and 0.6. As we mentioned earlier, constructing a pseudo power envelope is time consuming. We can consider reducing this time by restricted to intervals of 0.65 or less. If we use this restriction then we can discard sixty-four tests from a total of one hundred and seventy-one tests, a more than one-third time saving at least for this case.

In the following section we analyse the powers of each of the accepted tests against the pseudo power envelope.

5.3.3 Power results

Using the critical values k_i , $i = 1, 2, 3$, presented in Table 5.2, we calculated the powers of our test statistic through equation (4.13) for $\rho = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9 under H_a . The calculated powers of a total of one hundred and seventy-one tests for design matrix X_2 for $n = 20$ are presented in Table 5.4. From the table we observe that, under H_a , the calculated powers gradually increase as ρ increases from 0.1 to 0.9 except for two tests (test numbers 81 and 82), where powers first decrease and then increase. We also observe that when the interval Θ_1 is large (≥ 0.65), then calculated power for $\rho = 0.1$ is very small and then powers gradually increase as ρ increases.

The powers of the accepted tests with respect to the size conditions are indicated by a "double star" in Table 5.4. The power results of those one hundred and fifteen accepted tests were used to construct our pseudo power envelope. We calculated the AD and SSD of the power of each of the accepted tests under the pseudo power envelope and the test proposed in Chapter 4 (boundary points are chosen by the high size/low size technique). To calculate the AD and SSD of the power of each of the tests that have acceptable size, we first found the pseudo power envelope as outlined above (maxp) for $l (= 9)$ different points of ρ , namely, $\rho = 0.1i$, $i = 1, 2, \dots, l$, separately under H_a . We used the following two formulae: (i) $AD = \frac{1}{l} \sum_{i=1}^l (\max_{\rho=0.1i} - \text{power}_{\rho=0.1i})$, and (ii)

$$SSD = \sum_{i=1}^l (\max_{\rho=0.1i} - \text{power}_{\rho=0.1i})^2, \text{ to calculate the AD and SSD of the power of each}$$

of the acceptable tests, respectively under pseudo power envelope. The calculated AD and SSD of the power along with the rank of each of the accepted tests from smallest to the highest order are presented in Table 5.5.

From Table 5.5 we observe that the smallest AD is 0.005889, which occurs for test 64 for which the three subregions are $\Theta_1: 0 \leq \gamma \leq 0.2$, $\Theta_2: 0.2 < \gamma \leq 0.85$ and $\Theta_3: 0.85 < \gamma \leq 1$. We note that this test also has the smallest SSD, which fortunately means there is no conflict between these two criteria for assessing closeness to the pseudo power envelope. We can identify this as the best test with respect to closeness of power to the pseudo power envelope. The second smallest AD is 0.008333 for tests 45 and 74; in these cases the three subregions are $\Theta_1: 0 \leq \gamma \leq 0.15$, $\Theta_2: 0.15 < \gamma \leq 0.65$ and $\Theta_3: 0.65 < \gamma \leq 1$, and $\Theta_1: 0 \leq \gamma \leq 0.25$, $\Theta_2: 0.25 < \gamma \leq 0.65$ and $\Theta_3: 0.65 < \gamma \leq 1$, respectively. The worst AD is 0.042556 which occurs for test 76 whose three subregions are $\Theta_1: 0 \leq \gamma \leq 0.25$, $\Theta_2: 0.25 < \gamma \leq 0.75$ and $\Theta_3: 0.75 < \gamma \leq 1$. The results indicate that for different boundary points, the powers of our test cross and some depart significantly from this pseudo power envelope. The calculated AD of our test proposed in Chapter 4 is 0.023778 and the three subregions are $\Theta_1: 0 \leq \gamma \leq 0.23$, $\Theta_2: 0.23 < \gamma \leq 0.63$ and $\Theta_3: 0.63 < \gamma \leq 1$, which are very close to the three subregions of test 74, that has the second smallest AD. Also, after doing one month's extra work to calculate the pseudo power envelope, comparing the AD of our proposed test with the best test of the pseudo power envelope we see that there is only an average 1.8 percentage points increase in terms of power improvement. This small improvement in terms of power over the proposed test does not really justify all the extra computation that has gone into finding a best test through the pseudo power envelope.

Comparing the AD and SSD of the pseudo power envelope with the test proposed in Chapter 4, we observe that in both circumstances the boundary points of our proposed test are very similar to the boundary points of the test that has the second smallest AD

and SSD. Note that in some cases, the ranks for both AD and SSD are same, but in the other cases the ranks are different. Overall we claim that our high size/low size technique is an appropriate one for choosing the boundary points of the subregions under H_0 and our proposed test is also not too different from the best test.

5.4 Concluding remarks

In this chapter we have constructed a pseudo power envelope for the new test procedure, with the aim of investigating the results of choosing different boundary points of the three subregions under H_0 . We applied one hundred and seventy-one different versions of the new test procedure for testing MA(1) disturbances against AR(1) disturbances in the linear regression model to construct the pseudo power envelope. From the pseudo power envelope we find that the best test is test 64 which has the least AD and the second best test is 74. We compared the AD and SSD from the power envelope of the powers of our tests used to calculate this envelope with the AD and SSD of the power of the test proposed in Chapter 4. This comparison shows that the boundary points of our proposed test are very close to the boundary points of test 74. Also, as mentioned in the previous section, through identifying the best test by reference to the pseudo power envelope, we gained only 1.8 percentage points in terms of power improvement. Although we limited our experiment to a single design matrix because of computational and time constraints, the results are very encouraging. They indicate that without spending a lot of time for computation to construct the pseudo power envelope we can get reasonably close to the best test by using the high size/low size technique. Therefore the large amount of extra computation required finding the best test via a pseudo power envelope does not appear to be justified. Hence, we can say that for our proposed new

test procedure we can use the high size/low size technique to choose the boundary points of the subregions of Θ with some confidence.

Table 5.1

Ranges of three disjoint intervals and the corresponding values of j , where $\rho_j = \gamma_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$: X2 design matrix with $n = 20$.

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
1	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.1$ ($j=4, 5$)	$0.1 < \gamma \leq 1$ ($j=6, 7, \dots, 50$)
2	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.15$ ($j=4, 5, \dots, 8$)	$0.15 < \gamma \leq 1$ ($j=9, 10, \dots, 50$)
3	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.2$ ($j=4, 5, \dots, 10$)	$0.2 < \gamma \leq 1$ ($j=11, 12, \dots, 50$)
4	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.25$ ($j=4, 5, \dots, 13$)	$0.25 < \gamma \leq 1$ ($j=14, 15, \dots, 50$)
5	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.3$ ($j=4, 5, \dots, 15$)	$0.3 < \gamma \leq 1$ ($j=16, 17, \dots, 50$)
6	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.35$ ($j=4, 5, \dots, 18$)	$0.35 < \gamma \leq 1$ ($j=19, 20, \dots, 50$)
7	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.4$ ($j=4, 5, \dots, 20$)	$0.4 < \gamma \leq 1$ ($j=21, 22, \dots, 50$)
8	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.45$ ($j=4, 5, \dots, 23$)	$0.45 < \gamma \leq 1$ ($j=24, 25, \dots, 50$)
9	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.5$ ($j=4, 5, \dots, 25$)	$0.5 < \gamma \leq 1$ ($j=26, 27, \dots, 50$)
10	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.55$ ($j=4, 5, \dots, 28$)	$0.55 < \gamma \leq 1$ ($j=29, 30, \dots, 50$)
11	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.6$ ($j=4, 5, \dots, 30$)	$0.6 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
12	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.65$ ($j=4, 5, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
13	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.7$ ($j=4, 5, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
14	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.75$ ($j=4, 5, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
15	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.8$ ($j=4, 5, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
16	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.85$ ($j=4, 5, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
17	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.9$ ($j=4, 5, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
18	$0 \leq \gamma \leq 0.05$ ($j=1, 2, 3$)	$0.05 < \gamma \leq 0.95$ ($j=4, 5, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)

Table 5.1 (contr'n'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
19	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.15$ ($j=6, 7, 8$)	$0.15 < \gamma \leq 1$ ($j=9, 10, \dots, 50$)
20	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.2$ ($j=6, 7, \dots, 10$)	$0.2 < \gamma \leq 1$ ($j=11, 12, \dots, 50$)
21	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.25$ ($j=6, 7, \dots, 13$)	$0.25 < \gamma \leq 1$ ($j=14, 15, \dots, 50$)
22	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.3$ ($j=6, 7, \dots, 15$)	$0.3 < \gamma \leq 1$ ($j=16, 17, \dots, 50$)
23	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.35$ ($j=6, 7, \dots, 18$)	$0.35 < \gamma \leq 1$ ($j=19, 20, \dots, 50$)
24	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.4$ ($j=6, 7, \dots, 20$)	$0.4 < \gamma \leq 1$ ($j=21, 22, \dots, 50$)
25	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.45$ ($j=6, 7, \dots, 23$)	$0.45 < \gamma \leq 1$ ($j=24, 25, \dots, 50$)
26	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.5$ ($j=6, 7, \dots, 25$)	$0.5 < \gamma \leq 1$ ($j=26, 27, \dots, 50$)
27	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.55$ ($j=6, 7, \dots, 28$)	$0.55 < \gamma \leq 1$ ($j=29, 30, \dots, 50$)
28	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.6$ ($j=6, 7, \dots, 30$)	$0.6 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
29	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.65$ ($j=6, 7, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
30	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.7$ ($j=6, 7, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
31	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.75$ ($j=6, 7, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
32	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.8$ ($j=6, 7, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
33	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.85$ ($j=6, 7, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
34	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.9$ ($j=6, 7, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
35	$0 \leq \gamma \leq 0.1$ ($j=1, 2, \dots, 5$)	$0.1 < \gamma \leq 0.95$ ($j=6, 7, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
36	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.2$ ($j=9, 10$)	$0.2 < \gamma \leq 1$ ($j=11, 12, \dots, 50$)
37	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.25$ ($j=6, 7, \dots, 13$)	$0.25 < \gamma \leq 1$ ($j=14, 15, \dots, 50$)
38	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.3$ ($j=9, 10, \dots, 15$)	$0.3 < \gamma \leq 1$ ($j=16, 17, \dots, 50$)

Table 5.1 (contn'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
39	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.35$ ($j=9, 10, \dots, 18$)	$0.35 < \gamma \leq 1$ ($j=19, 20, \dots, 50$)
40	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.4$ ($j=9, 10, \dots, 20$)	$0.4 < \gamma \leq 1$ ($j=21, 22, \dots, 50$)
41	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.45$ ($j=9, 10, \dots, 23$)	$0.45 < \gamma \leq 1$ ($j=24, 25, \dots, 50$)
42	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.5$ ($j=9, 10, \dots, 25$)	$0.5 < \gamma \leq 1$ ($j=26, 27, \dots, 50$)
43	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.55$ ($j=9, 10, \dots, 28$)	$0.55 < \gamma \leq 1$ ($j=29, 30, \dots, 50$)
44	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.6$ ($j=9, 10, \dots, 30$)	$0.6 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
45	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.65$ ($j=9, 10, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
46	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.7$ ($j=9, 10, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
47	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.75$ ($j=9, 10, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
48	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.8$ ($j=9, 10, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
49	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.85$ ($j=9, 10, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
50	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.9$ ($j=9, 10, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
51	$0 \leq \gamma \leq 0.15$ ($j=1, 2, \dots, 8$)	$0.15 < \gamma \leq 0.95$ ($j=9, 10, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
52	$0 \leq \gamma \leq 0.2$ ($j=1, 2, \dots, 10$)	$0.2 < \gamma \leq 0.25$ ($j=11, 12, 13$)	$0.25 < \gamma \leq 1$ ($j=14, 15, \dots, 50$)
53	$0 \leq \gamma \leq 0.2$ ($j=1, 2, \dots, 10$)	$0.2 < \gamma \leq 0.3$ ($j=11, 12, \dots, 15$)	$0.3 < \gamma \leq 1$ ($j=16, 17, \dots, 50$)
54	$0 \leq \gamma \leq 0.2$ ($j=1, 2, \dots, 10$)	$0.2 < \gamma \leq 0.35$ ($j=11, 12, \dots, 18$)	$0.35 < \gamma \leq 1$ ($j=19, 20, \dots, 50$)
55	$0 \leq \gamma \leq 0.2$ ($j=1, 2, \dots, 10$)	$0.2 < \gamma \leq 0.4$ ($j=11, 12, \dots, 20$)	$0.4 < \gamma \leq 1$ ($j=21, 22, \dots, 50$)
56	$0 \leq \gamma \leq 0.2$ ($j=1, 2, \dots, 10$)	$0.2 < \gamma \leq 0.45$ ($j=11, 12, \dots, 23$)	$0.45 < \gamma \leq 1$ ($j=24, 25, \dots, 50$)
57	$0 \leq \gamma \leq 0.2$ ($j=1, 2, \dots, 10$)	$0.2 < \gamma \leq 0.5$ ($j=11, 12, \dots, 25$)	$0.5 < \gamma \leq 1$ ($j=26, 27, \dots, 50$)
58	$0 \leq \gamma \leq 0.2$ ($j=1, 2, \dots, 10$)	$0.2 < \gamma \leq 0.55$ ($j=11, 12, \dots, 28$)	$0.55 < \gamma \leq 1$ ($j=29, 30, \dots, 50$)

Table 5.1 (contn'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
59	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.6$ (j=11, 12, ..., 30)	$0.6 < \gamma \leq 1$ (j=31, 32, ..., 50)
60	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.65$ (j=11, 12, ..., 33)	$0.65 < \gamma \leq 1$ (j=34, 35, ..., 50)
61	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.7$ (j=11, 12, ..., 35)	$0.7 < \gamma \leq 1$ (j=36, 37, ..., 50)
62	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.75$ (j=11, 12, ..., 38)	$0.75 < \gamma \leq 1$ (j=39, 40, ..., 50)
63	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.8$ (j=11, 12, ..., 40)	$0.8 < \gamma \leq 1$ (j=41, 42, ..., 50)
64	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.85$ (j=11, 12, ..., 43)	$0.85 < \gamma \leq 1$ (j=44, 45, ..., 50)
65	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.9$ (j=11, 12, ..., 45)	$0.9 < \gamma \leq 1$ (j=46, 47, ..., 50)
66	$0 \leq \gamma \leq 0.2$ (j=1, 2, ..., 10)	$0.2 < \gamma \leq 0.95$ (j=11, 12, ..., 48)	$0.95 < \gamma \leq 1$ (j=49, 50)
67	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.3$ (j=14, 15)	$0.3 < \gamma \leq 1$ (j=16, 17, ..., 50)
68	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.35$ (j=14, 15, ..., 18)	$0.35 < \gamma \leq 1$ (j=19, 20, ..., 50)
69	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.4$ (j=14, 15, ..., 20)	$0.4 < \gamma \leq 1$ (j=21, 22, ..., 50)
70	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.45$ (j=14, 15, ..., 23)	$0.45 < \gamma \leq 1$ (j=24, 25, ..., 50)
71	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.5$ (j=14, 15, ..., 25)	$0.5 < \gamma \leq 1$ (j=26, 27, ..., 50)
72	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.55$ (j=14, 15, ..., 28)	$0.55 < \gamma \leq 1$ (j=29, 30, ..., 50)
73	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.6$ (j=14, 15, ..., 30)	$0.6 < \gamma \leq 1$ (j=31, 32, ..., 50)
74	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.65$ (j=14, 15, ..., 33)	$0.65 < \gamma \leq 1$ (j=34, 35, ..., 50)
75	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.7$ (j=14, 15, ..., 35)	$0.7 < \gamma \leq 1$ (j=36, 37, ..., 50)
76	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.75$ (j=14, 15, ..., 38)	$0.75 < \gamma \leq 1$ (j=39, 40, ..., 50)
77	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.8$ (j=14, 15, ..., 40)	$0.8 < \gamma \leq 1$ (j=41, 42, ..., 50)
78	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.85$ (j=14, 15, ..., 43)	$0.85 < \gamma \leq 1$ (j=44, 45, ..., 50)

Table 5.1 (contn'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
79	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.9$ (j=14, 15, ..., 45)	$0.9 < \gamma \leq 1$ (j=46, 47, ..., 50)
80	$0 \leq \gamma \leq 0.25$ (j=1, 2, ..., 13)	$0.25 < \gamma \leq 0.95$ (j=14, 15, ..., 48)	$0.95 < \gamma \leq 1$ (j=49, 50)
81	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.35$ (j=16, 17, 18)	$0.35 < \gamma \leq 1$ (j=19, 20, ..., 50)
82	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.4$ (j=16, 17, ..., 20)	$0.4 < \gamma \leq 1$ (j=21, 22, ..., 50)
83	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.45$ (j=16, 17, ..., 23)	$0.45 < \gamma \leq 1$ (j=24, 25, ..., 50)
84	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.5$ (j=16, 17, ..., 25)	$0.5 < \gamma \leq 1$ (j=26, 27, ..., 50)
85	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.55$ (j=16, 17, ..., 28)	$0.55 < \gamma \leq 1$ (j=29, 30, ..., 50)
86	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.6$ (j=16, 17, ..., 30)	$0.6 < \gamma \leq 1$ (j=31, 32, ..., 50)
87	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.65$ (j=16, 17, ..., 33)	$0.65 < \gamma \leq 1$ (j=34, 35, ..., 50)
88	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.7$ (j=16, 17, ..., 35)	$0.7 < \gamma \leq 1$ (j=36, 37, ..., 50)
89	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.75$ (j=16, 17, ..., 38)	$0.75 < \gamma \leq 1$ (j=39, 40, ..., 50)
90	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.8$ (j=16, 17, ..., 40)	$0.8 < \gamma \leq 1$ (j=41, 42, ..., 50)
91	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.85$ (j=16, 17, ..., 43)	$0.85 < \gamma \leq 1$ (j=44, 45, ..., 50)
92	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.9$ (j=16, 17, ..., 45)	$0.9 < \gamma \leq 1$ (j=46, 47, ..., 50)
93	$0 \leq \gamma \leq 0.3$ (j=1, 2, ..., 15)	$0.3 < \gamma \leq 0.95$ (j=16, 17, ..., 48)	$0.95 < \gamma \leq 1$ (j=49, 50)
94	$0 \leq \gamma \leq 0.35$ (j=1, 2, ..., 18)	$0.35 < \gamma \leq 0.4$ (j=19, 20)	$0.4 < \gamma \leq 1$ (j=21, 22, ..., 50)
95	$0 \leq \gamma \leq 0.35$ (j=1, 2, ..., 18)	$0.35 < \gamma \leq 0.45$ (j=19, 20, ..., 23)	$0.45 < \gamma \leq 1$ (j=24, 25, ..., 50)
96	$0 \leq \gamma \leq 0.35$ (j=1, 2, ..., 18)	$0.35 < \gamma \leq 0.5$ (j=19, 20, ..., 25)	$0.5 < \gamma \leq 1$ (j=26, 27, ..., 50)
97	$0 \leq \gamma \leq 0.35$ (j=1, 2, ..., 18)	$0.35 < \gamma \leq 0.55$ (j=19, 20, ..., 28)	$0.55 < \gamma \leq 1$ (j=29, 30, ..., 50)
98	$0 \leq \gamma \leq 0.35$ (j=1, 2, ..., 18)	$0.35 < \gamma \leq 0.6$ (j=19, 20, ..., 30)	$0.6 < \gamma \leq 1$ (j=31, 32, ..., 50)

Table 5.1 (contn'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
99	$0 \leq \gamma \leq 0.35$ ($j=1, 2, \dots, 18$)	$0.35 < \gamma \leq 0.65$ ($j=19, 20, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
100	$0 \leq \gamma \leq 0.35$ ($j=1, 2, \dots, 18$)	$0.35 < \gamma \leq 0.7$ ($j=19, 20, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
101	$0 \leq \gamma \leq 0.35$ ($j=1, 2, \dots, 18$)	$0.35 < \gamma \leq 0.75$ ($j=19, 20, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
102	$0 \leq \gamma \leq 0.35$ ($j=1, 2, \dots, 18$)	$0.35 < \gamma \leq 0.8$ ($j=19, 20, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
103	$0 \leq \gamma \leq 0.35$ ($j=1, 2, \dots, 18$)	$0.35 < \gamma \leq 0.85$ ($j=19, 20, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
104	$0 \leq \gamma \leq 0.35$ ($j=1, 2, \dots, 18$)	$0.35 < \gamma \leq 0.9$ ($j=19, 20, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
105	$0 \leq \gamma \leq 0.35$ ($j=1, 2, \dots, 18$)	$0.35 < \gamma \leq 0.95$ ($j=19, 20, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
106	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.45$ ($j=21, 22, 23$)	$0.45 < \gamma \leq 1$ ($j=24, 25, \dots, 50$)
107	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.5$ ($j=21, 22, \dots, 25$)	$0.5 < \gamma \leq 1$ ($j=26, 27, \dots, 50$)
108	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.55$ ($j=21, 22, \dots, 28$)	$0.55 < \gamma \leq 1$ ($j=29, 30, \dots, 50$)
109	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.6$ ($j=21, 22, \dots, 30$)	$0.6 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
110	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.65$ ($j=21, 22, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
111	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.7$ ($j=21, 22, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
112	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.75$ ($j=21, 22, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
113	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.8$ ($j=21, 22, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
114	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.85$ ($j=21, 22, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
115	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.9$ ($j=21, 22, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
116	$0 \leq \gamma \leq 0.4$ ($j=1, 2, \dots, 20$)	$0.4 < \gamma \leq 0.95$ ($j=21, 22, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
117	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.5$ ($j=24, 25$)	$0.5 < \gamma \leq 1$ ($j=26, 27, \dots, 50$)
118	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.55$ ($j=24, 25, \dots, 28$)	$0.55 < \gamma \leq 1$ ($j=29, 30, \dots, 50$)

Table 5.1 (contn'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
119	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.6$ ($j=24, 25, \dots, 30$)	$0.6 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
120	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.65$ ($j=24, 25, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
121	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.7$ ($j=24, 25, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
122	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.75$ ($j=24, 25, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
123	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.8$ ($j=24, 25, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
124	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.85$ ($j=24, 25, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
125	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.9$ ($j=24, 25, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
126	$0 \leq \gamma \leq 0.45$ ($j=1, 2, \dots, 23$)	$0.45 < \gamma \leq 0.95$ ($j=24, 25, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
127	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.55$ ($j=26, 27, 28$)	$0.55 < \gamma \leq 1$ ($j=29, 30, \dots, 50$)
128	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.6$ ($j=26, 27, \dots, 30$)	$0.6 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
129	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.65$ ($j=26, 27, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
130	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.7$ ($j=26, 27, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
131	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.75$ ($j=26, 27, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
132	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.8$ ($j=26, 27, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
133	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.85$ ($j=26, 27, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
134	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.9$ ($j=26, 27, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
135	$0 \leq \gamma \leq 0.5$ ($j=1, 2, \dots, 25$)	$0.5 < \gamma \leq 0.95$ ($j=26, 27, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
136	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.6$ ($j=29, 30$)	$0.6 < \gamma \leq 1$ ($j=31, 32, \dots, 50$)
137	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.65$ ($j=29, 30, \dots, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
138	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.7$ ($j=29, 30, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)

Table 5.1 (contn'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
139	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.75$ ($j=29, 30, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
140	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.8$ ($j=29, 30, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
141	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.85$ ($j=29, 30, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
142	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.9$ ($j=29, 30, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
143	$0 \leq \gamma \leq 0.55$ ($j=1, 2, \dots, 28$)	$0.55 < \gamma \leq 0.95$ ($j=29, 30, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
144	$0 \leq \gamma \leq 0.6$ ($j=1, 2, \dots, 30$)	$0.6 < \gamma \leq 0.65$ ($j=31, 32, 33$)	$0.65 < \gamma \leq 1$ ($j=34, 35, \dots, 50$)
145	$0 \leq \gamma \leq 0.6$ ($j=1, 2, \dots, 30$)	$0.6 < \gamma \leq 0.7$ ($j=31, 32, \dots, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
146	$0 \leq \gamma \leq 0.6$ ($j=1, 2, \dots, 30$)	$0.6 < \gamma \leq 0.75$ ($j=31, 32, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
147	$0 \leq \gamma \leq 0.6$ ($j=1, 2, \dots, 30$)	$0.6 < \gamma \leq 0.8$ ($j=31, 32, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
148	$0 \leq \gamma \leq 0.6$ ($j=1, 2, \dots, 30$)	$0.6 < \gamma \leq 0.85$ ($j=31, 32, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
149	$0 \leq \gamma \leq 0.6$ ($j=1, 2, \dots, 30$)	$0.6 < \gamma \leq 0.9$ ($j=31, 32, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
150	$0 \leq \gamma \leq 0.6$ ($j=1, 2, \dots, 30$)	$0.6 < \gamma \leq 0.95$ ($j=31, 32, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
151	$0 \leq \gamma \leq 0.65$ ($j=1, 2, \dots, 33$)	$0.65 < \gamma \leq 0.7$ ($j=34, 35$)	$0.7 < \gamma \leq 1$ ($j=36, 37, \dots, 50$)
152	$0 \leq \gamma \leq 0.65$ ($j=1, 2, \dots, 33$)	$0.65 < \gamma \leq 0.75$ ($j=34, 35, \dots, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
153	$0 \leq \gamma \leq 0.65$ ($j=1, 2, \dots, 33$)	$0.65 < \gamma \leq 0.8$ ($j=34, 35, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
154	$0 \leq \gamma \leq 0.65$ ($j=1, 2, \dots, 33$)	$0.65 < \gamma \leq 0.85$ ($j=34, 35, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
155	$0 \leq \gamma \leq 0.65$ ($j=1, 2, \dots, 33$)	$0.65 < \gamma \leq 0.9$ ($j=34, 35, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
156	$0 \leq \gamma \leq 0.65$ ($j=1, 2, \dots, 33$)	$0.65 < \gamma \leq 0.95$ ($j=34, 35, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
157	$0 \leq \gamma \leq 0.7$ ($j=1, 2, \dots, 35$)	$0.7 < \gamma \leq 0.75$ ($j=36, 37, 38$)	$0.75 < \gamma \leq 1$ ($j=39, 40, \dots, 50$)
158	$0 \leq \gamma \leq 0.7$ ($j=1, 2, \dots, 35$)	$0.7 < \gamma \leq 0.8$ ($j=36, 37, \dots, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)

Table 5.1 (contn'd)

Test number	Subregions		
	Θ_1	Θ_2	Θ_3
159	$0 \leq \gamma \leq 0.7$ ($j=1, 2, \dots, 35$)	$0.7 < \gamma \leq 0.85$ ($j=36, 37, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
160	$0 \leq \gamma \leq 0.7$ ($j=1, 2, \dots, 35$)	$0.7 < \gamma \leq 0.9$ ($j=36, 37, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
161	$0 \leq \gamma \leq 0.7$ ($j=1, 2, \dots, 35$)	$0.7 < \gamma \leq 0.95$ ($j=36, 37, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
162	$0 \leq \gamma \leq 0.75$ ($j=1, 2, \dots, 38$)	$0.75 < \gamma \leq 0.8$ ($j=39, 40$)	$0.8 < \gamma \leq 1$ ($j=41, 42, \dots, 50$)
163	$0 \leq \gamma \leq 0.75$ ($j=1, 2, \dots, 38$)	$0.75 < \gamma \leq 0.85$ ($j=39, 40, \dots, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
164	$0 \leq \gamma \leq 0.75$ ($j=1, 2, \dots, 38$)	$0.75 < \gamma \leq 0.9$ ($j=39, 40, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
165	$0 \leq \gamma \leq 0.75$ ($j=1, 2, \dots, 38$)	$0.75 < \gamma \leq 0.95$ ($j=39, 40, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
166	$0 \leq \gamma \leq 0.8$ ($j=1, 2, \dots, 40$)	$0.8 < \gamma \leq 0.85$ ($j=41, 42, 43$)	$0.85 < \gamma \leq 1$ ($j=44, 45, \dots, 50$)
167	$0 \leq \gamma \leq 0.8$ ($j=1, 2, \dots, 40$)	$0.8 < \gamma \leq 0.9$ ($j=41, 42, \dots, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
168	$0 \leq \gamma \leq 0.8$ ($j=1, 2, \dots, 40$)	$0.8 < \gamma \leq 0.95$ ($j=41, 42, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
169	$0 \leq \gamma \leq 0.85$ ($j=1, 2, \dots, 43$)	$0.85 < \gamma \leq 0.9$ ($j=44, 45$)	$0.9 < \gamma \leq 1$ ($j=46, 47, \dots, 50$)
170	$0 \leq \gamma \leq 0.85$ ($j=1, 2, \dots, 43$)	$0.85 < \gamma \leq 0.95$ ($j=44, 45, \dots, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
171	$0 \leq \gamma \leq 0.9$ ($j=1, 2, \dots, 45$)	$0.9 < \gamma \leq 0.95$ ($j=46, 47, 48$)	$0.95 < \gamma \leq 1$ ($j=49, 50$)
Our test	$0 \leq \gamma \leq 0.23$ ($j=1, 2, \dots, 12$)	$0.23 < \gamma \leq 0.63$ ($j=13, 14, \dots, 32$)	$0.63 < \gamma \leq 1$ ($j=33, 34, \dots, 50$)

Table 5.2

Final critical values k_1 , k_2 and k_3 when $m = 3$: X^2 design matrix with $n = 20$.

Test number	Critical values			Test number	Critical values		
	k_1	k_2	k_3		k_1	k_2	k_3
1	-0.0837	0.255	2.233	44	0.0706	1.7	0.4
2	-0.025	0.25	2.2	45	0.0815	1.8	0.1
3	-0.0544	0.437	2.18	46	0.0616	2.15	-0.01
4	-0.0252	0.49	2.195	47	0.0658	2.3	-0.18
5	-0.1167	0.93	1.59	48	0.0729	2.41	-0.23
6	-0.1464	1.2	1.2	49	0.0831	2.48	-0.37
7	-0.162	1.45	0.93	50	0.0758	2.74	-0.452
8	-0.1325	1.5	0.87	51	0.0761	2.91	-0.78
9	-0.1321	1.67	0.82	52	0.0155	0.61	1.8
10	-0.124	1.8	0.5	53	-0.0425	1.15	1.01
11	-0.1456	2.1	0.28	54	0.0222	1.15	0.95
12	-0.1257	2.17	0.07	55	0.123	1.0	1.1
13	-0.0845	2.05	0.07	56	0.0595	1.5	0.5
14	-0.0872	2.23	-0.1	57	0.1172	1.44	0.6
15	-0.081	2.34	-0.13	58	0.1278	1.59	0.35
16	-0.0814	2.5	-0.316	59	0.146	1.68	0.35
17	-0.0609	2.49	-0.33	60	0.1391	1.72	0.31
18	-0.0646	2.67	-0.606	61	0.1676	1.9	-0.001
19	-0.0123	0.21	2.21	62	0.1564	2.19	-0.19
20	-0.0228	0.39	2.1	63	0.1941	2.06	-0.174
21	-0.0051	0.475	2.22	64	0.2145	2.09	-0.3
22	-0.2373	1.33	1.05	65	0.2058	2.3	-0.39
23	-0.0805	1.05	1.25	66	0.1774	2.74	-0.785
24	-0.1044	1.33	0.93	67	0.029	1.58	0.8
25	-0.1079	1.52	0.63	68	0.1301	1.17	0.9
26	-0.0231	1.35	0.81	69	0.1778	1.23	0.82
27	-0.0526	1.65	0.5	70	0.1767	1.47	0.45
28	-0.0231	1.67	0.44	71	0.1798	1.75	0.37
29	-0.0679	2.06	0.065	72	0.2043	1.8	0.21
30	-0.074	2.27	-0.02	73	0.2344	1.8	0.2
31	-0.0067	2.05	-0.07	74	0.2715	1.71	0.003
32	-0.0217	2.3	-0.165	75	0.257	2.03	-0.09
33	-0.0103	2.38	-0.31	76	0.2432	2.4	-0.28
34	-0.0021	2.475	-0.36	77	0.2617	2.4	-0.3
35	-0.0113	2.7	-0.67	78	0.3054	2.21	-0.38
36	-0.0088	0.56	2.0	79	0.3063	2.38	-0.45
37	0.0139	0.595	2.05	80	0.2838	2.8	-0.87
38	-0.1833	1.5	0.87	81	0.1268	1.31	0.5
39	-0.1074	1.44	0.94	82	0.1633	1.7	0.45
40	0.0209	1.22	0.9	83	0.2727	1.31	0.43
41	0.0378	1.32	0.75	84	0.275	1.6	0.37
42	0.0155	1.64	0.62	85	0.3041	1.65	0.18
43	0.0578	1.6	0.49	86	0.3337	1.68	0.15

Table 5.2 (contn'd)

Test number	Critical values			Test number	Critical values		
	k_1	k_2	k_3		k_1	k_2	k_3
87	0.3463	1.8	-0.08	130	0.9659	1.51	-0.25
88	0.3497	2.0	-0.15	131	1.017	1.55	-0.35
89	0.365	2.1	-0.26	132	1.1196	1.42	-0.305
90	0.3733	2.26	-0.315	133	1.2003	1.22	-0.33
91	0.4023	2.25	-0.444	134	1.2142	1.32	-0.382
92	0.4363	2.3	-0.478	135	1.1393	1.83	-0.87
93	0.4922	2.25	-0.735	136	1.1383	1.5	-0.132
94	0.3644	1.44	0.66	137	1.1861	1.32	-0.25
95	0.3978	1.34	0.37	138	1.1621	1.6	-0.3313
96	0.4106	1.5	0.36	139	1.328	1.23	-0.31
97	0.4247	1.64	0.12	140	1.4252	1.01	-0.215
98	0.4246	1.98	-0.02	141	1.6307	0.54	-0.06
99	0.4488	2.0	-0.17	142	1.6087	0.77	-0.21
100	0.4813	2.0	-0.195	143	1.4991	1.26	-0.658
101	0.515	2.03	-0.297	144	1.2591	1.2	-0.37
102	0.5137	2.25	-0.357	145	1.347	1.35	-0.365
103	0.4915	2.6	-0.59	146	1.5806	0.68	-0.165
104	0.6851	1.96	-0.432	147	1.6489	0.65	-0.125
105	0.5457	2.72	-0.988	148	1.6687	0.65	-0.175
106	0.4494	1.27	0.24	149	1.5959	0.94	-0.345
107	0.525	1.3	0.35	150	1.521	1.63	-0.985
108	0.5605	1.38	0.17	151	1.5921	1.35	-0.397
109	0.5716	1.65	0.025	152	1.6158	1.2	-0.5
110	0.7117	1.35	-0.05	153	1.7768	0.54	-0.15
111	0.7229	1.5	-0.11	154	1.797	0.47	-0.155
112	0.654	1.95	-0.335	155	1.8559	0.3	-0.06
113	0.6889	2.02	-0.357	156	1.7598	1.005	-0.69
114	0.8187	1.7	-0.39	157	1.4471	1.9	-1.1
115	0.6967	2.405	-0.62	158	1.9115	0.2	-0.02
116	0.6913	2.7	-1.076	159	2.0243	-0.03	0.112
117	0.6716	1.54	0.3	160	1.9253	0.23	-0.063
118	0.6788	1.59	-0.046	161	1.8688	0.59	-0.48
119	0.9391	1.0	0.35	162	2.1168	-0.07	0.03
120	0.727	1.9	-0.27	163	2.1433	-0.197	0.157
121	0.8766	1.6	-0.21	164	2.137	-0.195	0.13
122	0.917	1.58	-0.29	165	2.0335	0.35	-0.38
123	0.9748	1.61	-0.3	166	2.0885	1.1	-1.09
124	1.0152	1.6	-0.409	167	2.123	0.78	-0.655
125	0.8707	2.45	-0.7	168	2.0932	1.4	-1.54
126	0.9889	2.08	-0.898	169	2.6104	-1.7	0.95
127	0.8507	1.1	-0.03	170	2.5871	-1.19	1.01
128	0.8327	1.62	-0.15	171	2.3732	0.8	-1.444
129	0.9224	1.43	-0.22	Our test	0.2098	1.84	0.055

Table 5.3

Calculated sizes of the test when $m = 3$ at the five percent level: X^2 design matrix with $n = 20$.

Parameter values	Sizes									
	Test number									
	1	2	3	4**	5	6	7	8	9**	10**
$\gamma = 0.0$	0.047	0.048	0.047	0.049	0.043	0.046	0.046	0.046	0.047	0.049
0.1	0.056	0.055	0.046	0.050	0.044	0.040	0.039	0.038	0.038	0.038
0.2	0.069*	0.068*	0.058	0.058	0.051	0.049	0.044	0.043	0.038	0.040
0.3	0.078*	0.077*	0.065*	0.063	0.063	0.060	0.057	0.055	0.050	0.050
0.4	0.080*	0.075*	0.067*	0.063	0.071*	0.070*	0.066*	0.063	0.058	0.060
0.5	0.071*	0.068*	0.059	0.056	0.067*	0.073*	0.074*	0.066*	0.059	0.062
0.6	0.056	0.053	0.051	0.047	0.057	0.065*	0.064*	0.060	0.054	0.058
0.7	0.046	0.044	0.043	0.040	0.046	0.052	0.055	0.050	0.046	0.052
0.8	0.032*	0.030*	0.028*	0.024*	0.037	0.043	0.045	0.042	0.040	0.044
0.9	0.022*	0.020*	0.019*	0.016*	0.024*	0.029*	0.036*	0.035*	0.029*	0.038
Test number										
	11**	12**	13**	14**	15**	16**	17	18**	19	20
$\gamma = 0.0$	0.050	0.047	0.047	0.046	0.046	0.046	0.044	0.045	0.049	0.047
0.1	0.037	0.037	0.043	0.042	0.042	0.041	0.044	0.042	0.052	0.045
0.2	0.036*	0.036*	0.048	0.045	0.046	0.044	0.052	0.050	0.063	0.059
0.3	0.045	0.045	0.057	0.053	0.055	0.053	0.060	0.058	0.072*	0.066*
0.4	0.052	0.053	0.061	0.057	0.056	0.056	0.063	0.060	0.072*	0.070*
0.5	0.057	0.057	0.063	0.061	0.060	0.059	0.064*	0.062	0.064*	0.061
0.6	0.058	0.058	0.059	0.059	0.056	0.056	0.060	0.058	0.053	0.052
0.7	0.054	0.060	0.059	0.055	0.053	0.052	0.052	0.050	0.044	0.043
0.8	0.046	0.055	0.051	0.053	0.045	0.048	0.046	0.044	0.029*	0.028*
0.9	0.043	0.053	0.047	0.053	0.044	0.051	0.048	0.048	0.020*	0.020*
Test number										
	21**	22	23	24	25	26	27**	28**	29**	30**
$\gamma = 0.0$	0.045	0.047	0.043	0.045	0.049	0.045	0.047	0.046	0.052	0.055
0.1	0.046	0.036*	0.040	0.038	0.040	0.042	0.040	0.043	0.041	0.041
0.2	0.050	0.043	0.047	0.041	0.041	0.051	0.041	0.047	0.041	0.039
0.3	0.057	0.058	0.059	0.057	0.054	0.061	0.051	0.056	0.048	0.043
0.4	0.058	0.070*	0.067*	0.064*	0.062	0.068*	0.061	0.063	0.052	0.049
0.5	0.054	0.077*	0.070*	0.072*	0.071*	0.068*	0.063	0.063	0.056	0.051
0.6	0.046	0.070*	0.061	0.063	0.066*	0.063	0.058	0.058	0.058	0.055
0.7	0.038	0.058	0.049	0.054	0.060	0.050	0.052	0.052	0.057	0.050
0.8	0.023*	0.046	0.042	0.044	0.049	0.042	0.042	0.042	0.053	0.049
0.9	0.015*	0.037	0.027*	0.035*	0.040	0.032*	0.038	0.036*	0.052	0.048

* denotes value significantly different from the five percent level at the 0.01 level.

** indicates an accepted test with respect to the size condition.

Table 5.3 (contn'd)

Parameter values	Sizes									
	Test number									
	31**	32**	33**	34**	35**	36	37**	38	39	40
$\gamma = 0.0$	0.044	0.047	0.045	0.045	0.044	0.050	0.053	0.072*	0.074*	0.055
0.1	0.043	0.042	0.044	0.044	0.044	0.048	0.046	0.049	0.047	0.047
0.2	0.050	0.047	0.050	0.049	0.049	0.055	0.047	0.045	0.042	0.048
0.3	0.058	0.050	0.054	0.057	0.053	0.063	0.054	0.059	0.044	0.061
0.4	0.059	0.055	0.057	0.058	0.056	0.069*	0.053	0.070*	0.058	0.065*
0.5	0.061	0.056	0.058	0.060	0.058	0.061	0.053	0.079*	0.061	0.073*
0.6	0.060	0.056	0.055	0.057	0.054	0.053	0.045	0.074*	0.058	0.064*
0.7	0.054	0.051	0.050	0.051	0.048	0.044	0.039	0.061	0.052	0.055
0.8	0.048	0.046	0.046	0.044	0.044	0.030*	0.026*	0.052	0.042	0.045
0.9	0.047	0.047	0.049	0.047	0.046	0.021*	0.018*	0.041	0.036*	0.036*
Test number										
	41	42**	43**	44**	45**	46**	47**	48**	49**	50**
$\gamma = 0.0$	0.059	0.061	0.053	0.052	0.053	0.052	0.053	0.052	0.052	0.050
0.1	0.048	0.045	0.047	0.049	0.050	0.045	0.047	0.048	0.048	0.049
0.2	0.048	0.040	0.049	0.050	0.053	0.044	0.045	0.046	0.049	0.047
0.3	0.058	0.047	0.051	0.053	0.054	0.046	0.045	0.047	0.051	0.047
0.4	0.063	0.054	0.058	0.057	0.060	0.049	0.050	0.050	0.054	0.051
0.5	0.067*	0.056	0.060	0.060	0.062	0.050	0.051	0.051	0.053	0.049
0.6	0.061	0.055	0.056	0.053	0.059	0.053	0.051	0.047	0.048	0.044
0.7	0.053	0.048	0.049	0.049	0.058	0.048	0.049	0.046	0.046	0.043
0.8	0.044	0.040	0.043	0.043	0.053	0.046	0.049	0.045	0.043	0.039
0.9	0.036*	0.036*	0.037	0.036*	0.047	0.046	0.049	0.048	0.047	0.043
Test number										
	51**	52**	53	54	55	56	57**	58**	59**	60
$\gamma = 0.0$	0.051	0.058	0.067*	0.064*	0.053	0.067*	0.057	0.057	0.054	0.096*
0.1	0.049	0.044	0.046	0.047	0.051	0.048	0.047	0.048	0.047	0.074*
0.2	0.047	0.044	0.047	0.045	0.053	0.043	0.046	0.046	0.047	0.066*
0.3	0.047	0.053	0.055	0.050	0.061	0.044	0.051	0.050	0.051	0.061
0.4	0.050	0.054	0.067*	0.060	0.064*	0.055	0.058	0.055	0.054	0.059
0.5	0.049	0.054	0.074*	0.062	0.065*	0.060	0.058	0.058	0.056	0.057
0.6	0.045	0.047	0.067*	0.060	0.057	0.062	0.057	0.058	0.052	0.051
0.7	0.041	0.041	0.056	0.052	0.047	0.058	0.049	0.053	0.047	0.044
0.8	0.039	0.030*	0.047	0.042	0.041	0.050	0.042	0.046	0.042	0.037
0.9	0.043	0.021*	0.037	0.036*	0.027*	0.044	0.036*	0.042	0.036*	0.033*

* denotes value significantly different from the five percent level at the 0.01 level.

** indicates an accepted test with respect to the size condition.

Table 5.3 (contn'd)

Parameter values	Sizes									
	Test number									
	61**	62**	63**	64**	65**	66**	67	68	69	70
$\gamma = 0.0$	0.052	0.053	0.046	0.037	0.042	0.050	0.066*	0.066*	0.067*	0.071*
0.1	0.050	0.048	0.051	0.046	0.047	0.049	0.048	0.050	0.050	0.052
0.2	0.051	0.046	0.055	0.056	0.056	0.051	0.045	0.044	0.048	0.045
0.3	0.055	0.047	0.059	0.063	0.062	0.053	0.048	0.048	0.052	0.047
0.4	0.057	0.050	0.061	0.063	0.062	0.054	0.063	0.057	0.060	0.055
0.5	0.058	0.050	0.059	0.062	0.061	0.051	0.069*	0.058	0.060	0.060
0.6	0.057	0.049	0.056	0.057	0.056	0.047	0.069*	0.058	0.058	0.061
0.7	0.051	0.048	0.051	0.051	0.049	0.043	0.060	0.050	0.049	0.057
0.8	0.048	0.046	0.045	0.044	0.043	0.040	0.051	0.042	0.042	0.052
0.9	0.046	0.048	0.046	0.046	0.040	0.043	0.042	0.036*	0.036*	0.043
Test number										
	71	72	73**	74**	75**	76**	77**	78**	79**	80**
$\gamma = 0.0$	0.070*	0.069*	0.063	0.054	0.061	0.063	0.061	0.041	0.041	0.047
0.1	0.052	0.054	0.054	0.055	0.057	0.053	0.055	0.048	0.048	0.050
0.2	0.042	0.045	0.050	0.058	0.054	0.049	0.054	0.056	0.055	0.056
0.3	0.041	0.040	0.046	0.057	0.054	0.046	0.052	0.062	0.062	0.057
0.4	0.048	0.046	0.050	0.061	0.055	0.042	0.053	0.060	0.061	0.055
0.5	0.052	0.049	0.051	0.061	0.051	0.042	0.046	0.057	0.057	0.050
0.6	0.052	0.052	0.050	0.060	0.050	0.042	0.043	0.051	0.049	0.046
0.7	0.052	0.054	0.048	0.059	0.047	0.043	0.041	0.046	0.046	0.039
0.8	0.047	0.048	0.045	0.058	0.047	0.045	0.041	0.041	0.039	0.037
0.9	0.041	0.041	0.039	0.056	0.050	0.052	0.048	0.044	0.043	0.041
Test number										
	81	82	83	84	85**	86**	87**	88**	89**	90**
$\gamma = 0.0$	0.074*	0.074*	0.064*	0.069*	0.063	0.055	0.052	0.053	0.042	0.039
0.1	0.051	0.049	0.053	0.052	0.054	0.053	0.052	0.053	0.049	0.047
0.2	0.040	0.036*	0.048	0.044	0.046	0.051	0.052	0.052	0.055	0.053
0.3	0.040	0.036*	0.047	0.040	0.044	0.051	0.054	0.052	0.056	0.056
0.4	0.047	0.041	0.054	0.046	0.047	0.053	0.054	0.054	0.052	0.054
0.5	0.058	0.052	0.059	0.050	0.050	0.051	0.053	0.049	0.049	0.048
0.6	0.060	0.056	0.060	0.052	0.051	0.050	0.052	0.047	0.046	0.045
0.7	0.060	0.055	0.056	0.051	0.053	0.050	0.055	0.046	0.044	0.041
0.8	0.054	0.052	0.051	0.044	0.048	0.045	0.056	0.049	0.044	0.042
0.9	0.048	0.045	0.044	0.040	0.042	0.041	0.058	0.055	0.051	0.046

* denotes value significantly different from the five percent level at the 0.01 level.

** indicates an accepted test with respect to the size condition.

Table 5.3 (contn'd)

Parameter values	Sizes									
	Test number									
	91**	92**	93**	94**	95**	96**	97**	98**	99**	100**
$\gamma = 0.0$	0.036*	0.030*	0.029*	0.053	0.046	0.052	0.050	0.050	0.041	0.037
0.1	0.045	0.044	0.041	0.047	0.044	0.048	0.048	0.040	0.046	0.046
0.2	0.054	0.051	0.053	0.045	0.045	0.050	0.047	0.045	0.051	0.052
0.3	0.060	0.060	0.061	0.047	0.047	0.051	0.049	0.047	0.050	0.054
0.4	0.058	0.060	0.063	0.054	0.051	0.051	0.048	0.043	0.048	0.055
0.5	0.052	0.055	0.060	0.058	0.056	0.052	0.048	0.041	0.047	0.050
0.6	0.048	0.050	0.053	0.056	0.057	0.051	0.049	0.043	0.045	0.047
0.7	0.042	0.045	0.048	0.048	0.055	0.049	0.051	0.047	0.046	0.046
0.8	0.040	0.039	0.041	0.042	0.051	0.044	0.047	0.047	0.049	0.048
0.9	0.044	0.043	0.043	0.037	0.044	0.040	0.043	0.049	0.056	0.053
Test number										
	101**	102**	103**	104**	105**	106**	107**	108**	109**	110**
$\gamma = 0.0$	0.034*	0.035*	0.040	0.024*	0.031*	0.062	0.037	0.033*	0.032*	0.025*
0.1	0.043	0.044	0.046	0.036*	0.043	0.055	0.044	0.042	0.041	0.037
0.2	0.053	0.054	0.053	0.050	0.051	0.048	0.050	0.047	0.046	0.049
0.3	0.056	0.056	0.054	0.059	0.057	0.043	0.053	0.050	0.049	0.054
0.4	0.053	0.052	0.050	0.063	0.052	0.045	0.054	0.054	0.049	0.059
0.5	0.052	0.048	0.044	0.059	0.049	0.049	0.054	0.051	0.048	0.060
0.6	0.046	0.044	0.040	0.053	0.044	0.051	0.050	0.048	0.045	0.061
0.7	0.043	0.039	0.034*	0.048	0.037	0.054	0.046	0.047	0.044	0.054
0.8	0.047	0.040	0.035*	0.042	0.032*	0.050	0.044	0.042	0.044	0.053
0.9	0.053	0.044	0.041	0.043	0.033*	0.044	0.038	0.040	0.044	0.053
Test number										
	111**	112**	113**	114**	115**	116**	117**	118**	119**	120**
$\gamma = 0.0$	0.026*	0.028*	0.027*	0.023*	0.027*	0.027*	0.031*	0.030*	0.020*	0.029*
0.1	0.038	0.039	0.039	0.035*	0.040	0.039	0.041	0.039	0.031*	0.039
0.2	0.048	0.050	0.049	0.049	0.049	0.048	0.047	0.047	0.044	0.048
0.3	0.055	0.054	0.056	0.059	0.056	0.055	0.051	0.048	0.055	0.049
0.4	0.059	0.052	0.052	0.063	0.051	0.049	0.054	0.049	0.058	0.051
0.5	0.057	0.049	0.051	0.059	0.049	0.046	0.052	0.048	0.059	0.047
0.6	0.056	0.045	0.047	0.054	0.045	0.044	0.049	0.046	0.053	0.043
0.7	0.050	0.042	0.040	0.048	0.038	0.035*	0.044	0.048	0.045	0.044
0.8	0.047	0.047	0.039	0.043	0.033*	0.031*	0.040	0.051	0.038	0.052
0.9	0.049	0.054	0.045	0.046	0.038	0.034*	0.037	0.053	0.034*	0.059

* denotes value significantly different from the five percent level at the 0.01 level.

** indicates an accepted test with respect to the size condition.

Table 5.3 (contn'd)

Parameter values	Sizes									
	Test number									
	121**	122**	123**	124**	125**	126**	127**	128**	129**	130**
$\gamma = 0.0$	0.023*	0.022*	0.020*	0.019*	0.023*	0.021*	0.030*	0.029	0.026*	0.024*
0.1	0.033*	0.033*	0.030*	0.031*	0.036*	0.030*	0.039	0.039	0.041	0.038
0.2	0.046	0.046	0.046	0.045	0.046	0.047	0.048	0.049	0.050	0.050
0.3	0.053	0.056	0.055	0.057	0.053	0.057	0.056	0.052	0.055	0.057
0.4	0.053	0.059	0.060	0.060	0.050	0.060	0.055	0.051	0.052	0.056
0.5	0.055	0.054	0.054	0.056	0.046	0.055	0.054	0.047	0.054	0.055
0.6	0.053	0.051	0.050	0.052	0.044	0.049	0.053	0.044	0.050	0.052
0.7	0.049	0.047	0.046	0.048	0.035*	0.044	0.050	0.043	0.047	0.047
0.8	0.049	0.046	0.042	0.041	0.031*	0.039	0.051	0.046	0.048	0.048
0.9	0.052	0.051	0.047	0.043	0.034*	0.042	0.052	0.049	0.052	0.052
Test number										
	131**	132**	133	134	135**	136**	137**	138**	139**	140**
$\gamma = 0.0$	0.022*	0.020*	0.019*	0.019*	0.020*	0.021*	0.021*	0.022*	0.016*	0.015*
0.1	0.036*	0.031*	0.031*	0.030*	0.033*	0.032*	0.030*	0.033*	0.027*	0.026*
0.2	0.047	0.047	0.049	0.050	0.049	0.044	0.043	0.045	0.043	0.044
0.3	0.056	0.058	0.062	0.063	0.061	0.051	0.053	0.055	0.053	0.056
0.4	0.058	0.061	0.068*	0.068*	0.063	0.054	0.054	0.055	0.058	0.060
0.5	0.054	0.057	0.064*	0.064*	0.058	0.052	0.053	0.051	0.054	0.061
0.6	0.050	0.051	0.060	0.062	0.052	0.052	0.050	0.048	0.051	0.057
0.7	0.045	0.048	0.053	0.052	0.048	0.048	0.047	0.044	0.047	0.051
0.8	0.046	0.044	0.045	0.045	0.042	0.048	0.047	0.047	0.044	0.045
0.9	0.051	0.046	0.046	0.046	0.043	0.049	0.051	0.054	0.050	0.046
Test number										
	141	142	143**	144**	145**	146**	147**	148	149	150**
$\gamma = 0.0$	0.014*	0.013*	0.014*	0.022*	0.019*	0.015*	0.014*	0.015*	0.015*	0.015*
0.1	0.022*	0.022*	0.022*	0.034*	0.030*	0.028*	0.024*	0.024*	0.026*	0.026*
0.2	0.039	0.039	0.041	0.045	0.045	0.044	0.042	0.042	0.044	0.042
0.3	0.055	0.054	0.054	0.054	0.055	0.056	0.055	0.056	0.058	0.055
0.4	0.065*	0.062	0.058	0.054	0.057	0.061	0.060	0.062	0.064*	0.059
0.5	0.069*	0.065*	0.061	0.052	0.052	0.063	0.063	0.065*	0.064*	0.056
0.6	0.063	0.060	0.056	0.048	0.050	0.060	0.058	0.059	0.061	0.051
0.7	0.059	0.057	0.051	0.046	0.046	0.055	0.055	0.055	0.053	0.046
0.8	0.052	0.050	0.045	0.048	0.049	0.049	0.046	0.047	0.045	0.040
0.9	0.047	0.049	0.049	0.057	0.056	0.050	0.044	0.042	0.047	0.042

* denotes value significantly different from the five percent level at the 0.01 level.

** indicates an accepted test with respect to the size condition.

Table 5.3 (contn'd)

Parameter values	Sizes							
	Test number							
	151**	152**	153	154	155	156**	157	158
$\gamma = 0.0$	0.016*	0.015*	0.015*	0.015*	0.014*	0.015*	0.026*	0.015*
0.1	0.026*	0.026*	0.025*	0.024*	0.024*	0.024*	0.047	0.027*
0.2	0.042	0.043	0.043	0.044	0.040	0.043	0.064*	0.044
0.3	0.054	0.054	0.056	0.057	0.055	0.056	0.072*	0.059
0.4	0.059	0.059	0.062	0.063	0.064*	0.062	0.062	0.070*
0.5	0.055	0.054	0.065*	0.066*	0.069*	0.063	0.053	0.069*
0.6	0.051	0.050	0.056	0.059	0.064*	0.059	0.045	0.064*
0.7	0.049	0.048	0.056	0.056	0.060	0.052	0.042	0.059
0.8	0.051	0.045	0.047	0.048	0.053	0.045	0.042	0.051
0.9	0.057	0.054	0.047	0.043	0.047	0.047	0.055	0.046
Test number								
	159	160	161	162	163	164	165	166**
$\gamma = 0.0$	0.014*	0.015*	0.016*	0.014*	0.014*	0.014*	0.016*	0.014*
0.1	0.025*	0.027*	0.028*	0.024*	0.024*	0.024*	0.026*	0.026*
0.2	0.039	0.044	0.045	0.041*	0.039	0.039	0.044	0.044
0.3	0.055	0.059	0.062	0.055	0.054	0.054	0.058	0.057
0.4	0.065*	0.068*	0.071*	0.062	0.062	0.062	0.068*	0.063
0.5	0.071*	0.071*	0.068*	0.067*	0.070*	0.069*	0.068*	0.060
0.6	0.066*	0.064*	0.067*	0.061*	0.064*	0.063	0.065*	0.053
0.7	0.062	0.059	0.058	0.060	0.061	0.062	0.059	0.050
0.8	0.057	0.051	0.052	0.052	0.054	0.055	0.053	0.043
0.9	0.046	0.046	0.053	0.047	0.046	0.046	0.053	0.048
Test number								
	167	168	169**	170**	171	Our test**		
$\gamma = 0.0$	0.015*	0.016*	0.012*	0.012*	0.016*	0.058		
0.1	0.026*	0.026*	0.019*	0.019*	0.027*	0.055		
0.2	0.045	0.045	0.030*	0.030*	0.046	0.049		
0.3	0.060	0.059	0.047	0.047	0.061	0.049		
0.4	0.065*	0.065*	0.055	0.054	0.069*	0.053		
0.5	0.064*	0.060	0.061	0.060	0.068*	0.052		
0.6	0.059	0.058	0.059	0.058	0.062	0.055		
0.7	0.052	0.049	0.060	0.059	0.054	0.053		
0.8	0.042	0.042	0.056	0.055	0.047	0.050		
0.9	0.044	0.043	0.051	0.056	0.047	0.046		

* denotes values significantly different from the five percent level at the 0.01 level.

** indicates an accepted test with respect to the size condition.

Table 5.4

Calculated powers of the test when $m = 3$ at the five percent level: X^2 design matrix with $n = 20$.

Parameter values	Powers									
	Test number									
	1	2	3	4**	5	6	7	8	9**	10**
$\rho = 0.1$	0.061	0.061	0.048	0.052	0.046	0.043	0.040	0.039	0.04	0.039
0.2	0.082	0.082	0.070	0.071	0.063	0.059	0.054	0.056	0.051	0.052
0.3	0.113	0.112	0.100	0.098	0.092	0.087	0.082	0.080	0.073	0.074
0.4	0.153	0.152	0.137	0.137	0.137	0.134	0.133	0.124	0.115	0.119
0.5	0.217	0.215	0.194	0.192	0.199	0.196	0.188	0.182	0.173	0.176
0.6	0.287	0.280	0.265	0.257	0.276	0.280	0.279	0.270	0.253	0.263
0.7	0.382	0.374	0.353	0.346	0.368	0.376	0.376	0.365	0.352	0.361
0.8	0.468	0.459	0.442	0.434	0.463	0.482	0.487	0.475	0.456	0.468
0.9	0.562	0.552	0.539	0.526	0.560	0.573	0.584	0.573	0.554	0.578
Test number										
	11**	12**	13**	14**	15**	16**	17	18**	19	20
$\rho = 0.1$	0.039	0.039	0.047	0.044	0.045	0.044	0.048	0.047	0.055	0.049
0.2	0.046	0.046	0.061	0.058	0.060	0.059	0.065	0.063	0.075	0.069
0.3	0.069	0.071	0.088	0.085	0.086	0.085	0.095	0.093	0.106	0.102
0.4	0.107	0.111	0.128	0.123	0.124	0.122	0.132	0.132	0.147	0.139
0.5	0.163	0.170	0.190	0.188	0.186	0.183	0.194	0.189	0.209	0.200
0.6	0.253	0.259	0.277	0.266	0.261	0.258	0.272	0.268	0.275	0.269
0.7	0.351	0.356	0.371	0.366	0.360	0.357	0.366	0.361	0.368	0.358
0.8	0.464	0.472	0.484	0.476	0.470	0.466	0.475	0.472	0.453	0.445
0.9	0.573	0.581	0.594	0.584	0.574	0.572	0.579	0.567	0.549	0.544
Test number										
	21**	22	23	24	25	26	27**	28**	29**	30**
$\rho = 0.1$	0.049	0.037	0.043	0.040	0.041	0.045	0.041	0.045	0.043	0.042
0.2	0.063	0.052	0.059	0.055	0.052	0.064	0.055	0.063	0.051	0.048
0.3	0.091	0.080	0.088	0.081	0.076	0.095	0.078	0.090	0.071	0.067
0.4	0.127	0.128	0.132	0.128	0.123	0.137	0.121	0.129	0.112	0.107
0.5	0.181	0.191	0.193	0.186	0.179	0.194	0.179	0.186	0.168	0.160
0.6	0.248	0.287	0.276	0.274	0.275	0.279	0.263	0.274	0.253	0.239
0.7	0.336	0.383	0.369	0.373	0.376	0.375	0.363	0.369	0.352	0.341
0.8	0.430	0.496	0.473	0.482	0.492	0.481	0.469	0.473	0.470	0.451
0.9	0.520	0.595	0.571	0.579	0.594	0.576	0.578	0.578	0.578	0.564

** indicate an accepted test.

Table 5.4 (contn'd)

Parameter values	Powers									
	Test number									
	31**	32**	33**	34**	35**	36	37**	38	39	40
$\rho = 0.1$	0.049	0.044	0.048	0.047	0.047	0.051	0.048	0.052	0.050	0.048
0.2	0.062	0.059	0.062	0.063	0.061	0.071	0.063	0.057	0.053	0.064
0.3	0.093	0.084	0.087	0.093	0.086	0.097	0.084	0.080	0.073	0.089
0.4	0.130	0.122	0.127	0.130	0.124	0.135	0.119	0.127	0.111	0.134
0.5	0.191	0.180	0.184	0.189	0.181	0.193	0.175	0.187	0.165	0.191
0.6	0.272	0.255	0.260	0.262	0.257	0.267	0.245	0.285	0.255	0.281
0.7	0.368	0.355	0.356	0.358	0.351	0.357	0.334	0.387	0.353	0.377
0.8	0.478	0.461	0.465	0.466	0.454	0.444	0.431	0.499	0.465	0.486
0.9	0.584	0.571	0.567	0.566	0.555	0.547	0.520	0.601	0.567	0.586
Test number										
	41	42**	43**	44**	45**	46**	47**	48**	49**	50**
$\rho = 0.1$	0.050	0.047	0.048	0.051	0.053	0.048	0.047	0.050	0.052	0.050
0.2	0.065	0.052	0.060	0.061	0.066	0.055	0.057	0.060	0.061	0.059
0.3	0.088	0.075	0.082	0.087	0.094	0.072	0.072	0.078	0.087	0.079
0.4	0.129	0.106	0.122	0.123	0.129	0.116	0.116	0.119	0.123	0.119
0.5	0.189	0.166	0.179	0.182	0.187	0.166	0.167	0.169	0.174	0.166
0.6	0.273	0.249	0.259	0.260	0.274	0.243	0.240	0.239	0.249	0.234
0.7	0.375	0.350	0.357	0.358	0.367	0.342	0.338	0.336	0.340	0.326
0.8	0.478	0.453	0.463	0.459	0.482	0.452	0.442	0.442	0.444	0.426
0.9	0.579	0.558	0.566	0.564	0.591	0.561	0.556	0.548	0.549	0.532
Test number										
	51**	52**	53	54	55	56	57**	58**	59**	60
$\rho = 0.1$	0.050	0.047	0.048	0.049	0.054	0.050	0.049	0.050	0.049	0.077
0.2	0.059	0.059	0.057	0.056	0.070	0.052	0.059	0.060	0.058	0.079
0.3	0.079	0.077	0.083	0.081	0.094	0.075	0.082	0.079	0.082	0.104
0.4	0.117	0.116	0.125	0.117	0.133	0.109	0.118	0.117	0.119	0.134
0.5	0.166	0.171	0.181	0.175	0.191	0.167	0.176	0.172	0.174	0.186
0.6	0.236	0.244	0.282	0.262	0.270	0.261	0.257	0.256	0.248	0.256
0.7	0.325	0.336	0.378	0.362	0.362	0.362	0.358	0.357	0.347	0.351
0.8	0.426	0.431	0.487	0.471	0.465	0.477	0.460	0.464	0.452	0.450
0.9	0.527	0.532	0.590	0.571	0.563	0.583	0.565	0.576	0.555	0.551

** indicates an accepted test.

Table 5.4 (contn'd)

Parameter values	Powers									
	Test number									
	61**	62**	63**	64**	65**	66**	67	68	69	70
$\rho = 0.1$	0.056	0.052	0.053	0.050	0.051	0.052	0.049	0.053	0.054	0.055
0.2	0.065	0.061	0.067	0.074	0.072	0.061	0.053	0.054	0.062	0.057
0.3	0.091	0.079	0.097	0.105	0.103	0.088	0.075	0.081	0.085	0.079
0.4	0.126	0.120	0.138	0.146	0.142	0.125	0.120	0.115	0.124	0.112
0.5	0.183	0.168	0.194	0.204	0.197	0.173	0.176	0.170	0.176	0.168
0.6	0.260	0.241	0.263	0.274	0.267	0.246	0.273	0.259	0.262	0.256
0.7	0.354	0.334	0.359	0.366	0.360	0.331	0.372	0.356	0.363	0.358
0.8	0.465	0.436	0.465	0.472	0.462	0.429	0.491	0.462	0.467	0.473
0.9	0.572	0.548	0.570	0.574	0.563	0.532	0.592	0.567	0.568	0.579
Test number										
	71	72	73**	74**	75**	76**	77**	78**	79**	80**
$\rho = 0.1$	0.053	0.057	0.057	0.06	0.06	0.058	0.059	0.052	0.051	0.052
0.2	0.054	0.058	0.065	0.069	0.065	0.064	0.064	0.075	0.077	0.065
0.3	0.071	0.076	0.078	0.097	0.085	0.074	0.086	0.105	0.105	0.095
0.4	0.099	0.102	0.118	0.138	0.127	0.111	0.124	0.144	0.144	0.127
0.5	0.153	0.157	0.167	0.191	0.172	0.159	0.168	0.192	0.189	0.174
0.6	0.230	0.231	0.237	0.270	0.246	0.222	0.240	0.265	0.263	0.245
0.7	0.333	0.331	0.339	0.362	0.340	0.314	0.323	0.349	0.348	0.328
0.8	0.443	0.443	0.442	0.473	0.440	0.411	0.423	0.452	0.447	0.418
0.9	0.556	0.555	0.553	0.583	0.555	0.522	0.528	0.555	0.549	0.519
Test number										
	81	82	83	84	85**	86**	87**	88**	89**	90**
$\rho = 0.1$	0.053	0.052	0.056	0.056	0.056	0.054	0.057	0.056	0.052	0.050
0.2	0.047	0.048	0.062	0.061	0.061	0.064	0.064	0.065	0.065	0.064
0.3	0.065	0.059	0.080	0.074	0.077	0.084	0.086	0.088	0.092	0.093
0.4	0.099	0.090	0.115	0.104	0.109	0.124	0.126	0.126	0.129	0.129
0.5	0.152	0.137	0.169	0.155	0.162	0.173	0.177	0.172	0.173	0.174
0.6	0.245	0.224	0.257	0.229	0.236	0.245	0.250	0.244	0.243	0.243
0.7	0.353	0.330	0.355	0.332	0.331	0.339	0.343	0.334	0.331	0.324
0.8	0.469	0.442	0.468	0.439	0.442	0.446	0.448	0.432	0.431	0.423
0.9	0.578	0.559	0.576	0.551	0.554	0.556	0.560	0.546	0.538	0.530

** indicates an accepted test.

Table 5.4 (contn'd)

Parameter values	Powers									
	Test number									
	91**	92**	93**	94**	95**	96**	97**	98**	99**	100**
$\rho = 0.1$	0.048	0.045	0.045	0.050	0.048	0.053	0.053	0.052	0.050	0.048
0.2	0.068	0.067	0.072	0.058	0.057	0.063	0.061	0.058	0.063	0.064
0.3	0.100	0.099	0.099	0.075	0.078	0.080	0.081	0.075	0.083	0.090
0.4	0.138	0.142	0.142	0.113	0.117	0.122	0.118	0.106	0.115	0.126
0.5	0.184	0.190	0.195	0.168	0.165	0.169	0.165	0.152	0.162	0.173
0.6	0.254	0.259	0.267	0.251	0.253	0.242	0.241	0.225	0.231	0.245
0.7	0.336	0.342	0.355	0.349	0.350	0.343	0.334	0.319	0.322	0.328
0.8	0.435	0.443	0.455	0.454	0.457	0.445	0.441	0.419	0.419	0.429
0.9	0.539	0.544	0.557	0.558	0.571	0.555	0.552	0.532	0.532	0.541
Test number										
	101**	102**	103**	104**	105**	106**	107**	108**	109**	110**
$\rho = 0.1$	0.045	0.046	0.050	0.038	0.046	0.057	0.048	0.045	0.044	0.037
0.2	0.066	0.068	0.069	0.065	0.069	0.059	0.062	0.060	0.058	0.058
0.3	0.093	0.093	0.094	0.095	0.098	0.076	0.086	0.086	0.084	0.094
0.4	0.133	0.133	0.125	0.137	0.134	0.105	0.125	0.124	0.121	0.133
0.5	0.180	0.176	0.166	0.192	0.175	0.158	0.173	0.171	0.164	0.190
0.6	0.248	0.243	0.235	0.266	0.243	0.236	0.245	0.242	0.237	0.263
0.7	0.330	0.324	0.313	0.354	0.325	0.332	0.344	0.337	0.330	0.358
0.8	0.430	0.422	0.398	0.457	0.419	0.445	0.443	0.439	0.428	0.466
0.9	0.538	0.523	0.502	0.560	0.514	0.555	0.556	0.550	0.541	0.576
Test number										
	111**	112**	113**	114**	115**	116**	117**	118**	119**	120**
$\rho = 0.1$	0.040	0.041	0.041	0.037	0.042	0.041	0.041	0.043	0.033	0.044
0.2	0.061	0.061	0.063	0.062	0.066	0.063	0.059	0.058	0.055	0.061
0.3	0.096	0.090	0.093	0.096	0.095	0.095	0.086	0.082	0.090	0.084
0.4	0.136	0.132	0.135	0.136	0.134	0.132	0.123	0.119	0.127	0.117
0.5	0.190	0.177	0.177	0.192	0.176	0.170	0.173	0.165	0.186	0.167
0.6	0.261	0.242	0.244	0.263	0.243	0.236	0.242	0.237	0.255	0.233
0.7	0.353	0.325	0.327	0.355	0.322	0.316	0.339	0.329	0.351	0.316
0.8	0.460	0.424	0.429	0.458	0.416	0.408	0.438	0.428	0.454	0.417
0.9	0.571	0.530	0.529	0.563	0.515	0.505	0.549	0.541	0.557	0.529

** indicates an accepted test.

Table 5.4 (contn'd)

Parameter values	Powers									
	Test number									
	121**	122**	123**	124**	125**	126**	127**	128**	129**	130**
$\rho = 0.1$	0.034	0.035	0.031	0.033	0.038	0.032	0.041	0.042	0.042	0.042
0.2	0.056	0.057	0.056	0.058	0.062	0.059	0.062	0.064	0.065	0.065
0.3	0.090	0.091	0.089	0.091	0.093	0.091	0.089	0.086	0.094	0.098
0.4	0.128	0.131	0.130	0.132	0.129	0.131	0.130	0.124	0.135	0.138
0.5	0.181	0.186	0.184	0.188	0.170	0.185	0.183	0.170	0.182	0.188
0.6	0.252	0.255	0.254	0.258	0.237	0.255	0.254	0.236	0.249	0.254
0.7	0.336	0.341	0.345	0.352	0.316	0.343	0.343	0.326	0.336	0.342
0.8	0.440	0.445	0.448	0.453	0.407	0.44	0.447	0.427	0.437	0.442
0.9	0.552	0.552	0.554	0.556	0.509	0.537	0.555	0.535	0.545	0.551
Test number										
	131**	132**	133	134	135**	136**	137**	138**	139**	140**
$\rho = 0.1$	0.038	0.033	0.033	0.032	0.034	0.032	0.033	0.034	0.028	0.027
0.2	0.062	0.060	0.061	0.062	0.061	0.054	0.053	0.057	0.052	0.051
0.3	0.095	0.093	0.094	0.093	0.093	0.087	0.087	0.091	0.082	0.083
0.4	0.134	0.133	0.136	0.137	0.135	0.126	0.127	0.129	0.127	0.127
0.5	0.183	0.189	0.199	0.198	0.193	0.179	0.178	0.178	0.181	0.187
0.6	0.254	0.260	0.270	0.273	0.259	0.250	0.248	0.245	0.254	0.262
0.7	0.334	0.353	0.369	0.374	0.351	0.337	0.334	0.329	0.345	0.359
0.8	0.440	0.455	0.474	0.475	0.448	0.439	0.435	0.429	0.452	0.468
0.9	0.545	0.560	0.578	0.575	0.547	0.552	0.547	0.538	0.554	0.573
Test number										
	141	142	143**	144**	145**	146**	147**	148	149	150**
$\rho = 0.1$	0.023	0.024	0.024	0.035	0.032	0.028	0.025	0.025	0.028	0.027
0.2	0.048	0.048	0.048	0.061	0.056	0.051	0.049	0.049	0.052	0.051
0.3	0.080	0.078	0.080	0.091	0.090	0.085	0.082	0.083	0.085	0.082
0.4	0.128	0.124	0.122	0.131	0.130	0.129	0.125	0.125	0.130	0.125
0.5	0.188	0.187	0.182	0.179	0.181	0.191	0.187	0.188	0.192	0.187
0.6	0.280	0.271	0.263	0.247	0.250	0.269	0.271	0.269	0.269	0.254
0.7	0.380	0.369	0.357	0.328	0.331	0.362	0.364	0.368	0.365	0.345
0.8	0.492	0.479	0.467	0.429	0.435	0.476	0.476	0.478	0.472	0.441
0.9	0.599	0.589	0.566	0.541	0.546	0.584	0.585	0.585	0.579	0.538

** indicates an accepted test.

Table 5.4 (contn'd)

Parameter values	Powers							
	Test number							
	151**	152**	153	154	155	156**	157	158
$\rho = 0.1$	0.028	0.028	0.026	0.026	0.025	0.026	0.048	0.028
0.2	0.053	0.053	0.051	0.050	0.050	0.050	0.074	0.053
0.3	0.082	0.082	0.084	0.084	0.083	0.083	0.109	0.087
0.4	0.127	0.126	0.127	0.129	0.129	0.126	0.142	0.135
0.5	0.183	0.181	0.190	0.189	0.190	0.187	0.188	0.195
0.6	0.254	0.256	0.272	0.274	0.280	0.265	0.255	0.279
0.7	0.345	0.342	0.367	0.371	0.379	0.361	0.331	0.380
0.8	0.450	0.443	0.479	0.482	0.495	0.469	0.416	0.491
0.9	0.553	0.546	0.587	0.588	0.601	0.570	0.511	0.601
Test number								
	159	160	161	162	163	164	165	166**
$\rho = 0.1$	0.025	0.028	0.029	0.025	0.025	0.025	0.029	0.027
0.2	0.049	0.052	0.053	0.050	0.049	0.050	0.052	0.053
0.3	0.082	0.086	0.087	0.083	0.083	0.083	0.087	0.082
0.4	0.130	0.135	0.134	0.127	0.126	0.126	0.132	0.125
0.5	0.189	0.194	0.201	0.188	0.189	0.189	0.196	0.188
0.6	0.284	0.281	0.278	0.277	0.277	0.277	0.276	0.259
0.7	0.383	0.381	0.380	0.375	0.377	0.377	0.376	0.355
0.8	0.500	0.493	0.488	0.490	0.492	0.492	0.486	0.448
0.9	0.605	0.602	0.595	0.599	0.600	0.600	0.593	0.549
Test number								
	167	168	169**	170**	171	Our test**	Maximum power	
$\rho = 0.1$	0.027	0.027	0.020	0.019	0.029	0.057	0.060	
0.2	0.051	0.053	0.037	0.038	0.054	0.063	0.077	
0.3	0.084	0.083	0.072	0.070	0.086	0.081	0.105	
0.4	0.130	0.128	0.116	0.115	0.132	0.120	0.146	
0.5	0.189	0.190	0.173	0.174	0.197	0.171	0.204	
0.6	0.267	0.263	0.262	0.261	0.272	0.248	0.277	
0.7	0.364	0.360	0.361	0.357	0.373	0.344	0.371	
0.8	0.466	0.455	0.483	0.476	0.473	0.454	0.484	
0.9	0.566	0.553	0.588	0.583	0.580	0.566	0.594	

** indicates an accepted test.

Table 5.5

Average deviations (AD), sum of squared deviations (SSD) of the power and the corresponding rank of accepted test with respect to size condition.

Test number	AD (rank)	SSD (rank)	Test number	AD (rank)	SSD (rank)
4	0.022778 (40)	0.008523 (72)	73	0.029111 (72)	0.009104 (75)
9	0.027889 (67)	0.007343 (59)	74	0.008333 (02)	0.000733 (02)
10	0.020889 (34)	0.004348 (32)	75	0.025333 (53)	0.007308 (58)
11	0.028111 (68)	0.007717 (62)	76	0.042556 (115)	0.021171 (114)
12	0.023667 (45)	0.005801 (42)	77	0.033667 (101)	0.014061 (103)
13	0.008667 (04)	0.001234 (05)	78	0.014333 (11)	0.003389 (22)
14	0.014222 (10)	0.002112 (10)	79	0.016111 (16)	0.004429 (33)
15	0.016889 (19)	0.002656 (14)	80	0.032778 (95)	0.014423 (105)
16	0.019111 (25)	0.003362 (21)	85	0.032222 (93)	0.010834 (89)
18	0.014000 (09)	0.001984 (09)	86	0.025889 (55)	0.007027 (56)
21	0.030333 (82)	0.011861 (94)	87	0.023000 (42)	0.005633 (40)
27	0.019000 (24)	0.003565 (24)	88	0.028333 (69)	0.009339 (78)
28	0.012333 (07)	0.001649 (06)	89	0.029333 (76)	0.010328 (85)
29	0.024444 (48)	0.005962 (46)	90	0.032000 (91)	0.012784 (98)
30	0.033222 (100)	0.010399 (86)	91	0.024000 (47)	0.007894 (66)
31	0.010111 (05)	0.001085 (04)	92	0.020778 (31)	0.005919 (45)
32	0.020778 (32)	0.003971 (28)	93	0.014556 (12)	0.002949 (17)
33	0.018000 (23)	0.003058 (18)	94	0.026889 (61)	0.007102 (57)
34	0.016000 (15)	0.002492 (12)	95	0.024667 (49)	0.005910 (44)
35	0.022444 (39)	0.005020 (36)	96	0.027333 (65)	0.007722 (63)
37	0.033222 (99)	0.013029 (101)	97	0.030222 (81)	0.009464 (79)
42	0.029111 (72)	0.008220 (68)	98	0.042222 (114)	0.019106 (111)
43	0.020222 (30)	0.003908 (26)	99	0.037889 (109)	0.016091 (109)
44	0.019222 (26)	0.003657 (25)	100	0.030444 (83)	0.010606 (88)
45	0.008333 (02)	0.000907 (03)	101	0.028333 (69)	0.009809 (83)
46	0.029222 (75)	0.008171 (67)	102	0.032222 (93)	0.013624 (102)
47	0.031333 (90)	0.009528 (81)	103	0.040667 (113)	0.023158 (115)
48	0.030778 (86)	0.009621 (82)	104	0.017111 (20)	0.003248 (20)
49	0.026556 (60)	0.007443 (60)	105	0.032778 (95)	0.015191 (107)
50	0.036333 (107)	0.014355 (104)	106	0.032778 (95)	0.011215 (91)
51	0.037000 (108)	0.015035 (106)	107	0.026222 (59)	0.007010 (55)
52	0.033889 (102)	0.012233 (95)	108	0.029333 (76)	0.008790 (73)
57	0.021556 (35)	0.004528 (34)	109	0.034556 (105)	0.012509 (96)
58	0.020778 (32)	0.004291 (31)	110	0.015889 (13)	0.002389 (11)
59	0.026000 (58)	0.006602 (50)	111	0.016667 (18)	0.002718 (15)
61	0.016222 (17)	0.002620 (13)	112	0.032889 (98)	0.012804 (99)
62	0.031000 (87)	0.010053 (84)	113	0.031111 (88)	0.011826 (93)
63	0.012444 (08)	0.001654 (07)	114	0.017333 (22)	0.003168 (19)
64	0.005889 (01)	0.000687 (01)	115	0.034333 (104)	0.015895 (108)
65	0.011222 (06)	0.001841 (08)	116	0.039111 (111)	0.020412 (113)
66	0.031222 (89)	0.011441 (92)	117	0.029778 (79)	0.008926 (74)

Table 5.5 (contn'd)

Test number	AD (rank)	SSD (rank)	Test number	AD (rank)	SSD (rank)
118	0.035111 (106)	0.012738 (97)	137	0.030667 (84)	0.009486 (80)
119	0.023333 (44)	0.005276 (37)	138	0.032000 (91)	0.011186 (90)
120	0.038889 (110)	0.016838 (110)	139	0.027000 (62)	0.006897 (52)
121	0.027667 (66)	0.007745 (64)	140	0.020111 (27)	0.003965 (27)
122	0.025000 (50)	0.006439 (48)	143	0.023222 (43)	0.005287 (38)
123	0.025222 (52)	0.006295 (47)	144	0.030667 (84)	0.010510 (87)
124	0.021889 (36)	0.004865 (35)	145	0.029667 (78)	0.009269 (77)
125	0.039667 (112)	0.020077 (112)	146	0.015889 (14)	0.002867 (16)
126	0.027222 (64)	0.008343 (71)	147	0.017111 (21)	0.003498 (23)
127	0.023778 (46)	0.005742 (41)	150	0.029778 (79)	0.009214 (76)
128	0.034222 (103)	0.01293 (100)	151	0.027000 (62)	0.006973 (54)
129	0.025889 (55)	0.007813 (65)	152	0.029000 (71)	0.008325 (69)
130	0.022000 (38)	0.005820 (43)	156	0.020111 (28)	0.004103 (29)
131	0.025889 (55)	0.007629 (61)	166	0.025778 (54)	0.006792 (51)
132	0.020222 (29)	0.004166 (30)	169	0.022889 (41)	0.006512 (49)
135	0.021889 (36)	0.005547 (39)	170	0.025000 (50)	0.006925 (53)
136	0.029111 (72)	0.008336 (70)	Our test	0.023778	0.005800

Chapter 6

Testing for Heteroscedastic Disturbances in the Linear Regression Model

6.1 Introduction

In Chapter 4 and Chapter 5, we applied our new test procedure to testing for MA(1) disturbances against AR(1) disturbances in the linear regression model. For this particular testing problem, both the null hypothesis and the alternative hypothesis parameter space were restricted to a positive unit interval. In this case, we saw that the new test procedure performs well for $m=3$ with respect to size conditions as well as having good power properties especially when the sample size is small. This good performance may be because the parameter spaces under both hypotheses are restricted to the unit interval. An interesting question is what is the test performance for other types of parameter spaces?

We do not know how well our test performs when both the null and the alternative hypothesis parameter space are extended to an infinite interval. In this chapter our aim is to investigate the performance of our new test for a testing problem which involves both the null hypothesis and the alternative hypothesis parameter spaces being a one-sided infinite interval. How well will our test perform in this case? How should m be chosen? These are some of the questions we are looking to answer in this chapter.

We will address these questions by considering the application of the test procedure to testing for heteroscedastic disturbances in the linear regression model.

The plan of this chapter is as follows. Section 6.2, describes and outlines the model, the testing problem and the test procedure for testing heteroscedastic disturbances in the linear regression model. A Monte Carlo experiment which aims to explore the properties of the test procedure, is outlined in Section 6.3. The design of experiment, size and power properties of the test when $m=1$ and $m=2$ are reported in Section 6.3.1, 6.3.2, 6.3.3, respectively. Also, test sizes controlled at $\lambda=0$ when $m=1$ are discussed in Section 6.4, and properties of this test procedure and results are reported in Section 6.4.1, 6.4.2, respectively. Section 6.5 contains some concluding remarks.

6.2 The model, testing problem and the test procedure

Consider the linear regression model with heteroscedastic disturbances

$$y = X\beta + v, \quad (6.1)$$

where y , X , β are as defined by (3.10) and v is $n \times 1$ disturbance vector whose components are generated as follows

$$v_t = z_t u_t + e_t, \quad t = 1, 2, \dots, n, \quad (6.2)$$

where z_t is a non-stochastic variable (which could be a regressor), $e_t \sim IN(0, \sigma^2)$, $u_t \sim IN(0, \lambda_0 \sigma^2)$, $\lambda_0 = \sigma_u^2 / \sigma^2$ and u_t is independent of e_t . Then the properties of v_t are that it is normally distributed with $E(v_t) = 0$, $\text{var}(v_t) = \sigma^2(1 + \lambda_0 z_t^2)$, and $\text{cov}(v_t, v_s) = 0$, $t \neq s$, that is, $v \sim N(0, \sigma^2 \Omega(\lambda_0))$, where $\Omega(\lambda_0)$ is the $n \times n$ diagonal

matrix with diagonal elements $(1 + \lambda_0 z_i^2)$. An example of this kind of model is the Hildreth-Houck (Hildreth and Houck (1968)) single random coefficient model.

An alternative hypothesis could be that v distributed as $v \sim N(0, \sigma^2 \Sigma(\tau_0))$, where $\Sigma(\tau_0)$ is the $n \times n$ diagonal matrix with diagonal elements $(1 + \tau_0 w_i^2)$, where w_i is another non-stochastic variable which may also be regressor.

Our interest is in testing

$$H_0: v \sim N(0, \sigma^2 \Omega(\lambda_0)), \quad 0 \leq \lambda_0 < \infty \quad (6.3)$$

against

$$H_a: v \sim N(0, \sigma^2 \Sigma(\tau_0)), \quad 0 < \tau_0 < \infty, \quad (6.4)$$

in the context of the linear regression model (6.1). In this testing problem, both the λ_0 and τ_0 parameters, lie between 0 and ∞ . For the simplicity of our calculations, we transform the parameters λ_0 and τ_0 to λ and τ by the functions $\lambda = \lambda_0 / (1 + \lambda_0)$ and $\tau = \tau_0 / (1 + \tau_0)$, respectively. The transformed ranges of λ and τ now become $[0, 1]$.

Using these transformations our testing problem now becomes one of testing

$$H_0: v \sim N(0, \sigma^2 \Omega(\lambda)), \quad 0 \leq \lambda < 1 \quad (6.5)$$

against

$$H_a: v \sim N(0, \sigma^2 \Sigma(\tau)), \quad 0 < \tau < 1, \quad (6.6)$$

where $\Omega(\lambda)$ is the $n \times n$ diagonal matrix with diagonal elements $(1 + [\lambda / (1 - \lambda)] z_i^2)$ and $\Sigma(\tau)$ is the $n \times n$ diagonal matrix with diagonal elements $(1 + [\tau / (1 - \tau)] w_i^2)$.

For this testing problem, β and σ^2 , are nuisance parameters. Their influence can be eliminated through the standard invariance arguments because (6.1), (6.5) and (6.6) are invariant to transformations of the form (3.13). Note that the linear regression models (3.10) and (6.1) are the same except for the disturbance term.

Now using the results of (3.16) and (3.17), our problem of testing (6.5) against (6.6) in the context of (6.1), becomes one of testing

$$H_0: h_0(v, \lambda) = \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{p/2} |\Omega(\lambda)|^{-p/2} |X'\Omega(\lambda)^{-1}X|^{-p/2} \left(\frac{\tilde{z}(\lambda)' \tilde{z}(\lambda)}{z'z} \right)^{-p/2},$$

$$0 \leq \lambda < 1, \quad (6.7)$$

against

$$H_a: h_a(v, \tau) = \frac{1}{2} \Gamma(p/2) \pi^{-p/2} |X'X|^{p/2} |\Sigma(\tau)|^{-p/2} |X'\Sigma(\tau)^{-1}X|^{-p/2} \left(\frac{\tilde{z}(\tau)' \tilde{z}(\tau)}{z'z} \right)^{-p/2},$$

$$0 < \tau < 1, \quad (6.8)$$

where $v = \frac{Pz}{(z'z)^{1/2}}$, P is any $p \times n$ matrix such that $P'P = M$ and $PP' = I_p$, $p = n - k$,

z is the OLS residual vector from (6.1) and $\tilde{z}(\lambda)$ is the OLS residual vector from the regression defined in (3.15), with $\Lambda(\lambda) = \Omega(\lambda)$ and also, $\tilde{z}(\lambda)$ denotes $\tilde{z}(\tau)$ with $\Lambda(\lambda) = \Sigma(\tau)$.

We observe that through the invariance arguments, we have again restricted the respective parameter spaces to $\Theta = \{\lambda: 0 \leq \lambda < 1\}$ and $\Phi = \{\tau: 0 < \tau < 1\}$. Now using equation (4.7), the test which maximizes average power subject to controlling average size over the entire null hypothesis parameter space, i.e., $m = 1$, can be based on the critical region of the form

$$\begin{aligned} & \int_0^1 |\Sigma(\tau)|^{-1/2} |X' \Sigma(\tau)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\tau)' \tilde{z}(\tau)}{z' z} \right)^{-n/2} p(\tau) d\tau \\ & \geq c_\alpha \int_0^1 |\Omega(\lambda)|^{-1/2} |X' \Omega(\lambda)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\lambda)' \tilde{z}(\lambda)}{z' z} \right)^{-n/2} p(\lambda) d\lambda \\ \text{or, } & \frac{\int_0^1 |\Sigma(\tau)|^{-1/2} |X' \Sigma(\tau)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\tau)' \tilde{z}(\tau)}{z' z} \right)^{-n/2} d\tau}{\int_0^1 |\Omega(\lambda)|^{-1/2} |X' \Omega(\lambda)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\lambda)' \tilde{z}(\lambda)}{z' z} \right)^{-n/2} d\lambda} \geq c_\alpha \end{aligned} \quad (6.9)$$

where $p(\tau)$ and $p(\lambda)$ are the weighting functions for τ and λ respectively, which are uniform distributions and c_α is the critical value calculated to control average size of the test. In inequality (6.9), $p(\tau)$ and $p(\lambda)$ have been replaced by uniform density functions.

If we denote the critical region given by (6.9) as $\omega(c_\alpha)$, then using equation (3.19), finding c_α involves solving

$$\int_{\omega(c_\alpha)} \int_0^1 h_0(v, \lambda) d\lambda dv = \int_0^1 \int_{\omega(c_\alpha)} h_0(v, \lambda) dv d\lambda = \alpha \quad (6.10)$$

for c_α , where α is the desired average size of the test.

Equations (6.9) and (6.10) involve integrals, which do not have any analytical solutions. Therefore, as discussed in Chapter 3, we can approximate the integrals in (6.9) by numerical integrals so that (6.9) can be approximated by (see Conte, 1965, p.120)

$$\frac{\sum_{j=1}^q |\Sigma(\tau_j)|^{-1/2} |X' \Sigma(\tau_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\tau_j)' \tilde{z}(\tau_j)}{z' z} \right)^{-n/2}}{\sum_{j=1}^q |\Omega(\lambda_j)|^{-1/2} |X' \Omega(\lambda_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\lambda_j)' \tilde{z}(\lambda_j)}{z' z} \right)^{-n/2}} \geq c_\alpha \quad (6.11)$$

where $\tau_j = \lambda_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$.

As we mentioned in Chapter 3, a potential problem is with controlling the average size of the test. Some sizes can be undesirably large in some parts of the null hypothesis parameter space while other sizes can be smaller than the desired size in some other parts of the null hypothesis parameter space. We conducted some simulation experiments, for $m = 1$ and $m = 2$ for testing (6.7) against (6.8) in the context of (6.1).

A detailed discussion of the test procedures, the size and the power properties of the test with $m = 1$ and $m = 2$, is given in the following section.

6.3 Monte Carlo experiment

In this section, we report a Monte Carlo experiment to investigate the behaviour of the small sample size and power properties of the test in the context of testing heteroscedastic disturbances in the linear regression model, when both the null and alternative hypothesis parameter spaces are one-sided infinite. Also, an aim is to find the appropriate choice of m for which the size conditions are satisfied.

6.3.1 Design of experiment

In the simulation experiments, two thousands replications were used to investigate the small sample size and power properties of the test. We calculated the exact five percent critical values. For this testing problem, we used the $X1$, $X2$, $X3$, $X4$ and $X5$ (defined in Chapter 4) design matrices which cover a variety of economic phenomena. Along with these design matrices two different sets of non-stochastic variables were used. The first set of non-stochastic variables take the values $z_{it} = t/n$ under H_0 and $w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, under H_a , i.e., w takes the reverse values of z . The second set of non-stochastic variables take the values $z_{it} = t/n$ under H_0 and $w_{it} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, under H_a . We used these two different sets of non-stochastic variables to see the variation in calculated sizes and powers of the test statistic. For both sets of non-stochastic variables, data were generated by randomly drawing from the null hypothesis parameter space in order to calculate the critical values (c_α) of the test statistic.

For the testing problem under consideration, the proposed new test under H_0 and H_a is invariant with respect to the nuisance parameters, β and σ^2 , and we have chosen $\beta = 0$ and $\sigma^2 = 1$ for all simulation experiments.

The first step involved deciding on the value of m for which the size of the test is approximately controlled. We carried out some simulation experiments to choose the value of m and conducted some experiments for $m=1$ and $m=2$ for the above mentioned two sets of non-stochastic variables. In the simulation experiment when $m=1$, we indeed find regions of larger sizes than average size and regions of lower sizes

than average size under the null hypothesis. For this reason, we divided the null hypothesis parameter space Θ into two disjoint intervals ($m=2$), determined by the boundary where the size in the case of $m=1$ as a function of λ crosses $\alpha = 0.05$ for the test given by (6.9). In this case we see that the sizes of the resultant test are reasonably acceptable. Thus for $m=1$, we see that the size performances of the test is not that good while for $m=2$, overall the test performs well. Detailed discussions are given below.

6.3.2 Size and power properties of the test with $m=1$

As mentioned in the previous section, two different sets of non-stochastic variables were used to calculate the sizes and powers of the test. For both sets of non-stochastic variables, λ is drawn randomly from (6.5), the null hypothesis parameter space (for data generation purposes), in order to calculate the critical values c_α for all design matrices. Thus based on (6.9) with $m=1$, we calculated the critical values which control the average size of the test over the entire null hypothesis parameter space. The calculated critical values are presented in Table 6.1.

For z_{it} and w_{it} , using the critical values presented in Table 6.1 through equation (6.11), we calculated the sizes of the test for different values of λ , namely, $\lambda = 0.0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9 under H_0 . The estimated sizes are presented in Table 6.2. For two thousand replications, estimated sizes outside the range $[0.037, 0.063]$ are significantly different from the five percent level at the 0.01 level. The values denoted by a 'star' in Table 6.2 are significantly different from five percent at the 0.01 level. From this table our interest is to see the behaviour of the estimated sizes of the test as we are controlling the average size at 0.05. We see that some estimated sizes are far from

the average size, i.e., significantly different from the five percent level. A few of them are smaller than the average size ($\alpha = 0.05$) while some others are undesirably larger than the average size. While we are controlling the average size of the test over the entire null hypothesis parameter space, we do not seem to be successfully controlling the probabilities of Type I errors across the parameter space under H_0 . This led us to consider $m = 2$ for z_{1t} and w_{1t} .

For z_{1t} and w_{1t} , we also calculated the powers of the test by using equation (6.11) for different values of τ , namely, $\tau = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9 under H_a and the estimated powers are presented in Table 6.3. From this table we see that the powers of the test gradually increase when τ increases both for small and large sample sizes for all design matrices. We also see that when the sample size is large, the powers of the test are much higher (more than double for the points $\tau = 0.3, 0.4, 0.5, 0.6, 0.7, 0.8$ and 0.9) compared to those for the smaller sample size.

Similarly, for z_{2t} and w_{2t} , taking the critical values calculated for this set of non-stochastic variables presented in Table 6.1, through equation (6.11), we calculated the sizes of the test for different values of λ under H_0 . Also, we calculated the powers of the test for different values of τ under H_a . The estimated sizes are presented in Table 6.4 and the powers are presented in Table 6.5. From Table 6.4 we see that there is a large number of estimated sizes that are far from the average size. Only a few sizes are smaller than the average size while a number of sizes are undesirably larger than the average size. This means our controlled sizes are not within an acceptable range. This motivated us to consider $m = 2$ for z_{2t} and w_{2t} .

From Table 6.5 we see that the powers of the test slowly increase as τ increases for all design matrices. We observe that when the sample size is large, the powers of the test are higher compared to those of the smaller sample size. For both sets of non-stochastic variables, we also observe that the powers of the test show a very similar pattern although the powers for z_{1t} and w_{1t} are slightly higher than the corresponding powers for z_{2t} and w_{2t} . This may be due to the fact that for z_{1t} and w_{1t} , the calculated sizes of the test for the starting boundary points $\lambda = 0.0, 0.1, 0.2$ are higher compared to the corresponding sizes for z_{2t} and w_{2t} . For z_{1t} and w_{1t} , the calculated sizes for the points $\lambda = 0.0, 0.1, 0.2$ are above the average size while for z_{2t} and w_{2t} , the sizes are below the average size.

In the following section, we discuss the size and power properties of the test with $m = 2$ for the two sets of non-stochastic variables.

6.3.3 Size and power properties of the test with $m = 2$

We desire to control the average estimated sizes of the test over suitable regions under H_0 . Bearing this in mind, for z_{1t} and w_{1t} , and based on the estimated sizes presented in Table 6.2 with $m = 1$, we divided the null hypothesis parameter space into two disjoint intervals. One subregion (Θ_1) is made up of the parameter values which have larger than average size while the other one (Θ_2) comprises those parameter values which resulted in smaller than average size under the null hypothesis for the $m = 1$ version of the test. When we control average size over these two regions, we are expecting that the estimated sizes will come very close to the average size. The boundary point of the two disjoint intervals is determined by the high size/low size technique. For z_{1t} and w_{1t} ,

ranges of the two disjoint intervals and the corresponding values of j , where $\lambda_j = \tau_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$, for all design matrices are presented in Table 6.6.

Thus for $m=2$, equation (3.7) gives the resultant test that maximizes average power subject to controlling two size conditions and involves a critical region of the form (6.11) which now becomes

$$\begin{aligned} & \sum_{j=1}^q |\Sigma(\tau_j)|^{-1/2} |X' \Sigma(\tau_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\tau_j)' \tilde{z}(\tau_j)}{z'z} \right)^{-n/2} \\ & \geq k_1 \sum_{j=1}^{q_1} |\Omega(\lambda_j)|^{-1/2} |X' \Omega(\lambda_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\lambda_j)' \tilde{z}(\lambda_j)}{z'z} \right)^{-n/2} \\ & + k_2 \sum_{j=q_1+1}^q |\Omega(\lambda_j)|^{-1/2} |X' \Omega(\lambda_j)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\lambda_j)' \tilde{z}(\lambda_j)}{z'z} \right)^{-n/2}, \end{aligned} \quad (6.12)$$

where k_1 and k_2 are the critical values and q_1 is determined by the boundary of the two subregions.

If we denote the critical region given by (6.12) as $\omega(k_1, k_2)$ and if $\Theta_1 = \{\lambda: \lambda_1 \leq \lambda \leq \lambda_2\}$ and $\Theta_2 = \{\lambda: \lambda_2 < \lambda \leq \lambda_3\}$, where $\lambda_1 = 0$, $\lambda_3 = 1$ and λ_2 is the chosen boundary of subregions of Θ , then finding k_1 and k_2 involves solving

$$\int_{\Theta_i} \int_{\omega(k_1, k_2)} h_0(v, \lambda) dv d\lambda = \alpha, \quad i = 1, 2, \quad (6.13)$$

jointly for k_1 and k_2 . In equation (6.12), q_1 is determined by rounding $(\lambda_2 q + (1/2))$, down to the nearest integer.

Now for z_{11} and w_{11} , the critical values, k_1 and k_2 , for all design matrices were calculated via trial and error, so that the two size conditions are simultaneously satisfied. These critical calculated values are presented in Table 6.7. From this table, we see that k_1 values are far larger than k_2 values for all cases. The values of k_1 are always positive while k_2 always takes small negative values. From the critical values, we also observe that k_1 always increases as the sample size increases, the one exception being for $X3$. On the other hand, k_2 decreases for $X1, X2, X3$ and increases for $X4, X5$ when the sample size increases.

Using the critical values presented in Table 6.7, we calculated the sizes of the test by using (6.12) for different values of λ under H_0 and calculated the powers of the test for different values of τ under H_a . The calculated sizes are presented in Table 6.8 and the calculated powers are presented in Table 6.9.

From Table 6.8 we see that there are a few sizes significantly different from five percent at the 0.01 significance level. There are only three sizes that are much larger than the average size when $n = 50$, these are at or near the first boundary point ($\lambda = 0.0$ for $X2$ and $\lambda = 0.0, 0.1$ for $X3$) of the null hypothesis parameter space. Overall we find that the sizes of the test are very acceptable, with $m=2$, for z_{11} and w_{11} , when the sample size is small.

From Table 6.9 we see that the powers of the test increase as τ increases for all design matrices. Comparing the powers presented in Table 6.3 for $m=1$ when the estimated sizes are not controlled for z_{11} and w_{11} , with those powers of the test for $m=2$, we see

that the powers of the test for $m=2$ are always lower than the powers of the test for $m=1$. The main reason may be because some of the calculated sizes of the test with $m=1$ are higher compared to those for $m=2$, particularly where the two parameter spaces meet at $\lambda = \tau = 0$.

Similarly, for z_{2i} and w_{2i} , based on the estimated sizes presented in Table 6.4 for $m=1$, we divided the null hypothesis parameter space into two disjoint intervals. One subregion (Θ_1) is a region which is made up of the parameter values that have smaller than the average size and the other one (Θ_2) is the region of values with larger than the average sizes under the null hypothesis parameter space for the $m=1$ version of the test. The boundary point of the two disjoint intervals is determined by the same technique mentioned above, i.e., the high size/low size technique. For z_{2i} and w_{2i} , ranges of the two disjoint intervals and the corresponding values of j , where $\lambda_j = \tau_j = (2j-1)/(2q)$, $j=1, 2, \dots, q$, $q=50$ for all design matrices are presented in Table 6.10.

Thus the test with $m=2$, involves a critical region given by (6.12) and using (6.12) we calculated the critical values k_1 and k_2 , via trial and error. The calculated critical values, k_1 and k_2 , for z_{2i} and w_{2i} , are presented in Table 6.11. From this table we see that k_2 values are far larger than k_1 values for all design matrices. The values of k_2 are always positive numbers whereas k_1 takes positive as well as small negative values.

For z_{2i} and w_{2i} , using the critical values presented in Table 6.11, by equation (6.12), we calculated the sizes of the test for different values of λ under H_0 and calculated the powers of the test for different values of τ under H_a . The calculated sizes are presented in Table 6.12 and the powers are presented in Table 6.13.

From Table 6.12 we see that all the estimated sizes of the test fall within the rejection probability range [0.037, 0.063] are not significantly different from five percent at the 0.01 level for all design matrices. It is pleasing to see that the sizes of the test are very acceptable for all design matrices, when $m=2$, for z_{2i} and w_{2i} , both for small and large sample sizes.

Also, from Table 6.13 we find that the powers of the test with $m=2$ increase as τ increases for all design matrices. It can be observed that the powers of the test with $m=1$ are slightly higher than the powers of the test with $m=2$. This is mainly because when $m=2$, we controlled the sizes of the test to the average size which makes the powers a bit lower than the powers of the test when the sizes are not so well controlled, that is, $m=1$.

Overall we can say that the size and power properties of the test are reasonably acceptable for testing for heteroscedastic disturbances in the linear regression model for both z_{1i} and w_{1i} , and z_{2i} and w_{2i} , when $m=2$.

6.4 Test sizes controlled at $\lambda = 0$ with $m=1$

We conducted a further simulation experiment to investigate the behaviour of the small sample size and power properties of the test when the test sizes are controlled at $\lambda = 0$ under the null hypothesis parameter space for z_{1i} and w_{1i} , when $m=1$. We carried out this simulation experiment to see how well our test procedure works in this case. It must be noted that in this case the resultant test is no longer optimal but it is of interest to see how much power is lost by using this slightly different test procedure.

6.4.1 Properties of the test when sizes controlled at $\lambda = 0$ with $m = 1$

In Section 6.3.3, we observed that the size and power properties of the test for $m = 2$ are acceptable for both z_{1r} and w_{1r} , and z_{2r} and w_{2r} . However, a lot of computational time was used to find the critical values k_1 and k_2 for which the size conditions are satisfied. Also, for z_{1r} and w_{1r} , when $m = 1$ we see from Table 6.2 that sizes of the test decrease as λ increase. Then it is very easy to control the maximum size of our test which always (at least in our experiments) occurs at $\lambda = 0$. For this case, in order to calculate the sizes and powers of the test, we calculated the critical value (c_α) for all design matrices based on (6.9), with $m = 1$, where the data are simulated by taking $\lambda = 0$. The calculated critical values are presented in Table 6.14.

6.4.2 Results

Using the critical values presented in Table 6.14, we calculated the sizes of the test through equation (6.11) for different values of λ under H_0 and calculated the powers of the test for different values of τ under H_a . The calculated sizes are presented in Table 6.15 and the powers are presented in Table 6.16.

From Table 6.15 we see that all the estimated sizes of the test for all design matrices are lower or equal to 0.05. This indicates that the size performance of the test is very satisfactory for all design matrices, when sizes of the test are controlled at $\lambda = 0$ with $m = 1$, for z_{1r} and w_{1r} , both for the smaller and the larger sample sizes. From this point of view, we can say that without spending a lot of computational time to calculate the sizes of the test which control average size over different regions under H_0 , we can get

the good results within a very short time by following the above-mentioned test procedure.

From Table 6.16 we see that the powers of the test increase as τ increases for all design matrices. For z_{1r} and w_{1r} , comparing the powers of the test presented in Table 6.9 for $m = 2$ when the estimated sizes are controlled, to those powers of the test for $m = 1$ controlled at $\lambda = 0$, we see that when the sample size is small, the powers of the test for $m = 1$ controlled at $\lambda = 0$ are always higher than the powers of the test for $m = 2$. On average, there is zero to 3.0 percentage points increase in terms of power improvement for all design matrices. Also for larger sample sizes, we see that the powers of the test for $m = 1$ controlled at $\lambda = 0$ are lower than the powers of the test for $m = 2$ with an on average 0.3 to 6.2 percentage points decrease. One reason for this case may be the fact that for the starting boundary point $\lambda = 0.0$, the calculated sizes of the test for $m = 2$ are always higher (above the average size) than the calculated sizes of the test for $m = 1$ with size controlled at $\lambda = 0$ where all estimated sizes are lower than or equal to 0.05. There is one case, namely, the X_4 with $n = 30$, where the power of the test coincides for the end boundary point ($\lambda = 0.9$). In general we can see that when the sample size is small, the size and power performance of the test for $m = 1$ when sizes are controlled at $\lambda = 0$ for z_{1r} and w_{1r} is very competitive and the test may be recommended. But when the sample size is large, it may be worthwhile for using the new procedure with $m = 2$ with respect to size and power properties.

6.5 Concluding remarks

In this chapter we have applied the new test approach to testing for heteroscedastic disturbances in the linear regression model. This testing problem is of interest because both the null and the alternative hypothesis parameter spaces are one-sided infinite intervals. For this testing problem we used two sets of non-stochastic variables and we got different results for these two different sets of variables. For both sets of non-stochastic variables when $m=2$, the test performs well with respect to size conditions. The power properties of the test are well behaved because we are testing for heteroscedastic disturbances. Further results for z_{it} and w_{it} when $m=1$ and the test sizes are controlled at $\lambda=0$ indicates that for small sample size, the test performance is fairly good but there is a clear loss of power when we used this test procedure for the larger sample size. It seems that extra computation for $m=2$ is rewarded in terms of power. From this viewpoint we conclude that the new test approach can work well for testing problems where both the null and alternative hypothesis parameter spaces are one-sided infinite intervals. We are able to work on this type of testing problem because of the parameter transformation used to transform the parameter spaces to finite intervals.

Table 6.1

Critical values (c_α) of the test for two different sets of non-stochastic regressor values for all design matrices when $m=1$, at the five percent level.

Design matrix	Sample Size (n)	Critical values (c_α)	
		For z_{it} and w_{it}	For z_{2t} and w_{2t}
$X1(n \times 3)$	20	4.1071	1.6210
	60	4.0562	1.5852
$X2(n \times 3)$	20	4.4477	1.5359
	60	4.2111	1.6391
$X3(n \times 5)$	20	4.3157	1.5522
	60	4.2352	1.5986
$X4(n \times 3)$	20	4.2945	1.5287
	30	3.8365	1.7021
$X5(n \times 3)$	20	4.4058	1.5145
	60	4.1852	1.6370

Table 6.2

Calculated sizes of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m=1$ at the five percent level.

Parameter values	Sizes									
	$X1$		$X2$		$X3$		$X4$		$X5$	
λ	$n=20$	60	20	60	20	60	20	30	20	60
0.0	0.079*	0.134*	0.079*	0.128*	0.079*	0.127*	0.075*	0.127*	0.075*	0.127*
0.1	0.069*	0.107*	0.068*	0.105*	0.070*	0.107*	0.065*	0.113*	0.065*	0.106*
0.2	0.061	0.090*	0.058	0.084*	0.064*	0.088*	0.055	0.106*	0.060	0.084*
0.3	0.056	0.072*	0.056	0.062	0.053	0.067*	0.049	0.087*	0.052	0.065*
0.4	0.048	0.054	0.048	0.053	0.043	0.053	0.044	0.071*	0.046	0.053
0.5	0.044	0.042	0.041	0.041	0.040	0.048	0.037	0.065*	0.040	0.041
0.6	0.041	0.034*	0.038	0.032*	0.036*	0.041	0.034*	0.056	0.035*	0.032*
0.7	0.035*	0.025*	0.034*	0.024*	0.029*	0.033*	0.030*	0.051	0.031*	0.026*
0.8	0.032*	0.019*	0.031*	0.021*	0.026*	0.029*	0.026*	0.041	0.027*	0.020*
0.9	0.027*	0.014*	0.025*	0.018*	0.024*	0.026*	0.023*	0.036*	0.023*	0.015*

* denotes value significantly different from the five percent level at the 0.01 level.

Table 6.3

Calculated powers of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 1$ at the five percent level.

Parameter values	Powers									
	X1		X2		X3		X4		X5	
τ	$n=20$	60	20	60	20	60	20	30	20	60
0.1	0.091	0.169	0.095	0.161	0.089	0.158	0.089	0.148	0.087	0.169
0.2	0.105	0.202	0.106	0.199	0.101	0.192	0.098	0.162	0.099	0.203
0.3	0.119	0.243	0.116	0.236	0.112	0.229	0.112	0.181	0.112	0.242
0.4	0.134	0.284	0.131	0.271	0.122	0.264	0.128	0.202	0.126	0.274
0.5	0.146	0.319	0.145	0.303	0.138	0.297	0.141	0.222	0.143	0.315
0.6	0.159	0.347	0.154	0.338	0.149	0.330	0.154	0.237	0.156	0.350
0.7	0.171	0.386	0.165	0.373	0.158	0.357	0.166	0.254	0.175	0.382
0.8	0.182	0.418	0.181	0.403	0.171	0.387	0.178	0.283	0.185	0.415
0.9	0.197	0.453	0.200	0.434	0.182	0.418	0.184	0.298	0.194	0.445

Table 6.4

Calculated sizes of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 1$ at the five percent level.

Parameter values	Sizes									
	X1		X2		X3		X4		X5	
λ	$n=20$	60	20	60	20	60	20	30	20	60
0.0	0.033*	0.040	0.045	0.040	0.053	0.034*	0.051	0.037	0.050	0.036*
0.1	0.036*	0.046	0.048	0.046	0.053	0.042	0.055	0.039	0.052	0.040
0.2	0.041	0.051	0.051	0.049	0.056	0.046	0.061	0.042	0.054	0.048
0.3	0.047	0.054	0.055	0.053	0.058	0.050	0.063	0.044	0.057	0.051
0.4	0.050	0.058	0.058	0.057	0.060	0.054	0.065*	0.045	0.060	0.056
0.5	0.052	0.062	0.061	0.060	0.063	0.057	0.067*	0.048	0.060	0.060
0.6	0.053	0.066*	0.061	0.064*	0.065*	0.061	0.070*	0.049	0.067*	0.064*
0.7	0.057	0.070*	0.065*	0.067*	0.067*	0.065*	0.074*	0.050	0.069*	0.065*
0.8	0.059	0.076*	0.066*	0.068*	0.069*	0.072*	0.077*	0.051	0.072*	0.068*
0.9	0.059	0.079*	0.070*	0.072*	0.072*	0.074*	0.080*	0.053	0.073*	0.071*

* denotes value significantly different from the five percent level at the 0.01 level.

Table 6.5

Calculated powers of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 1$ at the five percent level.

Parameter values	Powers									
	X1		X2		X3		X4		X5	
τ	$n=20$	60	20	60	20	60	20	30	20	60
0.1	0.051	0.062	0.059	0.062	0.059	0.057	0.066	0.046	0.058	0.059
0.2	0.059	0.088	0.071	0.079	0.070	0.082	0.079	0.054	0.072	0.081
0.3	0.070	0.117	0.080	0.106	0.081	0.112	0.091	0.061	0.086	0.106
0.4	0.084	0.146	0.099	0.132	0.091	0.140	0.102	0.077	0.095	0.134
0.5	0.097	0.179	0.110	0.164	0.102	0.166	0.123	0.093	0.107	0.174
0.6	0.109	0.212	0.127	0.197	0.112	0.195	0.136	0.105	0.121	0.204
0.7	0.119	0.242	0.141	0.228	0.125	0.227	0.151	0.123	0.136	0.232
0.8	0.131	0.279	0.154	0.259	0.134	0.257	0.162	0.140	0.153	0.265
0.9	0.143	0.310	0.171	0.294	0.139	0.287	0.177	0.158	0.166	0.294

Table 6.6

Ranges of the two disjoint intervals and the corresponding values of j , where $\lambda_j = \tau_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$, and $z_{it} = t/n$, $w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$.

Design matrix	Sample Size (n)	Subregions	
		Θ_1 (larger than average size for $m = 1$)	Θ_2 (smaller than average size for $m = 1$)
X1($n \times 3$)	20	$0 \leq \lambda \leq 0.38$ ($j = 1, 2, \dots, 19$)	$0.38 < \lambda < 1$ ($j = 20, 21, \dots, 50$)
	60	$0 \leq \lambda \leq 0.43$ ($j = 1, 2, \dots, 22$)	$0.43 < \lambda < 1$ ($j = 23, 24, \dots, 50$)
X2($n \times 3$)	20	$0 \leq \lambda \leq 0.38$ ($j = 1, 2, \dots, 19$)	$0.38 < \lambda < 1$ ($j = 20, 21, \dots, 50$)
	60	$0 \leq \lambda \leq 0.42$ ($j = 1, 2, \dots, 21$)	$0.42 < \lambda < 1$ ($j = 22, 23, \dots, 50$)
X3($n \times 5$)	20	$0 \leq \lambda \leq 0.33$ ($j = 1, 2, \dots, 17$)	$0.33 < \lambda < 1$ ($j = 18, 19, \dots, 50$)
	60	$0 \leq \lambda \leq 0.46$ ($j = 1, 2, \dots, 23$)	$0.46 < \lambda < 1$ ($j = 24, 25, \dots, 50$)
X4($n \times 3$)	20	$0 \leq \lambda \leq 0.23$ ($j = 1, 2, \dots, 12$)	$0.23 < \lambda < 1$ ($j = 13, 14, \dots, 50$)
	30	$0 \leq \lambda \leq 0.71$ ($j = 1, 2, \dots, 36$)	$0.71 < \lambda < 1$ ($j = 37, 38, \dots, 50$)
X5($n \times 3$)	20	$0 \leq \lambda \leq 0.33$ ($j = 1, 2, \dots, 17$)	$0.33 < \lambda < 1$ ($j = 18, 19, \dots, 50$)
	60	$0 \leq \lambda \leq 0.42$ ($j = 1, 2, \dots, 21$)	$0.42 < \lambda < 1$ ($j = 22, 23, \dots, 50$)

Table 6.7

Final critical values k_1 and k_2 when $z_{it} = t/n$ and $w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, for $m = 2$ at the five percent level.

Design matrix	Sample size (n)	Critical values	
		k_1	k_2
$X1(n \times 3)$	20	3.8946	-0.825
	60	3.9709	-0.71
$X2(n \times 3)$	20	3.5142	-0.70
	60	3.5379	-0.54
$X3(n \times 5)$	20	3.9040	-0.92
	60	3.0458	-0.50
$X4(n \times 3)$	20	3.3866	-0.74
	30	5.4805	-1.50
$X5(n \times 3)$	20	3.2769	-0.60
	60	3.9008	-0.62

Table 6.8

Calculated sizes of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level.

Parameter values	Sizes									
	$X1$		$X2$		$X3$		$X4$		$X5$	
	$n = 20$	60	20	60	20	60	20	30	20	60
λ										
0.0	0.041	0.063	0.058	0.065*	0.044	0.089*	0.043	0.059	0.052	0.062
0.1	0.036*	0.043	0.054	0.053	0.043	0.074*	0.041	0.052	0.051	0.048
0.2	0.035*	0.035*	0.050	0.044	0.042	0.059	0.043	0.044	0.044	0.037
0.3	0.033*	0.031*	0.047	0.036*	0.042	0.056	0.041	0.043	0.043	0.034*
0.4	0.035*	0.031*	0.045	0.033*	0.044	0.052	0.043	0.044	0.039	0.030*
0.5	0.035*	0.035*	0.044	0.035*	0.044	0.050	0.044	0.048	0.037	0.035*
0.6	0.038	0.041	0.044	0.041	0.045	0.051	0.044	0.049	0.039	0.040
0.7	0.042	0.046	0.047	0.049	0.047	0.053	0.045	0.052	0.038	0.044
0.8	0.045	0.050	0.049	0.051	0.048	0.055	0.049	0.055	0.041	0.049
0.9	0.048	0.057	0.047	0.055	0.052	0.057	0.049	0.060	0.043	0.056

* denotes value significantly different from the five percent level at the 0.01 level.

Table 6.9

Calculated powers of the test when $H_0: z_{it} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level.

Parameter values	Powers									
	$X1$		$X2$		$X3$		$X4$		$X5$	
	$n = 20$	60	20	60	20	60	20	30	20	60
τ										
0.1	0.046	0.078	0.066	0.091	0.047	0.115	0.048	0.070	0.059	0.081
0.2	0.054	0.102	0.070	0.111	0.052	0.137	0.049	0.081	0.066	0.099
0.3	0.058	0.117	0.077	0.131	0.057	0.165	0.056	0.096	0.074	0.121
0.4	0.066	0.143	0.087	0.160	0.064	0.194	0.060	0.113	0.082	0.144
0.5	0.072	0.172	0.097	0.191	0.073	0.223	0.069	0.118	0.086	0.179
0.6	0.077	0.200	0.104	0.220	0.079	0.253	0.078	0.130	0.095	0.208
0.7	0.084	0.229	0.111	0.252	0.087	0.284	0.084	0.142	0.105	0.240
0.8	0.089	0.263	0.119	0.275	0.093	0.311	0.090	0.154	0.113	0.263
0.9	0.100	0.288	0.133	0.306	0.094	0.337	0.098	0.164	0.122	0.286

Table 6.10

Ranges of the two disjoint intervals and the corresponding values of j , where $\lambda_j = \tau_j = (2j-1)/(2q)$, $j = 1, 2, \dots, q$, $q = 50$, and $z_{it} = t/n$, $w_{it} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$.

Design matrix	Sample Size (n)	Subregions	
		Θ_1	Θ_2
		(smaller than average size for $m = 1$)	(larger than average size for $m = 1$)
$X1(n \times 3)$	20	$0 \leq \lambda \leq 0.20$ ($j = 1, 2, \dots, 10$)	$0.20 < \lambda < 1$ ($j = 11, 12, \dots, 50$)
	60	$0 \leq \lambda \leq 0.18$ ($j = 1, 2, \dots, 9$)	$0.18 < \lambda < 1$ ($j = 10, 11, \dots, 50$)
$X2(n \times 3)$	20	$0 \leq \lambda \leq 0.17$ ($j = 1, 2, \dots, 9$)	$0.17 < \lambda < 1$ ($j = 10, 11, \dots, 50$)
	60	$0 \leq \lambda \leq 0.20$ ($j = 1, 2, \dots, 10$)	$0.20 < \lambda < 1$ ($j = 11, 12, \dots, 50$)
$X3(n \times 5)$	20	$0 \leq \lambda \leq 0.18$ ($j = 1, 2, \dots, 9$)	$0.18 < \lambda < 1$ ($j = 10, 11, \dots, 50$)
	60	$0 \leq \lambda \leq 0.20$ ($j = 1, 2, \dots, 10$)	$0.20 < \lambda < 1$ ($j = 11, 12, \dots, 50$)
$X4(n \times 3)$	20	$0 \leq \lambda \leq 0.20$ ($j = 1, 2, \dots, 10$)	$0.20 < \lambda < 1$ ($j = 11, 12, \dots, 50$)
	30	$0 \leq \lambda \leq 0.15$ ($j = 1, 2, \dots, 8$)	$0.15 < \lambda < 1$ ($j = 9, 10, \dots, 50$)
$X5(n \times 3)$	20	$0 \leq \lambda \leq 0.20$ ($j = 1, 2, \dots, 10$)	$0.20 < \lambda < 1$ ($j = 11, 12, \dots, 50$)
	60	$0 \leq \lambda \leq 0.27$ ($j = 1, 2, \dots, 14$)	$0.27 < \lambda < 1$ ($j = 15, 16, \dots, 50$)

Table 6.11

Final critical values k_1 and k_2 when $z_{2t} = t/n$ and $w_{2t} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, for $m = 2$ at the five percent level.

Design matrix	Sample size (n)	Critical values	
		k_1	k_2
$X1(n \times 3)$	20	-0.0349	1.60
	60	-0.0060	1.87
$X2(n \times 3)$	20	-0.2344	1.91
	60	0.0647	1.71
$X3(n \times 5)$	20	-0.1679	1.84
	60	-0.0036	1.90
$X4(n \times 3)$	20	-0.0985	1.72
	30	-0.2002	2.10
$X5(n \times 3)$	20	-0.1205	1.76
	60	0.2390	1.55

Table 6.12

Calculated sizes of the test when $H_0: z_{2t} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{2t} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level.

Parameter values	Sizes									
	$X1$		$X2$		$X3$		$X4$		$X5$	
λ	$n=20$	60	20	60	20	60	20	30	20	60
0.0	0.047	0.047	0.049	0.049	0.051	0.049	0.047	0.055	0.045	0.040
0.1	0.049	0.047	0.047	0.049	0.051	0.048	0.048	0.052	0.046	0.041
0.2	0.050	0.048	0.048	0.051	0.050	0.048	0.048	0.052	0.047	0.044
0.3	0.052	0.052	0.048	0.051	0.051	0.048	0.050	0.051	0.048	0.048
0.4	0.052	0.052	0.049	0.054	0.050	0.049	0.053	0.051	0.049	0.050
0.5	0.054	0.053	0.048	0.054	0.051	0.051	0.054	0.049	0.051	0.051
0.6	0.055	0.053	0.049	0.055	0.052	0.052	0.056	0.048	0.053	0.053
0.7	0.057	0.054	0.050	0.056	0.052	0.052	0.057	0.049	0.053	0.055
0.8	0.058	0.056	0.051	0.058	0.053	0.052	0.059	0.049	0.053	0.056
0.9	0.060	0.055	0.052	0.059	0.054	0.054	0.060	0.049	0.054	0.057

Table 6.13

Calculated powers of the test when $H_0: z_{2t} = t/n$, $0 \leq \lambda < 1$ is tested against $H_a: w_{2t} = 0$, if $t \leq n/2$, or 1, if $t > n/2$, $t = 1, 2, \dots, n$, $0 < \tau < 1$, for $m = 2$ at the five percent level.

Parameter values	Powers									
	$X1$		$X2$		$X3$		$X4$		$X5$	
τ	$n=20$	60	20	60	20	60	20	30	20	60
0.0	0.047	0.047	0.049	0.049	0.051	0.049	0.047	0.055	0.045	0.040
0.1	0.056	0.064	0.055	0.070	0.058	0.059	0.058	0.059	0.055	0.059
0.2	0.070	0.077	0.064	0.087	0.065	0.076	0.069	0.065	0.062	0.078
0.3	0.083	0.098	0.070	0.109	0.071	0.093	0.079	0.071	0.069	0.102
0.4	0.094	0.126	0.078	0.129	0.079	0.115	0.090	0.077	0.083	0.130
0.5	0.105	0.147	0.086	0.156	0.086	0.134	0.107	0.084	0.092	0.159
0.6	0.113	0.166	0.095	0.180	0.091	0.151	0.116	0.091	0.102	0.184
0.7	0.121	0.194	0.105	0.203	0.098	0.176	0.127	0.107	0.113	0.216
0.8	0.135	0.218	0.113	0.232	0.106	0.202	0.134	0.117	0.122	0.237
0.9	0.150	0.245	0.123	0.257	0.113	0.225	0.146	0.127	0.132	0.265

Table 6.14

Critical values (c_α) of the test for all design matrices for z_{2t} and w_{2t} when sizes are controlled at $\lambda = 0$, for $m = 1$ at the five percent level.

Design matrix	Sample Size (n)	Critical values (c_α)
$X1(n \times 3)$	20	5.3663
	60	9.0774
$X2(n \times 3)$	20	6.1441
	60	8.8586
$X3(n \times 5)$	20	5.5979
	60	8.7613
$X4(n \times 3)$	20	5.3611
	30	7.5320
$X5(n \times 3)$	20	5.8346
	60	9.2466

Table 6.15

Calculated sizes of the test when $H_0: z_{it} = t/n$, (sizes are controlled at $\lambda = 0$) is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, for $m = 1$ at the five percent level.

Parameter values	Sizes									
	X1		X2		X3		X4		X5	
λ	$n=20$	60	20	60	20	60	20	30	20	60
0.0	0.050	0.050	0.050	0.050	0.051	0.051	0.050	0.051	0.050	0.051
0.1	0.044	0.035*	0.042	0.038	0.041	0.044	0.044	0.041	0.041	0.035*
0.2	0.037	0.024*	0.036*	0.028*	0.037	0.034*	0.037	0.034*	0.035*	0.026*
0.3	0.033*	0.016*	0.030*	0.021*	0.029*	0.025*	0.032*	0.027*	0.028*	0.019*
0.4	0.027*	0.015*	0.026*	0.018*	0.026*	0.024*	0.028*	0.020*	0.024*	0.014*
0.5	0.023*	0.011*	0.023*	0.015*	0.024*	0.017*	0.022*	0.018*	0.020*	0.012*
0.6	0.020*	0.009*	0.019*	0.012*	0.022*	0.012*	0.020*	0.015*	0.016*	0.009*
0.7	0.016*	0.008*	0.017*	0.009*	0.019*	0.010*	0.017*	0.014*	0.013*	0.007*
0.8	0.014*	0.007*	0.014*	0.007*	0.018*	0.009*	0.015*	0.013*	0.012*	0.006*
0.9	0.011*	0.005*	0.013*	0.004*	0.017*	0.009*	0.012*	0.011*	0.010*	0.004*

* denotes value significantly different from the five percent level at the 0.01 level.

Table 6.16

Calculated powers of the test when $H_0: z_{it} = t/n$, (sizes are controlled at $\lambda = 0$) is tested against $H_a: w_{it} = (n+1-t)/n$, $t = 1, 2, \dots, n$, for $m = 1$ at the five percent level.

Parameter values	Powers									
	X1		X2		X3		X4		X5	
τ	$n=20$	60	20	60	20	60	20	30	20	60
0.1	0.060	0.071	0.057	0.068	0.062	0.071	0.058	0.062	0.057	0.069
0.2	0.070	0.093	0.063	0.096	0.072	0.093	0.070	0.074	0.065	0.089
0.3	0.075	0.110	0.076	0.114	0.081	0.114	0.081	0.091	0.074	0.108
0.4	0.088	0.135	0.087	0.135	0.089	0.133	0.089	0.108	0.083	0.129
0.5	0.100	0.158	0.096	0.163	0.098	0.159	0.098	0.117	0.094	0.161
0.6	0.107	0.191	0.105	0.193	0.105	0.183	0.112	0.129	0.105	0.190
0.7	0.116	0.218	0.115	0.221	0.113	0.209	0.122	0.140	0.114	0.223
0.8	0.127	0.246	0.128	0.250	0.125	0.239	0.133	0.153	0.122	0.249
0.9	0.136	0.278	0.137	0.272	0.133	0.264	0.141	0.164	0.138	0.274

Chapter 7

Testing Joint MA(1)-MA(4) against Joint AR(1)-AR(4) Disturbances in the Linear Regression Model

7.1 Introduction

Through the work of Thomas and Wallis (1971), and Wallis (1972), there is recognition that quarterly regression models may possess fourth-order autocorrelation because of seasonal variation. Wallis (1972) found a rigorous solution to the fourth-order autocorrelation testing problem by adapting the approach introduced by Durbin and Watson (1950, 1951) for the equivalent first-order testing problem. In fact Wallis (1972) and Vinod (1973) separately developed the fourth-order analogue of the Durbin and Watson test statistic which is now known as the Wallis test, while Webb (1973) suggested the generalisation of the Berenblut and Webb (1973) test. King (1984) considered the fourth-order analogue of his (1982b) locally optimal bounds test for first-order autoregressive disturbances. King (1989) noted that omitted or unobservable regressors with seasonal components might well lead to both first-order and fourth-order effects in the regression disturbances. He argued that the presence of first-order autocorrelation in a quarterly regression model is a good reason to suspect additional autocorrelation of a seasonal nature. This led him to propose a test for simple fourth-

order autoregressive [AR(4)] disturbances in the presence of first-order autoregressive [AR(1)] disturbances.

During last three decades, there has been an increased awareness of the moving average (MA) process in regression disturbances and the MA process has been seen as an important alternative [see, for example, Nicholls, Pagan and Terrell (1975)]. Perhaps also because the Durbin-Watson test is approximately LBI against either AR(1) or MA(1) disturbances in the linear regression model, interest has focused on testing AR(1) against MA(1) disturbances (see for example, Walker (1967), King (1983a), King and McAleer (1987), Burke et al. (1990), Smith and Tremayne (1990), Franses (1992), Silvapulle (1994a, 1994b), Baltagi and Li (1995)). The reverse problem of testing MA(1) against AR(1) disturbances has been investigated by Silvapulle and King (1991), Srikanthakumar (2000) and in Chapter 4. Walker (1967) developed Cox tests of $AR(p)$ processes against $MA(q)$ processes and vice versa. Silvapulle and King (1993) investigated the construction and application of POI tests of joint AR(1)-AR(4) disturbances against joint MA(1)-MA(4) disturbances in the linear regression model and they illustrated the testing problem with an application to a quarterly model of price inflation in the United Kingdom during 1947-1970. Thus the reverse testing problem is clearly also of interest. In addition, the increased awareness of MA processes and joint first-order and seasonal autocorrelation suggests the need for a test of joint MA(1)-MA(4) disturbances against joint AR(1)-AR(4) disturbances in quarterly regression models. Joint processes of the multiplicative type such as joint MA(1)-MA(4) and joint AR(1)-AR(4) processes have been found useful for empirical modelling of quarterly time series (for example, for joint MA(1)-MA(4) processes see Silvapulle and King (1993); for joint AR(1)-AR(4) processes see Box and Jenkins (1976), Prothero and

Wallis (1976), Wallis (1977), Burman (1980), Harvey (1981) and Granger and Newbold (1986)).

The aim of this chapter is to investigate the performance of the new test procedure introduced in Chapter 3, by applying it to the problem of testing joint MA(1)-MA(4) disturbances against joint AR(1)-AR(4) disturbances in the linear regression model. As our test procedure works well for the testing problem of MA(1) disturbances against AR(1) disturbances; and testing for heteroscedastic disturbances in the linear regression model, we now investigate how well it extends to testing two-dimensional parameter spaces under the null and alternative. Also, the aim of this chapter is to explore the properties of our testing procedure for multiplicative type models involving two-parameters. Particular issues are to see if the test procedure can be easily generalised to two-parameters and to see what an appropriate choice of m is in this case.

The plan of this chapter is as follows. Section 7.2 introduces the testing problem and the form of the test for testing joint MA(1)-MA(4) disturbances against joint AR(1)-AR(4) disturbances in the linear regression model. An outline of the simulation experiments is given in Section 7.3. Results of the simulations designed to explore the properties of the test procedure for $m = 1$, $m = 2$, $m = 2$ with special weights and $m = 3$, are reported in Sections 7.4, 7.5, 7.6 and 7.7, respectively. Section 7.8 contains some concluding remarks.

7.2 The testing problem and the form of the test

Consider the linear regression model (3.10). Suppose one possibility is that the components of the $n \times 1$ disturbance vector u are generated by the joint MA(1)-MA(4) process

$$u_t = (1 + \gamma_1 L)(1 + \gamma_4 L^4)e_t, \quad t = 1, 2, \dots, n, \quad (7.1)$$

where L is the lag operator such that $Le_t = e_{t-1}$, $0 < \gamma_i \leq 1$, $i = 1, 4$, are unknown parameters with at least one being non-zero and $e^* = (e_{-4}, e_{-3}, \dots, e_n)' \sim N(0, \sigma^2 I_{n+5})$.

The other possibility is that the components of u are generated by the stationary joint AR(1)-AR(4) process

$$(1 - \rho_1 L)(1 - \rho_4 L^4)u_t = e_t, \quad t = 1, 2, \dots, n, \quad (7.2)$$

where $0 \leq \rho_i < 1$, $i = 1, 4$, are unknown parameters and $e = (e_1, e_2, \dots, e_n)' \sim N(0, \sigma^2 I_n)$.

Now, consider the general fifth-order moving average (MA(5)) error process

$$u_t = e_t + \gamma_1 e_{t-1} + \gamma_2 e_{t-2} + \gamma_3 e_{t-3} + \gamma_4 e_{t-4} + \gamma_5 e_{t-5}.$$

By imposing restrictions on this process, that is, $\gamma_2 = \gamma_3 = 0$ and $\gamma_5 = \gamma_1 \gamma_4$, this reduces to

$$u_t = e_t + \gamma_1 e_{t-1} + \gamma_4 e_{t-4} + \gamma_1 \gamma_4 e_{t-5}. \quad (7.3)$$

Thus equation (7.1) is the special case of the general MA(5) process.

Similarly, by imposing similar restrictions on the general fifth-order autoregressive (AR(5)) error process yields

$$u_t = \rho_1 u_{t-1} + \rho_4 u_{t-4} - \rho_1 \rho_4 u_{t-5} + e_t, \quad (7.4)$$

which indicates that (7.2) is a special case of the stationary AR(5) process.

For equation (7.2), appropriate start up equations for the process following Silvapulle (1991) are

$$u_1 = b_{11}e_1,$$

$$u_2 = b_{21}u_1 + b_{22}e_2,$$

$$u_3 = b_{31}u_2 + b_{32}u_1 + b_{33}e_3,$$

$$u_4 = b_{41}u_3 + b_{42}u_2 + b_{43}u_1 + b_{44}e_4,$$

$$u_5 = b_{51}u_4 + b_{52}u_3 + b_{53}u_2 + b_{54}u_1 + b_{55}e_5,$$

where b_{ij} 's are functions of ρ_1 and ρ_4 , and we define

$$a_0 = \frac{(1 + \rho_1^2 + \rho_1^4 + \rho_1^6)}{(1 - \rho_1^8 \rho_4^2)((1 - \rho_1^8)(1 - \rho_4^2))} \sigma^2,$$

$$a_1 = \frac{\rho_1(1 + \rho_1^2 \rho_4)}{1 + \rho_1^4 \rho_4},$$

$$a_2 = \frac{\rho_1^2(1 + \rho_4)}{1 + \rho_1^4 \rho_4},$$

$$a_3 = \frac{\rho_1(\rho_1^2 + \rho_4)}{1 + \rho_1^4 \rho_4},$$

$$a_4 = \frac{\rho_1^4 + \rho_4}{1 + \rho_1^4 \rho_4},$$

$$b_{11} = \sqrt{a_0},$$

$$b_{21} = a_1,$$

$$b_{22} = b_{11} \sqrt{(1 - b_{21} a_1)},$$

$$b_{31} = \frac{a_1(1 - a_2)}{1 - a_1^2},$$

$$b_{32} = \frac{a_2 - a_1^2}{1 - a_1^2},$$

$$b_{33} = b_{11} \sqrt{(1 - b_{31} a_1 - b_{32} a_2)},$$

$$b_{41} = \frac{(1 - a_1^2)(a_1 - a_2 a_3) - a_1(1 - a_2)(a_2 - a_1 a_3)}{1 - 2a_1^2 + 2a_1^2 a_2 - a_2^2},$$

$$b_{42} = \frac{a_2 - a_1 a_3 - a_1(1 - a_2)b_{41}}{1 - a_1^2},$$

$$b_{43} = a_3 - b_{41} a_2 - b_{42} a_1,$$

$$b_{44} = b_{11} \sqrt{(1 - b_{41} a_1 - b_{42} a_2 - b_{43} a_3)},$$

$$b_{51} = \frac{c_1 c_6 - c_3 c_4}{c_6 c_2 - c_3 c_5},$$

$$b_{52} = \frac{c_1 - c_2 b_{51}}{c_6},$$

$$b_{53} = \frac{[a_3 - a_1 a_4 - (a_2 - a_1 a_3)b_{51} - (a_1 - a_1 a_2)b_{52}]}{1 - a_1^2},$$

$$b_{54} = a_4 - b_{51} a_3 - b_{52} a_2 - b_{53} a_1,$$

$$b_{55} = b_{11} \sqrt{(1 - b_{51} a_1 - b_{52} a_2 - b_{53} a_3 - b_{54} a_4)},$$

in which

$$c_1 = (1 - a_1^2)(a_1 - a_2 a_3) - (a_2 - a_1 a_3)(a_3 - a_1 a_4),$$

$$c_2 = (1 - a_1^2)(1 - a_2^2) - (a_2 - a_1 a_3)^2,$$

$$c_3 = (1 - a_1^2)(a_1 - a_2 a_3) - (a_2 - a_1 a_3)(a_1 - a_1 a_2),$$

$$c_4 = (1 - a_1^2)(a_2 - a_2 a_4) - (a_1 - a_1 a_2)(a_3 - a_1 a_4),$$

$$c_5 = (1 - a_1^2)(a_1 - a_2 a_3) - (a_1 - a_1 a_2)(a_2 - a_1 a_3),$$

$$c_6 = (1 - a_1^2)(1 - a_2^2) - (a_1 - a_1 a_2)^2,$$

(for details see Silvapulle, 1991, Appendix 4.1, p.88-92).

Let $\Sigma(\gamma)$ be the variance-covariance matrix of the MA(1) process, defined in Chapter 4 and $H(\gamma, n)$ be the Cholesky decomposition matrix of $\Sigma(\gamma)$, that is, by definition of Cholesky decomposition, $H(\gamma, n)H(\gamma, n)' = \Sigma(\gamma)$.

Thus $H(\gamma, n)$ is the $n \times n$ bi-diagonal matrix defined by Silvapulle (1991, p.70) as follows

$$H(\gamma, n) = \begin{bmatrix} h_{11} & 0 & 0 & \cdots & \cdots & 0 & 0 \\ h_{21} & h_{22} & 0 & \cdots & \cdots & 0 & 0 \\ 0 & h_{32} & h_{33} & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & h_{n-1,n-1} & 0 \\ 0 & 0 & 0 & \cdots & \cdots & h_{n,n-1} & h_{nn} \end{bmatrix}.$$

The non-zero elements of $H(\gamma, n)$ are generated by the recursive method as follows:

$$h_{11} = (1 + \gamma^2)^{1/2},$$

$$h_{i,j-1} = \gamma / h_{i-1,j-1}, \text{ and } h_{ii} = [(1 + \gamma^2) - h_{i,j-1}^2]^{1/2}, \quad i = 2, 3, \dots, n.$$

Suppose, $H_1(\gamma_1) = H(\gamma_1, n)$. If n is an integer multiple of 4, define

$$H_4(\gamma_4) = H(\gamma_4, n/4) \otimes I_4, \quad (7.5)$$

where \otimes is the Kronecker product. If n is not an integer multiple of 4, $H_4(\gamma_4)$ is defined by first rounding n up to the next integer multiple of 4, applying (7.5) and then deleting the extra number of bottom rows and right-hand columns to convert it back to an $n \times n$ matrix.

Now, the process defined by (7.1) can be expressed as

$$u = H_1(\gamma_1)H_4(\gamma_4)e,$$

and

$$\text{Var}(u) = \sigma^2 H_1(\gamma_1)H_4(\gamma_4)H_4(\gamma_4)'H_1(\gamma_1)'$$

$$= \sigma^2 \Sigma(\gamma_1, \gamma_4), \text{ say.}$$

The process (7.2) can be expressed as

$$G(\rho_1, \rho_4)u = e,$$

and

$$\text{Var}(u) = \sigma^2 G(\rho_1, \rho_4)^{-1} G(\rho_1, \rho_4)'^{-1}$$

$$= \sigma^2 \Omega(\rho_1, \rho_4), \text{ say,}$$

where $G(\rho_1, \rho_4)$ is the inverse of the Cholesky decomposition matrix of the variance-covariance matrix of the stationary joint AR(1)-AR(4) process, i.e., $\Omega(\rho_1, \rho_4) = G(\rho_1, \rho_4)^{-1} G(\rho_1, \rho_4)'^{-1}$. Silvapulle (1991), Wu (1991) and Ara (1995) derived the $G(\rho_1, \rho_4)$ matrix using different approaches. Following Silvapulle (1991), the $G(\rho_1, \rho_4)$ matrix can be written as

$$G(\rho_1, \rho_4) = \begin{bmatrix} \frac{1}{b_{11}} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\frac{b_{21}}{b_{22}} & \frac{1}{b_{22}} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\frac{b_{32}}{b_{33}} & -\frac{b_{31}}{b_{33}} & \frac{1}{b_{33}} & 0 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\frac{b_{43}}{b_{44}} & -\frac{b_{42}}{b_{44}} & -\frac{b_{41}}{b_{44}} & \frac{1}{b_{44}} & 0 & 0 & 0 & \cdots & 0 & 0 \\ -\frac{b_{54}}{b_{55}} & -\frac{b_{53}}{b_{55}} & -\frac{b_{52}}{b_{55}} & -\frac{b_{51}}{b_{55}} & \frac{1}{b_{55}} & 0 & 0 & \cdots & 0 & 0 \\ \rho_1 \rho_4 & -\rho_4 & 0 & 0 & -\rho_1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & \rho_1 \rho_4 & -\rho_4 & 0 & 0 & -\rho_1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & -\rho_1 & 1 \end{bmatrix}$$

where the b_{ij} 's are functions of ρ_1 and ρ_4 , are defined above.

It can be shown that $b_{42} = b_{32} = b_{33} = 0$, (for details see Silvapulle (1991, Appendix 4.1, p.88-92); Wu (1991, p.77-79); or Ara (1995, p.87-88)).

Our main concern is in testing

$$H_0: u \sim N(0, \sigma^2 \Sigma(\gamma_1, \gamma_4)), \quad 0 < \gamma_i \leq 1, \quad i = 1, 4, \quad (7.6)$$

against

$$H_a: u \sim N(0, \sigma^2 \Omega(\rho_1, \rho_4)), \quad 0 \leq \rho_i < 1, \quad i = 1, 4, \quad (7.7)$$

in the context of the linear regression model (3.10). Observe that we are assuming positive first and fourth-order autocorrelation in the disturbances because positive autocorrelation is the most likely outcome and therefore is the most interesting case to consider. The discussion below for the test procedure can easily be generalized to the case of negative autocorrelation.

The testing problem is invariant to transformations of the form (3.13). Now, using the results of (3.16) and (3.17), our problem of testing (7.6) against (7.7) in the context of (3.10), becomes one of testing (after cancellation of some constants)

$$H_0: h_0(v, \gamma_1, \gamma_4) \propto |\Sigma(\gamma_1, \gamma_4)|^{-1/2} |X' \Sigma(\gamma_1, \gamma_4)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_1, \gamma_4)' \tilde{z}(\gamma_1, \gamma_4)}{z'z} \right)^{-p/2},$$

$$0 < \gamma_i \leq 1, \quad i = 1, 4, \quad (7.8)$$

against

$$H_a: h_a(v, \rho_1, \rho_4) \propto |\Omega(\rho_1, \rho_4)|^{-1/2} |X' \Omega(\rho_1, \rho_4)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_1, \rho_4)' \tilde{z}(\rho_1, \rho_4)}{z'z} \right)^{-p/2},$$

$$0 \leq \rho_i < 1, \quad i = 1, 4, \quad (7.9)$$

where z is the OLS residual vector from (3.10) and $\tilde{z}(\gamma_1, \gamma_4)$ denotes $\tilde{z}(\lambda)$ with $\Lambda(\lambda) = \Sigma(\gamma_1, \gamma_4)$ and also, $\tilde{z}(\rho_1, \rho_4)$ denotes $\tilde{z}(\lambda)$ with $\Lambda(\lambda) = \Omega(\rho_1, \rho_4)$.

Observe that through invariance arguments and our assumption of positive autocorrelation, we have restricted the null and alternative hypothesis parameter spaces to $\Theta = \{(\gamma_1, \gamma_4): 0 < \gamma_i \leq 1, \quad i = 1, 4\}$, and $\Phi = \{(\rho_1, \rho_4): 0 \leq \rho_i < 1, \quad i = 1, 4\}$, respectively.

Let us first consider the test, for which we control average size over Θ , i.e., $m = 1$. The fundamental NP lemma implies that the test which maximizes average power subject to controlling average size over the entire null hypothesis parameter space, can be based on the critical region of the form

$$\int_0^1 \int_0^1 |\Omega(\rho_1, \rho_4)|^{-1/2} |X' \Omega(\rho_1, \rho_4)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_1, \rho_4)' \tilde{z}(\rho_1, \rho_4)}{z'z} \right)^{-p/2} p(\rho_1, \rho_4) d\rho_1 d\rho_4$$

$$\geq c_\alpha \int_0^1 \int_0^1 |\Sigma(\gamma_1, \gamma_4)|^{-1/2} |X' \Sigma(\gamma_1, \gamma_4)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_1, \gamma_4)' \tilde{z}(\gamma_1, \gamma_4)}{z'z} \right)^{-p/2} p(\gamma_1, \gamma_4) d\gamma_1 d\gamma_4$$

$$\text{or, } \frac{\int_0^1 \int_0^1 |\Omega(\rho_1, \rho_4)|^{-1/2} |X' \Omega(\rho_1, \rho_4)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_1, \rho_4)' \tilde{z}(\rho_1, \rho_4)}{z' z} \right)^{-p/2} d\rho_1 d\rho_4}{\int_0^1 \int_0^1 |\Sigma(\gamma_1, \gamma_4)|^{-1/2} |X' \Sigma(\gamma_1, \gamma_4)^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_1, \gamma_4)' \tilde{z}(\gamma_1, \gamma_4)}{z' z} \right)^{-p/2} d\gamma_1 d\gamma_4} \geq c_\alpha \quad (7.10)$$

where $p(\rho_1, \rho_4)$ and $p(\gamma_1, \gamma_4)$ are the weighting functions for the joint distribution of ρ_1 and ρ_4 , and γ_1 and γ_4 , respectively and c_α is the calculated critical value which controls the average size of the test. In inequality (7.10), $p(\rho_1, \rho_4)$ and $p(\gamma_1, \gamma_4)$ have been replaced by the uniform density functions.

If we denote the critical region given by (7.10) as $\omega(c_\alpha)$, then finding c_α involves solving

$$\int_{\omega(c_\alpha)} \int_0^1 \int_0^1 h_0(v, \gamma_1, \gamma_4) d\gamma_1 d\gamma_4 dv = \int_0^1 \int_0^1 \int_{\omega(c_\alpha)} h_0(v, \gamma_1, \gamma_4) dv d\gamma_1 d\gamma_4 = \alpha \quad (7.11)$$

for c_α , where α is the desired average size of the test.

Inequality (7.10) and equation (7.11) both involve integrals, which have no analytical solutions. For a numerical solution, we can approximate the integrals given in (7.10) by numerical integrals so that it is approximately equivalent to

$$\frac{\sum_{j=1}^q \sum_{l=1}^q |\Omega(\rho_{1j}, \rho_{4l})|^{-1/2} |X' \Omega(\rho_{1j}, \rho_{4l})^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\rho_{1j}, \rho_{4l})' \tilde{z}(\rho_{1j}, \rho_{4l})}{z' z} \right)^{-p/2}}{\sum_{j=1}^q \sum_{l=1}^q |\Sigma(\gamma_{1j}, \gamma_{4l})|^{-1/2} |X' \Sigma(\gamma_{1j}, \gamma_{4l})^{-1} X|^{-1/2} \left(\frac{\tilde{z}(\gamma_{1j}, \gamma_{4l})' \tilde{z}(\gamma_{1j}, \gamma_{4l})}{z' z} \right)^{-p/2}} \geq c_\alpha \quad (7.12)$$

where $\gamma_{1j} = \rho_{1j} = (2j-1)/2q$, and $\gamma_{4l} = \rho_{4l} = (2l-1)/2q$, $j, l = 1, 2, \dots, q$.

As discussed in Chapter 3, a potential problem is with controlling the average size of the test. We conducted some simulation experiments, for $m=1$, $m=2$, $m=2$ with special weights and $m=3$ for testing (7.6) against (7.7) in the context of (3.10).

Detailed discussion of the test procedures, the size and the power properties of the test with $m=1$, $m=2$, $m=2$ with special weights and $m=3$ are given in the following sections.

7.3 Outline of the simulation experiments

In this section, we report the results of simulation experiments to assess the small sample size and power properties of our test in the linear regression model. The main objectives of the experiments are to judge the performance of the test procedure for the joint distribution of multiplicative type processes with an appropriate choice of m which approximately controls the size of the test.

In this simulation experiment, 2000 replications were used to calculate the exact five percent critical values. Sizes and powers of the test were also calculated at the five percent level of significance. We calculated the critical values with the disturbances of (3.10) being generated following process (7.1). The sizes and powers of the test were

calculated at each of the 25 grid-points, defined as $\{(\gamma_1, \gamma_4): \gamma_1, \gamma_4 = 0.1, 0.3, 0.5, 0.7, 0.9\}$, under the null hypothesis parameter space Θ and $\{(\rho_1, \rho_4): \rho_1, \rho_4 = 0.1, 0.3, 0.5, 0.7, 0.9\}$, under the alternative hypothesis parameter space Φ , respectively.

For this testing problem, in order to calculate the size and power properties of the test we used the $X2$, $X3$ and $X5$ design matrices defined in Chapter 4 and design matrices $X6$ which is as follows:

$X6: (n \times 6; n = 20, 60)$. A constant dummy, both current and lagged quarterly Australian private capital movements, current and lagged quarterly Australian Government capital movements and Australian retail trade commencing 1968(1).

These design matrices were chosen to cover a range of economic applications with different degrees of seasonality. Moderate and steady seasonality are found in the Australian retail trade series. Also, large fluctuations and strong seasonality with two seasonal peaks per year are found in the two capital movement series.

We mentioned in the previous section that the testing problem is invariant with respect to the nuisance parameters, β and σ^2 , and we have chosen, $\beta = 0$ and $\sigma^2 = 1$ for all simulation experiments.

The first step involved deciding on the value of m for which the estimated size of the test is approximately controlled. We pointed out earlier that we conducted some simulation experiments to choose the value of m and we tried $m = 1$, $m = 2$, $m = 2$ with special weights and $m = 3$. Finally we found that the test performs well for $m = 3$. Detailed discussions follow.

7.4 Results for $m = 1$

Considering the entire null hypothesis parameter space Θ i.e., $m = 1$, we conducted simulation experiments to calculate the critical values c_α of the test statistic for all design matrices based on (7.10), which control average size over Θ . The calculated critical values are presented in Table 7.1. Using these critical values, we calculated the sizes and powers of the test through equation (7.12), respectively for 25 grid-points of (γ_1, γ_4) under H_0 and for 25 grid-points of (ρ_1, ρ_4) under H_a defined in the previous section. For design matrices $X2$, $X3$, $X5$ and $X6$, the calculated sizes and powers of the test are presented in Tables 7.2a, 7.2b, 7.2c and 7.2d, respectively. A value denoted by a 'star' is significantly different from five percent at the 0.01 level under H_0 .

From the size analysis of the test for $m = 1$, we see from Tables 7.2a-7.2d that there are some sizes which are significantly different from the five percent significant level, when the sample size is small. We also observe that among these significant sizes, a few sizes are far larger than the average size, for example, for the grid-points $\{(\gamma_1, \gamma_4) = (0.1, 0.5), (0.5, 0.5)\}$ for $X2$, $\{(\gamma_1, \gamma_4) = (0.1, 0.3), (0.1, 0.5), (0.3, 0.3), (0.3, 0.5), (0.5, 0.5)\}$ for $X3$ and $\{(\gamma_1, \gamma_4) = (0.3, 0.3)\}$ for $X5$ and $X6$. When the sample size is large, we find that almost all sizes are significantly different from the five percent level but for the grid-points $\{(\gamma_1, \gamma_4): \gamma_1, \gamma_4 = 0.1, 0.3, 0.5\}$ the sizes are undesirably larger than the average size while others are smaller than the average size. This shows that the test does not successfully control the average size over Θ with $m = 1$. This motivates us to consider two subregions over Θ , i.e., $m = 2$.

Also, from Tables 7.2a-7.2d we see that the calculated powers of the test for $m = 1$, always increase as the grid-point under H_a moves further away from $(\rho_1, \rho_4) = (0, 0)$. When the sample size is large, the powers are typically much higher than the powers for the smaller sample size. These power properties of the test are almost the same for all design matrices.

In the following section we discuss the results, i.e. the size and power properties of the test for $m = 2$.

7.5 Results for $m = 2$

We want to control the estimated sizes of the test to be 0.05 over suitable subregions under H_0 . When we controlled the average size over Θ for the $m = 1$ version of the test, we found some sizes are above average and some sizes are below average. Using the high size/low size approach we divided Θ into two subregions. Accordingly, based on the calculated sizes of the test for $m = 1$ presented in Tables 7.2a-7.2d, we divided the null hypothesis parameter space $\{0 < \gamma_i \leq 1, i = 1, 4\}$ into two disjoint regions, with the boundary points calculated through the high size/low size approach. For example, in the case of the X^2 design matrix with $n = 20$, smaller than average sizes lie in the regions $(0 < \gamma_1, \gamma_4 \leq 0.2)$ and $(0.6 < \gamma_1 \leq 1 \cap 0 < \gamma_4 \leq 1)$, and larger than average sizes lie in the regions $(0 < \gamma_1 \leq 0.2 \cap 0.2 < \gamma_4 \leq 1)$ and $(0.2 < \gamma_1 \leq 0.6 \cap 0 < \gamma_4 \leq 1)$, and similarly $n = 60$ for X^2 , smaller than average sizes lie in the regions $(0 < \gamma_1 \leq 0.6 \cap 0.6 < \gamma_4 \leq 1)$ and $(0.6 < \gamma_1 \leq 1 \cap 0 < \gamma_4 \leq 1)$, and larger than average sizes lie in the region $(0 < \gamma_1, \gamma_4 \leq 0.6)$; see Table 7.2a. Ranges of the two disjoint regions and the corresponding values of j and l , where $\gamma_{1j} = (2j-1)/2q$, and $\gamma_{4l} = (2l-1)/2q$,

$j = l = 1, 2, \dots, q, q = 50$, for all design matrices are presented in Table 7.3. In each case the Θ_1 region (smaller than average size) is shaded in green and the Θ_2 region (higher than average size) is shaded in red.

Thus for $m = 2$, equation (3.7) gives the resultant test that maximizes average power subject to two size conditions and involves a critical region of the form (7.12) which now becomes

$$\begin{aligned} & \sum_{j=1}^q \sum_{l=1}^q \left| \Omega(\rho_{1j}, \rho_{4l}) \right|^{-1/2} \left| X' \Omega(\rho_{1j}, \rho_{4l})^{-1} X \right|^{-1/2} \left(\frac{\tilde{z}(\rho_{1j}, \rho_{4l})' \tilde{z}(\rho_{1j}, \rho_{4l})}{z'z} \right)^{-p/2} \\ & \geq \left[k_1 \sum_{(\gamma_{1j}, \gamma_{4l}) \in \Theta_1} \left| \Sigma(\gamma_{1j}, \gamma_{4l}) \right|^{-1/2} \left| X' \Sigma(\gamma_{1j}, \gamma_{4l})^{-1} X \right|^{-1/2} \left(\frac{\tilde{z}(\gamma_{1j}, \gamma_{4l})' \tilde{z}(\gamma_{1j}, \gamma_{4l})}{z'z} \right)^{-p/2} \right. \\ & \quad \left. + k_2 \sum_{(\gamma_{1j}, \gamma_{4l}) \in \Theta_2} \left| \Sigma(\gamma_{1j}, \gamma_{4l}) \right|^{-1/2} \left| X' \Sigma(\gamma_{1j}, \gamma_{4l})^{-1} X \right|^{-1/2} \left(\frac{\tilde{z}(\gamma_{1j}, \gamma_{4l})' \tilde{z}(\gamma_{1j}, \gamma_{4l})}{z'z} \right)^{-p/2} \right] \end{aligned} \quad (7.13)$$

where k_1 and k_2 are critical values calculated to control average size over the two subregions of Θ .

If we denote the critical region given by (7.13) as $\omega(k_1, k_2)$ and if Θ_1 are the set of (γ_1, γ_4) points indicated in Table 7.3 shaded in green and Θ_2 are the set of (γ_1, γ_4) points indicated in Table 7.3 shaded in red, then finding k_1 and k_2 involves solving

$$\iint_{\Theta_i} \int_{\omega(k_1, k_2)} h_0(v, \gamma_1, \gamma_4) dv d\gamma_1 d\gamma_4 = \alpha, \quad i = 1, 2, \quad (7.14)$$

jointly for k_1 and k_2 .

The critical values k_1 and k_2 were calculated via the trial and error method, and are presented in Table 7.4, for all design matrices. From this table we observe that k_2 values are far larger than k_1 values for all cases. The values of k_2 are always positive while k_1 takes positive values when the sample size is small and negative values when the sample size is large. From the critical values we also observe that both k_1 and k_2 values decrease as the sample size increases for all design matrices.

Using the critical values k_1 and k_2 presented in Table 7.4, we calculated the sizes and powers of the test through equation (7.13), for 25 grid-points of (γ_1, γ_4) under H_0 and for 25 grid-points of (ρ_1, ρ_4) under H_a , respectively. For all design matrices, the calculated sizes and powers of the test are presented in Tables 7.5a-7.5d.

Now if we analyse the calculated sizes of the test for $m = 2$, we see from Tables 7.5a-7.5d that our test successfully controls the average size over the two regions under H_0 , when the sample size is small. In other words, when the sample size is small, the calculated sizes of the test fall within the rejection probabilities range $[0.037, 0.063]$ at the five percent significance level or are less than 0.05. But when the sample size is large we observe that there are a few sizes which are significantly different from five percent. Among these significant sizes, in particular, one calculated size which is much smaller than the average size is that for grid-point $\{(\gamma_1, \gamma_4) = (0.1, 0.1)\}$ and the calculated size for grid-point $\{(\gamma_1, \gamma_4) = (0.9, 0.9)\}$ is rather larger than the average size. This indicates that for large samples, our test does not successfully control size over Θ for $m = 2$.

From the analysis of powers of the test for $m = 2$, we see that the powers of the test are well behaved for all cases. As the sizes of the test for $m = 1$ and $m = 2$ are different, so the powers for $m = 1$ and $m = 2$ cannot be compared in a meaningful way.

7.6 Results for $m = 2$ with special weights

In this simulation experiment we have chosen only the $X5$ design matrix with $n = 60$, to look at whether our test procedure works well or not with some special weights. We restricted our choice to a single design matrix, because we have to calculate the critical values, k_1 and k_2 , so that the two size conditions are simultaneously satisfied, via the trial and error method. The calculation for one trial takes about two days. Typically, calculating the critical values requires 10-30 trials, i.e., 20-60 days. Then follows the calculation of sizes and powers of the test. Therefore, calculating critical values and the required sizes and powers of the test is laborious and time consuming. So our choice had to be limited to one design matrix. This is a clear drawback of our simulation experiment but we can at least gauge the performance of the test procedure in this case.

In the previous section, we mentioned that the calculated sizes of the test for $m = 2$ with $n = 60$, behave badly for the grid-points $\{(\gamma_1, \gamma_4) = (0.1, 0.1)\}$ and $\{(\gamma_1, \gamma_4) = (0.9, 0.9)\}$. Suppose we give some extra weight to the troublesome parts of the null hypothesis parameter space then the size might be controlled and the largest and smallest sizes might be close to 0.05. This led us to consider $m = 2$, with special weights for troublesome parts of parameter space such as $\{(\gamma_1, \gamma_4): \gamma_1 \in [0, 0.2], \gamma_4 \in [0, 0.2]\}$ i.e., for the grid-points $\{(\gamma_1, \gamma_4) = (0.1, 0.1)\}$ and $\{(\gamma_1, \gamma_4): \gamma_1 \in [0.8, 1], \gamma_4 \in [0.8, 1]\}$ i.e., for the grid-points $\{(\gamma_1, \gamma_4) = (0.9, 0.9)\}$.

In this experiment, we used the same two disjoint subregions of Θ and the corresponding values of j and l , for the $X5$ design matrix for $m=2$ with $n=60$ which were presented in Table 7.3, but with special weights on $\{(\gamma_1, \gamma_4): \gamma_1 \in [0, 0.2], \gamma_4 \in [0, 0.2]\}$ and $\{(\gamma_1, \gamma_4): \gamma_1 \in [0.8, 1], \gamma_4 \in [0.8, 1]\}$ (four-times the normal weight).

Applying the four-times weights for the grid-points $\{(\gamma_1, \gamma_4) = (0.1, 0.1)\}$ and $\{(\gamma_1, \gamma_4) = (0.9, 0.9)\}$, we calculated the critical values k_1 and k_2 which turned out to be $k_1 = -0.000006$ and $k_2 = 1.0$. If we compare these critical values with the previous critical values for $m=2$ with equal (normal) weights, we see that new values are quite different to the old values.

Using the critical values $k_1 = -0.000006$ and $k_2 = 1.0$, through equation (7.13) we calculated the sizes and powers of the test which are presented in Table 7.6. From this table we observe that there is no great improvement in the calculated sizes of the test. While some calculated sizes improve, other calculated sizes are worst than previously. The results indicate that giving special heavy weights to some part of the null parameter space is helping to control size in that part of the space but at the expense of losing some control in other parts of the space. We conclude that the test procedure does not work that well with special weights for $m=2$. The test procedure may be able to be tuned to make it work by changing weights but we do not know for which weights it will work well. It would take a very large amount of computation to check this out. Therefore we considered $m=3$.

7.7 Results for $m=3$

In this section, we have also chosen only the $X5$ design matrix with $n=60$, to conduct the simulation experiments for $m=3$. The reason is as explained in the previous section. For $m=3$, calculating critical values and the required sizes and powers of the test is much more laborious and time consuming than for $m=2$. Therefore we have restricted our choice to a single design matrix.

Thus, based on the calculated sizes of the test for $m=1$ and $m=2$ presented in Table 7.2c and Table 7.5c, respectively, we divided the null hypothesis parameter space into three disjoint regions and controlled average size over each of the subregions. We used the high size/low size technique for division into two ($m=2$) regions of the test for $m=1$. In this case ($m=2$) there was one troublesome region (low size) and we divided the low size region into two with the troublesome high size parts of this region certainly the third region. That is, one subregion (Θ_1) is that for which the smaller than average sizes resulted, sizes larger than average size determined the second subregion (Θ_2) and the third subregion (Θ_3) covered those regions which resulted very high size (see Table 7.5c). Ranges of the three disjoint regions and the corresponding values of j and l , where $\gamma_{1j} = (2j-1)/2q$, and $\gamma_{4l} = (2l-1)/2q$, $j, l = 1, 2, \dots, q$, are given in Table 7.7. The Θ_1 region is shaded in green, the Θ_2 region is shaded in red and the Θ_3 region is shaded in yellow.

Thus for $m=3$, equation (3.7) gives the resultant test that maximizes average power subject to three size conditions and involves a critical region of the form (7.12) which now becomes

$$\begin{aligned}
& \sum_{j=1}^q \sum_{l=1}^q \left| \Omega(\rho_{1j}, \rho_{4l}) \right|^{-1/2} \left| X' \Omega(\rho_{1j}, \rho_{4l})^{-1} X \right|^{-1/2} \left(\frac{\tilde{z}(\rho_{1j}, \rho_{4l})' \tilde{z}(\rho_{1j}, \rho_{4l})}{z' z} \right)^{-p/2} \\
& \geq \left[k_1 \sum_{(y_1, y_4) \in \Theta_1} \left| \Sigma(y_1, y_4) \right|^{-1/2} \left| X' \Sigma(y_1, y_4)^{-1} X \right|^{-1/2} \left(\frac{\tilde{z}(y_1, y_4)' \tilde{z}(y_1, y_4)}{z' z} \right)^{-p/2} \right. \\
& \quad + k_2 \sum_{(y_1, y_4) \in \Theta_2} \left| \Sigma(y_1, y_4) \right|^{-1/2} \left| X' \Sigma(y_1, y_4)^{-1} X \right|^{-1/2} \left(\frac{\tilde{z}(y_1, y_4)' \tilde{z}(y_1, y_4)}{z' z} \right)^{-p/2} \\
& \quad \left. + k_3 \sum_{(y_1, y_4) \in \Theta_3} \left| \Sigma(y_1, y_4) \right|^{-1/2} \left| X' \Sigma(y_1, y_4)^{-1} X \right|^{-1/2} \left(\frac{\tilde{z}(y_1, y_4)' \tilde{z}(y_1, y_4)}{z' z} \right)^{-p/2} \right] \quad (7.15)
\end{aligned}$$

where k_1 , k_2 and k_3 are critical values calculated to control average size over the three subregions of Θ .

If we denote the critical region given by (7.15) as $\omega(k_1, k_2, k_3)$ and if Θ_1 is the set of (y_1, y_4) points shaded in green, Θ_2 is the set of (y_1, y_4) points shaded in red and Θ_3 is the set of (y_1, y_4) points shaded in yellow in Table 7.7, then finding k_1 , k_2 and k_3 involves solving

$$\iint_{\Theta_i} \int_{\omega(k_1, k_2, k_3)} h_0(v, y_1, y_4) dv dy_1 dy_4 = \alpha, \quad i = 1, 2, 3, \quad (7.16)$$

jointly for k_1 , k_2 and k_3 .

The critical values k_1 , k_2 and k_3 were calculated via trial and error, so that the three size conditions are satisfied. The calculated critical values are $k_1 = -0.000845$, $k_2 = 1.6$ and $k_3 = 0.000012$. Using these critical values through equation (7.15), we calculated the sizes and the powers of the test, which are presented in Table 7.8. From this table, we observe that there is a large improvement in the size performance of the test for $m = 3$ but yet a few sizes are significantly different from five percent. There are only two sizes which are higher than the upper limit of the rejection probability range, these are for the grid-points $\{(y_1, y_4) = (0.3, 0.5)\}$ and $\{(y_1, y_4) = (0.9, 0.9)\}$. The sizes for the grid-points $\{(y_1, y_4) = (0.3, 0.5)\}$ and $\{(y_1, y_4) = (0.9, 0.9)\}$ were 0.062 and 0.194 for $m = 2$, now these sizes are controlled to 0.064 and 0.074, respectively. Also we observe that where the sizes of the test for $m = 2$ were smaller than 0.05, now these sizes come very close to 0.05. Overall we observe that the sizes of the test for $m = 3$ are better controlled than the sizes of the test for $m = 2$, for $X5$ and $n = 60$. If we compare the powers of the test for $m = 3$ with $m = 2$, we see that the powers for $m = 3$ are bit higher than for $m = 2$, but not remarkably higher. In other words, the power properties of the test are almost unchanged.

7.8 Conclusions

This chapter investigated the problem of testing for joint MA(1)-MA(4) disturbances against joint AR(1)-AR(4) disturbances in the linear regression model. We are interested in this testing problem because the models we are considering both under the null and alternative hypothesis are two-dimensional parameter space models. In this testing problem, we have applied the new test approach as an example of an application to a multidimensional parameter space under the null hypothesis. From the results, we see

that the size and power performances of the test are very good for $m=2$, when the sample size is small. For the larger sample size, the estimated average sizes of the test do not perform well for $m=2$, i.e., all calculated sizes do not fall within the rejection probability ranges. Also, further results for $m=2$ with special weights indicate that the size properties of the test for $n=60$ do not perform as well. There is no improvement of the size performance of the test for $m=2$ with the special weights perhaps because we did not follow a system to determine the weights for different parts of the null parameter space. From the simulation results of the test for $m=3$ with $n=60$, we can say that the size properties of the test are close to acceptable. The results reported for $m=3$ with $n=60$, suggested that, the extra computation of the test procedure is well rewarded at least in terms of size. As a result, we can conclude that our new test procedure can be made to work for testing problems involving two-dimensional parameter space models but it is time consuming to apply in practice, and may require future advances in computer technology to make it a feasible test in this case.

Table 7.1

Critical values (c_α) of the test with $m=1$, for testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances for all design matrices at the five percent level.

Design matrix	Sample sizes (n)	Critical values (c_α)
$X2(n \times 3)$	20	2.9115
	60	1.8080
$X3(n \times 5)$	20	2.6663
	60	1.8128
$X5(n \times 3)$	20	3.3635
	60	1.9682
$X6(n \times 6)$	20	2.5588
	60	1.9434

Table 7.2a

Selected calculated sizes and powers of the test for the $X2$ design matrix when $m=1$, at the five percent level.

$n=20 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1=9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1=1$	0.3	0.5	0.7	0.9
0.1	0.016*	0.031*	0.054	0.058	0.041	0.1	0.038	0.078	0.162	0.297	0.498
0.3	0.018*	0.038	0.058	0.061	0.063	0.3	0.067	0.102	0.184	0.335	0.548
0.5	0.017*	0.027*	0.066*	0.058	0.066*	0.5	0.155	0.173	0.231	0.376	0.602
0.7	0.016*	0.035*	0.054	0.060	0.064*	0.7	0.402	0.392	0.405	0.490	0.672
0.9	0.018*	0.032*	0.056	0.058	0.068*	0.9	0.834	0.821	0.803	0.785	0.835

$n=60 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1=9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1=1$	0.3	0.5	0.7	0.9
0.1	0.001*	0.025*	0.096*	0.152*	0.106*	0.1	0.111	0.292	0.619	0.893	0.988
0.3	0.001*	0.030*	0.112*	0.181*	0.164*	0.3	0.316	0.412	0.662	0.910	0.990
0.5	0.003*	0.025*	0.084*	0.121*	0.112*	0.5	0.636	0.646	0.765	0.931	0.995
0.7	0.001*	0.010*	0.029*	0.036*	0.033*	0.7	0.911	0.909	0.929	0.979	0.999
0.9	0.000*	0.004*	0.006*	0.008*	0.004*	0.9	0.997	0.993	0.994	0.999	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.2b

Selected calculated sizes and powers of the test for the $X3$ design matrix when $m = 1$, at the five percent level.

$n = 20 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.015*	0.032*	0.053	0.062	0.045	0.1	0.044	0.077	0.138	0.243	0.422
0.3	0.017*	0.033*	0.057	0.073*	0.065*	0.3	0.067	0.088	0.151	0.280	0.471
0.5	0.017*	0.045	0.068*	0.079*	0.079*	0.5	0.149	0.156	0.200	0.323	0.528
0.7	0.017*	0.042	0.064*	0.071*	0.073*	0.7	0.386	0.359	0.365	0.438	0.594
0.9	0.025*	0.046	0.066*	0.073*	0.081*	0.9	0.816	0.800	0.772	0.745	0.778

$n = 60 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.002*	0.024*	0.104*	0.155*	0.106*	0.1	0.107	0.277	0.606	0.885	0.985
0.3	0.003*	0.034*	0.113*	0.177*	0.159*	0.3	0.299	0.396	0.634	0.897	0.986
0.5	0.003*	0.031*	0.084*	0.123*	0.109*	0.5	0.614	0.625	0.747	0.925	0.994
0.7	0.001*	0.013*	0.035*	0.039	0.039	0.7	0.899	0.895	0.914	0.971	0.999
0.9	0.001*	0.003*	0.012*	0.010*	0.007*	0.9	0.996	0.995	0.994	0.999	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.2c

Selected calculated sizes and powers of the test for the $X5$ design matrix when $m = 1$, at the five percent level.

$n = 20 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.006*	0.029*	0.072*	0.073*	0.044	0.1	0.048	0.099	0.169	0.253	0.339
0.3	0.009*	0.031*	0.063	0.067*	0.054	0.3	0.072	0.102	0.167	0.251	0.343
0.5	0.011*	0.035*	0.056	0.059	0.056	0.5	0.157	0.175	0.212	0.282	0.370
0.7	0.011*	0.032*	0.049	0.051	0.053	0.7	0.391	0.385	0.386	0.409	0.450
0.9	0.008*	0.027*	0.046	0.046	0.043	0.9	0.829	0.822	0.802	0.753	0.681

$n = 60 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.002*	0.022*	0.093*	0.142*	0.093*	0.1	0.098	0.278	0.606	0.870	0.977
0.3	0.002*	0.025*	0.108*	0.173*	0.150*	0.3	0.299	0.385	0.633	0.877	0.980
0.5	0.001*	0.023*	0.079*	0.115*	0.108*	0.5	0.631	0.629	0.747	0.909	0.982
0.7	0.001*	0.010*	0.033*	0.033*	0.032*	0.7	0.906	0.897	0.912	0.959	0.992
0.9	0.001*	0.002*	0.005*	0.005*	0.002*	0.9	0.996	0.995	0.993	0.997	0.999

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.2d

Selected calculated sizes and powers of the test for the $X6$ design matrix when $m = 1$, at the five percent level.

$n = 20 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.037	0.050	0.063	0.061	0.054	0.1	0.065	0.081	0.134	0.234	0.433
0.3	0.034*	0.047	0.059	0.068*	0.058	0.3	0.084	0.086	0.137	0.256	0.490
0.5	0.033*	0.041	0.055	0.058	0.058	0.5	0.136	0.135	0.171	0.293	0.545
0.7	0.035*	0.044	0.053	0.054	0.063	0.7	0.311	0.291	0.285	0.360	0.611
0.9	0.033*	0.043	0.049	0.055	0.062	0.9	0.739	0.699	0.648	0.600	0.710

$n = 60 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.002*	0.020*	0.094*	0.139*	0.091*	0.1	0.095	0.257	0.558	0.848	0.981
0.3	0.005*	0.028*	0.111*	0.164*	0.146*	0.3	0.276	0.357	0.601	0.873	0.985
0.5	0.004*	0.027*	0.082*	0.113*	0.108*	0.5	0.580	0.590	0.728	0.908	0.993
0.7	0.002*	0.013*	0.035*	0.039	0.035*	0.7	0.890	0.878	0.895	0.960	0.998
0.9	0.000*	0.004*	0.010*	0.008*	0.008*	0.9	0.994	0.994	0.993	0.997	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.3

Ranges of the two disjoint intervals and the corresponding values of j and l , where $\gamma_{lj} = (2j-1)/2q$, $j = 1, 2, \dots, q$ and $\gamma_{li} = (2l-1)/2q$, $l = 1, 2, \dots, q$, $q = 50$.

For X2 design matrix	$n = 20$	50	$\gamma_4 = 1.0$	High Size	High Size	High Size	Low Size	Low Size
		40	0.8	High Size	High Size	High Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	Low Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50
	$n = 60$	50	$\gamma_4 = 1.0$	Low Size	Low Size	Low Size	Low Size	Low Size
		40	0.8	Low Size	Low Size	Low Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	High Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50
For X3 design matrix	$n = 20$	50	$\gamma_4 = 1.0$	High Size	High Size	High Size	Low Size	Low Size
		40	0.8	High Size	High Size	High Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	Low Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50
	$n = 60$	50	$\gamma_4 = 1.0$	Low Size	Low Size	Low Size	Low Size	Low Size
		40	0.8	Low Size	Low Size	Low Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	High Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50

Table 7.3 (contn'd)

For X5 design matrix	$n = 20$	50	$\gamma_4 = 1.0$	Low Size	Low Size	Low Size	Low Size	Low Size
		40	0.8	High Size	High Size	Low Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	Low Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50
	$n = 60$	50	$\gamma_4 = 1.0$	Low Size	Low Size	Low Size	Low Size	Low Size
		40	0.8	Low Size	Low Size	Low Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	High Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50
For X6 design matrix	$n = 20$	50	$\gamma_4 = 1.0$	High Size	High Size	Low Size	Low Size	Low Size
		40	0.8	High Size	High Size	High Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	High Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50
	$n = 60$	50	$\gamma_4 = 1.0$	Low Size	Low Size	Low Size	Low Size	Low Size
		40	0.8	Low Size	Low Size	Low Size	Low Size	Low Size
		30	0.6	High Size	High Size	High Size	Low Size	Low Size
		20	0.4	High Size	High Size	High Size	Low Size	Low Size
		10	0.2	High Size	High Size	High Size	Low Size	Low Size
		$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
		$j=1$		10	20	30	40	50

Table 7.4

Critical values k_1 and k_2 when $m = 2$, for testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances at the five percent level.

Design matrix	Sample size (n)	Critical values	
		k_1	k_2
$X2(n \times 3)$	20	0.1395	2.5
	60	-0.000081	1.6
$X3(n \times 5)$	20	0.1557	2.45
	60	-0.000135	1.58
$X5(n \times 3)$	20	0.0779	2.5
	60	-0.000082	1.65
$X6(n \times 6)$	20	0.5714	1.91
	60	-0.000148	1.685

Table 7.5a

Selected calculated sizes and powers of the test for the $X2$ design matrix when $m = 2$, at the five percent level.

$n = 20 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1 = .9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.050	0.058	0.058	0.039	0.023*	0.1	0.023	0.054	0.146	0.334	0.559
0.3	0.044	0.053	0.054	0.041	0.029*	0.3	0.039	0.069	0.158	0.342	0.588
0.5	0.044	0.050	0.045	0.035*	0.031*	0.5	0.101	0.119	0.185	0.368	0.621
0.7	0.043	0.047	0.046	0.036*	0.035*	0.7	0.326	0.316	0.343	0.464	0.684
0.9	0.042	0.049	0.051	0.043	0.035*	0.9	0.809	0.790	0.780	0.772	0.848

$n = 60 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1 = .9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.028*	0.035*	0.045	0.039	0.010*	0.1	0.012	0.087	0.414	0.866	0.985
0.3	0.030*	0.039	0.056	0.059	0.042	0.3	0.094	0.172	0.484	0.882	0.989
0.5	0.032*	0.037	0.058	0.065*	0.052	0.5	0.455	0.479	0.672	0.925	0.996
0.7	0.055	0.033*	0.048	0.053	0.045	0.7	0.888	0.883	0.924	0.979	1.000
0.9	0.155*	0.033*	0.040	0.039	0.041	0.9	0.997	0.997	0.998	1.000	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.5b

Selected calculated sizes and powers of the test for the $X3$ design matrix when $m = 2$, at the five percent level.

$n = 20 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1 = .9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.051	0.059	0.054	0.032*	0.021*	0.1	0.021	0.049	0.114	0.263	0.471
0.3	0.049	0.051	0.046	0.038	0.026*	0.3	0.033	0.052	0.120	0.274	0.503
0.5	0.050	0.046	0.046	0.038	0.030*	0.5	0.086	0.092	0.139	0.292	0.535
0.7	0.044	0.045	0.042	0.038	0.037	0.7	0.293	0.275	0.283	0.391	0.591
0.9	0.047	0.049	0.050	0.044	0.039	0.9	0.781	0.759	0.736	0.711	0.771

$n = 60 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1 = .9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.032*	0.040	0.049	0.036*	0.010*	0.1	0.011	0.086	0.401	0.855	0.985
0.3	0.037	0.045	0.062	0.061	0.043	0.3	0.091	0.161	0.463	0.876	0.989
0.5	0.040	0.047	0.065*	0.071*	0.060	0.5	0.428	0.462	0.665	0.917	0.996
0.7	0.061	0.041	0.059	0.058	0.052	0.7	0.878	0.878	0.918	0.974	0.999
0.9	0.146*	0.040	0.049	0.050	0.045	0.9	0.997	0.997	0.998	1.000	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.5c

Selected calculated sizes and powers of the test for the $X5$ design matrix when $m = 2$, at the five percent level.

$n = 20 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1 = .9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.052	0.058	0.061	0.044	0.023*	0.1	0.028	0.063	0.139	0.264	0.367
0.3	0.047	0.050	0.054	0.045	0.028*	0.3	0.041	0.068	0.137	0.242	0.348
0.5	0.050	0.052	0.054	0.046	0.037	0.5	0.124	0.134	0.182	0.270	0.380
0.7	0.052	0.057	0.058	0.048	0.047	0.7	0.375	0.377	0.392	0.418	0.474
0.9	0.049	0.059	0.058	0.050	0.050	0.9	0.838	0.834	0.827	0.788	0.745

$n = 60 \quad H_0: u_t = (1 + \gamma_1 L)(1 + \gamma_4 L)e_t$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_t = e_t$					
γ_4	$\gamma_1 = .9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.023*	0.030*	0.046	0.039	0.010*	0.1	0.016	0.096	0.408	0.824	0.975
0.3	0.024*	0.034*	0.055	0.060	0.041	0.3	0.097	0.176	0.458	0.838	0.977
0.5	0.030*	0.038	0.059	0.062	0.056	0.5	0.451	0.467	0.633	0.889	0.984
0.7	0.059	0.031*	0.046	0.048	0.043	0.7	0.883	0.875	0.897	0.968	0.992
0.9	0.194*	0.030*	0.040	0.039	0.038	0.9	0.998	0.997	0.996	0.998	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.5d

Selected calculated sizes and powers of the test for the X_6 design matrix when $m = 2$, at the five percent level.

$n = 20 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.055	0.062	0.063	0.054	0.041	0.1	0.043	0.063	0.128	0.256	0.459
0.3	0.048	0.056	0.058	0.052	0.041	0.3	0.056	0.073	0.125	0.265	0.506
0.5	0.046	0.054	0.052	0.049	0.044	0.5	0.103	0.115	0.156	0.298	0.566
0.7	0.045	0.050	0.054	0.047	0.044	0.7	0.263	0.258	0.267	0.360	0.630
0.9	0.046	0.052	0.052	0.048	0.047	0.9	0.705	0.675	0.624	0.608	0.719

$n = 60 \quad H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.033*	0.038	0.048	0.032*	0.009*	0.1	0.012	0.079	0.356	0.814	0.982
0.3	0.035*	0.044	0.057	0.053	0.039	0.3	0.084	0.156	0.415	0.840	0.988
0.5	0.034*	0.043	0.061	0.072*	0.056	0.5	0.399	0.430	0.621	0.899	0.996
0.7	0.046	0.038	0.058	0.059	0.051	0.7	0.865	0.857	0.890	0.971	1.000
0.9	0.093*	0.037	0.049	0.050	0.048	0.9	0.997	0.996	0.998	0.999	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.6

Selected calculated sizes and powers of the test for the X_5 design matrix with $n = 60$ when special weights given for $m = 2$, at the five percent level.

$H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.033*	0.047	0.080*	0.081*	0.036*	0.1	0.046	0.182	0.544	0.877	0.983
0.3	0.036*	0.054	0.092*	0.111*	0.095*	0.3	0.196	0.296	0.569	0.886	0.985
0.5	0.034*	0.049	0.086*	0.104*	0.094*	0.5	0.575	0.577	0.730	0.927	0.989
0.7	0.031*	0.040	0.065*	0.067*	0.066*	0.7	0.922	0.910	0.935	0.977	0.996
0.9	0.044	0.035*	0.051	0.054	0.051	0.9	0.999	0.998	0.997	0.999	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Table 7.7

Ranges of the three disjoint intervals and the corresponding values of j and l , where $\gamma_{1j} = (2j-1)/2q$, $j = 1, 2, \dots, q$ and $\gamma_{4l} = (2l-1)/2q$, $l = 1, 2, \dots, q$, $q = 50$ for the X_5 design matrix with $n = 60$.

50	$\gamma_4 = 1.0$	Low Size	Low Size	Low Size	Low Size	High-Low Size
40	0.8	Low Size	Low Size	Low Size	Low Size	High-Low Size
30	0.6				Low Size	Low Size
20	0.4				Low Size	Low Size
10	0.2				Low Size	Low Size
$l=1$	0	$\gamma_1 = 0.2$	0.4	0.6	0.8	1.0
	$j=1$	10	20	30	40	50

Table 7.8

Selected calculated sizes and powers of the test for the X_5 design matrix with $n = 60$ when $m = 3$, at the five percent level.

$H_0: u_i = (1 + \gamma_1 L)(1 + \gamma_4 L)e_i$						$H_a: (1 - \rho_1 L)(1 - \rho_4 L)u_i = e_i$					
γ_4	$\gamma_1 = 9$	0.7	0.5	0.3	0.1	ρ_4	$\rho_1 = 1$	0.3	0.5	0.7	0.9
0.1	0.047	0.031*	0.047	0.040	0.010*	0.1	0.018	0.100	0.413	0.829	0.975
0.3	0.048	0.036*	0.056	0.062	0.044	0.3	0.101	0.182	0.463	0.841	0.977
0.5	0.046	0.039	0.062	0.064*	0.057	0.5	0.455	0.473	0.637	0.892	0.985
0.7	0.041	0.032*	0.047	0.049	0.044	0.7	0.885	0.877	0.901	0.969	0.993
0.9	0.074*	0.050	0.043	0.040	0.040	0.9	0.998	0.997	0.996	0.998	1.000

* denotes value significantly different from the five percent level at the 0.01 level.

Chapter 8

Summary and Conclusions

Econometric modelling can be a powerful tool for economic analysis and forecasting and a meaningful evaluation of an econometric model depends on the use of accurate statistical test procedures. Another issue related to statistical testing is the presence of nuisance parameters which can lead to misleading inferences in small samples. Proper handling of nuisance parameters may result in a better test. Regression coefficients and autocorrelated disturbances are common in econometrics and many techniques have been suggested to eliminate or deal with nuisance parameters, such as, sufficient statistics, invariant methods, marginal likelihoods, etc. A lack of observations also implies there is a need for tests which are as accurate as possible in small samples.

In this thesis we have proposed a new optimal test procedure which is based on the generalized NP lemma, called 'the average power test', for testing a composite null against a composite alternative. It maximizes the average power function of the test subject to the average size being controlled over different subregions of the null hypothesis parameter space. The new testing approach involves the finding of several critical values. These critical values are calculated by Monte Carlo simulation

experiment via trial and error so that the average size conditions are simultaneously satisfied.

This thesis has investigated the performance of the above mentioned new testing approach with respect to testing for MA(1) disturbances against AR(1) disturbances, testing the form of heteroscedastic disturbances and for testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances in the context of the linear regression model. The first aim was to investigate the performances of our new approach to see how well it works for a single parameter testing problem; the second aim was to see the performance of the test procedure for a heteroscedastic disturbances testing problem involving an infinite parameter space; and the third aim was to investigate the test's performance in a two-dimensional parameter testing problem. The following are a detailed discussion of the findings of the thesis.

A literature survey related to hypothesis testing was presented in Chapter 2. This chapter reviewed the literature involving optimal, PO and APO tests in the context of linear regression model for composite hypotheses testing problems. Because all the applications we considered in this thesis are non-nested, the review also focused on contributions to tests of non-nested hypotheses and some popular non-nested tests were discussed briefly and their limitations noted. Other topics briefly discussed include nuisance parameters, invariance arguments and marginal likelihood methods. This literature review reveals the importance of developing optimal tests which have excellent finite-sample properties, such as, PO tests, rather than large-sample based tests. When testing composite hypotheses, PO tests cannot always be constructed. In situations where they cannot be constructed, King (1987a) suggested the construction of an APO

test. Studies involving APO tests have indicated that it is not always suitable for all composite hypothesis testing problems and in some cases it does not work well at all. Almost all existing tests in the context of non-nested testing are large-sample based tests and many of them perform poorly in finite samples. Srikanthakumar (2000) constructed the g test which is an APOI test for testing a composite null against a composite alternative. The g test is constructed using the generalized NP lemma to obtain a PO test for a composite testing problem which required a finite number of representative densities under the null and a single representative density under the alternative. Thus this approach depends very much on the choice of representative densities. Our new approach does not use representative densities but rather maximizes average power subject to controlling average size over different subregions of the null hypothesis parameter space.

In Chapter 3, we introduced our new approach to testing a composite null against a composite alternative. We discussed the theory behind this new approach and how the testing procedure can be applied to the problem of testing the structure of the disturbances in the linear regression model. Controlling average size over the entire null parameter space, namely Θ , may result in undesirably large sizes in some parts of Θ and smaller sizes in other parts of Θ . To minimize this possibility, we allow for Θ to be partitioned into m disjoint subregions, namely, $\Theta_1, \Theta_2, \dots, \Theta_m$ so that $\Theta = \Theta_1 \cup \Theta_2 \cup \dots \cup \Theta_m$. As mentioned earlier, this new approach involves finding several critical values for which the average size conditions are fulfilled simultaneously. We discussed the standard invariance technique which was used to simplifying the testing problem by eliminating some nuisance parameters. We outlined the steps involved in calculating the critical values by Monte Carlo simulation.

In Chapter 4, we illustrated the new test procedure by applying it to the problem of testing for MA(1) disturbances against AR(1) disturbances in the linear regression model. For this testing problem, Silvapulle and King (1991) could not find a POI test and were forced to use an APOI test. This chapter explored a number of practical issues. An important question was whether controlling average size did a good job in controlling size overall. Typically we found that the sizes of our test were very acceptable for small sample sizes when $m=3$, that is, controlling average size over three subregions of the null hypothesis parameter space. We compared the performance of our new approach with three different versions of Silvapulle and King's (1991) APOI tests, namely, $s(0.3, \gamma_0^*)$, $s(0.5, \gamma_0^*)$ and $s(0.75, \gamma_0^*)$, and found that the powers of the new test are typically higher than those of the APOI tests. The new test has a very clear power advantage. When the sample size is small, on average the new test gained 0.7 to 5.7 percentage points in terms of power for all design matrices and when the sample size is large, the test gained 0.7 to 1.9 percentage points in terms of power improvement except in one case. We concluded that at least for this testing problem, the new test is better than the APOI test which Silvapulle and King (1991) found out performed the Lagrange multiplier test and an analogous test to Burke et al.'s (1990) asymptotic test. Using Silvapulle and King's (1991) APOI tests as benchmarks, the new test was found to have fairly good size and power properties. Thus the new procedure has considerable potential, particularly for problems in which the effective null hypothesis parameter space can be kept small.

In Chapter 5, we investigated the choice of boundary points of the subregions under the null hypothesis parameter space. In Chapter 4, we used the high size/low size technique

to determine the boundary points of the subregions under the null hypothesis parameter space. This chapter constructed a pseudo power envelope to trace out the maximum attainable power for a given problem in order to provide a benchmark against which the new test procedure was evaluated. This analysis indicated that without spending a lot of time on computation to construct the pseudo power envelope, we can get reasonably close to the best test by using the high size/low size technique. That is, when we identified the best test by reference to the pseudo power envelope, we gained only 1.8 percentage points in terms of power improvement over the test constructed using the high size/low size method. The large amount of extra computation required for finding the best test via a pseudo power envelope does not appear to be justified. Thus, we conclude that for our proposed new test procedure we can use the high size/low size technique to choose the boundary points of the subregions of Θ with some confidence.

Chapter 6 investigated the performance of the new test procedure by applying it to testing for heteroscedastic disturbances in the linear regression model. This is a testing problem where both the null and alternative hypothesis parameter spaces are one sided infinite intervals in contrast to the unit intervals for the problem of testing MA(1) disturbances against AR(1) disturbances. Two different sets of non-stochastic variables were used to investigate the performances of the test procedure. Our results indicate that for both sets of non-stochastic variables when $m = 2$, the test performs well with respect to size conditions and the power properties of the test are very acceptable. Also, the results for z_{it} and w_{it} , when test sizes for $m = 1$ are controlled at $\lambda = 0$ indicated that for small sample sizes, the size and power performance of the test are fairly good but there is a clear loss of power for the larger sample size. This indicates that extra computation for $m = 2$ could be worthwhile. From this viewpoint, we conclude that the

new test approach can work well for testing problems where both the null and alternative hypothesis parameter spaces are one-sided infinite intervals.

In Chapter 7, we explored the problem of testing joint MA(1)-MA(4) against joint AR(1)-AR(4) disturbances in the linear regression model. As mentioned earlier, the new test procedure works well for the testing problem of MA(1) disturbances against AR(1) disturbances; and also for testing heteroscedastic disturbances in the linear regression model. This chapter investigated the performance of the test when it extends to testing two-dimensional parameter spaces under the null and alternative. The results indicate that the size and power performances of the test are very good for $m = 2$, when the sample size is small. For the larger sample size, the estimated average sizes of the test do not perform well for $m = 2$, that is, a few calculated sizes fall outside the rejection probability ranges.

Also, further results for $m = 2$ with heavier weights applied for troublesome parts of the null parameter space indicated that the size properties of the test for the larger sample size do not perform well. With the heavier weights applied to troublesome parts, the sizes are controlled for those parts of the parameter space but can be uncontrolled in other parts of the null parameter space. There was no improvement in the size performance of the test for $m = 2$ with the heavier weights, perhaps because we did not follow a system to determine the weights for different parts of the null parameter space. From the results of the test for $m = 3$ with $n = 60$, we conclude that the size properties of the test are close to acceptable and power properties are well behaved. The results reported for $m = 3$ with $n = 60$, suggest that the extra computation of the test procedure is well rewarded at least in terms of size properties. As a result, we conclude that our

new test procedure can be made to work for testing problems involving two-dimensional parameter space models but it is time consuming to apply in practice, and may require future advances in computer technology to make it a feasible test in this case.

The overall recommendation of this thesis is that the new test procedure has considerable potential, particularly for problems in which the effective null hypothesis parameter space can be kept as small as possible. A small sample power comparison reveals that the new test can be more powerful than various versions of the APOI test recommended by Silvapulle and King (1991). In situations where PO and APOI cannot be constructed for composite null hypothesis testing problem, the new optimal test procedure can be considered for general testing solution. In this new test approach, certainly we can use the high size/low size method to choose the boundary points of the subregions under the null parameter space. As the new test approach works well for a testing problem when both null and alternative hypothesis parameter spaces are both one-sided infinite intervals, it is recommended for testing problems when both the null and alternative hypothesis parameter spaces are one-sided infinite intervals. This is possible because of our transformation of the parameter space used in Chapter 6. The new test procedure can be applied to two-dimensional parameter space models, the only limitation is that it is time consuming but in the future, advances in computer technology could make the test realistic for multi-dimensional parameter space models.

Finally, there are a few areas in which further research could be undertaken. In this thesis, we applied the new optimal test procedure to three different composite testing situations, with encouraging results. Clearly it is of interest to investigate its behaviour for other testing problems. In our investigations, we only used the standard uniform

distribution as a weighting function. It would be of interest to investigate the procedure's behaviour for other weighting functions.

In this thesis, the new approach was applied to one parameter testing problems when both null and alternative hypothesis parameter spaces are the unit interval $[0, 1]$ and the one-sided infinite interval $(0, \infty)$. An obvious question is, can this approach be applied to testing problems involving hypotheses defined over two-sided infinite intervals, i.e., $(-\infty, \infty)$? For this, a logistic function might be a useful way to reduce the infinite interval to a finite interval $[-1, 1]$. How well the new approach performs in this case is a worthy research question. An important issue of the new approach was that of computation. We took a simple approach to the problem of calculating the value of various integrals, i.e., a numerical approximation method that works for one-dimensional parameter testing problem. An obvious question is, are there better ways of calculating these integrals that can reduce computation, particularly for multi-dimensional problems.

The new approach could also be used for testing a composite null against a simple alternative hypothesis. The new test would maximize power under the simple alternative subject to control of average size under the null. The robustness of the testing procedure to non-normal errors is another area that could also be explored.

In summary, this thesis has investigated a new approach to testing composite hypotheses. The new approach shows considerable promise when testing one-dimensional parameter spaces, but future work is needed to understand its full potential.

References

- Ai, C. and Cassou, S.P. (1993), A diagnostic test without numerical integration, *Economics Letters*, 42, 129-132.
- Andrews, D.W.K. (1988), Chi-square diagnostic tests for econometric models: Theory, *Econometrica*, 56, 1419-1453.
- Andrews, D.W.K., Lee, I. and Ploberger, W. (1996), Optimal changepoint tests for normal linear regression, *Journal of Econometrics*, 70, 9-38.
- Andrews, D.W.K. and Ploberger, W. (1994), Optimal tests when a nuisance parameter is present only under the alternative, *Econometrica*, 62, 1383-1414.
- Andrews, D.W.K. and Ploberger, W. (1995), Admissibility of the likelihood ratio test when a nuisance parameter is present only under the alternative, *The Annals of Statistics*, 23, 1609-1629.
- Andrews, D.W.K. and Ploberger, W. (1996), Testing for serial correlation against an ARMA(1, 1) process, *Journal of the American Statistical Association*, 91, 1331-1342.
- Anselin, L., Bera, A.K., Florax, R. and Yoon, M.J. (1996), Simple diagnostic tests for spatial dependence, *Regional Science and Urban Economics*, 26, 77-104.
- Ara, I. (1995), Marginal Likelihood Based Tests of Regression Disturbances, unpublished doctoral dissertation, Monash University.
- Ara, I. and King, M.L. (1993), Marginal likelihood based tests of regression disturbances, *Mimeo*, Monash University.
- Ara, I. and King, M.L. (1995), Marginal likelihood based tests of a subvector of the parameter vector of linear regression model, in: C.S. Frobos, P. Kofman and T.R.L. Fry, eds., *Proceedings of the 1995 Econometrics Conference at Monash University*, Monash University, Melbourne, 69-106.
- Arthanari, T.S. and Dodge, Y. (1980), *Mathematical Programming in Statistics*, John Wiley and Sons; New York.
- Atkinson, A.C. (1970), A method for discriminating between models, *Journal of the Royal Statistical Society*, B, 32, 323-353.
- Baltagi, B.H. and Li, Q. (1995), Testing AR(1) against MA(1) disturbances in an error component model, *Journal of Econometrics*, 68, 133-151.
- Barndorff-Nielsen, O.E. (1983), On a formula for the distribution of the maximum likelihood estimator, *Biometrika*, 70, 343-365.
- Beggs, J.J. (1988), Diagnostic testing in applied econometrics, *Economic Record*, 64, 81-101.
- Bera, A.K. and Premaratne, G. (2001), General hypothesis testing, in: B.H. Baltagi, eds., *A Companion to Theoretical Econometrics*, Blackwell, Oxford, 38-61.
- Berenblut, I.I. and Webb, G.I. (1973), A new test for autocorrelated errors in the linear regression model, *Journal of the Royal Statistical Society*, B, 35, 33-50.
- Bewley, R. (1986), *Allocation Models: Specification, Estimation and Applications*, Ballinger, Boston, MA.

- Bhargava, A. (1986), On the theory of testing unit roots in observed time series, *Review of Economic Studies*, 53, 369-384.
- Bhatti, M.I. (1992a), Testing Regression Models Based on Sample Survey Data, unpublished doctoral dissertation, Monash University.
- Bhatti, M.I. (1992b), Optimal testing for serial correlation in a large number of small samples, *Biometrical Journal*, 34, 57-67.
- Bhatti, M.I. and King, M.L. (1990), A beta-optimal test of equicorrelation coefficient, *Australian Journal of Statistics*, 32, 87-97.
- Bhatti, M.I. and King, M.L. (1993), Testing for subblock effects in multistage linear regression models, *Mimeo*, Monash University.
- Bhatti, M.I. and King, M.L. (1994), Testing for subblock effects in multistage linear regression models, Paper presented at the European Meeting of the Econometric Society, Maastricht, Netherlands.
- Bierens, H.J. (1988), Towards a theory of point optimal testing: Comment, *Econometric Reviews*, 6, 231-233.
- Bjerkholt, O. (1995), *Foundations of Modern Econometrics: The Selected Essays of Ragnar Frisch*, vol. I, Edward Elgar, Aldershot.
- Bjerkholt, O. (1998), Ragnar Frisch and the Foundation of the Econometric Society and *Econometrica*, in: S. Strom, ed., *Econometrics and Economic Theory in the 20th Century*, Cambridge University Press, 26-57.
- Box, G.E.P. and Jenkins, G.M. (1976), *Time Series Analysis: Forecasting and Control*, 2nd edition, Holden-Day; San Francisco.

- Brooks, R.D. (1993), Alternative point-optimal tests for regression coefficient stability, *Journal of Econometrics*, 57, 365-376.
- Brooks, R.D. (1995), The robustness of point optimal testing for Rosenberg random regression coefficients, *Econometric Reviews*, 14, 35-53.
- Brooks, R.D. and King, M.L. (1994a), Testing Hildreth-Houck against return to normalcy random regression coefficients, *Journal of Quantitative Economics*, 10, 33-52.
- Brooks, R.D. and King, M.L. (1994b), Hypothesis testing of varying coefficient regression models: Procedures and applications, *Pakistan Journal of Statistics*, 10, 301-358.
- Burke, S.P., Godfrey, L.G. and Tremayr, A.R. (1990), Testing AR(1) against MA(1) disturbances in the linear regression model: An alternative procedure, *Review of Economic Studies*, 57, 135-145.
- Burman, J.P. (1980), Seasonal adjustment by signal extraction, *Journal of the Royal Statistical Society, A*, 143, 321-337.
- Chen, D. (2001), Median-Unbiased Estimation in Linear Autoregressive Time Series Models, unpublished doctoral dissertation, Monash University.
- Chen, Y.T. and Kuan, C.M. (2002), The pseudo-true score encompassing test for non-nested hypotheses, *Journal of Econometrics*, 106, 271-295.
- Chesher, A. and Austin, G. (1991), The finite sample distributions of heteroskedasticity robust Wald statistics, *Journal of Econometrics*, 47, 153-173.
- Chow, G.C. (1957), *Demand for Automobiles in the United States*, North-Holland; Amsterdam.

- Comiffe, D. (1987), Expected maximum log likelihood estimation, *The Statistician*, 36, 317-329.
- Conte, S.D. (1965), *Elementary Numerical Analysis: An Algorithmic Approach*, McGraw-Hill Book Company; New York, Toronto.
- Corduas, M. (1986), The Use of the Marginal Likelihood in Testing for Serial Correlation in Time Series Regression, unpublished M.Phil dissertation, University of Lancaster.
- Cox, D.R. (1961), Tests of separate families of hypotheses, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley: University of California Press, 1, 105-123.
- Cox, D.R. (1962), Further results on tests of separate families of hypotheses, *Journal of the Royal Statistical Society, B*, 24, 406-424.
- Cox, D.R. and Hinkley, D.V. (1974), *Theoretical Statistics*, Chapman and Hall; London.
- Cox, D.R. and Reid, N. (1987), Parameter orthogonality and approximate conditional inference (with discussion), *Journal of the Royal Statistical Society, B*, 49, 1-39.
- Cox, D.R. and Reid, N. (1993), A note on the calculation of adjusted profile likelihood, *Journal of the Royal Statistical Society, B*, 55, 467-471.
- Dantzig, G.B. and Wald, A. (1951), On the fundamental lemma of Neyman and Pearson, *Annals of Mathematical Statistics*, 22, 87-93.
- Darnell, A.C. (1984), Economic statistics and econometrics, in: J. Creedy and D.P. O'Brien, eds., *Economic Analysis in Historical Perspective*, Butterworth, London, 152-185.

- Darnell, A.C. (1994), *The History of Econometrics*, vol. II, Edward Elgar, Aldershot.
- Dastoor, N.K. (1981), A note on the interpretation of the Cox procedure for non-nested hypotheses, *Economics Letters*, 8, 113-119.
- Dastoor, N.K. and Fisher, G. (1988a), On point-optimal Cox tests, *Econometric Theory*, 4, 97-107.
- Dastoor, N.K. and Fisher, G. (1988b), The theory and practice of point optimal testing, *Econometric Reviews*, 6, 219-229.
- Davidson, R. and MacKinnon, J.G. (1981), Several tests for model specification in the presence of alternative hypotheses, *Econometrica*, 49, 781-793.
- Davidson, R. and MacKinnon, J.G. (1982), Some non-nested hypothesis tests and the relations among them, *Review of Economic Studies*, 49, 551-565.
- Davidson, R. and MacKinnon, J.G. (1996), The size distortion of bootstrap tests, Discussion paper 936, Queen's University.
- Davidson, R. and MacKinnon, J.G. (2002), Bootstrap *J* tests of nonnested linear regression models, *Journal of Econometrics*, 109, 167-193.
- Davies, R.B. (1969), Beta-optimal tests and an application to the summary evaluation of experiments, *Journal of the Royal Statistical Society, B*, 31, 524-538.
- Davies, R.B. (1977), Hypothesis testing when a nuisance parameter is present only under the alternative, *Biometrika*, 64, 247-254.
- Davies, R.B. (1980), Algorithm AS155: The distribution of a linear combination of χ^2 random variables, *Applied Statistics*, 29, 323-333.

- Deaton, A.S. (1982), Model selection procedures, or, does the consumption function exist? in: G.C. Chow and P. Corsi, eds., *Evaluating the Reliability of Macroeconomic Models*, John Wiley and Sons, New York, 43-65.
- Delgado, M.A. and Stengos, T. (1994), Semiparametric specification testing of non-nested econometric models, *Review of Economic Studies*, 61, 291-303.
- Dufour, J.-M. and King, M.L. (1991), Optimal invariant tests for the autocorrelation coefficient in linear regressions with stationary or nonstationary AR(1) errors, *Journal of Econometrics*, 47, 115-143.
- Durbin, J. and Watson, G.S. (1950), Testing for serial correlation in least squares regression I, *Biometrika*, 37, 409-428.
- Durbin, J. and Watson, G.S. (1951), Testing for serial correlation in least squares regression II, *Biometrika*, 38, 159-178.
- Durbin, J. and Watson, G.S. (1971), Testing for serial correlation in least squares regression III, *Biometrika*, 58, 1-19.
- Efron, B. (1975), Defining the curvature of a statistical problem (with application to second order efficiency), *Annals of Statistics*, 3, 1189-1242.
- Evans, G.A. and Deaton, A. (1980), Testing linear versus logarithmic regression models, *Review of Economic Studies*, 47, 275-291.
- Evans, M.A. and King, M.L. (1985a), Critical value approximations for tests of linear regression disturbances, *Australian Journal of Statistics*, 27, 68-83.
- Evans, M.A. and King, M.L. (1985b), A point optimal test for heteroscedastic disturbances, *Journal of Econometrics*, 27, 163-178.

- Evans, M.A. and King, M.L. (1988), A further class of tests for heteroscedasticity, *Journal of Econometrics*, 37, 265-276.
- Fan, Y. and Li, Q. (1995), Bootstrapping J-type tests for non-nested regression models, *Economics Letters*, 48, 107-112.
- Ferguson, T.S. (1967), *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York and London.
- Fisher, G. and McAleer, M. (1979), On the interpretation of the Cox test in econometrics, *Economics Letters*, 4, 145-150.
- Fisher, G.R. and McAleer, M. (1981), Alternative procedures and associated tests of significance for non-nested hypotheses, *Journal of Econometrics*, 16, 103-119.
- Franses, P.H. (1992), A model selection test for an AR(i) versus an MA(1) model, *Statistics and Probability Letters*, 15, 281-284.
- Franzini, L. and Harvey, A.C. (1983), Testing for deterministic trend and seasonal components in time series models, *Biometrika*, 70, 673-682.
- Fraser, D.A.S. (1967), Data transformations and the linear model, *Annals of Mathematical Statistics*, 38, 1456-1465.
- Frisch, R. (1926), Sur un problème d'économie pure, *Norsk Matematisk Forenings Skrifter*, 1, 1-40.
- GAUSS (1998), GAUSS for Windows NT/95 Version 3.2.35, Aptech Systems, Inc., Maple Valley, WA.
- Godfrey, L.G. (1998), Tests of non-nested regression models: Some results on small sample behaviour and the bootstrap, *Journal of Econometrics*, 84, 59-74.

- Godfrey, L.G. and Pesaran, M.H. (1983), Tests of non-nested regression models: Small sample adjustments and Monte Carlo evidence, *Journal of Econometrics*, 21, 133-154.
- Godfrey, L.G. and Tremayne, A.R. (1988), Checks of model adequacy for univariate time series models and their application to econometric relationships, *Econometric Reviews*, 7, 1-42.
- Godfrey, L.G. and Tremayne, A.R. (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.
- Gourieroux, C. and Monfort, A. (1995), Testing encompassing and simulating dynamic econometric models, *Econometric Theory*, 11, 195-228.
- Gourieroux, C., Monfort, A. and Trognon, A. (1983), Testing nested or non-nested hypotheses, *Journal of Econometrics*, 21, 83-115.
- Granger, C.W.J. and Newbold, P. (1986), *Forecasting Economic Time Series*, 2nd edition, Academic Press; New York.
- Grosc, S. (1998), Marginal Likelihood Methods in Econometrics, unpublished doctoral dissertation, Monash University.
- Hall, A.D. and McAleer, M. (1989), A Monte Carlo study of some tests of model adequacy in time series analysis, *Journal of Business and Economic Statistics*, 7, 95-106.
- Harvey, A.C. (1981), *Time Series Models*, Philip Allan; Oxford.

- Hausman, J.J. (1978), Specifications tests in econometrics, *Econometrica*, 46, 1215-1272.
- Hendry, D.F. (2001), Achievements and challenges in econometric methodology, *Journal of Econometrics*, 100, 7-10.
- Hendry, D.F. and Richard, J.F. (1982), On the formulation of empirical models in dynamic econometrics, *Journal of Econometrics*, 20, 3-34.
- Hildreth, C. and Houck, J.P. (1968), Some estimators for a linear model with random coefficients, *Journal of the American Statistical Association*, 63, 584-595.
- Hillier, G.H. (1987), Classes of similar regions and their power properties for some econometric testing problems, *Econometric Theory*, 3, 1-44.
- Hillier, G.H. (1991), On multiple diagnostic procedures for the linear model, *Journal of Econometrics*, 47, 47-66.
- Hillier, G.H. and King, M.L. (1991), Editors' introduction, *Journal of Econometrics*, 47, 1-4.
- Hinde, J. and Aitkin, M. (1987), Canonical likelihoods: A new likelihood treatment of nuisance parameters, *Biometrika*, 74, 45-58.
- Honda, Y. (1988), A size correction to the Lagrange multiplier test for heteroscedasticity, *Journal of Econometrics*, 38, 375-386.
- Honda, Y. (1989), On the optimality of some tests of the error covariance matrix in the linear regression model, *Journal of the Royal Statistical Society, B*, 51, 71-79.
- Horowitz, J.L. (1994), Bootstrap-based critical values for the information matrix test, *Journal of Econometrics*, 61, 395-411.

- Horowitz, J.L. (2001), The bootstrap and hypothesis tests in econometrics, *Journal of Econometrics*, 100, 37-40.
- Horowitz, J.L. and Savin, N.E. (2000), Empirically relevant critical values for hypothesis tests: A bootstrap approach, *Journal of Econometrics*, 95, 375-389.
- Hwang, J. and Schmidt, P. (1993), On the power of point optimal tests of the trend stationarity hypothesis, *Economics Letters*, 43, 143-147.
- Imhof, P.J. (1961), Computing the distribution of quadratic forms in normal variables, *Biometrika*, 48, 419-426.
- Isaacson, S.L. (1951), On the theory of unbiased tests of simple statistical hypotheses specifying the values of two or more parameters, *Annals of Mathematical Statistics*, 22, 217-234.
- Jaeyoun, H. and Schmidt, P. (1996), Alternative methods of detrending and the power of unit root tests, *Journal of Econometrics*, 71, 227-248.
- Kadiyala, K.R. (1970), Testing for the independence of regression disturbances, *Econometrica*, 38, 97-117.
- Kalbfleisch, J.D. and Sprott, D.A. (1970), Application of likelihood methods to models involving large numbers of parameters (with discussion), *Journal of the Royal Statistical Society, B*, 32, 175-208.
- Kang, H. and Inder, B. (1996), Diagnostic test for structural change in cointegrated regression models, *Economics Letters*, 50, 179-187.
- Kariya, T. and Eaton, M.L. (1977), Robust tests for spherical symmetry, *Annals of Statistics*, 9, 210-213.

- King, M.L. (1979), Some Aspects of Statistical Inference in the Linear Regression Model, unpublished doctoral dissertation, University of Canterbury.
- King, M.L. (1980), Robust tests for spherical symmetry and their application to least squares regression, *Annals of Statistics*, 8, 1265-1271.
- King, M.L. (1981a), The alternative Durbin-Watson test: An assessment of Durbin and Watson's choice of test statistic, *Journal of Econometrics*, 17, 51-66.
- King, M.L. (1981b), A small sample property of the Cliff-Ord test for spatial correlation, *Journal of the Royal Statistical Society, B*, 43, 263-264.
- King, M.L. (1982a), Testing for a serially correlated component in regression disturbances, *International Economic Review*, 23, 577-582.
- King, M.L. (1982b), A locally optimal bounds test for autoregressive disturbances, Paper presented to the European Meeting of the Econometric Society, Dublin.
- 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 regression disturbances, *Australian Journal of Statistics*, 25, 23-34.
- King, M.L. (1984), A new test for fourth-order autoregressive disturbances, *Journal of Econometrics*, 24, 269-277.
- King, M.L. (1985a), A point optimal test for autoregressive disturbances, *Journal of Econometrics*, 27, 21-37.
- King, M.L. (1985b), A point optimal test for moving average regression disturbances, *Econometric Theory*, 1, 211-222.

- King, M.L. (1986), Efficient estimation and testing of regressions with a serially correlated error component, *Journal of Quantitative Economics*, 2, 231-247.
- King, M.L. (1987a), Towards a theory of point optimal testing, *Econometric Reviews*, 6, 169-218.
- 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. (1989), Testing for fourth-order autocorrelation in regression disturbances when first-order autocorrelation is present, *Journal of Econometrics*, 41, 285-301.
- 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 Evans, M.A. (1984), A joint test for serial correlation and heteroscedasticity, *Economics Letters*, 16, 297-302.
- King, M.L. and Evans, M.A. (1986), Testing for block effects in regression models based on survey data, *Journal of the American Statistical Association*, 81, 677-679.
- King, M.L. and Hillier, G.H. (1985), Locally best invariant tests of the error covariance matrix of the linear regression model, *Journal of the Royal Statistical Society, B*, 47, 98-102.
- King, M.L. and Inder, B. (1986), Testing the covariance matrix of the linear regression model: Some further results, *Mimeo*, Department of Econometrics, Monash University.

- King, M.L. and McAleer, M. (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 Smith, M.D. (1986), Joint one-sided tests of linear regression coefficients, *Journal of Econometrics*, 32, 367-383.
- King, M.L. and Wu, P.X. (1997), Locally optimal one-sided tests for multiparameter hypotheses, *Econometrics Reviews*, 16, 131-156.
- Klein, L.R. (1947), The use of econometric models as a guide to economic policy, *Econometrica*, 15, 111-151.
- Koerts, J. and Abrahamse, A.P.J. (1969), *On the Theory and Application of the General Linear Model*, Rotterdam: Rotterdam University Press.
- Laskar, M.R. (1998), Estimation and Testing of Linear Regression Disturbances Based on Modified Likelihood and Message Length Functions, unpublished doctoral dissertation, Monash University.
- Laskar, M.R. and King, M.L. (1995), Parameter orthogonality and likelihood functions, in: C.S. Frobos, P. Kofman and T.R.L. Fry, eds., *Proceedings of the 1995 Econometrics Conference at Monash University*, Monash University, Melbourne, 253-289.
- Laskar, M.R. and King, M.L. (1998), Estimation and testing of regression disturbances based on modified likelihood functions, *Journal of Statistical Planning and Inference*, 71, 75-92.
- Lehmann, E.L. (1947), On families of admissible tests, *Annals of Mathematical Statistics*, 18, 97-104.

- Lehmann, E.L. (1959), *Testing Statistical Hypotheses*, John Wiley and Sons, New York.
- Lehmann, E.L. (1986), *Testing Statistical Hypotheses*, 2nd edition, John Wiley and Sons, New York.
- Lehmann, E.L. and Stein, C. (1948), Most powerful tests of composite hypotheses, I. Normal distributions, *Annals of Mathematical Statistics*, 19, 495-516.
- Levenbach, H. (1973), The estimation of heteroscedasticity from a marginal likelihood function, *Journal of the American Statistical Association*, 68, 436-439.
- Ljung, G.M. (1986), Diagnostic testing of univariate time series models, *Biometrika*, 73, 725-730.
- Majumder, A.K. (1999), One-Sided and Partially One-Sided Hypothesis Testing in Econometrics, unpublished doctoral dissertation, Monash University.
- Mao, C.X. and Lindsay, B.G. (2002), Tests and diagnostics for heterogeneity in the species problem, *Computational Statistics and Data Analysis* (forthcoming).
- McAleer, M. (1987), Specification tests for separate models: A survey, in: M.L. King and D.E.A. Giles, eds., *Specification Analysis in the Linear Model*, Routledge and Kegan Paul, London, 146-195.
- McAleer, M. (1995), The significance of testing empirical non-nested models, *Journal of Econometrics*, 67, 149-171.
- McAleer, M., McKenzie, C.R. and Hall, A.D. (1988), Testing separate time series models, *Journal of Time Series Analysis*, 9, 169-189.
- McAleer, M. and Pesaran, M.H. (1986), Statistical inference in non-nested econometric models, *Applied Mathematics and Computation*, 20, 271-311.

- McAleer, M., Pesaran, M.H. and Bera, A.K. (1990), Alternative approaches to testing non-nested models with autocorrelated disturbances: An application to models of US unemployment, *Communication in Statistics: Theory and Methods*, 19, 3619-3644.
- McCullagh, P. and Tibshirani, R. (1990), A simple method for the adjustment of profile likelihoods, *Journal of the Royal Statistical Society, B*, 52, 325-344.
- McKenzie, C.R., McAleer, M. and Gill, L. (1999), Simple procedures for testing autoregressive versus moving average errors in regression models, *Japanese Economic Review*, 50, 239-252.
- Michelis, L. (1999), The distribution of the J and Cox non-nested tests in regression models with weakly correlated regressors, *Journal of Econometrics*, 93, 369-401.
- Mizon, G.E. (1984), The encompassing approach in econometrics, in: D. Hendry and K.F. Wallis, eds., *Econometrics and Quantitative Economics*, Basil Blackwell, Oxford, 135-172.
- Mizon, G.E. and Richard, J.F. (1986), The encompassing principle and its application to testing non-nested hypotheses, *Econometrica*, 54, 657-678.
- Moulton, B.R. and Randolph, W.C. (1989), Alternatives tests of the error components model, *Econometrica*, 57, 685-693.
- Newey, W. (1985), Maximum likelihood specification testing and conditional moment tests, *Econometrica*, 53, 1047-1070.
- Neyman, J. (1935), Sur la vérification des hypothèses statistiques composées, *Bulletin de la société Mathématique de France*, 63, 246-266.

- Neyman, J. and Pearson, E.S. (1933), On the problem of the most efficient tests of statistical hypotheses, *Philosophical Transactions of the Royal Society, A*, 231, 289-337.
- Neyman, J. and Pearson, E.S. (1936), Contribution to the theory of testing statistical hypotheses I: Unbiased critical regions of Type A and Type A_1 , *Statistical Research Memoirs*, 1, 1-37.
- Neyman, J. and Pearson, E.S. (1938), Contribution to the theory of testing statistical hypotheses II: Certain theorems on unbiased critical regions, *Statistical Research Memoirs*, 2, 25-57.
- Nicholls, D.F., Pagan, A.R. and Terrell, R.D. (1975), The estimation and use of models with moving average disturbance terms: A survey, *International Economic Review*, 16, 113-134.
- Pagan, A. and Vella, F. (1989), Diagnostic tests for models based on individual data: A survey, *Journal of Applied Econometrics*, 4, s29-s59.
- Pearson, K. (1900), On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Philosophical Magazine*, 5, 50, 157-175.
- Pesaran, M.H. (1974), On the general problem of model selection, *Review of Economic Studies*, 41, 153-171.
- Pesaran, M.H. and Deaton, A.S. (1978), Testing non-nested nonlinear regression models, *Econometrica*, 46, 677-694.

- Pesaran, M.H. and Pesaran, B. (1993), A simulation approach to the problem of computing Cox's statistic for testing non-nested models, *Journal of Econometrics*, 57, 377-392.
- Pesaran, M.H. and Weeks, M. (2001), Non-nested hypothesis testing: An overview, in: B.H. Baltagi, eds., *A Companion to Theoretical Econometrics*, Blackwell, Oxford, 279-309.
- Poskitt, D.S. and Tremayne, A.R. (1981), An approach to testing linear time series models, *Annals of Statistics*, 9, 974-986.
- Poskitt, D.S. and Tremayne, A.R. (1982), Diagnostic tests for multiple time series models, *Annals of Statistics*, 10, 114-120.
- Pötscher, B.M. (1988), Towards a theory of point optimal testing: Comment, *Econometric Reviews*, 6, 235-239.
- Prothero, D.L. and Wallis, K.F. (1976), Modelling macroeconomic time series, *Journal of the Royal Statistical Society, A*, 139, 468-500.
- Quandt, R.E. (1974), A comparison of methods for testing nonnested hypotheses, *Review of Economics and Statistics*, 56, 92-99.
- Rahman, S. and King, M.L. (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 King, M.L. (1998), Marginal likelihood score-based tests of regression disturbances in the presence of nuisance parameters, *Journal of Econometrics*, 82, 81-106.

- Ramalho, J.J.S. and Smith, R.J. (2002), Generalized empirical likelihood non-nested tests, *Journal of Econometrics*, 107, 99-125.
- Rothenberg, T.J. and Stock, J.H. (1997), Inference in a nearly integrated autoregressive model with nonnormal innovations, *Journal of Econometrics*, 80, 269-286.
- Rowley, J.C.D. and Wilton, D.A. (1973), Quarterly models of wage determination: Some new efficient estimates, *The American Economic Review*, 63, 380-389.
- Saikkonen, P. and Luukkonen, R. (1993), Point optimal tests for testing the order of differencing in ARIMA models, *Econometric Theory*, 9, 343-362.
- Sargan, J.D. and Bhargava, A. (1983), Testing residuals from least squares regression for being generated by the Gaussian random walk, *Econometrica*, 51, 153-174.
- Sen Gupta, A. (1987), On tests for equicorrelation coefficient of a standard symmetric multivariate normal distribution, *Australian Journal of Statistics*, 29, 49-59.
- Shively, T.S. (1986), Issues in the Analysis of Stochastic Coefficient Regression Models, unpublished doctoral dissertation, University of Chicago.
- Shively, T.S. (1988a), An analysis of tests for regression coefficient stability, *Journal of Econometrics*, 39, 367-386.
- Shively, T.S. (1988b), An exact test for a stochastic coefficient in a time series regression model, *Journal of Time Series Analysis*, 9, 81-88.
- Shively, T.S. (1993), Testing for autoregressive disturbances in a time series regression with missing observations, *Journal of Econometrics*, 57, 233-255.
- Silvapulle, P. (1991), Non-Nested Testing of Regression Disturbances, unpublished doctoral dissertation, Monash University.

- Silvapulle, P. (1994a), Testing for AR(1) against IMA(1,1) disturbances in the linear regression model, *Communication in Statistics: Theory and Methods*, 23, 701-720.
- Silvapulle, P. (1994b), Testing AR(1) against MA(1) disturbances in the dynamic linear regression model, *Pakistan Journal of Statistics*, 10, 395-422.
- Silvapulle, P. and King, M.L. (1991), Testing moving average against autoregressive disturbances in the linear regression model, *Journal of Business and Economic Statistics*, 9, 329-335.
- Silvapulle, P. and King, M.L. (1993), Non-nested testing for autocorrelation in the linear regression model, *Journal of Econometrics*, 58, 295-314.
- Smith, M.D. and Tremayne, A.R. (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.
- Spiegelhalter, D.J. (1983), Diagnostic tests of distributional shape, *Biometrika*, 70, 401-409.
- Sriananthakumar, S. (2000), Contributions to the Theory and Practice of Hypothesis Testing, unpublished doctoral dissertation, Monash University.
- Stone, R. (1954), *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom 1920-1938*, I, Cambridge University Press.
- Strom, S. (1998), *Econometrics and Economic Theory in the 20th Century*, Cambridge University Press.
- Tauchen, G. (1985), Diagnostic testing and evaluation of maximum likelihood models, *Journal of Econometrics*, 30, 415-443.

- Thomas, J.J. and Wallis, K.F. (1971), Seasonal variation in regression analysis, *Journal of the Royal Statistical Society, A*, 134, 57-72.
- Thompson, S.B. (2002), Optimal versus robust inference in nearly integrated non Gaussian models, Unpublished paper, Internet address: <http://www.ruf.rice.edu/~econ/seminars/bin/S.Thompson...>
- Verbyla, A.P. (1990), A conditional derivation of residual maximum likelihood, *Australian Journal of Statistics*, 32, 227-230.
- Vinod, H.D. (1973), Generalization of the Durbin-Watson statistic for higher order autoregressive processes, *Communications in Statistics*, 2, 115-144.
- 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.
- Walker, A.M. (1967), Some tests of separate families of hypotheses in time series analysis, *Biometrika*, 54, 39-68.
- Wallis, K.F. (1972), Testing for fourth-order autocorrelation in quarterly regression equations, *Econometrica*, 40, 617-636.
- Wallis, K.F. (1977), Multiple time series analysis and the final form of econometric models, *Econometrica*, 45, 1481-1498.
- Webb, G.I. (1973), Autocorrelations and the General Linear Model, unpublished doctoral dissertation, the City University.
- White, H. (1982), Maximum likelihood estimation of misspecified models, *Econometrica*, 50, 1-25.

- Wooldridge, J.M. (1990), An encompassing approach to conditional mean tests with applications to testing non-nested hypotheses, *Journal of Econometrics*, 45, 331-350.
- Wooldridge, J.M. (2001), Diagnostic testing, in: B.H. Baltagi, eds., *A Companion to Theoretical Econometrics*, Blackwell, Oxford, 180-200.
- Wu, P.X. (1991), One-Sided and Partially One-Sided Multiparameter Hypothesis Testing in Econometrics, unpublished doctoral dissertation, Monash University.
- Wu, P.X. and King, M.L. (1994), One-sided hypothesis testing in econometrics: A survey, *Pakistan Journal of Statistics*, 10, 261-300.