

The copyright © of this thesis belongs to its rightful author and/or other copyright owner. Copies can be accessed and downloaded for non-commercial or learning purposes without any charge and permission. The thesis cannot be reproduced or quoted as a whole without the permission from its rightful owner. No alteration or changes in format is allowed without permission from its rightful owner.



***SURE-AUTOMETRICS* ALGORITHM FOR MODEL SELECTION
IN MULTIPLE EQUATIONS**



NORHAYATI YUSOF

UUM

Universiti Utara Malaysia

**DOCTOR OF PHILOSOPHY
UNIVERSITI UTARA MALAYSIA
2016**



Awang Had Salleh
Graduate School
of Arts And Sciences

Universiti Utara Malaysia

PERAKUAN KERJA TESIS / DISERTASI
(Certification of thesis / dissertation)

Kami, yang bertandatangan, memperakukan bahawa
(We, the undersigned, certify that)

NORHAYATI YUSOF

calon untuk Ijazah PhD
(candidate for the degree of)

telah mengemukakan tesis / disertasi yang bertajuk:
(has presented his/her thesis / dissertation of the following title):

"SURE-AUTOMETRIC\$ ALGORITHM FOR MODEL SELECTION IN MULTIPLE EQUATIONS"

seperti yang tercatat di muka surat tajuk dan kulit tesis / disertasi.
(as it appears on the title page and front cover of the thesis / dissertation).

Bahawa tesis/disertasi tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan, sebagaimana yang ditunjukkan oleh calon dalam ujian lisan yang diadakan pada: **26 April 2016.**

That the said thesis/dissertation is acceptable in form and content and displays a satisfactory knowledge of the field of study as demonstrated by the candidate through an oral examination held on:

April 26, 2016.

Pengerusi Viva:
(Chairman for VIVA)

Assoc. Prof. Dr. Sharipah Soaad Syed Yahaya

Tandatangan
(Signature)

Pemeriksa Luar:
(External Examiner)

Assoc. Prof. Dr. Ahmad Mahir Razali

Tandatangan
(Signature)

Pemeriksa Luar:
(External Examiner)

Dr. Siti Meriam Zahari

Tandatangan
(Signature)

Pemeriksa Dalam:
(Internal Examiner)

Dr. Muhammad Mat Yusof

Tandatangan
(Signature)

Nama Penyelia/Penyelia-penyelia:
(Name of Supervisor/Supervisors)

Assoc. Prof. Dr. Suzilah Ismail

Tandatangan
(Signature)

Tarikh:

(Date) April 26, 2016

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences

UUM College of Arts and Sciences

Universiti Utara Malaysia

06010 UUM Sintok

Abstrak

Ketidaktentuan dalam proses pembinaan model dapat dijelaskan oleh pakar pemodelan kerana pengetahuan tersirat yang diperoleh melalui pengalaman menjalankan penyelidikan. Sementara itu, pengamal yang kebiasaannya bukan pakar dan kurang pengetahuan statistik akan berhadapan dengan kesukaran semasa proses pemodelan. Maka, algoritma yang disertai panduan langkah demi langkah adalah bermanfaat dalam pembinaan, pengujian dan pemilihan model. Bagaimanapun, kebanyakan algoritma pemilihan model seperti *Autometrics* hanya tertumpu pada pemodelan persamaan tunggal yang aplikasinya adalah terhad. Oleh itu, kajian ini bertujuan membangunkan algoritma bagi pemilihan model dalam persamaan berganda yang memfokuskan kepada model *Seemingly Unrelated Regression Equations (SURE)*. Algoritma tersebut dibangunkan dengan menyepadukan model *SURE* dan strategi carian oleh *Autometrics*; maka dinamakan *SURE-Autometrics*. Prestasinya dinilai dengan menggunakan ujikaji simulasi Monte Carlo berdasarkan lima model spesifikasi, tiga tahap kekuatan korelasi antara ralat, dan dua saiz sampel. Dua set *General Unrestricted Models (GUMS)* kemudiannya diformulasi dengan menambah beberapa pemboleh ubah tidak relevan terhadap model spesifikasi tersebut. Prestasi tersebut ditentukan melalui peratusan keupayaan algoritma *SURE-Autometrics* berupaya menyingkirkan pemboleh ubah tidak relevan dalam GUMS awalan yang terdiri daripada dua, empat dan enam persamaan. *SURE-Autometrics* juga ditentusahkan menggunakan dua set data sebenar melalui perbandingan ramalan ukuran ralat telahan dengan lima algoritma pemilihan model dan tiga prosedur bukan algoritma. Dapatan daripada uji kaji simulasi mencadangkan bahawa *SURE-Autometrics* berprestasi baik apabila bilangan persamaan dan bilangan pemboleh ubah relevan dalam model spesifikasi sebenar adalah minima. Aplikasi terhadap data sebenar menunjukkan bahawa beberapa model mampu meramal dengan tepat jika data tidak mempunyai masalah kualiti. Algoritma pemilihan model secara automatik ini adalah lebih baik berbanding prosedur bukan algoritma yang memerlukan pengetahuan dan masa tambahan. Kesimpulannya, prestasi pemilihan model bagi persamaan berganda menggunakan *SURE-Autometrics* bergantung pada kualiti data dan kompleksiti dalam model *SURE*.

Kata kunci: Pemilihan model, Algoritma *SURE-Autometrics*, *Seemingly unrelated regression equations*.

Abstract

The ambiguous process of model building can be explained by expert modellers due to their tacit knowledge acquired through research experiences. Meanwhile, practitioners who are usually non-experts and lack of statistical knowledge will face difficulties during the modelling process. Hence, algorithm with a step by step guidance is beneficial in model building, testing and selection. However, most model selection algorithms such as *Autometrics* only concentrate on single equation modelling which has limited application. Thus, this study aims to develop an algorithm for model selection in multiple equations focusing on seemingly unrelated regression equations (SURE) model. The algorithm is developed by integrating the SURE model with the *Autometrics* search strategy; hence, it is named as *SURE-Autometrics*. Its performance is assessed using Monte Carlo simulation experiments based on five specification models, three strengths of correlation disturbances and two sample sizes. Two sets of general unrestricted models (GUMS) are then formulated by adding a number of irrelevant variables to the specification models. The performance is measured by the percentages of *SURE-Autometrics* algorithm that are able to eliminate the irrelevant variables from the initial GUMS of two, four and six equations. The *SURE-Autometrics* is also validated using two sets of real data by comparing the forecast error measures with five model selection algorithms and three non-algorithm procedures. The findings from simulation experiments suggested that *SURE-Autometrics* performed well when the number of equations and number of relevant variables in the true specification model were minimal. Its application on real data indicated that several models are able to forecast accurately if the data has no quality problem. This automatic model selection algorithm is better than non-algorithm procedure which requires knowledge and extra time. In conclusion, the performance of model selection in multiple equations using *SURE-Autometrics* is dependent upon data quality and complexities of the SURE model.

Keywords: Model selection, *SURE-Autometrics* algorithm, Seemingly unrelated regression equations.

Acknowledgement

In the name of Allah, the Most Gracious and the Most Merciful.

Alhamdulillah, all praises to Allah for the strengths and His blessing for completing this thesis. I would like to express my sincere appreciation to my supervisor, Dr. Suzilah Ismail for her extraordinary support in this study, for her patience, motivation, enthusiasm, and immense knowledge. Her guidance really helps me in writing of this thesis.

Besides my supervisor, I would like to thanks my friends and colleagues for their encouragement and insightful comments.

My sincere acknowledgement also goes to my husband, Tengku Mohd Nazli Tengku Mansor, my families, and those who indirectly contributed to my research. I love you all.

Finally, I am very grateful with my employer, Universiti Utara Malaysia for giving me chances to pursue this study. Without their financial support and scholarship, this research would never end with success. Thank you very much.

Declaration Associated with the Thesis

Parts of this thesis have been presented and published for the following occasions:

1. Independence Test in *SURE-Autometrics* Algorithm
Presented at the *International Symposium on Forecasting (ISF)*, Prague, Czech Republic, June 27 – 29, 2011.
2. Lag Variable Reduction in Multiple Models Selection Algorithm
Published in the Proceedings of *International Conference on the Analysis and Mathematical Applications in Engineering and Science (AMAES)*, Curtin University, Sarawak Malaysia, January 19 – 22, 2014.
3. Assessing the Simulation Performances of Multiple Models Selection Algorithm
Published in the Proceedings of *5th International Conference on Computing and Informatics (ICOCI)*, Istanbul, Turkey, August 11 – 13, 2015.
4. Algorithmic Approaches in Model Selection of the Air Passengers Flows Data
Published in the Proceedings of *5th International Conference on Computing and Informatics (ICOCI)*, Istanbul, Turkey, August 11 – 13, 2015.
5. Analysis of *SURE-Autometrics* Algorithm Performance using Simulation Experiment
Journal of Information and Communication Technology (forthcomings)
6. Empirical Study of *SURE-Autometrics* via Air Passengers Flow Data
Journal of Advanced Digital Technology (forthcomings)

Table of Contents

| | |
|---|-----------|
| Permission to Use | ii |
| Abstrak..... | iii |
| Abstract..... | iv |
| Acknowledgement | v |
| Declaration Associated with the Thesis..... | vi |
| Table of Contents..... | vii |
| List of Tables | x |
| List of Figures | xii |
| List of Appendices | xiii |
| List of Abbreviations | xiv |
| CHAPTER ONE INTRODUCTION | 1 |
| 1.1 Background of the Study | 1 |
| 1.2 Problem Statement..... | 4 |
| 1.3 Objectives of the Study..... | 6 |
| 1.4 Significance of the Study..... | 7 |
| 1.5 Scope of the Study..... | 7 |
| 1.6 Thesis Outline..... | 7 |
| CHAPTER TWO MODEL SELECTION ALGORITHM WITHIN THE GENERAL-TO-SPECIFIC APPROACH..... | 9 |
| 2.1 Model Selection | 9 |
| 2.2 Algorithm of Model Selection | 13 |
| 2.3 General-to-Specific Modelling Approach..... | 16 |
| 2.4 <i>PcGets</i> and <i>Autometrics</i> Algorithm | 19 |
| 2.5 Multiple Equations Model | 23 |
| 2.6 <i>SURE-PcGets</i> Algorithm | 27 |
| CHAPTER THREE <i>SURE-AUTOMETRICS</i> ALGORITHM..... | 32 |
| 3.1 Framework of Algorithm Development..... | 32 |
| 3.2 Phase 1: Initial General Unrestricted Model (GUMS) | 34 |
| 3.2.1 Series of Diagnostic Tests..... | 34 |

| | |
|--|-----------|
| 3.2.1.1 Test of normal distribution..... | 35 |
| 3.2.1.2 Test of parameters constancy | 36 |
| 3.2.1.3 Test of autocorrelation..... | 36 |
| 3.2.1.4 Test of unconditional homoscedasticity | 37 |
| 3.2.1.5 Test of conditional homoscedasticity | 38 |
| 3.2.2 Test of Correlation Disturbances amongst Equations..... | 38 |
| 3.2.2.1 Monte Carlo–quasi likelihood ratio test | 40 |
| 3.2.2.2 Multivariate independence test..... | 41 |
| 3.3 Phase 2: Pre-search Reduction..... | 42 |
| 3.3.1 Closed Lag | 44 |
| 3.3.2 Common Lag | 45 |
| 3.3.3 Common– X Lag..... | 46 |
| 3.3.4 Encompassing Test | 46 |
| 3.4 Phase 3: Variable Reduction over Root Branches | 47 |
| 3.4.1 The reduction principles | 48 |
| 3.4.1.1 Pruning | 48 |
| 3.4.1.2 Bunching | 49 |
| 3.4.1.3 Chopping | 50 |
| 3.5 Phase 4: Search for Nested Terminals | 50 |
| 3.6 Phase 5: Selection of the Final Model | 51 |
| | |
| CHAPTER FOUR SIMULATION ASSESSMENT OF <i>SURE-AUTOMETRICS</i> | |
| | 53 |
| 4.1 Experimental Frames | 53 |
| 4.1.1 Artificial Dependent Variables | 54 |
| 4.1.2 Initial General Unrestricted Model (GUMS)..... | 59 |
| 4.2 Measurement of <i>SURE-Autometrics</i> Performance..... | 60 |
| 4.3 Simulation Results | 61 |
| 4.3.1 Two Equations Model..... | 61 |
| 4.3.2 Four Equations Model | 67 |
| 4.3.3 Six Equations Model..... | 72 |
| 4.4 Summary of Findings..... | 76 |

| | |
|---|------------|
| CHAPTER FIVE EVALUATION OF <i>SURE-AUTOMETRICS</i> USING REAL DATA | 83 |
| 5.1 Model Selection Procedures | 83 |
| 5.2 Measures of Forecasting Errors | 86 |
| 5.2.1 Error Measures..... | 87 |
| 5.2.2 Equality Test..... | 88 |
| 5.3 Air Passenger Flows Data..... | 90 |
| 5.3.1 Estimated Models of the Six Routes..... | 95 |
| 5.3.2 Air Passengers Flows excluding the UK-Japan Route | 111 |
| 5.3.3 Air Passengers Flows excluding the UK-Japan and UK-Sweden Routes | 114 |
| 5.4 National Growth Rates Data | 116 |
| 5.4.1 Estimated Models of National Growth Rates | 120 |
| 5.4.2 Forecast Accuracy of National Growth Rates | 128 |
| 5.5 Summary of Findings..... | 136 |
| CHAPTER SIX CONCLUSION AND FUTURE RESEARCH..... | 140 |
| 6.1 Conclusion | 140 |
| 6.2 Suggestions for Further Research | 151 |
| REFERENCES..... | 154 |
| Appendix A <i>Autometrics</i> Algorithm..... | 160 |
| Appendix B <i>SURE-PcGets</i> Algorithm..... | 162 |
| Appendix C <i>SURE-Autometrics</i> Algorithm..... | 165 |
| Appendix D Data of Air Passengers Flows | 168 |
| Appendix E Data of National Growth Rates..... | 174 |

List of Tables

| | |
|---|----|
| Table 3.1 Pre-test Significance Values for $k_p = 1, 2, \dots, 5$ | 44 |
| Table 3.2 Significance Values using Bunching Principle for $k_b = 1, 2, \dots, 5$ | 49 |
| Table 4.1 True Specification Models used by <i>SURE-PcGets</i> | 54 |
| Table 4.2 True Specification Models of Two Equations ($m = 2$) by <i>SURE-Autometrics</i> | 55 |
| Table 4.3 True Specification Models of Four Equations ($m = 4$) by <i>SURE-Autometrics</i> | 56 |
| Table 4.4 True Specification Models of Six Equations ($m = 6$) by <i>SURE-Autometrics</i> | 57 |
| Table 4.5 Categories of Simulation Outcomes | 60 |
| Table 4.6 Simulation Results for $m = 2, n = 146, k = 39$ | 63 |
| Table 4.7 Simulation Results for $m = 2, n = 146, k = 18$ | 63 |
| Table 4.8 Simulation Results for $m = 2, n = 73, k = 39$ | 64 |
| Table 4.9 Simulation Results for $m = 2, n = 73, k = 18$ | 64 |
| Table 4.10 Simulation Results for $m = 4, n = 146, k = 39$ | 68 |
| Table 4.11 Simulation Results for $m = 4, n = 146, k = 18$ | 68 |
| Table 4.12 Simulation Results for $m = 4, n = 73, k = 39$ | 69 |
| Table 4.13 Simulation Results for $m = 4, n = 73, k = 18$ | 69 |
| Table 4.14 Simulation Results for $m = 6, n = 146, k = 39$ | 73 |
| Table 4.15 Simulation Results for $m = 6, n = 146, k = 18$ | 73 |
| Table 4.16 Simulation Results for $m = 6, n = 73, k = 39$ | 74 |
| Table 4.17 Simulation Results for $m = 6, n = 73, k = 18$ | 74 |
| Table 4.18 Simulation Experiment Conditions | 77 |
| Table 4.19 Percentages of Finding Correct Specification for $m = 2$ and $\rho = 0.9$ | 78 |
| Table 4.20 Percentages of Finding Correct Specification for $m = 4$ and $\rho = 0.9$ | 79 |
| Table 4.21 Percentages of Finding Correct Specification for $m = 6$ and $\rho = 0.9$ | 80 |
| Table 5.1 The Selected Airports to Determine Price Tickets between Countries | 92 |
| Table 5.2 Estimated GUMS of Air Passengers' Flows using FGLS and OLS | 94 |
| Table 5.3 Estimated Models for UK-Germany Route | 97 |
| Table 5.4 Estimated Models for UK-Sweden Route | 98 |
| Table 5.5 Estimated Models for UK-Italy Route | 99 |

| | |
|--|-----|
| Table 5.6 Estimated Models for UK-Japan Route | 100 |
| Table 5.7 Estimated Models for UK-US Route | 101 |
| Table 5.8 Estimated Models for UK-Canada Route | 102 |
| Table 5.9 Air Passengers' Forecasting Performances based on RMSE..... | 104 |
| Table 5.10 Air Passengers' Forecasting Performances based on GRMSE..... | 105 |
| Table 5.11 Estimated Models for UK-Germany without UK-Japan Route..... | 106 |
| Table 5.12 Estimated Models for UK-Sweden without UK-Japan Route..... | 107 |
| Table 5.13 Estimated Models for UK-Italy without UK-Japan Route | 108 |
| Table 5.14 Estimated Models for UK-US without UK-Japan Route..... | 109 |
| Table 5.15 Estimated Models for UK-Canada without UK-Japan Route..... | 110 |
| Table 5.16 Forecasting Performances (without UK-Japan) based on RMSE..... | 112 |
| Table 5.17 Forecasting Performances (without UK-Japan) based on GRMSE..... | 113 |
| Table 5.18 Forecasting Performances (without UK-Japan and UK-Sweden) based on RMSE..... | 115 |
| Table 5.19 Forecasting Performances (without UK-Japan and UK-Sweden) based on GRMSE..... | 115 |
| Table 5.20 Estimated GUMS of National Growth Rates using FGLS and OLS | 119 |
| Table 5.21 Estimated Models of Denmark Growth Rate..... | 121 |
| Table 5.22 Estimated Models of Ireland Growth Rate | 122 |
| Table 5.23 Estimated Models of Italy Growth Rate | 123 |
| Table 5.24 Estimated Models of Netherland Growth Rate..... | 124 |
| Table 5.25 Estimated Models of UK Growth Rate..... | 125 |
| Table 5.26 Estimated Models of US Growth Rate | 126 |
| Table 5.27 National Growth Rates' Forecasting Performances based on RMSE | 129 |
| Table 5.28 National Growth Rates' Forecasting Performances based on GRMSE... | 130 |
| Table 5.29 Measures of Forecast Error for Italy Growth Rate | 131 |
| Table 5.30 Forecasting Performances (without Italy) based on RMSE..... | 133 |
| Table 5.31 Forecasting Performances (without Italy) based on GRMSE..... | 134 |
| Table 5.32 Forecasting Performances (without Italy and US) based on RMSE..... | 135 |
| Table 5.33 Forecasting Performances (without Italy and US) based on GRMSE..... | 136 |
| Table 6.1 Experiment Conditions with Probabilities above 80% | 145 |

List of Figures

| | |
|--|----|
| Figure 2.1. Concept of GETS Modelling Approach. Adapted from “The Theory of Reduction” by D. F. Hendry and J. A. Doornik, 2014, Empirical Model Discovery and Theory Evaluation, p. 96. Copyright 2014 by the MIT Press..... | 17 |
| Figure 2.2. Search strategy in Autometrics. Adapted from “Autometrics” by J. A. Doornik, 2009, The Methodology and Practice of Econometrics, p. 93. Copyright 2009 by the Oxford University Press..... | 22 |
| Figure 3.1. SURE-Autometrics Development Framework..... | 33 |
| Figure 4.1. Overall Performances on Two Equations Model using Large Sample Size | 65 |
| Figure 4.2. Overall Performances on Two Equations Model using Small Sample Size | 66 |
| Figure 4.3. Overall Performances on Four Equations Model using Large Sample Size | 71 |
| Figure 4.4. Overall Performances on Four Equations Model using Small Sample Size | 71 |
| Figure 4.5. Overall Performances on Six Equations Model using Large Sample Size | 75 |
| Figure 4.6. Overall Performances on Six Equations Model using Small Sample Size | 76 |
| Figure 4.7. Overall Performances in Finding Correct Specifications | 82 |
| Figure 5.1. Number of Air Passengers for Six Routes..... | 91 |

List of Appendices

| | |
|--|-----|
| Appendix A <i>Autometrics</i> Algorithm | 160 |
| Appendix B <i>SURE-PcGets</i> Algorithm..... | 162 |
| Appendix C <i>SURE-Autometrics</i> Algorithm | 165 |
| Appendix D Data of Air Passengers Flows | 168 |
| Appendix E Data of National Growth Rates..... | 174 |



UUM
Universiti Utara Malaysia

List of Abbreviations

| | |
|--------|---|
| DGP | data-generating process |
| e.g. | for example |
| et al. | and others |
| etc. | and so forth |
| FGLS | feasible generalised least squares |
| GDP | gross domestic product |
| GETS | general-to-specific |
| GLS | generalised least squares |
| GRMSE | geometric root mean square error |
| GUM | general unrestricted model of single equation |
| GUMS | general unrestricted model of multiple equations |
| i.e. | that is |
| LM | Lagrange Multiplier |
| MC-QLR | Monte Carlo-Quasi Likelihood Ratio |
| MI | Multivariate independent |
| OLS | ordinary least squares |
| RMSE | root mean square error |
| SUM | specific unrestricted model of single equation |
| SUMS | specific unrestricted model of multiple equations |
| SURE | seemingly unrelated regression equations |
| UK | United Kingdom |
| US | United States |

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Statistical modelling normally has inexplicit processes due to a tacit or personal knowledge. This can be gained through experience where modellers combined their judgmental knowledge and theoretical studies at some point in the modelling process (Magnus & Morgan, 1999). Generally, the process commence with a model formulation which involves specification of identified variables and followed by estimation procedure. Then it is validated through a series of evaluations where re-specification will be required according to certain criteria such as diagnostic testing, goodness of fit and hypothesis testing of the parameters.

The specification of model involves choosing which variables to include or exclude from the model while maintaining the consistencies with the observed data. According to Magnus (1999), the selection of predictor variables could be based on two basic modelling approaches where it can possibly starts from a simple model and expand it, or from a general model which subsequently reduce to a more simplified form. The first approach is known as specific-to-general or bottom-up where it uses the theory to provide an initial specification. Then, it is refined by adding or subtracting the variables or substitutes the coefficients estimator according to modeller's prior belief or data exploration techniques such as Cochrane-Orcutt transformation. On the contrary, the second approach starts with a general model formulated based on information collected from theories, previous empirical research evidence, institutional knowledge, and common sense (Hendry & Doornik, 2014). This initial model which comprises of all the candidate variables is then refined by the

process of removing unimportant variables, leading to a simple model that better represents the data. Thus, it is called a general-to-specific (GETS) or top-bottom approach. However, experts acknowledged that there is no best way amongst these two approaches in specifying the model (Granger, 1999; Hendry, 1980).

Once the model has been specified by either way of approach, it will be estimated and tested for diagnostic to ensure it is consistent with the observed data. Practically, there is no assurance that a given model is correctly specified especially in a non-experimental study such as econometric modelling (Hendry, 2001). The misspecification occurs whenever the relevant variables are omitted from the model, the irrelevant variables included in the model, incorrect choices of functional form or the model failed any diagnostic tests. These specification errors will influence the properties of estimation technique, the quality of inferences, and the accuracy of the forecasting.

Thus, if the initial model is already mis-specified, then the attempts of refining the model on the basis of statistical tests are comparable to improvement based on erroneous statistical procedure. According to Granger and Hendry (2005), there is insufficient theoretical justification for adding new variable in order to improve a model with an evidence of serial correlation in the errors. This state commonly occurs to those who employed the specific-to-general approach. For instance, Bhatti, Al-Shanfari and Hossain (2006) found that the estimated model sometime failed to advocate an adequate functional form. Hence, choosing the general-to-specific modelling approach would overcome this weakness where the models will be revised as a response to any indications of model misspecification.

However, to formulate and simplify such a general model requires resource demanding, and time consuming (Magnus, 1999). Moreover, with k number of candidate variables will produce 2^k number of possible specifications. Thus, model simplification leads to difficulties in determining which set of predictors that best explain the observed responses because there could be more than one possible specification for one modelling situation.

A procedure of choosing an adequate model rather than making an arbitrary choice of model is known as model selection. The specification will solely determine by the data through the process of estimating the performance of different models in order to choose the 'best' one that described the process under study (Hastie, Tibshirani, & Friedman, 2001). It involves the inclusion or removal of variables until some termination criterion is satisfied. The criterion is used as an indicator for the goodness of fit for a particular model while taking into account the number of estimated parameters. The criteria that can be used for this purpose are adjusted R^2 , Akaike's information criterion (AIC), Schwartz's criterion (SC), final prediction error (FPE), etc. As a guide for selecting the 'best' model, Doornik (2008) favoured the termination based on a principle of minimizing the residual sum of squares or pre-specified significance level.

In practice, the modellers frequently used the significance values of t and F tests to decide whether certain variable or a group of variables should be included or excluded from the model. Thus, the 'best' model is presented with all the significant predictors. Moreover, Hendry and Doornik (2014) showed that a good model not only the 'best'

model but also consistent with the observed data which means that it is not violated any diagnostic tests such as normality, autocorrelation and homoscedasticity.

Overall, the formulation of a model involves the process of finding the right specification by selecting appropriate variables without violating any assumptions about the error component in order to be a good model as a representation of a particular phenomenon under study. Hence, the model will give accurate forecasts.

1.2 Problem Statement

Commonly, expert modellers are unable to clearly explain the process of finding the right variables in a good model determined by the data. The experimental studies by Magnus and Morgan (1999) proved the involvement of ‘tacit’ or personal knowledge along the process of modelling. This type of knowledge cannot be articulated and can only be learned through research experiences since it combined both theories and intuitive judgments. Therefore, it is difficult to master (Pindyck & Rubinfeld, 1998) especially for practitioners who are usually non-experts and lack of statistical knowledge. The gap can be bridged through an automatic modelling approach which provided an algorithm to guide the modellers a step-by-step procedure of formulating and testing the model in order to find the appropriate one for forecasting or policy making purposes. As a consequence, different researchers should obtain the same results by following the same algorithm for a given data set.

The model selection algorithm executes the process of inclusion or exclusion of variables from or into the model according to some termination criteria. These algorithms search the ‘best’ model that fits the data from a set of alternative possible models to be a good forecasting model or to enhance the reliability of coefficient

estimates. The most notable algorithm is *Stepwise Regression* which introduced by Efroymson (1960). The procedure starts from simple model and expand it by adding variables. Even though it has been widely used in many fields, stepwise also received various criticisms (see among others, Lovell (1983), Derksen and Keselman (1992), Foster and Stine (2004) and Whittingham, Stephens, Bradbury, and Freckleton (2006)) particularly due to the consequence of misspecification in the initial model where the criterion used to decide the additional variable will be tainted by the biases. Moreover, the search strategy leads to the algorithm that stuck in a sequence of tests that accidentally eliminates a variable which matters, and thereby retains ‘spurious’ variables (Hendry & Krolzig, 2003).

Unlike *Stepwise*, the algorithm that explores more than one path and includes the diagnostic tests as one of the termination criteria is *PcGets* (Hendry & Krolzig, 2001). It has been replaced by *Autometrics* (Doornik, 2009) in order to introduce more paths searches by implementing tree search strategy which uses a systematic reduction process and to improve the computational efficiency. Both algorithms able to solve the problem faced in *Stepwise* by employing general-to-specific (GETS) approach while finding the best model. Hence, the algorithm developed within the GETS has found to be more advanced due to an exhaustive search strategy (Castle, Doornik, & Hendry, 2011; Granger, Hendry, & Hansen, 2005; Hendry & Krolzig, 2005).

Previous studies indicate that most of the model selection algorithms are established for single equation (Castle, Qin, & Reed, 2013; Hendry & Doornik, 2014; Santos, Hendry, & Johansen, 2007) but limited for multiple equations model (Ismail, 2005; Krolzig, 2001). This leads to more opportunities on extending and exploring the

algorithm for multiple equations. Specifically model selection within the GETS approach since the properties have been well developed (Castle et al., 2011; Hendry & Doornik, 2014; Hendry & Krolzig, 2005).

Hence, this study aims to develop a model selection algorithm on the basis of GETS approach for multiple equations. The motivation is to advance the specification process during the model formulation through a model selection technique by providing an algorithm for automatic procedure. Moreover, this study is inspired to implement the model selection process within the GETS approach, as well as to broaden the theme of algorithm development which most previously favours on the single equations modelling.

1.3 Objectives of the Study

Based on the aim, this study embarks on the following objectives,

1. To develop an algorithm of model selection for seemingly unrelated regression equations model. The new algorithm is named as *SURE-Autometrics*.
2. To assess the performances of *SURE-Autometrics* using simulation experiments based on five specification models with three different number of equations, three strength of correlation disturbances, two sample sizes, two sets of initial models and two significance levels.
3. To evaluate the performances of *SURE-Autometrics* using two sets of real data based on forecasting accuracy.

1.4 Significance of the Study

Most modellers implement tacit knowledge throughout the modelling process. Pindyck and Rubinfeld (1998) believed that this type of knowledge is very difficult to master especially inexperienced researchers. Hence, an automatic modelling approach can be a formal way of bridging the gap. It provides an algorithm to guide the modellers a step-by-step procedure of formulating and testing the model in order to find the most appropriate for forecasting and policy making purposes.

Substantially, this study offers a reliable guidance steps in developing automatic model selection for multiple equations model. The new algorithm is established for the seemingly unrelated regression equations (SURE) model type. Thus, it will lessen the dispute amongst modellers in the process of multiple equations modelling. Since real data possess different characteristics, assessments of the new algorithm are crucial to highlight the performances of the new algorithm.

1.5 Scope of the Study

This study concentrates on the econometric modelling using the general-to-specific (GETS) approach where the data are time series. While there are many types of multiple equations model, the focus was on the initial specification of seemingly unrelated regression equations (SURE) proposed by Zellner (1962).

1.6 Thesis Outline

This thesis is organized in six chapters as follows:

Chapter one provides an overview of the study regarding the background, statement of the problem, objectives as well as the scope, the significance of this study and finally the outline of thesis.

Chapter two reviews all the previous studies concerning the issues in statistical modelling approaches which lead to the implementation of automatic procedure. In general, this chapter discusses topics regarding model selection algorithm within the general-to-specific approach for single and multiple equations.

Chapter three contains the main contributions of this study where it describes the development of new model selection algorithm for multiple equations within the general-to-specific approach. The descriptions are based on five development phases.

Chapter four assesses the performances of *SURE-Autometrics* through various experimental simulations. It starts with the experimental frames involving data generation process and several conditions that are considered during the model selection procedure. The simulation results are separated according to number of equations.

Chapter five illustrates the application of the *SURE-Autometrics* using air passengers' flows data and national growth rates data. The chapter begins with descriptions of several model selection using algorithm (automatic) and non-algorithm (manual) procedures. The forecasts from model selected by these procedures are compared using the error measures and equality tests.

Chapter six presents the overall conclusions. It also specifies the suggestions for further study.

CHAPTER TWO

MODEL SELECTION ALGORITHM WITHIN THE GENERAL- TO-SPECIFIC APPROACH

2.1 Model Selection

The model selection technique is a statistical procedure of searching for the 'best' model that fits the data from a set of alternative possible models (Bhatti et al., 2006; Miller, 2002). Generally, the techniques can be classified into two broad categories. The first category is based on hypothesis testing procedures (Leeb & Pötscher, 2005) where multiple tests such as t test and F test are required to determine the significance of variables in the model, and diagnostic tests for checking any violation in model's assumptions. Thus, the data set is reused for the purpose of model selection. This situation is known as data mining (Lovell, 1983) or data snooping (White, 2000) leading to the problem of mass significance where incorrect significance level is actually reported by the modellers. Moreover, Denton (1985) showed the probability of false inclusion increases with the number of researchers involved (or number of hypotheses tested) which implied that data mining also occur when a single researcher tests on several hypotheses and selects the best model, or many researchers test one hypothesis each and only significant results are reported. For instance, 40% of overall significance is gained if repeated testing is done at 5% nominal level for ten different tests. This means, a nominal level of 0.1% to 1% is needed to obtain an overall level of 5%, depending on the number of tests perform. Besides, this category only appropriate for nested case where the set of competing models are ordered by the inclusion of relation (Leeb & Pötscher, 2009).

The second category is another way of selecting models by optimizing an information criterion (IC). There are many criteria have been proposed over the years where R^2 and adjusted R^2 are amongst the earliest one. Some other examples include final prediction error criterion (FPE), Akaike information criterion (AIC), Mallows' C_p criterion, Schwarz's Bayesian information criterion (SC), and Hannan and Quinn's (HQ) criterion. Among these, AIC and SC are the most commonly used in econometrics studies (Bhatti et al., 2006). The main aim of these criteria is to select a model that gives the maximum or minimum information by using risk measure such as mean square error of prediction and Kullback-Leiber discrepancy. However, this category seems infeasible for large number of predictors because the information from all the 2^k possible models (including empty model) need to be computed in order to select the 'optimum' specification, yet the result is either asymptotically efficient or consistent. Moreover, the criteria are not particularly concerned about whether a true model exists or not (Leeb & Pötscher, 2009).

In recent years, shrinkage estimators and Bayesian model averaging (BMA) also can be considered for model selection (Leeb & Pötscher, 2009; Miller, 2002). Shrinkage methods can be seen as continuous versions of model selection where different weights are allocated to the regression coefficients in relation to their significance. Methods that have been proposed which essentially give shrunken estimates are ridge regression, nonnegative garrotte (Breiman, 1995) and LASSO (least absolute shrinkage and selection operator) (Tibshirani, 1996). The first is continuously shrinks the coefficients by minimizing a penalized residual of sum squares while the rests are both impose a limit on the sum of absolute values of the regression coefficients, and

lower this limit until some kind of optimum is reached. For further discussions, see Hastie, Tibshirani, and Friedman (2001), and Miller (2002).

Unlike shrinkage methods, BMA gives weight to the models. Leamer (1978) appeared to be the first pointed out the fundamental idea of BMA as a means of combining models so that it will account for the uncertainty involved in selecting the model or uncertainty about statistical structure when making inferences. Instead of choosing one model, this technique considered an average over all possible models in a given model class. The weight are the posterior probabilities that are strongly influenced by the two components of the prior probabilities, that is the prior for the model and the prior for the parameters (regression coefficients and residual variance) in that model. Theoretically, BMA provides better average predictive performance which Hoeting, Madigan, Raftery, and Volinsky (1999) supported by practically implementing to four different classes of model: linear regression models; generalized linear models; survival analysis; and graphical models.

However, shrinkage and BMA procedures have some disadvantages. The shrinkage methods trading-off decreased variance for increased bias in the estimates of regression coefficients estimators as well as estimates which have been shrunk (Miller, 2002). In other words, the model has too many variables to make sure that the relevant one is not dropped. BMA also gives a similar outcome with shrinkage methods where the final model includes all of the predictors, though usually some of them are given so little weight that they are effectively excluded. Hendry and Reade (2008) suggested model averaging will be much more effective after a good model selection is carried out. Thus, model averaging is complementary to model selection.

According to all the selection techniques, it seems sequential testing is very crucial in econometric empirical modelling because the formulation involves simultaneous activities of hypothesis testing and estimation method (Hendry, 1980; Magnus, 1999). The mass significance problem though can be avoided if we are able to control the probability of making at least one false rejection or also known as Family-Wise Error (FWE) (Hansen, 2005; Romano, Shaikh, & Wolf, 2008; White, 2000). Since the procedure requires a large number of tests and time consuming, it would be better if we employ an automatic procedure rather than manual procedure. Besides, it will reduce human interference in the process of re-specification.

For a given data set, the manual process frequently leads to different final models due to specialization of interest and beliefs, specific or mix of methodological approach, and general scientific principles (Magnus & Morgan, 1999). This situation particularly will affect the choice of response and predictor variables, the measurement of the variables, and the effect of including dynamics. For instance, most of the modellers involved in the experimental study conducted by Magnus and Morgan (1999) showed that re-specification was based on the diagnostic results where the model is re-specified to overcome heteroscedasticity and non-normality problem, or because of the presence of outliers.

The disagreements among the modellers are not only about the approach or the results obtained, but most importantly about the justification for intermediate steps which were seen to embody methodology, assumptions, and technical skills. In most of empirical studies, the results are reported without explicitly explaining the modelling process. Thus, it will be hard for other practitioners to capture the methodology,

unless they read on theoretical articles and refer examples through applied works which will lead to different interpretations. As for those who have experienced, personal knowledge would be involved in the model specification (Magnus & Morgan, 1999).

2.2 Algorithm of Model Selection

Utilizing the computerized procedure particularly for model selection provide a rule or algorithm that will be the guide for practitioner in empirical modelling where the data-generating process (DGP) and coefficients are unknown (Hendry & Doornik, 2014). Recently, Castle, Qin and Reed (2013) reviewed several model selection algorithms and concluded that these algorithm were developed with different purposes. The aims of selecting a model(s) is to find the best that represents the true DGP, to advance the properties of inference such as reliable coefficient estimates, and finally to find a good model with ability to forecast out-of-sample observations. At the same time, algorithm could help diminishing the role of tacit knowledge as well as eliminating computational burden with the increasing large amounts of information which is very rare in the past (Granger & Hendry, 2005).

A model selection algorithm requires a strategy in searching the best model. The search strategies can be divided into two groups which are 'cheap' methods and exhaustive methods. The first group is called the 'cheap' strategy due to its inconsistency in finding the best fitting model (Miller, 1984). The algorithms in this group are usually referred to as stepwise multiple regression (Miller, 2002). These include forward selection, backward elimination and stepwise regression. The most commonly used by practitioners is the stepwise regression (Whittingham et al., 2006).

The forward selection operates by successively adding a significant variable according to the highest correlation with the response in the initial step, and partial correlation on the following steps. On the contrary, backward elimination starts with all the candidate variables included in the model. Any insignificant variables are removed consecutively based on t value at each steps. Combining both procedures, stepwise regression employs forward selection first, followed by backward elimination to examine whether the included variables could be removed along the process.

The stepwise regression is basically developed within the specific-to-general approach. Even for manual procedure, this modelling approach shows disadvantages that cause most of econometric modellers preferred to commence from the other way around (Hendry, 1980; Leamer, 1978; Sims, 1980). The main problem is when the simple model has incorrect specification. For instance, the model has two or more failure tests of diagnostic assumptions such as heteroscedasticity and autocorrelated disturbances. It can be refined based on correction of one or both conditions, or sought other factors. According to Hendry and Krolzig (2001), if several tests are computed sequentially, and a later one rejects, then it will invalidate all the earlier inferences which indicated that the approach is inefficient.

Therefore, studies have shown that the algorithm developed within the general-to-specific (GETS) modelling approach are more advanced in this field (Granger & Hendry, 2005; Hendry & Krolzig, 2003). The algorithm implements an exhaustive search strategy to ensure the finding of the best fitting model. The idea is to formulate a general model that captures the characteristics of the data including previous

empirical findings and knowledge of economic theories, and then proceed with the process of variables elimination via sequential testing to select a parsimonious model that encompasses rival models, while retaining the congruency. According to Hendry (1995), a model is congruent if it passes all the misspecification tests such as normality, structural change, autocorrelation, and heteroscedasticity. Every insignificant variable in the general model defines a path. Instead of searching for one path as in 'cheap' method, *PcGets* (Krolzig & Hendry, 2001) implements the multiple paths search where more than one simplified model could be resulted which all are valid reduction of the general model. However, Doornik (2009) said that most of the paths possibly will turn out to be the same and the rest are left unsearched. Thus, algorithm known as *Autometrics* (Doornik, 2009) embedded in *PcGive* software is developed to introduce more paths by searching the whole space of models generated by the variables in the initial model.

Overall, the stepwise, *PcGets* and *Autometrics* are widely used since the algorithms have been transformed into software that is easily applied. However there are other algorithms for model selection. As for example, *ModelBuilder* (Mycielski & Kurcewicz, 2004) is the extended version of *PcGets* that accounts for cointegration vectors, Relevant Transformations of the Inputs Network Approach (*RETINA*) (Pérez-Amaral, Gallo, & White, 2003) is developed for selecting nonlinear representations, and *PIC* model selection (Philips, 2003) used the Bayesian approach to select forecasting models. Their algorithms rely on automated significance tests in conjunction with model selection rules for dealing with rival specifications that are unresolved by significance testing.

Finally, the drawbacks of model selection technique could be reduced by properly implementing it during the model formulation phases. The best strategy is via algorithm where we can simulate the process in order to study the cost of selection (Hendry & Krolzig, 1999; Hoover & Perez, 1999; Lovell, 1983). The cost is important to determine the accuracy of recovering the true model if its specification were known, and to measure the Type I and Type II error rates. Compared to commonly practice model selection algorithm, the algorithms within GETS approach have shown their success in producing low costs of selection (Hendry & Krolzig, 2005).

2.3 General-to-Specific Modelling Approach

The general-to-specific (GETS) method of modelling also known as 'Hendry' methodology was promoted by David Hendry in early 1980. It is developed based on model specification philosophy in the LSE framework. The abbreviation 'LSE' is derived from the fact that there is a tradition of time-series econometrics that arose in the 1960s at the London School of Economics. A brief history of this philosophy can be found in Mizon (1995) while the description about the origins was provided by Hendry (2003).

Basically, Hendry (1980, 1995) viewed a model specifically the econometric model as a representation of the process or probability distribution generating the sample data that is data-generating process (DGP). It begins with a very general parameterisation that is acceptable to a range of plausible theoretical positions representing the DGP. Then it is reduced to the 'local' DGP (LDGP) which defined as the joint distribution of the subset of variables by operations such as aggregation, marginalization,

conditioning, sequentially factorization and transformation. This is a theory of reduction (Hendry, 1995). In other words, the potentially vast initial of information set is reduced to the small transformed subset that is relevant for the problem in question. The general intention is to counteract the fundamental problem in the classical economic research which naturally using the specific-to-general approach.

Figure 2.1 shows this modelling concept.

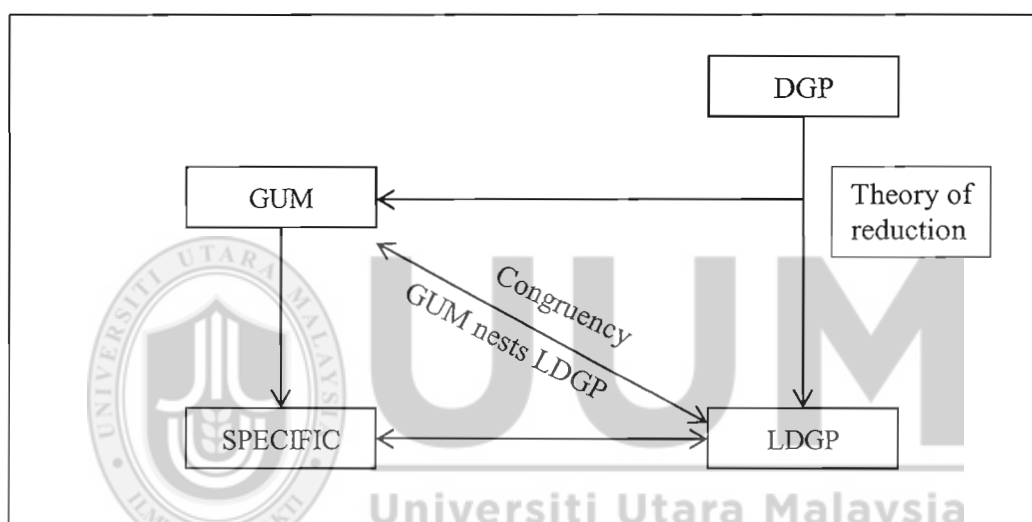


Figure 2.1. Concept of GETS Modelling Approach. Adapted from “The Theory of Reduction” by D. F. Hendry and J. A. Doornik, 2014, *Empirical Model Discovery and Theory Evaluation*, p. 96. Copyright 2014 by the MIT Press.

Practically, a general unrestricted model (GUM) that captures the characteristics of the data including previous empirical findings and knowledge of economic theories is derived. Then statistical procedures are applied to the GUM whereas diagnostic tests are used to retain the congruency. According to Hendry (1995), a model is congruent if it passes all the misspecification tests such as normality, structural change, autocorrelation, and heteroscedasticity. In this way, the empirical model is determined only via estimation and testing procedures (Bond & Harrison, 1992). Hendry (1980)

also admitted that this approach is ad hoc and used economic theory as a guideline and relied heavily on data.

Hoover and Perez (1999) was the first who evaluated the performances of GETS modelling using an algorithm. They simulated GETS selection for dynamic linear regression models and found that their algorithm performs well in re-mining the 'Lovell database' (1983). The search commenced from a congruent general model, followed by a number of reduction paths which can be terminated by either no further feasible reductions or significant diagnostic tests occurred. Models survived from the reduction process were selected based on which one fits best. They showed that with a structured approach of GETS method of econometric modelling, the results are better than any method Lovell considered. Note that the methods are 'max-min-t' selection strategy, maximizing R^2 through exhaustive search, and a stepwise regression procedure. Additionally, the size (type I error rates) of their selection procedure was close to expectation, and the power (type II error rates) of test was reasonable.

Subsequently, *PcGets* (Hendry & Krolzig, 1999) algorithm adopted the multiple paths strategy, exploring the consequences of all initially feasible paths, and collecting the 'terminal' models resulting from each search. If many 'terminal' models are found, these are tested for parsimoniously encompassing their union, namely the smallest model that nests all the contenders. The algorithm also able to achieve a better size and power of test using the similar Lovell's experiments (Hendry & Krolzig, 1999; Krolzig & Hendry, 2001). With multiple series of studies, Hendry and Krolzig (2001, 2003, 2004, 2005) demonstrated the major and unique strength of GETS approach are within their guidelines regarding specification search via model selection algorithm.

Essentially, *PcGets* has exemplified a more robust approach by adding the specification tests in the process of model selection.

Ultimately in 2007, Doornik and Hendry have introduced *Autometrics* as a new model selection algorithm. This is an improvement that the algorithm within GETS approach showed progressively enhancing through the years. The new algorithm advanced on several aspects such as, new search method by considering more paths; avoid repeated estimation of the same model; delay diagnostic testing; and recall terminals between iterations (Doornik, 2009). Hence, the computational efficiency has improved and better results are obtained in the operational studies of cost of selections.

2.4 *PcGets* and *Autometrics* Algorithm

PcGets is an algorithm developed by Hendry and Krolzig (2001) for implementation of an automatic selection for linear, dynamic, regression models. The algorithm was embedded in an econometric software package known as *GiveWin* 2.10. The algorithm was based on the principles discussed in Hendry (1995) which is the general-to-specific (GETS) approach. It also successfully executes a consistent automatic model selection for single equation econometric modelling (Hendry & Krolzig, 2005) which has been the main problem faced by stepwise regression. Unlike stepwise, the GETS algorithm implements a multiple paths search strategy within a general-to-specific approach with diagnostic checking for each selection stages. The path is determined by each of insignificant variable in the general model (Hoover & Perez, 1999) and a number of variables that is grouped according to pre-specified significance level (Hendry & Krolzig, 1999). Since *PcGets* implements multiple path searches, there is possibility that more than one candidate model (denoted terminal)

survived the reduction process. Thus the terminals are tested by encompassing against the general model. If several terminals remain acceptable, so the reduction process recommences from their union, till a unique outcome is obtained. The search is terminated whenever all selected simplifications re-appear and the final model is chosen using the Schwarz (1978) information criterion. Finally, *PcGets* employs a sub-sample insignificance in order to identify ‘spuriously significant’ predictors.

Every estimated model including the initial GUM in GETS modelling is subjected to a battery of diagnostic tests to ensure the validity of the model along the reduction process. Hendry (2000) described that a model is congruent if it is fulfil all criteria resulted from empirical model selection. The criteria are homoscedastic innovation errors, weakly exogenous conditioning variables for the parameters of interest, constant or invariant parameters of interest, theory consistent, identifiable structures, and data admissible formulations on accurate observations. Specifically, the model should pass the entire following tests,

1. Normality
2. Autocorrelation (AR)
3. Autocorrelated conditional heteroscedasticity (ARCH)
4. Heteroscedasticity
5. Parameter constancy

Section 3.2.1 will describe these tests explicitly.

The modellers should decide whether to ignore the rejected tests, continue or revise the GUM when one or more these tests rejects the GUM (Hendry & Krolzig, 1999).

While on the reduction process, *PcGets* will adjust the significant level to prevent them from rejection test. For example, 1% level is reduced to 0.1%.

However, *Autometrics* (Doornik, 2009) behaved in different way where the diagnostic validity is restored along the process of reduction by giving the preference to terminal models that pass on the original level (say 1%) during the search. The algorithm is embedded within the current version of *PcGive* which is Version 12. It is computer software for econometric data analysis where the 'GIVE' stands for Generalised Instrumental Variables Estimators. *Autometrics* is considered as a new generation of *PcGets* where it contemplates all the properties in general-to-specific model selection (Hendry & Krolzig, 2005). It has three main features that are aimed to improve the predecessor. First, *Autometrics* is able to function with or without pre-search reduction process because it was an ad-hoc procedure in *PcGets* (Doornik, 2009). Second, instead of multiple paths search, *Autometrics* introduced tree search strategy to find all the possible sets of variables in the general model which will lead to more paths can be explored in a systematic way. By employing pruning, bunching, and chopping, more terminal models are available to be selected by the information criteria. Finally, *Autometrics* aimed to improve the computational efficiency in GETS modelling which is achieved by avoiding repeated estimation of the same model, delayed the diagnostic testing, and memorize terminals between iterations. Consequently, it has been successfully applied for single equation modelling (Ericsson & Kamin, 2009; Reade, 2007).

Hence, *Autometrics* is the most recent model selection algorithm. It is a successor of *PcGets* and can be considered as the third generation of GETS model selection

algorithm. The program embedded with this algorithm was already available to practitioners since 2007, though the description is in Doornik (2009). An overview of the algorithm is provided in Appendix A.

The algorithm has all the properties in automated GETS modelling. Thus it implements all the main elements in *PcGets* namely, the general model, diagnostic testing, presearch reduction prior to multiple path searches, encompassing tests, and iterative procedure.

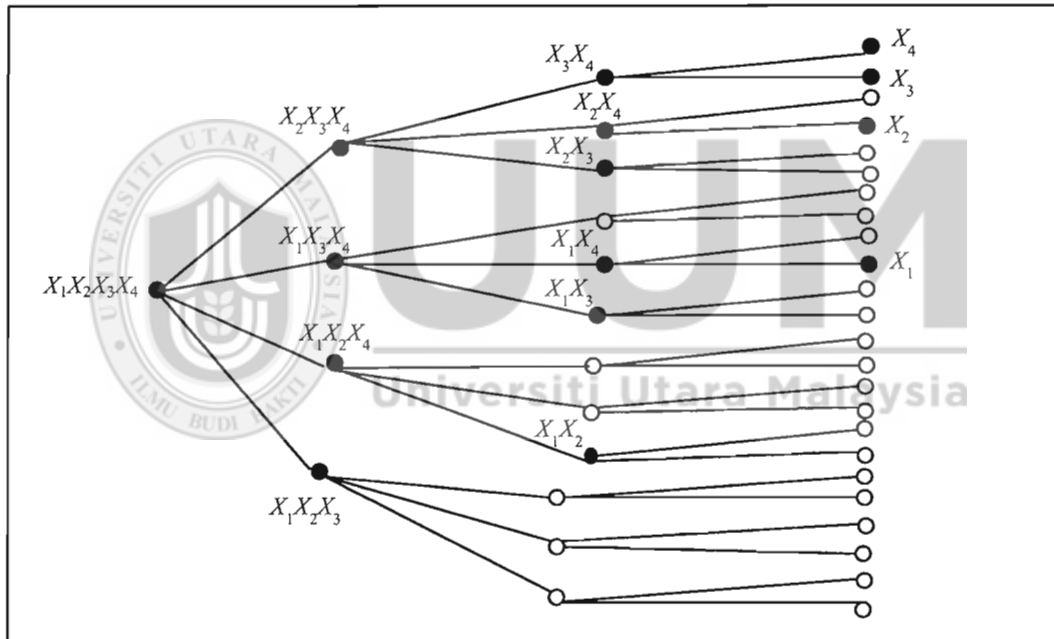


Figure 2.2. Search strategy in *Autometrics*. Adapted from “*Autometrics*” by J. A. Doornik, 2009, *The Methodology and Practice of Econometrics*, p. 93. Copyright 2009 by the Oxford University Press.

The resulting tree is a unique representation of the model space. Precisely, all possible models would be estimated if we move from left to the right, and top to the bottom. In other words, we have considered all possible subsets starting from the GUM, where the ordering is defined by increasing significance. Note that the first path in the tree

corresponds to the first of the paths considered in Hoover and Perez (1999) and Hendry and Krolzig (1999, 2001) but after that it diverges. Doornik (2009) admitted that each of subset is not feasible. Therefore, he implemented systematic strategies to move efficiently through pruning, bunching, and chopping scheme.

2.5 Multiple Equations Model

Most of the algorithms were successfully developed for single equation modelling (Castle et al., 2013; Doornik & Hendry, 2007; Hendry & Krolzig, 2001). A model too might contain a series of equations that are independent of each other such as multivariate linear regressions. In this situation, *Autometrics* treats the equation individually by executing a model selection separately for multiple times. This procedure however is inappropriate if these equations have contemporaneously correlated disturbances amongst the equation. Zellner (1962) presented this type of model as a seemingly unrelated regression equations (SURE) model for simultaneously estimating the equations.

Most of SURE model applications arises in econometric, financial and sociological modelling (Fildes, Wei, & Ismail, 2011; Srivastava & Giles, 1987; Zellner, 1962). However, it can also be applied to other areas such as human genetics (Verzilli, Stallard, & Whittaker, 2005) and behavioural science (Fernandez, Smith, & Wenger, 2007; Schwartz, 2006). These examples advocate that SURE model is appropriate and useful for wide range of applications.

The main goal of modelling with SURE specification is to improve efficiency in estimation by combining information on different equations. It also aims to impose or

to test restrictions that involve parameters in different equations. The series of equations are specified as follows,

$$\begin{aligned}
 y_{1t} &= \beta_{11}x_{1t,1} + \beta_{12}x_{1t,2} + \dots + \beta_{1k_1}x_{1t,k_1} + \varepsilon_{1t} \\
 y_{2t} &= \beta_{21}x_{2t,1} + \beta_{22}x_{2t,2} + \dots + \beta_{2k_2}x_{2t,k_2} + \varepsilon_{2t} \\
 &\vdots \\
 y_{mt} &= \beta_{m1}x_{mt,1} + \beta_{m2}x_{mt,2} + \dots + \beta_{mk_m}x_{mt,k_m} + \varepsilon_{mt}
 \end{aligned} \tag{2.1}$$

which can be written in general form,

$$\underset{T \times 1}{\mathbf{y}_i} = \underset{T \times k_i}{\mathbf{X}_i} \underset{k_i \times 1}{\boldsymbol{\beta}_i} + \underset{T \times 1}{\boldsymbol{\varepsilon}_i} \quad i = 1, 2, \dots, m \tag{2.2}$$

where \mathbf{y}_i is vector of T identically distributed observations for each random variable, \mathbf{X}_i is a non-stochastic matrix of fixed variables of rank k_i , $\boldsymbol{\beta}_i$ is vector of unknown coefficients, and $\boldsymbol{\varepsilon}_i$ is a vector of disturbances.

As usual, it is assumed that the disturbances has a multivariate normal distribution with mean and covariance structure,

$$E(\boldsymbol{\varepsilon}_{it}) = 0, \quad E(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{jt}) = \sigma_{ij}, \quad \text{and } E(\boldsymbol{\varepsilon}_{it}\boldsymbol{\varepsilon}_{jl}) = 0 \text{ if } t \neq l \tag{2.3}$$

where $i, j = 1, 2, \dots, m$ and $t, l = 1, 2, \dots, T$. If the disturbances are contemporaneously correlated (Kontoghiorghes, 2004) it can be expressed as,

$$E(\boldsymbol{\varepsilon}_i) = 0 \quad \text{and} \quad E(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j') = \sigma_{ij}\mathbf{I}_T \tag{2.4}$$

where $\sigma_{ij}^2 = \sigma_i^2$ is the variance of disturbances for the i^{th} equation if $i = j$, and $E(\cdot)$ is the expectation operator.

Further simplification of equations in 2.1 can be accomplished in vector concatenation form by stacking the m vector equations together (Beasley, 2008; Kontoghiorghes, 2004; Timm & Al-Subaihi, 2001). The T observations vectors \mathbf{y}_i and the corresponding disturbances vectors \mathbf{u}_i are stacked one upon another to form the single observation vectors $\mathbf{y}'_{(T \times 1)} = (\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_m)$ and $\boldsymbol{\varepsilon}'_{(T \times 1)} = (\boldsymbol{\varepsilon}'_1, \boldsymbol{\varepsilon}'_2, \dots, \boldsymbol{\varepsilon}'_m)$, respectively. Additionally, the design matrix for the m stacked observation vectors \mathbf{y} is allowed to be the matrix $\mathbf{X}_{(T \times k)} = \bigoplus_{i=1}^m \mathbf{X}_i$, which is the direct sum of the individual design matrices \mathbf{X}_i for the m equations where $k = \sum_{i=1}^m k_i$ is the total number of parameters over the m equations. It can be defined as,

$$\bigoplus_i^m \mathbf{X}_i = \mathbf{X}_1 \oplus \dots \oplus \mathbf{X}_m = \begin{bmatrix} \mathbf{X}_1 & 0 & \dots & 0 \\ 0 & \mathbf{X}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{X}_m \end{bmatrix} \quad (2.5)$$

Then, the SURE model can be written in compact form as the following linear model,

$$\underset{T \times 1}{\mathbf{y}} = \underset{T \times k}{\mathbf{X}} \underset{k \times 1}{\boldsymbol{\beta}} + \underset{T \times 1}{\boldsymbol{\varepsilon}} \quad (2.6)$$

The vector $\boldsymbol{\beta}$ in 2.6 is constructed in a same manner as vector \mathbf{y} and $\boldsymbol{\varepsilon}$, is a stacked vector containing stacked elements the parameter vectors $\boldsymbol{\beta}_i$ defined in 2.2. Thus, equation 2.4 becomes,

$$E(\boldsymbol{\varepsilon}) = 0 \text{ and } E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \boldsymbol{\Sigma} \otimes \mathbf{I}_T = \boldsymbol{\Omega}_{Tm \times Tm} \quad (2.7)$$

where \otimes denotes a Kronecker product, \mathbf{I}_T is $T \times T$ identity matrix, and $\boldsymbol{\Sigma} = [\sigma_{ij}]$ is a $m \times m$ positive definite symmetric covariance matrix.

The specification of SURE model in econometric analysis can be distinguished through four cases (Greene, 2012):

- i. The equations are contemporaneously uncorrelated disturbances.
- ii. The equations have identical predictor but contemporaneously correlated disturbances.
- iii. The equations are contemporaneously correlated disturbances with different predictors across equations.
- iv. The equations are related and predictors in one block of equations are a subset of those in another (contemporaneously correlated disturbances, with a subset of identical predictors across the equations).

The first case is actually the classical multivariate linear regression model if it assumed that the same predictors are associated with all the response variables. Thus, the efficient and best linear unbiased estimators are given by ordinary least squares (OLS). For the second case, it is well known that the generalised least square (GLS) is more efficient since the OLS estimators are no longer efficient if the disturbances are correlated. However, in most cases, the covariance of disturbances $\boldsymbol{\Omega}$ is unknown, and hence GLS is not feasible. Therefore, Zellner (1962, 1963) proposed feasible

generalised least squares (FGLS) where Ω is replaced by a consistent estimator. The rest of the cases are considered in Srivastava and Dwivedi (1979).

The FGLS estimator of β is given by,

$$\hat{\beta}_{FGLS} = (\mathbf{X}'\hat{\Omega}^{-1}\mathbf{X})^{-1} \mathbf{X}'\hat{\Omega}^{-1}\mathbf{y} \quad (2.8)$$

where $\hat{\Omega}$ is a consistent estimator based on the residuals. Thus, the elements of matrix $\hat{\Omega}$ can be computed as follows,

$$\hat{\sigma}_{ij} = \frac{\hat{\varepsilon}_i \hat{\varepsilon}_j}{T}, \quad i, j = 1, 2, \dots, m \quad (2.9)$$

$$\hat{\varepsilon}_i = \mathbf{y}_i - \mathbf{X}_i \hat{\beta}_i, \quad i = 1, 2, \dots, m \quad (2.10)$$

where T is total observations for each of m equation. The covariance matrix of the estimated parameters is,

$$V(\hat{\beta}_{FGLS}) = (\mathbf{X}'\hat{\Omega}^{-1}\mathbf{X})^{-1} \quad (2.11)$$

2.6 SURE-PcGets Algorithm

The *PcGets* and *Autometrics* were implemented for the multiple equations model but particularly on the Vector Autoregressive (VAR) model (Doornik & Hendry, 2007; Hendry & Krolzig, 2001; Krolzig, 2001). Hendry and Krolzig (2005) also showed the possibility of implementing automatic GETS selection for simultaneous equations model but conditional to endogenous predictors as in the VAR model.

The VAR model is introduced by Sims in 1980. Since then, it is widely used in econometrics studies due to minimal theoretical demands on the structure of a model where the dynamic relations between variables is unrestricted (Greene, 2012). Moreover, the equations are constrained to be linear thus the modellers need not be concerned with the functional forms. The specification of this model requires the endogenous and exogenous variables that are believed to interact, and the largest number of lags needed to capture most of the effects that the variables have on each other (Pindyck & Rubinfeld, 1998). Thus, it indicates that both algorithms have been extended to multiple equations particularly for VAR model selection.

The reduction process of VAR within the general-to-specific modelling initiated from an unrestricted VAR with assumptions that the VAR is covariance-stationary and the variance-covariance matrix also unrestricted (Doornik & Hendry, 2007; Hendry & Krolzig, 2001). In the case of a vector system, there are two phases of reduction: (i) joint reductions of the system and (ii) reductions of the individual equations. The first reduction imposed parameters restriction which will determine whether the predictors should be excluded from all equations of the system. The process is a sequential simplification and testing procedure corresponds to the reduction process of single equation modelling, but the diagnostics are constructed to test the properties of the vector of residuals. Second phase concerned on the test of causal relationships where the result will decide whether the equations are considered as a system or separately treat as individual. Contemporaneous non-causality implies that the equations of the VAR are unrelated thus model selection procedures can be applied equation-by-equation without a loss in efficiency.

The equations in the VAR are only seemingly unrelated to each other if the predictors found to be weakly exogenous (Krolzig, 2001). Thus eliminating a variable in one equation will affect the others which indicate that single equation model selection is inefficient. In the current form, *Autometrics* does not offer application for this condition. However, Ismail (2005) had made an attempt for this application by introduced *SURE-PcGets* algorithm as an extended version of *PcGets*.

The GETS algorithm has been applied to multiple equations model such as seemingly unrelated regression equations (SURE) model known as *SURE-PcGets* (Ismail, 2005). The basic concept of this algorithm is combining the selection stages in *PcGets* with SURE models. The *SURE-PcGets* selection steps are similar to those of *PcGets*, but the differences are in the first step where the GUM is formulated and there is an extra step, i.e. testing for contemporaneous correlation disturbances.

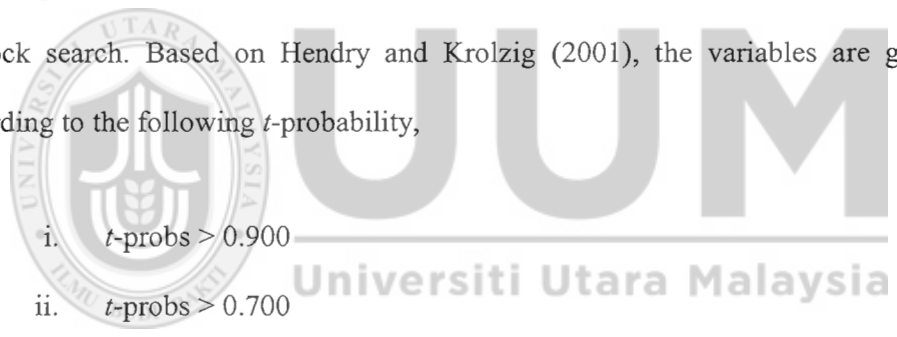
Designing the GUM in the first step is very crucial in determining the success of *SURE-PcGets*, because the steps that follow in the simplification process are done automatically. A poor general framework is unlikely to lead to a good final model choice. The formulation of the GUM based on subject matter theory, institutional knowledge, historical contingencies, data availability and measurement information, ensuring the resulting model encompasses previous evidence, with a relatively orthogonal parameterisation of the k candidate predictors.

The diagnostic tests in the second step play an important role in identifying congruence model since SURE model can be used when the classical assumptions of linear regression are satisfied (Srivastava & Maekawa, 1995). In order to conduct the

test, each equation is estimated using OLS. If the test is not fulfilled in one or more of the equation, the GUM needs to be reformulated.

Once the GUM is congruent, pre-search testing of lags, top down (removed groups of irrelevant variables) and bottom up (include groups of relevant variables) are then conducted where groups of predictors in each equation are tested and removed in the order of their absolute t -values. Then, the simplified model obtain for each equation are estimated using FGLS.

At the next step of *SURE-PcGets*, the models are simplified using the multiple search paths implemented in *PcGets* where the removal of variables are determined through a block search. Based on Hendry and Krolzig (2001), the variables are grouped according to the following t -probability,

- 
- i. t -probs $>$ 0.900
 - ii. t -probs $>$ 0.700
 - iii. t -probs $>$ 0.500
 - iv. t -probs $>$ 0.250
 - v. t -probs $>$ 0.100
 - vi. t -probs $>$ 0.050
 - vii. t -probs $>$ 0.010
 - viii. t -probs $>$ 0.001

If the model is congruent after removing this group, then a terminal model is found. If more than one terminal model is found after the block search, then encompassing testing starts.

A model encompasses another if it contains all the information conveyed by another model. Encompassing tests select between candidate congruent models at the end of each path search procedure. Each contender is tested against the union. If a unique model results, it is selected; otherwise, if some are rejected, *SURE-PcGets* forms the union of the remaining models. That union then constitutes a new starting point and the path-searching algorithm repeats until the union remain unchanged between successive rounds. If more than one model survives this test, the final model is selected based on information criteria.

The simplified model is tested for sub-sample reliability, which helps to identify ‘spuriously significant’ predictors. The sample is split into two (possibly overlapping) sub-samples. The sub-samples are the first 75% (denoted Split 1) and the last 75% (denoted Split 2) of the full sample. The two sub-samples overlap in the middle 50% of the full sample. Once the sub-samples are identified, FGLS is used to estimate the model for full sample, and the two sub-samples.

Last but not least, a Monte Carlo-quasi likelihood ratio test (MC-QLR) was conducted to test the correlation in the errors across the equations. *SURE-PcGets* algorithm was developed based on the selection steps described above. An overview of the algorithm is shown in Appendix B.

CHAPTER THREE

SURE-AUTOMETRICS ALGORITHM

3.1 Framework of Algorithm Development

Achievement of *SURE-PcGets* motivates this study to develop an algorithm of model selection for multiple equations model. Similarly to *PcGets*, *Autometrics* algorithm also possesses the model selection properties within GETS approach. Hence, the name of the new algorithm is *SURE-Autometrics* since it implements the selection strategies in *Autometrics*.

Basically, the *SURE-Autometrics* is developed in five phases. The search of the ‘best’ SURE model starts with the initial formulation of general unrestricted model in the first phase. Second phase aims to reduce the complexities in the initial model by concentrate at reduction of lagged variables. Two strategies are used in removing the highest insignificant lagged variables where each will end with one reduced model. Thus, a new general model is formulated based on the encompassing test results. Third phase involves the implementation of tree search procedure where the variables in each equation are reduced to simplify the model. Fourth phase intends to achieve more reduced model. This phase and the previous phase are iterated whenever it is possible to create another new general model that is less complex than the previous general model. Since there are many possible specific models survived the reduction processes, the final phase chooses one of them to be the ‘best’ simplified models using information criterion. This development framework is summarised in Figure 3.1. The details of each procedure are described in Section 3.2 to 3.6.

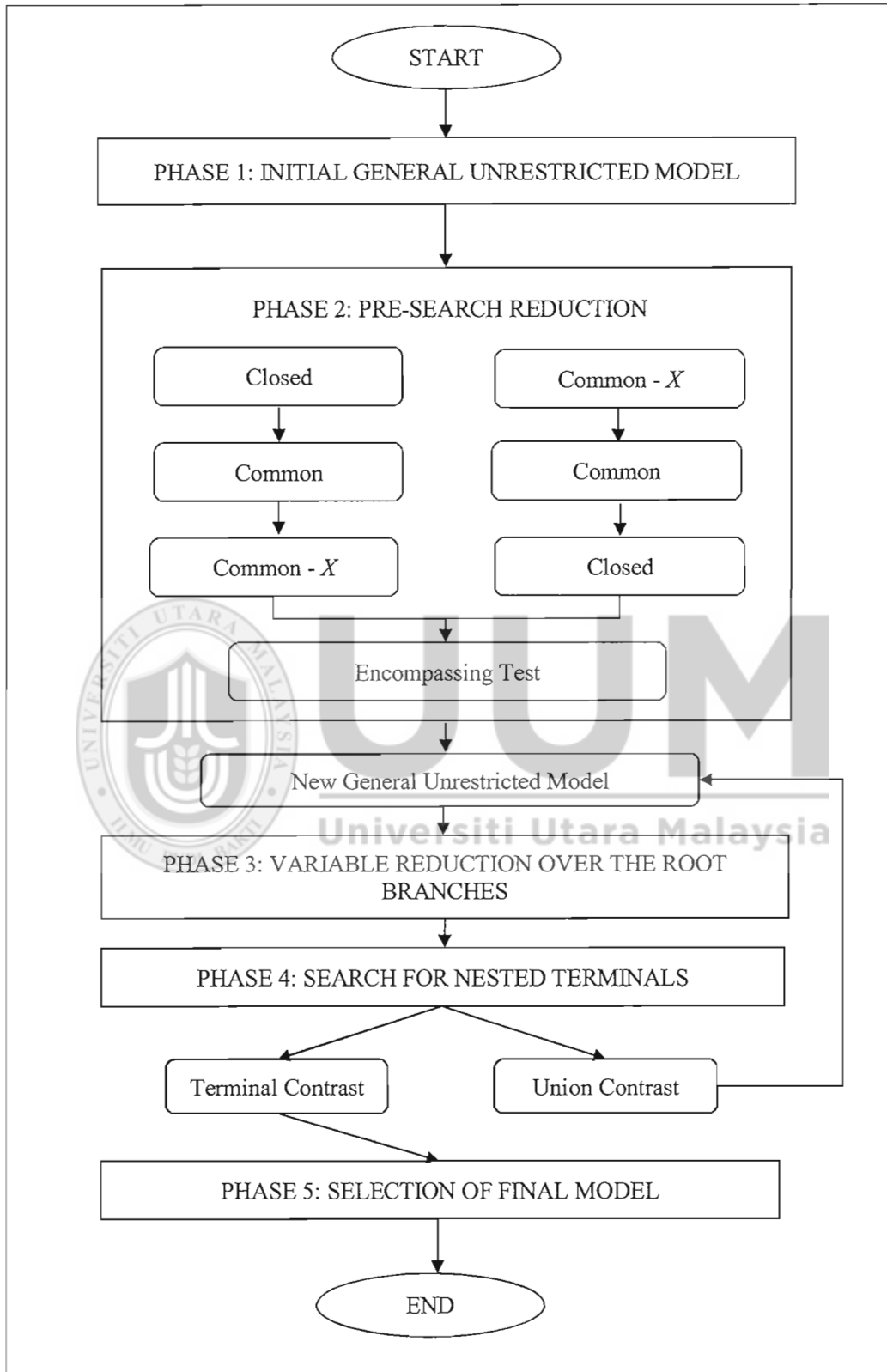


Figure 3.1. SURE-Autometrics Development Framework

3.2 Phase 1: Initial General Unrestricted Model (GUMS)

The *SURE-Autometrics* algorithm is developed within the GETS modelling approach. Thus, the model selection for SURE begins with an initial formulation of general unrestricted model which is denoted as GUMS. The addition of 'S' in the acronym 'GUM' is to reflect the multiple equations in the model, unlike the 'GUM' in *Autometrics*. The specification of each equation involves all the potential relevant variables including their lags which have been reviewed from previous theoretical and empirical findings. It starts off by setting up the main significance level for selection at $p_a = 5\%$ or $p_a = 1\%$. In order to ensure the congruency of the GUMS, each equation is run through a series of diagnostic tests. Details of these tests are explained in Section 3.2.1. Then, the model is tested for dependency of disturbances amongst equations. Section 3.2.2 describes more on this test. Finally, the initial GUMS is estimated using the feasible generalised least squares (FGLS) method.

3.2.1 Series of Diagnostic Tests

Model adequacy checking analysis comprises a similar series of diagnostic tests as in *Autometrics*. In this part, these tests are implemented as a signal to modellers about the congruency of the formulated initial GUMS. If there is any of the tests fail, then it is up to the modellers to reformulate or continue with the model selection. The series are again used throughout the reduction process to ensure the simplified models are also congruent.

All the equations are tested separately. Hence, this part involves the OLS residuals of each equation and the significance level is $p_d = 0.01$. The tests are checking the assumption of a normal distribution errors, parameter constancy, autocorrelation

errors, unconditional and conditional homoscedasticity. Descriptions of these tests are briefly explained in the following sub-sections.

3.2.1.1 Test of normal distribution

The test is for checking whether the skewness and kurtosis of the residuals correspond to the assumption of a normal distribution where the hypotheses are,

H_0 : Errors are normally distributed

The skewness and kurtosis are defined respectively as,

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2} \quad (3.1)$$

where μ denote the mean and $\mu_i = E[x^i - \mu]^i$. σ^2 is the variance of variable x_t , so that $\sigma^2 = \mu^2$. A normal variate will have $\sqrt{\beta_1} = 0$ and $\beta_2 = 3$.

The sample estimates of these parameters are given by,

$$\bar{x} = \frac{1}{T} \sum_{t=1}^T x_t, \quad m_1 = \frac{1}{T} \sum_{t=1}^T (x_t - \bar{x})^i, \quad \sqrt{b_1} = \frac{m_3}{m_2^{3/2}} \quad \text{and} \quad b_2 = \frac{m_4}{m_2^2} \quad (3.2)$$

where T is the number of observations. Based on Doornik and Hansen (1994), the test statistic is,

$$e_2 = z_1^2 + z_2^2 \sim \chi^2(2) \quad (3.3)$$

where z_1 and z_2 represent the transformed of skewness and kurtosis which are defined as,

$$z_1 = \sqrt{\frac{T}{6}}\beta_1 \quad \text{and} \quad z_2 = \sqrt{\frac{T}{24}}\beta_2 \quad (3.4)$$

3.2.1.2 Test of parameters constancy

Chow predictive test is used to examine whether the coefficients in two models on different data sets are equal. The equation is divided into two where the first is based on the whole sample, T and the second is the sub-sample, T_1 determined at 70% breakpoint. The null hypothesis is,

H_0 : The coefficients are identical in both equations

The test statistic has the following form,

$$\eta_3 = \frac{RSS_T - RSS_{T_1} / (T - T_1)}{RSS_{T_1} / (T_1 - k)} \sim F(T - T_1, T_1 - k) \quad (3.5)$$

where k is the number of parameters in the equation, T_1 is number of observations in sub-sample, RSS_T is the residual sum of squares for the whole sample, and RSS_{T_1} is the residual sum of squares for the sub-sample.

3.2.1.3 Test of autocorrelation

The test is an auxiliary regression of residuals on all the predictors of the original equation and the lagged residuals for lags p to r where missing residuals are set to zero. The hypothesis is,

$H_0: \rho = 0$ (there is no autocorrelation i.e. errors are white noise)

Hence the statistic is,

$$TR^2 \sim \chi^2(r-p+1) \quad (3.6)$$

which equivalent to,

$$\frac{R^2}{1-R^2} \cdot \frac{T-k-r+p-1}{r-p+1} \sim F(r-p+1, T-k-r+p-1) \quad (3.7)$$

where p is the minimum lag number, r is the maximum lag number and R^2 is the goodness of fit measure. All results from auxiliary regression in *Autometrics* come in the form of F distribution (Doornik & Hendry, 2007, p. 216). Thus, *SURE-Autometrics* follows this setting including test for unconditional and conditional homoscedasticity.

3.2.1.4 Test of unconditional homoscedasticity

The test is checking for assumption of constant error variance or homoscedasticity. The condition of unequal variances is known as heteroscedasticity. In *Autometrics*, the test also involves an auxiliary regression of squared residuals on the original predictors, x_{it} and all their squares, x_{it}^2 . The null hypothesis is,

$$H_0: \text{Var}(\varepsilon_t) = \sigma^2 \text{ is constant (unconditional homoscedasticity)}$$

The statistic is based on White (1980),

$$\frac{R^2}{1-R^2} \cdot \frac{T-k-q}{q} \sim F(q, T-k-q) \quad (3.8)$$

where q is number of original regressors and all their cross-product.

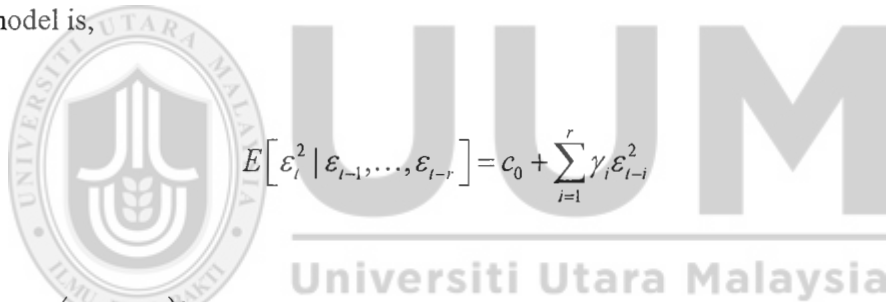
3.2.1.5 Test of conditional homoscedasticity

This is the autoregressive conditional heteroscedasticity (ARCH) test (Engle, 1982)

which involves the following hypothesis:

$$H_0: \gamma = 0 \text{ (conditional homoscedasticity)}$$

The model is,



$$E[\varepsilon_t^2 | \varepsilon_{t-1}, \dots, \varepsilon_{t-r}] = c_0 + \sum_{i=1}^r \gamma_i \varepsilon_{t-i}^2 \quad (3.9)$$

where $\gamma = (\gamma_1, \dots, \gamma_r)'$ and r is the maximum lag number. Again, the F form is,

$$\frac{R^2}{1-R^2} \cdot \frac{T-k-r}{r} \sim F(r, T-k-r) \quad (3.10)$$

from the regression of $\hat{\varepsilon}_t^2$ on a constant and $\hat{\varepsilon}_{t-1}^2$ up to $\hat{\varepsilon}_{t-r}^2$.

3.2.2 Test of Correlation Disturbances amongst Equations

Basically, the decision whether to apply FGLS or OLS on SURE model is relies on the covariance of disturbances. Both estimators are equivalent if the disturbances between equations are uncorrelated i.e. independence. Explicitly,

$$\sigma_{ij} = 0 \quad \text{for } \forall i \neq j \quad (3.11)$$

Under the null hypothesis, the disturbances in multiple equations are contemporaneously uncorrelated where the covariance will be a diagonal matrix i.e.,

$$H_0 : \sum_{\varepsilon} = D_m(\sigma_i^2) = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_m^2) \quad (3.12)$$

Initially, this study has two choices of dependence test of the correlation disturbances. One is the Monte Carlo–quasi likelihood ratio (MC-QLR) (Dufour & Khalaf, 2002) test, and the other is the multivariate independence (MI) (Tsay, 2004) test. The MC-QLR test is able to control the probability of type I errors while the MI test is specially developed to control the serial dependence of disturbances. Yusof and Ismail (2011) compared these tests using data from Fisher (1993) in order to choose the one that is fit in the algorithm. The findings suggested that MC-QLR is more appropriate because MI test produced inconsistent results. Furthermore, the *SURE-Autometrics* already has the diagnostic test for autocorrelation.

Therefore, *SURE-Autometrics* only considers the MC-QLR to indicate the dependencies of correlation disturbances amongst the equations in SURE model. Moreover, this test also was implemented in *SURE-PcGets* algorithm. The significance level is setup at 0.10. The value is quite high as compared to 0.01 in the diagnostics testing series because the efficiency of FGLS estimator not only depends of the correlation disturbances, but also relies on the number of equations and the sample sizes (Srivastava & Maekawa, 1995; Timm, 2002). The following sections describe both tests in details.

3.2.2.1 Monte Carlo–quasi likelihood ratio test

The Monte Carlo (MC) test allows us to obtain provably exact randomized tests in finite samples using very small numbers of MC replications of the original test statistic under the null hypothesis. Dufour and Khalaf (2002) showed that the performance of Monte-Carlo-quasi likelihood ratio (MC-QLR) was outstanding in testing the dependency of correlation disturbances amongst equations in SURE model. Thus, this study employs the MC test with 999 replications of quasi-likelihood ratio (QLR) test statistic. The statistic is based on FGLS estimators where the OLS residuals are used to estimate the disturbance covariance matrix. The test statistic is,

$$QLR = T \ln \left(\frac{|D_m(\hat{\sigma}_i^2)|}{|\hat{\Sigma}^{(h)}|} \right), \quad i = 1, 2, \dots, m; \quad h = 1, 2, \dots \quad (3.13)$$

where $D_m(\hat{\sigma}_i^2)$ is the diagonal matrix whose diagonal elements are $\hat{\sigma}_1^2, \dots, \hat{\sigma}_m^2$, m is the number of equations, T is the number of observations and $\hat{\Sigma}^{(h)}$ is the partially iterated estimators of disturbances covariance matrix.

If QLR is large, then H_0 as in equation 3.12 will be rejected. To compute the p -value for the MC-QLR test, the survival function under H_0 is denoted by $G(x) = P[QLR \geq x]$. Suppose QLR_0 be the test statistic computed from the observed data. Then the associated critical region of size α is expressed as $G(QLR_0) \leq \alpha$. The MC methods is used to generate N independent realizations $QLR_1, QLR_2, \dots, QLR_N$ of QLR under H_0 and computes the randomized p -value as follows,

$$\hat{p}_N(x) = \frac{N\hat{G}_N(x) + 1}{N + 1} \quad (3.14)$$

where

$$\hat{G}_N(x) = \frac{1}{N} \sum_{i=1}^N I_{[0,\infty)}(QLR_i - x), \quad I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} \quad (3.15)$$

This test requires the variables in the model to be fixed or strictly exogenous and independent of disturbances. Additionally, it also assumes that the disturbances of each equation are independent and identically normal distribution.

3.2.2.2 Multivariate independence test

As compared to previous test, multivariate independence (MI) test is an asymptotic test proposed by Tsay (2004). This test is considered in the development of *SURE-Autometrics* since it allows the disturbances to be weakly dependent processes. Hence, in addition to assumptions in MC-QLR, this test assumes that the variables are dependent observations. This assumption is useful for application using time series data. From the OLS residuals ($\hat{\varepsilon}_{it}$) of m regression model, the statistic is defined as,

$$MI = \hat{\Lambda}' \hat{\Omega}^{-1} \hat{\Lambda} \quad (3.16)$$

where

$$\hat{\Lambda} = \frac{1}{\sqrt{T}} \sum_{t=1}^T \hat{Z}_t \quad \hat{Z}_{ij,t} = \hat{\varepsilon}_{i,t} \hat{\varepsilon}_{j,t} \quad (3.17)$$

$$\hat{\Omega} = \begin{bmatrix} \hat{\Omega}_{12} & 0 & \cdots & 0 \\ 0 & \hat{\Omega}_{13} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\Omega}_{m(m-1)} \end{bmatrix} \quad (3.18)$$

$\hat{\Omega}$ is a $\frac{m(m-1)}{2} \times \frac{m(m-1)}{2}$ matrix such that,

$$\hat{\Omega}_{ij} = \sum_{h=1-T}^{T-1} C_{i,h} C_{j,h}, \quad C_{i,h} = \frac{1}{T} \sum_{t(h)} e_{i,t} e_{i,t+h}, \quad C_{j,h} = \frac{1}{T} \sum_{t(h)} e_{j,t} e_{j,t+h} \quad (3.19)$$

Here, $\sum_{t(h)}$ is a sum over $1 \leq t, t+h \leq T$. The MI statistic has χ^2 distribution with

$m(m-1)/2$ degrees of freedom.

3.3 Phase 2: Pre-search Reduction

The second phase is known as the pre-search reduction aims to remove a group of variables from the initial GUMS at loose significance levels. It is defined as follows,

$$p_p = \frac{5p_a^{0.8}}{1+4p_a^{0.8}} \quad (3.20)$$

where p_a is the main significance level (0.01 or 0.05). The value of pre-test significance value, p_p are 0.1141 and 0.3337, respectively. If only one variable need to be removed, the pre-test significance value is given by,

$$p_{p,1} = \max \left\{ \frac{1}{2} p_a^{1/2}, p_a^{3/4} \right\} \quad (3.21)$$

The reduction procedures will be focused on the lagged variables only. There are three types of reduction that will be executed in two strategies. The types are,

1. Closed lag reduction to remove the group of variable with the largest lag downwards and stop when the removal is failed.
2. Common lag reduction to remove the group of variable with the highest insignificant value.
3. Common- X lag reduction. It has similar removal procedure as in common lag except that it excludes the lagged dependent variable (Y).

Details of these procedures are explained in Sections 3.3.1 to 3.3.3. Sequentially, each equation attempts for the lag variable reduction while estimating as a multiple equations using FGLS method. Meanwhile, the two strategies are,

1. Starts with closed lag, and then followed by common lag and ends with common- X lag. The reduced model is denoted as Model 1 (M1).
2. Starts with common- X lag, and then followed by common lag and ends with closed lag. The reduced model is denoted as Model 2 (M2).

Both resulted models are tested using encompassing test to select either one or the union of both models to be the current GUMS. The encompassing test is described in Section 3.3.4.

3.3.1 Closed Lag

In this procedure, the removal of group lag variable is based on the lag number starting from the highest value. The process begins with identification of the group according to their lag number. Suppose q is the highest lag in each equation. Then, the variables with lag q are grouped and k_p is denoted as the number of variables involved. The removal of this group is based on four conditions.

The first condition concentrates on the p -value of each variable in the group. The individual p -values are checked against a specified significance value which determined by the following,

$$p_{p,1}^*(k_p) = \max \left\{ 1 - (1 - p_{p,1})^{k_p}, 0.2p_a \right\} \quad (3.22)$$

The $p_{p,1}^*$ value will change according to the number of variables involved (k_p). The possible value with respect to the correspond k_p are shown in Table 3.1.

Table 3.1
Pre-test Significance Values for $k_p = 1, 2, \dots, 5$

| Main significance level, p_a | Pre-test significance level, p_p | $p_{p,1}$ | No. of variables, k_p | $p_{p,1}^*(k_p)$ |
|--------------------------------|------------------------------------|-----------|-------------------------|------------------|
| 0.01 | 0.1141 | 0.0500 | 1 | 0.0500 |
| | | | 2 | 0.0975 |
| | | | 3 | 0.1426 |
| | | | 4 | 0.1855 |
| | | | 5 | 0.2262 |
| 0.05 | 0.3337 | 0.1118 | 1 | 0.1118 |
| | | | 2 | 0.2111 |
| | | | 3 | 0.2993 |
| | | | 4 | 0.3776 |
| | | | 5 | 0.4472 |

If all the individual p -values are found to be larger than $p_{p,l}^*$, then the group of variable is removed from the equation. For instance, if the main significance level is setup at 0.05 and there are four variables with the highest lag, then their individual p -values should be more than 0.3776 so that all these variables are able to be removed from the equation.

The second condition focuses on the model after the removal of the group variables for each equation. The reduced model is estimated using FGLS and the joint p -value obtained through F test is compared against the model before reduction. If the value is greater than the pre-test significance value, p_p , then the process is continued.

The third condition is similar to the second condition but the joint p -value is compared against the initial GUMS. Both tests are important to make sure that the removal of variables is a valid reduction from the general model.

The fourth condition inspects whether the removal of variables affect the congruency of model. Each of equation is tested through a series of diagnostic tests. All the variables in the group are returned back to the equation if any of these conditions has failed, and the process of reduction is stopped. Otherwise, the removal will continue with group of lag $(q-1)$ until lag of one.

3.3.2 Common Lag

In this part of reduction, the lag variables are grouped according to their lag number, and their joint significance are calculated and arranged in descending order. The process of reduction starts with the group of lag with the highest insignificance value.

The procedure starts with the collection of all predictors at lag one until q in each equation. Then, the group is removed starting from the highest lag, q . The reduced model is estimated using FGLS and test against current model (model before reduction) to determine their joint p -values. The processes of finding these values are repeated for all lag number.

The reduction starts with the group of lag that has the highest insignificant joint p -value. The decision whether it is able to be removed or not is similar to all the conditions described in previous section.

3.3.3 Common- X Lag

The reduction procedure is similar to common lag procedure except that lag of Y is excluded from the process.

3.3.4 Encompassing Test

A model is said to encompass another if it contains all the information conveyed by another model (Hoover & Perez, 1999). Hence, the aim of this test is to determine which model resulted from the two strategies (M1 or M2) is better than the other. Each model is tested against the union of both models.

Suppose M1 has E_{11} as the first equation with k_1+k_2 predictors (x_{1t}, x_{2t}), whereas E_{12} is the first equation from M2 has k_2+k_3 (x_{2t}, x_{3t}). Thus, x_{2t} is in common and their union, E_{1U} comprises of $k = k_1+k_2+k_3$ a non-redundant set for x_{1t}, x_{2t} , and x_{3t} . Meanwhile RSS_{11} , RSS_{12} , and RSS_{1U} denote the residual sum of squares from E_{11} , E_{12} , and E_{1U} respectively.

The hypotheses are,

$$H_0: E_{11} \in_p E_{12} \text{ (} E_{11} \text{ parsimonious encompassing } E_{12}\text{)}$$

$$H_1: E_{11} \text{ does not encompass}$$

The test statistic is,

$$\eta_4 = \frac{(RSS_{11} - RSS_{1U})/k_3}{RSS_{1U}/(T - k)} \sim F_{(k_3, T-k)} \quad (3.23)$$

where k_3 is the number of non-redundant predictors for E_{11} and E_{1U} . The tests are conducted at main significance level (p_a) for each of equation in the model. The result could either be only one of the equations or the union of both equations. Then, the equations resulted from this test becomes the first equation in the reduced GUMS. This new GUMS will be tested for adequacy for each equation and determined for dependency of correlation disturbances amongst the equations.

3.4 Phase 3: Variable Reduction over Root Branches

This phase contains the main procedure in *SURE-Autometrics* which is known as tree search. The tree is meant to discover all the unique models generated from the variables in the GUMS.

The reduction procedure in this phase starts with an attempt of removing all the variables in the model. If it failed, then the current GUMS is denoted as GUMS 0. Reduction from root branches means that the removal of variables begins with the highest insignificance variable in each equation in the model. Then the process of

reduction continues by implementing three principles which are pruning, bunching, and chopping. These principles are described in Section 3.4.1.

Similar to previous phase, the processes of reduction will be done sequentially amongst the equation. The FGLS method is employed whenever estimation is needed. The model that cannot be reduced anymore is known as terminal model. The terminal has to go through the series of diagnostic test as described in Section 3.2.1. Indication of any failure due to significance level, $p_d = 0.01$ will let the level to be reduced at 0.005.

3.4.1 The reduction principles

The aim of employing the tree search is to find all the possible models according to variables in the GUMS. However, it would be computationally inefficient to find all the possibility of variables combinations. Hence, the reduction principles are important to advance the search in a systematic way so that unnecessary path could be skipped so that only unique models are obtained. Following *Autometrics*, the principles are pruning, bunching and chopping.

3.4.1.1 Pruning

In this principle, the subsequent path are pruned or ignored whenever a removal of one variable is failed due to either p -value is greater than p_a , or violation in any of diagnostic tests.

3.4.1.2 Bunching

Instead of removing a variable one by one, this principle allows the removal of a group or bunch of variables in a single step. The significance test of the bunch is done at p_a while the p -value p_b determine the amount of bunching. The variables are grouped according to their individual insignificance as long as their smallest p -value in the bunch is larger than $p_b^*(k_b)$ which given by,

$$p_b^*(k_b) = p_b^{1/2} \left[1 - \left(1 - p_b^{1/2} \right)^{k_b} \right] \quad (3.24)$$

$$p_b = \max \left\{ \frac{1}{2} p_a^{1/2}, p_a^{3/4} \right\} \quad (3.25)$$

where k_b is the size of bunch. If removal fails, the bunch is shrunk until size of one. Table 3.2 shows the significance values used in this principle with respect to the size or number of variables involved in the group.

Table 3.2

Significance Values using Bunching Principle for $k_b = 1, 2, \dots, 5$

| Main significance level, p_a | p_b | k_b | $p_b^*(k_b)$ |
|--------------------------------|--------|-------|--------------|
| 0.01 | 0.0500 | 1 | 0.0500 |
| | | 2 | 0.0888 |
| | | 3 | 0.1190 |
| | | 4 | 0.1424 |
| | | 5 | 0.1605 |
| 0.05 | 0.1118 | 1 | 0.1118 |
| | | 2 | 0.1862 |
| | | 3 | 0.2358 |
| | | 4 | 0.2687 |
| | | 5 | 0.2907 |

For instance, at 5% level of significance, a bunch of four variables are removed if their individual p -value is larger than 0.2687. The group removal fails when at least one of individual p -value is smaller than 0.2687. Then the size of bunch is reduced to three variables where individual p -value must be greater than 0.2358 in order to be removed from the equation. The process of reducing the size of bunch is continued to make sure that at least one variable can be removed by using this principle.

3.4.1.3 Chopping

Reduction of variables through this principle is a way of removing one or more variables permanently from the search procedure if it is highly insignificant. The reduction is determined by the specified significance value p_c which is defined as,

$$p_c = p_b \quad (3.26)$$

3.5 Phase 4: Search for Nested Terminals

This phase focuses on finding further terminals using model contrast technique. Since the tree is uniquely ordered, then findings a similar terminal is already void. This phase however aims to determine different terminal models through the deletion of minimal bunch along the current path. Unlike previous phases, the search concentrates more on the removal of variables where the interest is to find variables that should be in the equation within the GUMS. Two strategies are employed which are union and terminal contrast.

1. Union contrast determines the contrasting bunch with respect to the union of the current set of terminal models. This type of contrast is used while current GUMS still changes between iterations.
2. Terminal contrast determines the smallest bunch that would yield a model that is different from any of the current terminals. This mode is used at the end when the current GUMS is fixed.

3.6 Phase 5: Selection of the Final Model

The procedures in Phases 3 and 4 will produce more than one terminal model. These terminals are valid reduction from the GUMS when variables are significant, each equation passed every diagnostic test and disturbances are contemporaneously correlated amongst equations. As an algorithm aims for the automatically model selection procedure, only one terminal will be selected as the ‘best’ model. The chosen terminal is the final model and known as a specific unrestricted model (SUMS) since the algorithm initiated from a very general model.

The selection is based on information criteria. Following *SURE-PcGets*, three information criteria are considered i.e. Akaike criterion (AIC), Hannan-Quinn criterion (HQ), and Schwartz criterion (SC). These criteria respectively defined as,

$$AIC = \ln \tilde{\sigma}^2 + \frac{2k}{T} \quad (3.27)$$

$$HQ = \ln \tilde{\sigma}^2 + \frac{2k [\ln(\ln T)]}{T} \quad (3.28)$$

$$SC = \ln \tilde{\sigma}^2 + \frac{k(\ln T)}{T} \quad (3.29)$$

where the maximum likelihood estimate of $\tilde{\sigma}^2$ is given by,

$$\tilde{\sigma}^2 = \frac{1}{T} \sum_{i=1}^T \hat{\varepsilon}_i^2 \quad (3.30)$$

with $\hat{\varepsilon}_i$ is the disturbances, T is the number of observations and k is the number of predictors. All these criteria are calculated for each of equations and averaged across the terminals. However, only SC with the smallest average value is used to select the terminal that will be the SUMS. The criterion was chosen because it could lead to a consistent model selection (Hendry & Krolzig, 2001; Judge, Hill, Griffiths, Lütkepohl, & Lee, 1988). Moreover, it is also used by the *Autometrics*.

Details of the *SURE-Autometrics* algorithm are presented at Appendix C. Subsequently, the algorithm is transformed into a computer program via *GAUSS* (Version 9.0) programming language. This algorithm is suitable for SURE model since it simultaneously selects all the equations while using the FGLS to maintain the efficiency of the estimators. The estimation method used in *Autometrics* ignores the correlation of disturbances amongst the equations in SURE model. Hence, the standard error of selected model is large because the selection are done separately equation by equation.

CHAPTER FOUR

SIMULATION ASSESSMENT OF *SURE-AUTOMETRICS*

4.1 Experimental Frames

Upon completion of the computer programme for the *SURE-Autometrics* algorithm, its performances have to be assessed using various simulation experimental conditions. The goal is to measure the ability of algorithm in finding the true SURE model specification when data-generating process (DGP) is known. Irrelevant variables are added in the true models to form the general unrestricted model (GUMS) which will be simplified using the algorithm. The algorithm is well performed if it would be able to remove the irrelevant variables during the model selection processes.

The simulation experiments involve 100 replications of *SURE-Autometrics* simplifying the initial GUMS based on several conditions. Most of the experimental conditions considered in this study were adopted from the evaluation of *SURE-PcGets* (Ismail, 2005) and *Autometrics* (Doornik, 2009). The analyses start by generating artificial dependent variables according to five models specification (S1, S2, S3, S4, S5), three strengths of correlation disturbances amongst equations ($\rho = 0.9, 0.6, 0.2$) and two sample sizes ($n = 146, 73$). The number of specifications which different in terms of number of variables are based on Doornik (2009), while the other two conditions were chosen according to Ismail (2005). Then, the *SURE-Autometrics* is run at two significance levels ($\alpha = 0.05, 0.01$) to simplify two sets of initial GUMS ($k = 18, 39$). Since the evaluation of *SURE-PcGets* (Ismail, 2005) only involved with equations model, this study varied the multiple equations number by assessing the simulation on model of two, four and six equations ($m = 2, 4, 6$). Hence, the

performances of algorithm are measured for model contains small to large number of equations.

4.1.1 Artificial Dependent Variables

In simulation study of *SURE-Autometrics* performances, the artificial dependent variables (y_{it}) are generated according to true specification models. Ismail (2005) used three specifications of six equations model to assess the performances of *SURE-PcGets* as shown in Table 4.1.

Table 4.1

True Specification Models used by SURE-PcGets

| | | |
|-----|---|---|
| HP1 | $y_{1t} = 0.017 + 0.026\varepsilon_{1t}$ $y_{2t} = 0.025 + 0.032\varepsilon_{2t}$ $y_{3t} = 0.018 + 0.028\varepsilon_{3t}$ | $y_{4t} = 0.043 + 0.061\varepsilon_{4t}$ $y_{5t} = 0.023 + 0.029\varepsilon_{5t}$ $y_{6t} = 0.018 + 0.024\varepsilon_{6t}$ |
| HP2 | $y_{1t} = 0.005 + 0.7092y_{1t-1} + 0.026\varepsilon_{1t}$ $y_{2t} = 0.010 + 0.5924y_{2t-1} + 0.032\varepsilon_{2t}$ $y_{3t} = 0.007 + 0.6295y_{3t-1} + 0.028\varepsilon_{3t}$ | $y_{4t} = 0.022 + 0.4963y_{4t-1} + 0.061\varepsilon_{4t}$ $y_{5t} = 0.009 + 0.6218y_{5t-1} + 0.029\varepsilon_{5t}$ $y_{6t} = 0.006 + 0.6896y_{6t-1} + 0.024\varepsilon_{6t}$ |
| HP7 | $y_{1t} = 0.005 + 0.7156y_{1t-1} + 0.2960x_{21t} + 0.0627x_{21t-1} + 0.014\varepsilon_{1t}$ $y_{2t} = 0.010 + 0.5946y_{2t-1} + 0.0379x_{22t} - 0.009x_{22t-1} + 0.010\varepsilon_{2t}$ $y_{3t} = 0.005 + 0.7074y_{3t-1} + 0.4093x_{23t} + 0.0532x_{23t-1} + 0.028\varepsilon_{3t}$ $y_{4t} = 0.018 + 0.5301y_{4t-1} + 0.6493x_{24t} - 0.4794x_{24t-1} + 0.014\varepsilon_{4t}$ $y_{5t} = 0.008 + 0.6555y_{5t-1} + 0.3724x_{25t} + 0.0630x_{25t-1} + 0.002\varepsilon_{5t}$ $y_{6t} = 0.005 + 0.7299y_{6t-1} + 0.2993x_{26t} + 0.0372x_{26t-1} + 0.019\varepsilon_{6t}$ | |

Note. Adapted from “*Algorithmic Approaches to Multiple Time Series Forecasting* (Doctoral dissertation),” by S. Ismail, 2005, University of Lancaster, Lancaster.

Originally, there were nine specifications initiated by Hoover and Perez (1999) denoted HP1 to HP9. However only three were chosen in the evaluation of *PcGets* (Hendry & Krolzig, 1999). Since both studies have focused on the single equation model, Ismail (2005) revised the three specifications to fit the six equations of SURE model. Based on the table, HP1 does not has any variable while HP2 contains lag one of dependent variable, and HP7 has three variables which are lag one of dependent

variable, one independent variable including the lag one of independent variable. Basically, the inclusions of variables in the specifications are instigated through Hendry and Krolzig (1999) and the coefficients' values are estimated by FGLS using a real data.

Meanwhile the simulation assessment of *Autometrics* (Doornik, 2009) also relied on *PcGets* with additional of HP8 and HP9 from Hoover and Perez (1999). Both added models have one lag of the dependent variable. HP8 is different from HP7 in terms of independent variable that was included, whereas HP9 contains both independent variables considered in HP7 and HP8 including their lag of one. Sharing a similar evaluation approach with Ismail (2005), five models are specified in this study according to two, four and six equations model as shown in Table 4.2, 4.3 and 4.4, respectively.

Table 4.2
True Specification Models of Two Equations (m = 2) by SURE-Autometrics

| | |
|----|--|
| S1 | $y_{1t} = 0.0230 + 0.0293\varepsilon_{1t}$ $y_{2t} = 0.0182 + 0.0240\varepsilon_{2t}$ |
| S2 | $y_{1t} = 0.0087 + 0.6170y_{1t-1} + 0.0229\varepsilon_{1t}$ $y_{2t} = 0.0058 + 0.6825y_{2t-1} + 0.0173\varepsilon_{2t}$ |
| S3 | $y_{1t} = 0.0078 + 0.6340y_{1t-1} + 0.3685x_{2t} - 0.3020x_{2t-1} + 0.0201\varepsilon_{1t}$ $y_{2t} = 0.0060 + 0.6915y_{2t-1} + 0.2811x_{2t} - 0.2224x_{2t-1} + 0.0151\varepsilon_{2t}$ |
| S4 | $y_{1t} = 0.0049 + 0.5966y_{1t-1} + 0.4820x_{4t} - 0.2072x_{4t-1} + 0.0221\varepsilon_{1t}$ $y_{2t} = 0.0028 + 0.6517y_{2t-1} + 0.1273x_{4t} + 0.1053x_{4t-1} + 0.0171\varepsilon_{2t}$ |
| S5 | $y_{1t} = 0.0049 + 0.6154y_{1t-1} + 0.3376x_{2t} - 0.2881x_{2t-1} + 0.3429x_{4t} - 0.1237x_{4t-1} + 0.0197\varepsilon_{1t}$ $y_{2t} = 0.0038 + 0.6720y_{2t-1} + 0.2742x_{2t} - 0.2268x_{2t-1} + 0.0278x_{4t} + 0.1377x_{4t-1} + 0.0149\varepsilon_{2t}$ |

The variables inclusion were based on HP1, HP2, HP7, HP8 and HP9 where each was rename with S1, S2, S3, S4 and S5, respectively. The coefficients' values in the models were obtained by using a real data through FGLS estimation method. The data

represents the annual number of air passengers between United Kingdom (UK) and six different countries from 1961 to 1997 which was acquired from the study of *SURE-PcGets* (Ismail, 2005). The countries are Germany, Sweden, Italy, Japan, United States (US), and Canada. The numbers of equations in the model are represented by the routes connecting the UK and the countries. Model with six equations involved all the routes, whereas four equations model are UK-Germany, UK-Italy, UK-US, and UK-Canada. Since US and Canada are from the same region, then routes UK-US and UK-Canada are selected for a two equations model.

Table 4.3

True Specification Models of Four Equations (m = 4) by SURE-Autometrics

| | | |
|----|--|--|
| S1 | $y_{1t} = 0.0175 + 0.0259\varepsilon_{1t}$ $y_{2t} = 0.0179 + 0.0279\varepsilon_{2t}$ | $y_{3t} = 0.0230 + 0.0293\varepsilon_{3t}$ $y_{4t} = 0.0182 + 0.0240\varepsilon_{4t}$ |
| S2 | $y_{1t} = 0.0058 + 0.6653y_{1t-1} + 0.0182\varepsilon_{1t}$ $y_{2t} = 0.0077 + 0.5711y_{2t-1} + 0.0216\varepsilon_{2t}$ | $y_{3t} = 0.0094 + 0.5864y_{3t-1} + 0.0229\varepsilon_{3t}$ $y_{4t} = 0.0063 + 0.6563y_{4t-1} + 0.0173\varepsilon_{4t}$ |
| S3 | $y_{1t} = 0.0052 + 0.6359y_{1t-1} + 0.3652x_{21t} - 0.2681x_{21t-1} + 0.0169\varepsilon_{1t}$ $y_{2t} = 0.0059 + 0.6325y_{2t-1} + 0.4056x_{22t} - 0.3251x_{22t-1} + 0.0171\varepsilon_{2t}$ | $y_{3t} = 0.0084 + 0.6114y_{3t-1} + 0.3008x_{23t} - 0.2532x_{23t-1} + 0.0203\varepsilon_{3t}$ $y_{4t} = 0.0063 + 0.6725y_{4t-1} + 0.2933x_{24t} - 0.2360x_{24t-1} + 0.0151\varepsilon_{4t}$ |
| S4 | $y_{1t} = 0.0026 + 0.6406y_{1t-1} + 0.3480x_{41t} - 0.1096x_{41t-1} + 0.0177\varepsilon_{1t}$ $y_{2t} = 0.0047 + 0.5600y_{2t-1} + 0.1442x_{42t} + 0.0602x_{42t-1} + 0.0215\varepsilon_{2t}$ | $y_{3t} = 0.0052 + 0.5724y_{3t-1} + 0.4896x_{43t} - 0.1932x_{43t-1} + 0.0221\varepsilon_{3t}$ $y_{4t} = 0.0030 + 0.6296y_{4t-1} + 0.1358x_{44t} + 0.1128x_{44t-1} + 0.0171\varepsilon_{4t}$ |
| S5 | $y_{1t} = 0.0027 + 0.6147y_{1t-1} + 0.3459x_{21t} - 0.2600x_{21t-1} + 0.2118x_{41t} - 0.0151x_{41t-1} + 0.0167\varepsilon_{1t}$ $y_{2t} = 0.0042 + 0.6272y_{2t-1} + 0.4040x_{22t} - 0.3283x_{22t-1} + 0.0315x_{42t} + 0.0874x_{42t-1} + 0.0170\varepsilon_{2t}$ | $y_{3t} = 0.0051 + 0.5957y_{3t-1} + 0.2757x_{23t} - 0.2429x_{23t-1} + 0.3750x_{43t} - 0.1255x_{43t-1} + 0.0198\varepsilon_{3t}$ $y_{4t} = 0.0039 + 0.6551y_{4t-1} + 0.2865x_{24t} - 0.2398x_{24t-1} + 0.0307x_{44t} + 0.1469x_{44t-1} + 0.0149\varepsilon_{4t}$ |

The dependent variable (y_{it}) is the number of air passengers' and there are seven independent variables that can be associated with this variable. The aim of this chapter is to assess whether *SURE-Autometrics* able to find the specifications used in

generating the artificial data. Hence, the details about the dependent variable as well as the independent variables are not explained in this chapter. The applications of the real data are discussed in Section 5.3.

Table 4.4

True Specification Models of Six Equations (m = 6) by SURE-Autometrics

| | | |
|----|---|---|
| S1 | $y_{1t} = 0.0175 + 0.0259\varepsilon_{1t}$ | $y_{4t} = 0.0428 + 0.0605\varepsilon_{4t}$ |
| | $y_{2t} = 0.0245 + 0.0322\varepsilon_{2t}$ | $y_{5t} = 0.0230 + 0.0293\varepsilon_{5t}$ |
| | $y_{3t} = 0.0179 + 0.0279\varepsilon_{3t}$ | $y_{6t} = 0.0182 + 0.0240\varepsilon_{6t}$ |
| S2 | $y_{1t} = 0.0054 + 0.6891y_{1t-1} + 0.0182\varepsilon_{1t}$ | $y_{4t} = 0.0184 + 0.5694y_{4t-1} + 0.0526\varepsilon_{4t}$ |
| | $y_{2t} = 0.0103 + 0.5900y_{2t-1} + 0.0261\varepsilon_{2t}$ | $y_{5t} = 0.0095 + 0.5855y_{5t-1} + 0.0229\varepsilon_{5t}$ |
| | $y_{3t} = 0.0073 + 0.5934y_{3t-1} + 0.0216\varepsilon_{3t}$ | $y_{6t} = 0.0067 + 0.6317y_{6t-1} + 0.0174\varepsilon_{6t}$ |
| S3 | $y_{1t} = 0.0049 + 0.6600y_{1t-1} + 0.3067x_{21t} - 0.2224x_{21t-1} + 0.0169\varepsilon_{1t}$ | |
| | $y_{2t} = 0.0101 + 0.5944y_{2t-1} + 0.0230x_{22t} + 0.0082x_{22t-1} + 0.0260\varepsilon_{2t}$ | |
| | $y_{3t} = 0.0059 + 0.6326y_{3t-1} + 0.3603x_{23t} - 0.2793x_{23t-1} + 0.0172\varepsilon_{3t}$ | |
| | $y_{4t} = 0.0154 + 0.5926y_{4t-1} + 0.7001x_{24t} - 0.5538x_{24t-1} + 0.0466\varepsilon_{4t}$ | |
| | $y_{5t} = 0.0085 + 0.6079y_{5t-1} + 0.2968x_{25t} - 0.2420x_{25t-1} + 0.0203\varepsilon_{5t}$ | |
| | $y_{6t} = 0.0073 + 0.6266y_{6t-1} + 0.2852x_{26t} - 0.2012x_{26t-1} + 0.0151\varepsilon_{6t}$ | |
| S4 | $y_{1t} = 0.0025 + 0.6735y_{1t-1} + 0.3361x_{41t} - 0.1268x_{41t-1} + 0.0177\varepsilon_{1t}$ | |
| | $y_{2t} = 0.0067 + 0.5844y_{2t-1} + 0.3125x_{42t} - 0.0727x_{42t-1} + 0.0257\varepsilon_{2t}$ | |
| | $y_{3t} = 0.0045 + 0.5875y_{3t-1} + 0.1345x_{43t} + 0.0528x_{43t-1} + 0.0215\varepsilon_{3t}$ | |
| | $y_{4t} = 0.0169 + 0.5693y_{4t-1} - 0.8238x_{44t} + 0.9356x_{44t-1} + 0.0511\varepsilon_{4t}$ | |
| | $y_{5t} = 0.0052 + 0.5720y_{5t-1} + 0.4898x_{45t} - 0.1929x_{45t-1} + 0.0221\varepsilon_{5t}$ | |
| | $y_{6t} = 0.0031 + 0.6052y_{6t-1} + 0.1451x_{46t} + 0.1212x_{46t-1} + 0.0171\varepsilon_{6t}$ | |
| S5 | $y_{1t} = 0.0025 + 0.6451y_{1t-1} + 0.2992x_{21t} - 0.2223x_{21t-1} + 0.2198x_{41t} - 0.0435x_{41t-1} + 0.0166\varepsilon_{1t}$ | |
| | $y_{2t} = 0.0069 + 0.5827y_{2t-1} + 0.0135x_{22t} + 0.0069x_{22t-1} + 0.3045x_{42t} - 0.0799x_{42t-1} + 0.0257\varepsilon_{2t}$ | |
| | $y_{3t} = 0.0042 + 0.6283y_{3t-1} + 0.3582x_{23t} - 0.2801x_{23t-1} + 0.0396x_{43t} + 0.0778x_{43t-1} + 0.0171\varepsilon_{3t}$ | |
| | $y_{4t} = 0.0142 + 0.5900y_{4t-1} + 0.6638x_{24t} - 0.5258x_{24t-1} - 0.5428x_{44t} + 0.6469x_{44t-1} + 0.0458\varepsilon_{4t}$ | |
| | $y_{5t} = 0.0051 + 0.5933y_{5t-1} + 0.2701x_{25t} - 0.2278x_{25t-1} + 0.3759x_{45t} - 0.1311x_{45t-1} + 0.0198\varepsilon_{5t}$ | |
| | $y_{6t} = 0.0047 + 0.6090y_{6t-1} + 0.2709x_{26t} - 0.1943x_{26t-1} + 0.0485x_{46t} + 0.1423x_{46t-1} + 0.0150\varepsilon_{6t}$ | |

A constant term in S1 was based on the mean of y_{it} and the coefficient for disturbances is represented by the standard deviation of y_{it} for each equation. As for other specifications (S2, S3, S4 and S5), the coefficients were FGLS estimators

obtained by using the real data where the standard error of estimated equation be the coefficients for the disturbances. Meanwhile, the S3 or S4 model contains one out of seven independent variables. Similar to *Autometrics*, S3 considers variable with the highest correlation with y_{it} and S4 includes the second highest correlation variable.

Since the annual series only contains 37 observations, it was converted to quarterly data using quadratic-match sum frequency conversion in *EViews* for the purpose of simulation study. Subsequently the quarterly data were log-transformed as well as differenced two times in order to achieve stationarity. Therefore, the total number of observations turns into 146.

Noticeably, the specification models also contains the disturbances terms. Another requirement for generating the artificial dependent variable is to allow the contemporaneous correlation disturbances amongst the equations. Hence, the disturbances are simulated using standard normal distribution and allowed to correlate with other equations based on three levels of correlation strength. The levels are 0.9 to indicate strong correlation, 0.6 to represent moderate correlation, and 0.2 for weak correlation.

In general, generating the artificial dependent variables require the real data, the specification models and the simulated random error variable. The real data will be used to find the coefficient values in the models as well as the variables during the generating process.

4.2 Measurement of *SURE-Autometrics* Performance

Basically, each specification model has 300 artificial dependent variables generated according to three different levels of correlation disturbances strength since they will be replicated 100 times. As a result, each different number of equations has 1500 artificial data sets which lead to a total of 4500 sets since each true specification has three different levels of strength. These data sets are used for the simulation assessment of *SURE-Autometrics*. Subsequently, numerous irrelevant variables were added to the specifications during the formulation of the initial GUMS in the first phase of *SURE-Autometrics*. The algorithm will reduce the variables in searching the ‘best’ multiple equations model. The algorithm is performing well if it able to remove the irrelevant variables and retain the relevant variables since the artificial data were generated based on the specification model. Hence, the true specification model is obtained. The simulation outcomes of *SURE-Autometrics* model selection are classified into four categories where the criteria were adapted from Ismail (2005). The categories are described in Table 4.5. The outcomes fall to the designated category if all the equations in the model possessed the criteria.

Table 4.5

Categories of Simulation Outcomes

| Category | Criteria | Explanation |
|----------|--|--|
| 1 | TRUE = FINAL | The true specification is chosen. |
| 2 | TRUE \subset Final | The true specification is nested in the final specification. |
| 3 | TRUE $\not\subset$ Final | An incorrect specification is chosen, the true specification is not nested in the final specification. |
| 4 | At least one of the equations failed to fall under the same category | |

The performance of *SURE-Autometrics* is indicated by high percentage of outcomes in Category 1 where all the equations have similar variables as in the true specification. Category 4 is designated for model with outcomes of each equation are different.

4.3 Simulation Results

The simulation experiments have various conditions varying from different GUMS to the different sample sizes. Hence, each table of the results contain the percentages of outcomes in finding S1 to S5 when the initial GUMS comprises of 18 and 39 number of variables using large ($n = 146$) and small ($n = 73$) sample size. The results tables also indicated the outcomes when changing the setting of significance level in the *SURE-Autometrics* algorithm. The results were classified into three parts according to the number of equations in the model. The performances of *SURE-Autometrics* for each specification within any number of equations is assessed through 24 conditions comprised of combinations between two sets of initial GUMS, three strengths of correlation disturbances, two sample sizes and two significance levels. Hence a total of 120 conditions used to assess different multiple equations. The performances are measured by the percentage of similarities among the selected final model and the specification in terms of the variables which indicated by the outcomes in Category 1. This measure also equivalent as calculating the probability of the algorithm finding the correct specification models.

4.3.1 Two Equations Model

The simulation experiments of *SURE-Autometrics* started with assessment of two equations model as initial indicator for the algorithm performances in small number of

multiple equations model. Each of 120 conditions discussed in Section 4.1 were replicated 100 times using the artificial data sets. The outcomes of simulating *SURE-Autometrics* on two equations initiated from large ($k = 39$) and small ($k = 18$) GUMS for full sample size ($n = 146$) are shown in Table 4.6 and 4.7, respectively. While Tables 4.8 and 4.9 indicate the simulation outcomes for half of the sample size ($n = 73$). Each table represents the percentages of outcomes in Category 1 to 4 for five true specifications (S1, S2, S3, S4, and S5). Each specification has three levels (0.9, 0.6, and 0.2) of disturbances correlation strength. The results in Tables 4.6 to 4.9 also have been classified into two significance levels (0.05 and 0.01).

All the tables indicate that most of the outcomes were in Category 1 which implied that the true specifications for multiple equations are similar to the final selected model. Unlike other true models, only S1 resulted in two categories which are Category 1 and 2 for all experiment conditions. This is because the model simply contains a constant value. Thus S1 is always nested in the final selected model. Outcomes in Category 4 reveals that less than 10% achieved models contain of equations with different outcome's category. Generally, the overall results were high for full sample size, small initial GUMS and 5% level of significance.

In particular, Table 4.6 shows overall outcomes of 83.87% and 78.87% using 0.05 and 0.01 level of significance, respectively. Both S1 and S3 with the strongest (0.9) correlation disturbances are able to achieve the highest percentage of similarities (89%) although significance level is different, whereas S5 received the lowest percentage (67%).

Table 4.6

Simulation Results for $m = 2, n = 146, k = 39$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 89 | 83 | 85 | 88 | 85 | 86 | 83 | 80 | 81 | 87 | 84 | 82 | 83 | 82 | 80 | 83.87 |
| | 2 | 11 | 17 | 15 | 8 | 8 | 7 | 7 | 8 | 7 | 8 | 7 | 9 | 10 | 8 | 9 | 9.27 |
| | 3 | - | - | - | 2 | 3 | 4 | 6 | 7 | 6 | 1 | 3 | 4 | 1 | 1 | 3 | 2.73 |
| | 4 | - | - | - | 2 | 4 | 3 | 4 | 5 | 6 | 4 | 6 | 5 | 6 | 9 | 8 | 4.13 |
| 1% | 1 | 85 | 82 | 82 | 88 | 80 | 79 | 89 | 78 | 77 | 80 | 79 | 78 | 70 | 67 | 69 | 78.87 |
| | 2 | 15 | 18 | 18 | 8 | 9 | 8 | 6 | 7 | 9 | 8 | 8 | 9 | 9 | 8 | 7 | 9.80 |
| | 3 | - | - | - | 2 | 7 | 8 | 2 | 9 | 6 | 5 | 8 | 6 | 12 | 13 | 13 | 6.06 |
| | 4 | - | - | - | 2 | 4 | 5 | 3 | 6 | 8 | 7 | 5 | 7 | 9 | 12 | 11 | 5.27 |

3

Table 4.7

Simulation Results for $m = 2, n = 146, k = 18$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----------|-----|-----|-----|-----|-----|-----|-----|-----------|-----------|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 96 | 97 | 94 | 91 | 92 | 85 | 88 | 90 | 88 | 92 | 94 | 89 | 90 | 87 | 85 | 90.53 |
| | 2 | 4 | 3 | 6 | 5 | 4 | 7 | 6 | 3 | 5 | 4 | 2 | 7 | 4 | 3 | 5 | 4.53 |
| | 3 | - | - | - | 2 | 2 | 5 | 2 | 4 | 4 | 2 | 1 | 1 | 2 | 4 | 5 | 2.27 |
| | 4 | - | - | - | 2 | 2 | 3 | 4 | 3 | 3 | 2 | 3 | 3 | 4 | 6 | 5 | 2.67 |
| 1% | 1 | 85 | 85 | 84 | 85 | 86 | 83 | 88 | 85 | 82 | 89 | 89 | 87 | 84 | 81 | 82 | 85.00 |
| | 2 | 15 | 15 | 16 | 6 | 6 | 5 | 3 | 8 | 5 | 4 | 5 | 4 | 3 | 5 | 4 | 6.93 |
| | 3 | - | - | - | 6 | 4 | 7 | 4 | 2 | 7 | 3 | 1 | 2 | 5 | 5 | 6 | 3.47 |
| | 4 | - | - | - | 3 | 4 | 5 | 5 | 5 | 6 | 4 | 5 | 7 | 8 | 9 | 8 | 4.60 |

Table 4.8

Simulation Results for $m = 2, n = 73, k = 39$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 88 | 85 | 87 | 85 | 84 | 84 | 82 | 84 | 83 | 84 | 86 | 85 | 79 | 80 | 82 | 83.87 |
| | 2 | 12 | 15 | 13 | 7 | 7 | 6 | 5 | 4 | 6 | 5 | 4 | 4 | 5 | 3 | 3 | 6.60 |
| | 3 | - | - | - | 3 | 5 | 6 | 7 | 8 | 7 | 6 | 4 | 5 | 7 | 7 | 4 | 4.60 |
| | 4 | - | - | - | 5 | 4 | 4 | 6 | 4 | 4 | 5 | 6 | 6 | 9 | 10 | 11 | 4.93 |
| 1% | 1 | 83 | 82 | 82 | 80 | 79 | 80 | 82 | 80 | 79 | 79 | 78 | 80 | 65 | 66 | 63 | 77.20 |
| | 2 | 17 | 18 | 18 | 9 | 9 | 10 | 9 | 9 | 9 | 9 | 12 | 10 | 10 | 12 | 14 | 11.67 |
| | 3 | - | - | - | 6 | 4 | 4 | 2 | 2 | 4 | 4 | 1 | 1 | 11 | 10 | 10 | 3.93 |
| | 4 | - | - | - | 5 | 8 | 6 | 7 | 9 | 8 | 8 | 9 | 9 | 14 | 12 | 13 | 7.20 |

64

Table 4.9

Simulation Results for $m = 2, n = 73, k = 18$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----------|-----|-----|-----|-----|-----|-----|-----------|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 91 | 92 | 91 | 89 | 89 | 88 | 84 | 82 | 84 | 89 | 89 | 88 | 87 | 87 | 87 | 87.80 |
| | 2 | 9 | 8 | 9 | 4 | 4 | 5 | 6 | 5 | 7 | 5 | 4 | 7 | 6 | 4 | 4 | 5.80 |
| | 3 | - | - | - | 4 | 4 | 3 | 6 | 8 | 4 | 1 | 3 | 2 | 2 | 4 | 3 | 2.93 |
| | 4 | - | - | - | 3 | 3 | 4 | 4 | 5 | 5 | 5 | 4 | 3 | 5 | 5 | 6 | 3.47 |
| 1% | 1 | 83 | 84 | 82 | 84 | 83 | 84 | 85 | 84 | 86 | 84 | 81 | 82 | 80 | 81 | 80 | 82.87 |
| | 2 | 17 | 16 | 18 | 4 | 3 | 4 | 5 | 5 | 4 | 6 | 4 | 5 | 7 | 8 | 8 | 7.60 |
| | 3 | - | - | - | 6 | 9 | 7 | 3 | 5 | 2 | 3 | 8 | 7 | 3 | 3 | 3 | 3.93 |
| | 4 | - | - | - | 6 | 5 | 5 | 7 | 6 | 8 | 7 | 7 | 6 | 10 | 8 | 9 | 5.60 |

Meanwhile, the overall values are slightly higher for the small sets of initial GUMS as displayed in Table 4.7 which are 90.53% and 85%. Again S1 and S5 obtained the highest (97%) and the lowest (81%) percentages. These values are indeed higher than the results from the large sets of initial GUMS. Significance level of 0.05 consistently yields high percentages as compared to 0.01 based on Category 1 and 2.

The results obtained by using half of the sample size according to Table 4.8 and 4.9 show similar occurrences as in utilising full sample. At main significance level of 0.05, the highest percentage accomplished by S1 for both sets of initial GUMS while S5 is unable to achieve more than 82% if selected from large set of GUMS. The values were dropped tremendously when the level is set at 1% where the percentages are around 65%. These results are the lowest amongst others.

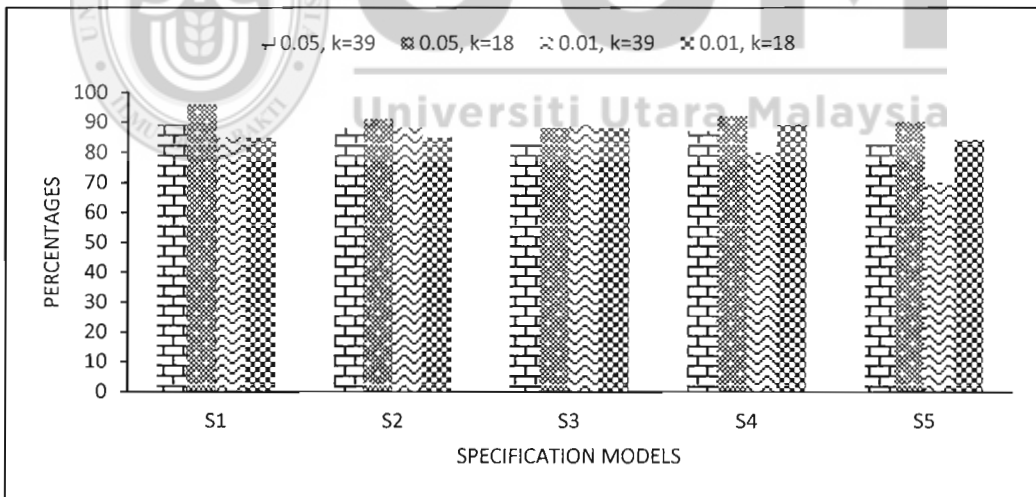


Figure 4.1. Overall Performances on Two Equations Model using Large Sample Size

Figure 4.1 above shows the performances of *SURE-Autometrics* correctly found the five specification models based on four combinations of two initial GUMS and two significance levels using all the observations in the sample. Obviously it can be seen

that the algorithm is poorly performed in finding the S5 specification when initiated from large set of GUMS using 1% level of significance. While there is not much different performances in finding other specification models based on the four combinations of two initial GUMS and two levels of significance.

Figure 4.2 below displays the performances using the small sample size which is half of the observations. The figure seems similar to previous but the pattern in S2 and S3 are slightly different. In finding S2 at small significance level (0.01), the lowest percentages obtained by large set of GUMS using half of sample size as compared to small GUMS for full sample size. *SURE-Autometrics* has similarly performed for all combinations in finding S3 specification using small size, whereas the percentages are marginally lower for the combination of 0.05 significance level and large GUMS using large sample size. In general, both figures are almost comparable indicating that the sample sizes did not affect the algorithm performances.

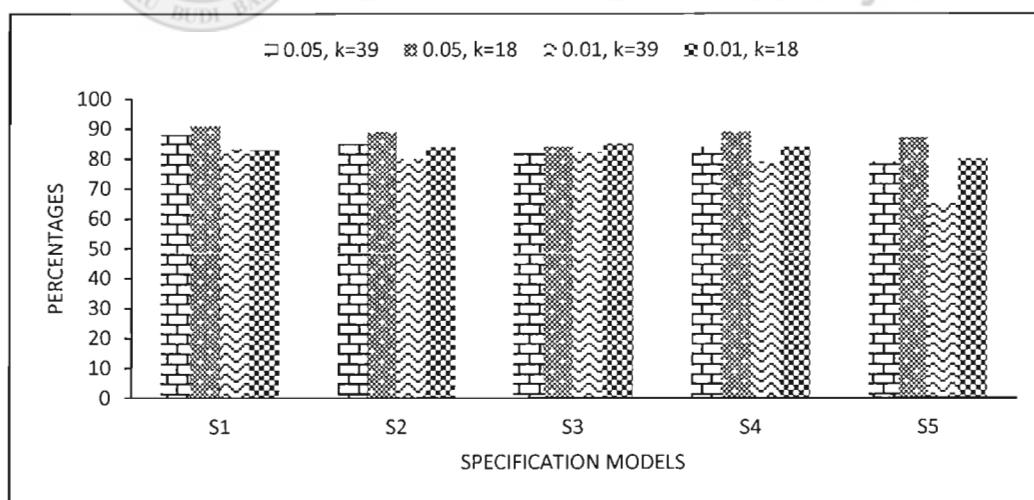


Figure 4.2. Overall Performances on Two Equations Model using Small Sample Size

4.3.2 Four Equations Model

The simulation assessment of *SURE-Autometrics* on four equations used similar procedures as in two equations model. The data set consists of artificial dependent variables that were generated for the five true specification models (S1, S2, S3, S4 and S5) based on Table 4.3. Subsequently *SURE-Autometrics* is employed at two significance levels (0.05 and 0.01) to search the best models from two sets of initial GUMS comprises of 39 and 18 predictor variables using full and half of the sample sizes (146 and 73).

The outcomes are categorised according to criteria indicated in Table 4.5. The performances are measured by the percentages of outcome has the criteria for the corresponding category's. The results of simulation are divided into four tables where Tables 4.10 and 4.11 show percentages obtained by searching model from two sets of GUMS with all observations, whereas Table 4.12 and 4.13 display results using half observations in the sample.

On average, the overall percentages resulted in Category 1 ranged from 77% to 84% for selection at 5% level of significance, while changing to smaller level (1%) causes the percentages reduced slightly in the range of 2% to 3%. Sharing similar patterns as in two equations model, none of the outcomes in Category 3 for S1 because all the final selected models contained a constant value. Thus, it will always nest the true specification models. Moreover, the outcomes of simulation in assessing S1 show that the four equations have similar criteria which are why none of the outcomes are in Category 4.

Table 4.10

Simulation Results for $m = 4, n = 146, k = 39$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----------|-----|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 87 | 88 | 85 | 85 | 82 | 84 | 86 | 87 | 85 | 82 | 83 | 81 | 79 | 77 | 78 | 83.27 |
| | 2 | 13 | 12 | 15 | 6 | 6 | 6 | 3 | 4 | 4 | 6 | 7 | 7 | 7 | 9 | 6 | 7.40 |
| | 3 | - | - | - | 4 | 7 | 6 | 5 | 2 | 6 | 5 | 4 | 7 | 5 | 7 | 7 | 4.33 |
| | 4 | - | - | - | 5 | 5 | 4 | 6 | 7 | 5 | 7 | 6 | 5 | 9 | 7 | 9 | 5.00 |
| 1% | 1 | 85 | 87 | 88 | 89 | 88 | 84 | 78 | 76 | 79 | 84 | 81 | 79 | 76 | 77 | 74 | 81.67 |
| | 2 | 15 | 13 | 12 | 3 | 3 | 4 | 6 | 6 | 5 | 6 | 5 | 6 | 7 | 6 | 7 | 6.93 |
| | 3 | - | - | - | 1 | 4 | 4 | 7 | 9 | 8 | 2 | 7 | 6 | 6 | 4 | 9 | 4.47 |
| | 4 | - | - | - | 7 | 5 | 8 | 9 | 9 | 8 | 8 | 7 | 9 | 11 | 13 | 10 | 6.93 |

89

Table 4.11

Simulation Results for $m = 4, n = 146, k = 18$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 89 | 92 | 90 | 81 | 80 | 79 | 77 | 77 | 75 | 78 | 75 | 79 | 79 | 75 | 74 | 80.00 |
| | 2 | 11 | 8 | 10 | 8 | 9 | 9 | 8 | 7 | 9 | 8 | 9 | 7 | 5 | 5 | 7 | 8.00 |
| | 3 | - | - | - | 6 | 5 | 8 | 8 | 9 | 9 | 6 | 9 | 6 | 6 | 8 | 5 | 5.67 |
| | 4 | - | - | - | 5 | 6 | 4 | 7 | 7 | 7 | 8 | 7 | 8 | 10 | 12 | 14 | 6.33 |
| 1% | 1 | 85 | 87 | 84 | 80 | 78 | 80 | 77 | 78 | 75 | 78 | 76 | 79 | 74 | 72 | 75 | 78.53 |
| | 2 | 15 | 13 | 16 | 7 | 8 | 9 | 9 | 8 | 9 | 9 | 7 | 6 | 9 | 7 | 7 | 9.27 |
| | 3 | - | - | - | 7 | 8 | 6 | 7 | 8 | 6 | 7 | 9 | 8 | 4 | 6 | 7 | 5.53 |
| | 4 | - | - | - | 6 | 6 | 5 | 7 | 6 | 8 | 8 | 8 | 7 | 13 | 15 | 11 | 6.67 |

Table 4.12

Simulation Results for $m = 4, n = 73, k = 39$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----|-----|-----------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 81 | 81 | 82 | 84 | 85 | 82 | 81 | 79 | 82 | 81 | 80 | 78 | 78 | 75 | 77 | 80.40 |
| | 2 | 19 | 19 | 18 | 6 | 5 | 6 | 9 | 7 | 5 | 6 | 8 | 8 | 6 | 9 | 6 | 9.13 |
| | 3 | - | - | - | 4 | 3 | 7 | 4 | 7 | 8 | 5 | 6 | 5 | 9 | 9 | 8 | 5.00 |
| | 4 | - | - | - | 6 | 7 | 5 | 6 | 7 | 5 | 8 | 6 | 9 | 7 | 7 | 9 | 5.47 |
| 1% | 1 | 79 | 82 | 80 | 84 | 82 | 83 | 75 | 74 | 72 | 79 | 75 | 73 | 75 | 72 | 71 | 77.07 |
| | 2 | 21 | 18 | 20 | 8 | 7 | 5 | 9 | 9 | 9 | 3 | 9 | 7 | 5 | 6 | 6 | 9.46 |
| | 3 | - | - | - | 2 | 4 | 4 | 9 | 6 | 9 | 8 | 7 | 9 | 10 | 10 | 12 | 6.00 |
| | 4 | - | - | - | 6 | 7 | 8 | 7 | 11 | 10 | 10 | 9 | 11 | 10 | 12 | 11 | 7.47 |

69

Table 4.13

Simulation Results for $m = 4, n = 73, k = 18$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 84 | 85 | 84 | 79 | 80 | 81 | 71 | 75 | 76 | 72 | 76 | 73 | 77 | 75 | 78 | 77.73 |
| | 2 | 16 | 15 | 16 | 9 | 8 | 7 | 9 | 9 | 9 | 9 | 8 | 10 | 10 | 9 | 9 | 10.20 |
| | 3 | - | - | - | 6 | 6 | 5 | 11 | 8 | 7 | 10 | 8 | 8 | 5 | 7 | 4 | 5.67 |
| | 4 | - | - | - | 6 | 6 | 7 | 9 | 8 | 8 | 9 | 8 | 9 | 8 | 9 | 9 | 6.40 |
| 1% | 1 | 83 | 80 | 81 | 76 | 74 | 77 | 71 | 69 | 68 | 75 | 72 | 73 | 73 | 75 | 72 | 74.60 |
| | 2 | 17 | 20 | 19 | 9 | 9 | 9 | 10 | 10 | 10 | 8 | 9 | 9 | 10 | 9 | 9 | 11.13 |
| | 3 | - | - | - | 7 | 9 | 6 | 10 | 11 | 9 | 6 | 5 | 6 | 6 | 7 | 5 | 5.80 |
| | 4 | - | - | - | 8 | 8 | 8 | 9 | 10 | 13 | 11 | 14 | 12 | 11 | 9 | 14 | 8.47 |

However, the outcomes in Category 4 for other specifications shows an increment indicating that the percentages of equations obtained different category within the same model had increased as compared to model with two equations. Regardless the condition of experiments, the lowest percentage in this category is 5% which has increased by 2.33%, whereas the highest percentage (8.47%) is also increased by 2.38%.

By considering the highest outcomes in the experiments of four equations model amongst all the conditions indicated in the tables, S1 with full observations was able to accomplish the highest outcomes (92%) when searched from small set of initial GUMS ($k = 18$) at 0.05 significance level. The results also show that by using 0.01 level of significance value, S2 obtained 89% and 84% for both sample sizes. On contrary, percentages for S5 were constantly lowest except for small sample size of GUMS with 13 irrelevant variables. Once more, the percentages of outcomes are generally equivalent between three correlation disturbances and generally the overall results for two levels of significance are quite similar with difference of 1% to 2%.

The overall performances of *SURE-Autometrics* on four equations based on combinations of two initial GUMS and two significance levels are summarised in Figure 4.3 and 4.4 using full and half of the sample sizes, respectively. By using all the observations ($n = 146$), the algorithm has well performed for all the combinations in finding S1 as compared to S5 specification. Regardless the combinations, the percentages are similar for S3, S4 and S5, while S2 obtained a little bit higher percentage values.

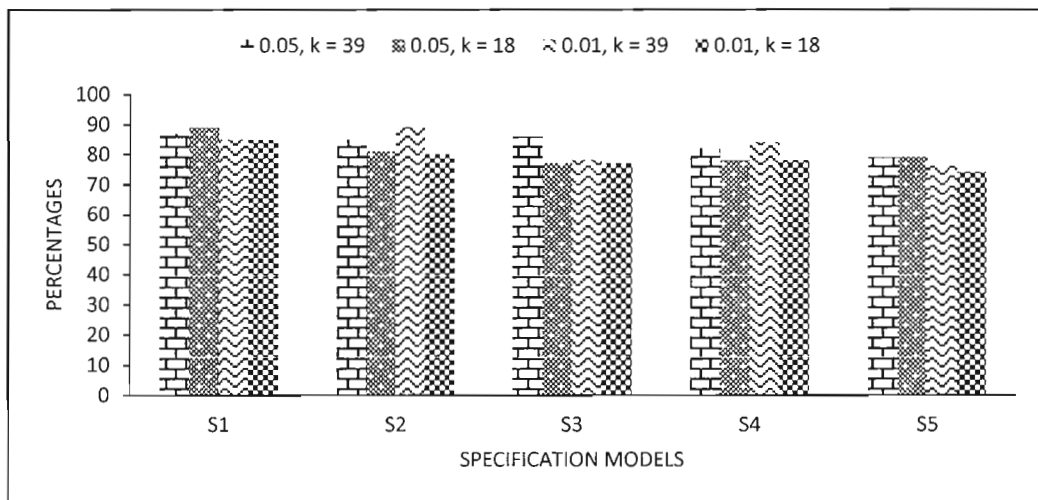


Figure 4.3. Overall Performances on Four Equations Model using Large Sample Size

The performances varied differently for finding S1 and S2 when using half of the sample size with only S3 displays similar pattern. Generally, the percentage values are declined for all the combinations of set of GUMS and significance levels.

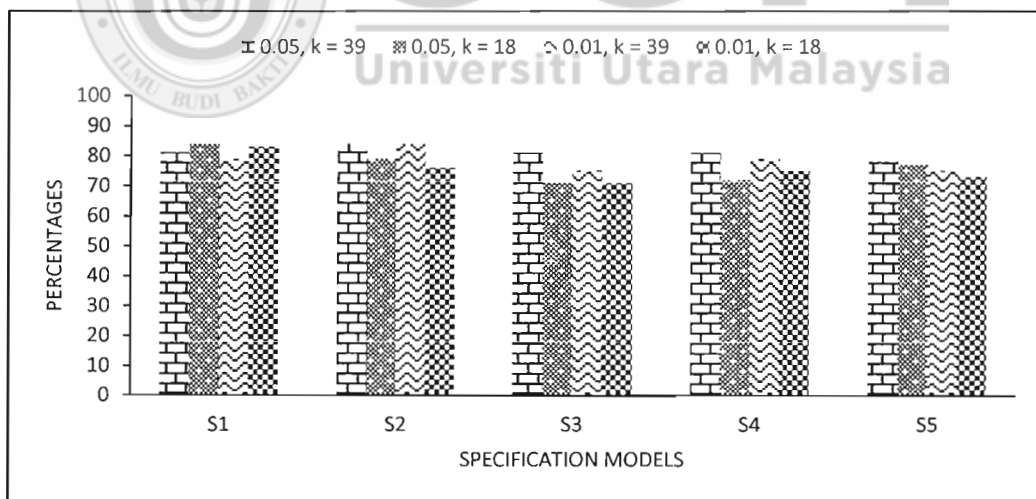


Figure 4.4. Overall Performances on Four Equations Model using Small Sample Size

4.3.3 Six Equations Model

The simulation experiments involving six equations model also have all the conditions tested in models of two and four equations. The true specification models (S1, S2, S3, S4 and S5) were based on Table 4.4. The performances of *SURE-Autometrics* are assessed by computing the percentages of simulation outcomes fall to any category described in Table 4.5.

Two sets of initial GUMS were formulated in the first phase of *SURE-Autometrics* which then reduced to the final model which also known as the specific-to-general model (SUMS) at 5% and 1% levels of significance. The simulation processes and assessment are similarly used as in two and four equations model. However, the results from previous model of two and four equations implied that the performances of *SURE-Autometrics* deteriorated as the number of equations changed, including the number of variables in the specification such as S3, S4 and S5 which have at least three variables.

The results are shown in Table 4.14 to Table 4.17. These tables display that most of the outcomes were in Category 1 which means that all the six equations in the final selected model have similar specification as in the true models. Specifically, the overall percentages have reduced with large difference against the results in two equations model and slightly dissimilar from four equations model. At 0.05 and 0.01 level of significance, the overall values ranged from 72% to 76%, and 70% to 75%, respectively.

Table 4.14

Simulation Results for $m = 6, n = 146, k = 39$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----|-----------|-----|-----|-----|-----|-----------|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 79 | 78 | 80 | 78 | 79 | 77 | 74 | 72 | 75 | 74 | 74 | 73 | 71 | 70 | 72 | 75.07 |
| | 2 | 21 | 22 | 20 | 6 | 6 | 7 | 8 | 8 | 6 | 7 | 8 | 9 | 9 | 9 | 10 | 10.40 |
| | 3 | - | - | - | 5 | 5 | 4 | 7 | 9 | 9 | 7 | 7 | 6 | 9 | 11 | 5 | 5.60 |
| | 4 | - | - | - | 11 | 10 | 12 | 11 | 11 | 10 | 12 | 11 | 12 | 11 | 10 | 13 | 8.93 |
| 1% | 1 | 73 | 75 | 74 | 76 | 75 | 77 | 78 | 79 | 75 | 72 | 73 | 71 | 69 | 67 | 70 | 73.60 |
| | 2 | 27 | 25 | 26 | 9 | 8 | 8 | 8 | 7 | 7 | 9 | 7 | 10 | 10 | 8 | 8 | 11.80 |
| | 3 | - | - | - | 2 | 3 | 1 | 2 | 2 | 3 | 5 | 5 | 7 | 6 | 10 | 8 | 3.60 |
| | 4 | - | - | - | 13 | 14 | 14 | 12 | 12 | 15 | 14 | 15 | 12 | 15 | 15 | 14 | 11.00 |

73

Table 4.15

Simulation Results for $m = 6, n = 146, k = 18$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----------|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 79 | 80 | 78 | 77 | 75 | 76 | 74 | 75 | 72 | 74 | 76 | 75 | 72 | 70 | 71 | 74.93 |
| | 2 | 21 | 20 | 22 | 7 | 8 | 9 | 9 | 8 | 9 | 8 | 7 | 8 | 9 | 11 | 10 | 11.07 |
| | 3 | - | - | - | 6 | 8 | 4 | 6 | 5 | 9 | 6 | 6 | 6 | 7 | 9 | 7 | 5.27 |
| | 4 | - | - | - | 10 | 9 | 11 | 11 | 12 | 10 | 12 | 11 | 11 | 12 | 10 | 12 | 8.73 |
| 1% | 1 | 76 | 77 | 74 | 76 | 75 | 78 | 74 | 75 | 73 | 78 | 79 | 76 | 74 | 73 | 71 | 75.27 |
| | 2 | 24 | 23 | 26 | 6 | 5 | 6 | 6 | 8 | 9 | 7 | 5 | 5 | 9 | 6 | 9 | 10.26 |
| | 3 | - | - | - | 8 | 8 | 7 | 9 | 6 | 9 | 5 | 7 | 6 | 4 | 10 | 5 | 5.60 |
| | 4 | - | - | - | 10 | 12 | 9 | 11 | 11 | 9 | 10 | 9 | 13 | 13 | 11 | 15 | 8.87 |

Table 4.16

Simulation Results for $m = 6, n = 73, k = 39$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----------|-----|-----|-----|-----|-----------|-----|-----|-----|-----|-----------|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 77 | 78 | 75 | 75 | 75 | 76 | 70 | 69 | 71 | 73 | 71 | 69 | 71 | 67 | 69 | 72.40 |
| | 2 | 23 | 22 | 25 | 6 | 8 | 9 | 9 | 8 | 10 | 9 | 9 | 9 | 9 | 10 | 10 | 11.73 |
| | 3 | - | - | - | 7 | 7 | 4 | 10 | 11 | 10 | 8 | 8 | 11 | 9 | 9 | 8 | 6.80 |
| | 4 | - | - | - | 12 | 10 | 11 | 11 | 12 | 9 | 10 | 12 | 11 | 11 | 14 | 13 | 9.07 |
| 1% | 1 | 72 | 74 | 71 | 70 | 69 | 69 | 74 | 71 | 73 | 71 | 71 | 74 | 66 | 68 | 65 | 70.53 |
| | 2 | 28 | 26 | 29 | 10 | 10 | 9 | 9 | 10 | 9 | 10 | 11 | 9 | 9 | 10 | 11 | 13.34 |
| | 3 | - | - | - | 8 | 7 | 11 | 4 | 7 | 4 | 7 | 6 | 3 | 11 | 10 | 9 | 5.80 |
| | 4 | - | - | - | 12 | 14 | 11 | 13 | 12 | 14 | 12 | 12 | 14 | 14 | 12 | 15 | 10.33 |

74

Table 4.17

Simulation Results for $m = 6, n = 73, k = 18$

| Level | Category | S1 | | | S2 | | | S3 | | | S4 | | | S5 | | | Overall (%) |
|-------|----------|-----|-----|-----|-----------|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| | | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | 0.9 | 0.6 | 0.2 | |
| 5% | 1 | 75 | 72 | 74 | 76 | 74 | 77 | 72 | 70 | 69 | 70 | 71 | 69 | 71 | 68 | 72 | 72.00 |
| | 2 | 25 | 28 | 26 | 9 | 8 | 7 | 9 | 9 | 10 | 10 | 9 | 11 | 8 | 7 | 6 | 12.14 |
| | 3 | - | - | - | 6 | 6 | 6 | 9 | 10 | 7 | 8 | 9 | 7 | 10 | 11 | 9 | 6.53 |
| | 4 | - | - | - | 9 | 12 | 10 | 10 | 11 | 14 | 12 | 11 | 13 | 11 | 14 | 13 | 9.33 |
| 1% | 1 | 74 | 70 | 72 | 74 | 75 | 72 | 73 | 71 | 72 | 74 | 72 | 74 | 71 | 69 | 69 | 72.13 |
| | 2 | 26 | 30 | 28 | 7 | 6 | 9 | 9 | 9 | 8 | 9 | 9 | 8 | 9 | 10 | 11 | 12.54 |
| | 3 | - | - | - | 7 | 8 | 9 | 5 | 10 | 8 | 5 | 8 | 7 | 10 | 6 | 6 | 5.93 |
| | 4 | - | - | - | 12 | 11 | 10 | 13 | 10 | 12 | 12 | 11 | 11 | 10 | 15 | 14 | 9.40 |

The tables also indicated that overall outcomes in Category 4 have slightly increased showing that with large number of multiple equations, different outcome categories are achieved amongst the equations. The results in this category ranged from 8% to 11% suggested that difficulties in getting similar specification for all six equations. The condition with the highest percentage of this is GUMS with 39 variables selected at 1% level where both sample sizes produce almost similar values (12.92% and 13.75%).

The S1 specification is once again accomplished the highest percentages for all type of conditions where none of the outcomes resulted in Category 3 and 4. Similar to other multiple equations, the reason is S1 always nested in the final selected model since it only contains constant value. Meanwhile, S5 specification is easily comparable because it keeps achieving the lowest percentages except in two conditions. Both conditions are different at significance level for small set of GUMS ($k = 18$) and small sample size ($n = 73$). These conditions are the lowest obtained by S4 specification. Despite of the severe results, the range of percentages in Category 1 is from 65% to 80%. The outcomes within the same specification are not affected by different strength of correlation disturbances.

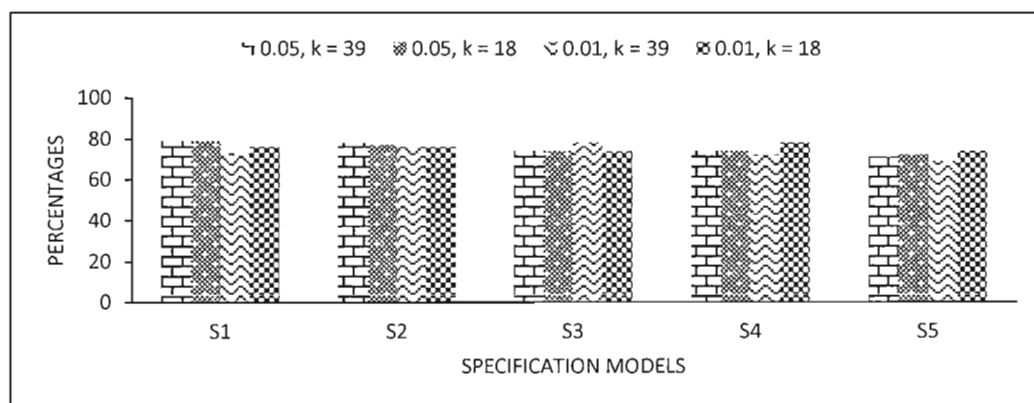


Figure 4.5. Overall Performances on Six Equations Model using Large Sample Size

Figure 4.5 summarised the overall performances for each specification model based on the combination of initial GUMS and the significance level using all the observations. The multiple bar charts indicated that minor differences amongst the four combinations within each specification. Meanwhile, the overall performances using half sample size are displayed in Figure 4.6. Similarly, there are not much different between the combinations of two sets of GUMS and two levels of significance in finding each true specification model. However, the figure showed that the bars are slightly lower than previous figure implying that the performances are much better using large sample size.

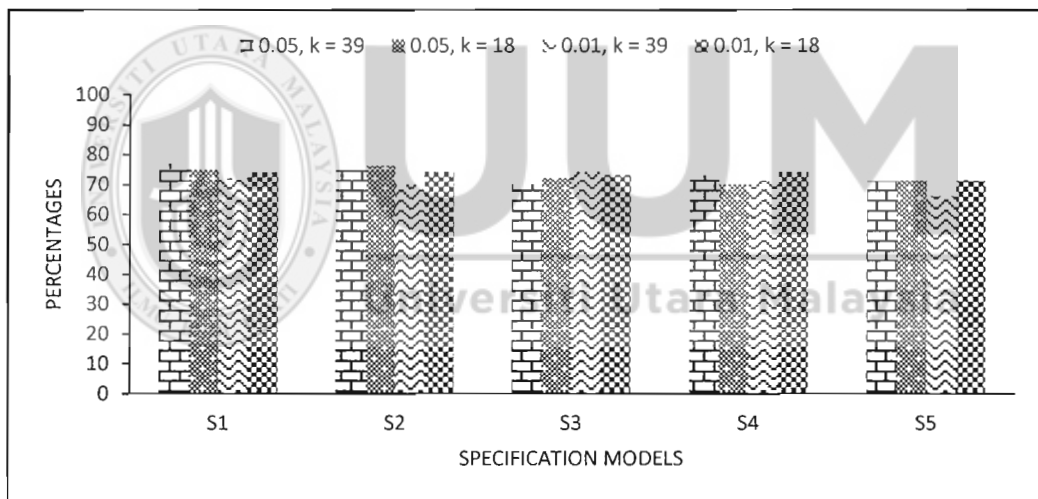


Figure 4.6. Overall Performances on Six Equations Model using Small Sample Size

4.4 Summary of Findings

The purpose of experimental simulation is to assess the performances of *SURE-Autometrics* in finding the true specification models since the data-generating process is known. The measure also indicates the probability of finding the true multiple equations model from the GUMS that is formulated during the initial phase of *SURE-Autometrics*. The specification searches involved 120 experiment conditions of three

different numbers of equations with a total of 360 conditions. The conditions arise during the artificial data generation and simulation of *SURE-Autometrics*. Specifically, it based on the multiple numbers of equations models, the specifications model used in generating data, the strengths of correlation disturbances amongst the equations, the initial sets of GUMS, the sample sizes and the levels of significance. These conditions were shown in Table 4.18.

Table 4.18

Simulation Experiment Conditions

| Conditions of experiment | Level |
|---|---|
| 1. Number of equations in the model | Small, $m = 2$ Medium, $m = 4$ Large, $m = 6$ |
| 2. True specification model | S1 = without any relevant variables S2 = one relevant variable S3 = three relevant variables S4 = three relevant variables S5 = five relevant variables |
| 3. Strength of correlation disturbances | Strong, $\rho = 0.9$ Moderate, $\rho = 0.6$ Weak, $\rho = 0.2$ |
| 4. Initial GUMS | Small set, $k = 18$ variables (13 to 18 irrelevant) Large set, $k = 39$ variables (34 to 39 irrelevant) |
| 5. Sample sizes | Large, $n = 146$ Small, $n = 73$ |
| 6. Main significance level | $\alpha = 0.05$ $\alpha = 0.01$ |

Most of these conditions have been adopted from the simulation study of *SURE-PcGets* (Ismail, 2005) and *Autometrics* (Doornik, 2009). *SURE-PcGets* used similar conditions but only focused on the selection of six equations model with only three

Based on the table, at least 80% of the final selected multiple equations models have similar specification as the true models. It was obtained by all except for one condition which resulted in less than 70% for both sample sizes. Specifically, S5 already has five variables which are relevant to the final model and there were another 34 variables added in the formulation of initial GUMS. Hence the total of variables is 39 variables. The additional variables are irrelevant to be retained in the final model. Moreover, the percentages from small set of GUMS (at most 18 irrelevant variables) were considerably higher compared to large set.

Table 4.20

Percentages of Finding Correct Specification for $m = 4$ and $\rho = 0.9$

| True specification | Sample sizes, n | $\alpha = 5\%$ | | $\alpha = 1\%$ | |
|--------------------|-------------------|----------------|-----------|----------------|-----------|
| | | $k = 39$ | $k = 18$ | $k = 39$ | $k = 18$ |
| S1 | 146 | 87 | 89 | 85 | 85 |
| | 73 | 81 | 84 | 79 | 83 |
| S2 | 146 | 85 | 81 | 89 | 80 |
| | 73 | 84 | 79 | 84 | 76 |
| S3 | 146 | 86 | 77 | 78 | 77 |
| | 73 | 81 | 71 | 75 | 71 |
| S4 | 146 | 82 | 78 | 84 | 78 |
| | 73 | 81 | 72 | 79 | 75 |
| S5 | 146 | 79 | 79 | 76 | 74 |
| | 73 | 78 | 77 | 75 | 73 |

As compared to the results from model of two equations, Table 4.20 indicated that *SURE-Autometrics* performed well by obtaining 70% to 90% similarities to the true specification four equations model. However, it also indicated that the algorithm were not able to achieve 80% in finding the true specification of S5 model. The highest percentage (89%) was from the same condition as in two equations model, whereas the lowest percentage (71%) attained at two conditions by S3. On average, outcomes

from large set of GUMS for these models are higher than the small set and there are only slight differences between two levels of significance.

Meanwhile, the result from Table 4.21 proved that as number of equations in the model increases, the probability of *SURE-Autometrics* finds the true specification becomes lower indicating that the performances are deteriorated. The algorithm started to show inability of achieving higher than 80% for most of the experimental conditions for model with four equations (Table 4.20) where the highest probability to find S5 is 79%. This could be due to number of variables that should be retained in S5 is more than other specifications. Moreover, all four equations must be able to reduce all the irrelevant variables in order to be counted as successful. However, the percentages were not that bad because it was at least 70%. Therefore, it was expected that outcomes for the simulation experiment using six equations will be slightly decreased from these results.

Table 4.21

Percentages of Finding Correct Specification for $m = 6$ and $\rho = 0.9$

| True specification | Sample sizes, n | $\alpha = 5\%$ | | $\alpha = 1\%$ | |
|--------------------|-------------------|----------------|----------|----------------|----------|
| | | $k = 39$ | $k = 18$ | $k = 39$ | $k = 18$ |
| S1 | 146 | 79 | 79 | 73 | 76 |
| | 73 | 77 | 75 | 72 | 74 |
| S2 | 146 | 78 | 77 | 76 | 76 |
| | 73 | 75 | 76 | 70 | 74 |
| S3 | 146 | 74 | 74 | 78 | 74 |
| | 73 | 70 | 72 | 74 | 73 |
| S4 | 146 | 74 | 74 | 72 | 78 |
| | 73 | 73 | 70 | 71 | 74 |
| S5 | 146 | 71 | 72 | 69 | 74 |
| | 73 | 71 | 71 | 66 | 71 |

In general, the *SURE-Autometrics* is able to achieve the true specification model with high percentages of simulation outcomes for SURE model with two and four equations. However, the number of variables in the true models appears to affect the algorithm performances. This can be seen from the results where there is high percentage in finding S1 as compared to low achievement in finding S5 for three different numbers of equations.

Based on these tables, the *SURE-Autometrics* performed well in the simulation analysis when the number of equations and number of variables in the true specification models were as minimal as possible. These variables are relevant to retain in the final model while the additional variables included in the initial GUMS are irrelevant and should be removed during the model selection procedure. Hence, it will be difficult to have similar results for all the equations within the model since S3, S4 and S5 contain of at least three relevant variables while removing 34 to 36 or 13 to 15 irrelevant variables.

Meanwhile, different setting of significance levels seems to have small effect on the algorithm performance. Figure 4.7 shows the performances of each number of equations in finding the correct specification based on two levels of significance. It can be seen that the effects of significance levels within the equation are marginal.

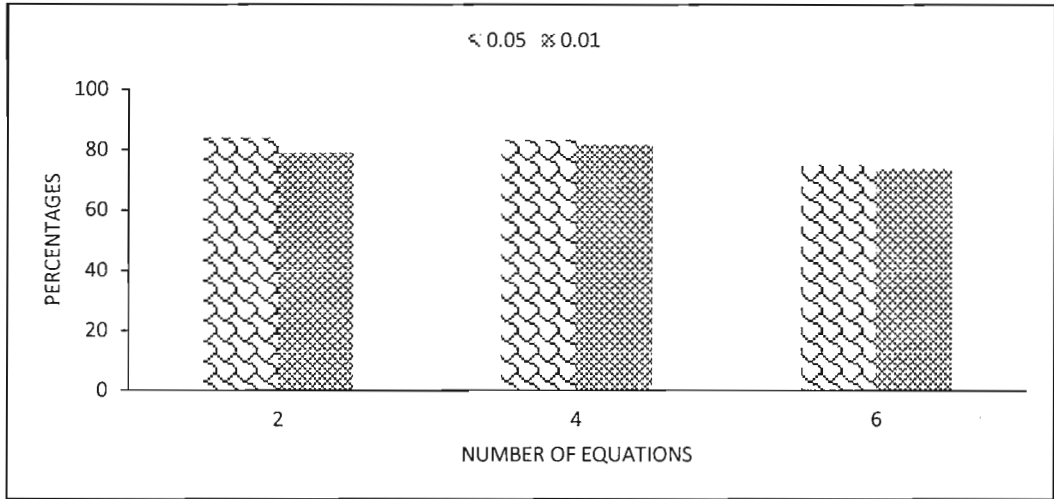


Figure 4.7. Overall Performances in Finding Correct Specifications

The sample size has some effects on the performances but it also related to the specification models which characterise by the number of variables in the model.



UUM
Universiti Utara Malaysia

CHAPTER FIVE

EVALUATION OF *SURE-AUTOMETRICS* USING REAL DATA

5.1 Model Selection Procedures

Evaluation using real data is very crucial in identifying whether the *SURE-Autometrics* algorithm exhibits ‘data mining’ characteristics. This problem is common amongst model builders since data mining permitted the selection of best models within-sample fitted model and able to satisfy all measures of goodness of fits. However, the data mining models might fail when it comes to forecasting. In this chapter, *SURE-Autometrics* and several other model selection procedures were applied on two sets of real data. Section 5.3 illustrates the analysis of the air passengers’ flows data, whereas Section 5.4 analyses the national growth rates data. Subsequently, the selected models were used in forecasting where the error measures were compared to determine which procedure offers a model with better forecasts. The data were also utilised in the validation of *SURE-PcGets*. Hence it were obtained from Ismail (2005).

The model selection procedures considered in this study involving algorithm and non-algorithm procedures. Implementation with algorithms imply that the selection are done automatically based on the steps provided in the algorithms, while non-algorithm indicates manual selection procedures using my own personal knowledge on theory, judgement and experience in statistical modelling. Additionally, these procedures varied with two estimation methods which are feasible generalised least squares (FGLS) and ordinary least squares (OLS). Model selection procedures

through OLS method means that the equation is selected separately within the model, whereas FGLS indicates the multiple equations are selected simultaneously.

Specifically, there are nine model selection procedures that can be classified into four categories as follows,

1. *Autometrics* and *Stepwise* are the algorithms for single equation model. Since the model has multiple equations, each is estimated using OLS and individually selected for multiple times. Model selection through *Autometrics* is applied using *PcGive* software and *Stepwise* is employed by *SAS Enterprise Guide*.
2. *Autometrics-SURE* and *Stepwise-SURE* are procedures that used previous algorithms in the model selection where each equation separately selected with OLS estimation method. However the final model is estimated using FGLS.
3. *SURE-Autometrics* and *SURE-PcGets* are the algorithms for automatic model selection procedures focus on the multiple equations model. The selection is implemented simultaneously with FGLS method of estimation.
4. *Mine*, *Mine-SURE* and *SURE-Mine* are non-algorithm model selection procedures which means the selection is a process of trial and error based on my personal judgment. Each of these procedures are fitted into each previous category, respectively.

The *SURE-Mine* used FGLS as a method of estimation and the inspection of variables are done simultaneously within the model according to the rules above. Meanwhile,

the *Mine-SURE* and *Mine* selects the equation by equation with OLS estimators and FGLS is used to estimate the final model of *Mine-SURE*.

By the inclusion of the fourth categories amongst the model selection procedures, this study aims not to overlook the performances of manual selection procedures as compared to algorithm procedures. Hence the existence of tacit knowledge can be ascertained. The selections of models are determined manually by inspecting the p -values and decision is made based on personal knowledge.

In the manual selection, insignificant variables are determined from the GUMS using the p -values based on 5% significance level. The standard errors of estimated model, as well as the correlations between the independent and dependent variables are also considered in the process of making decision whether to remove or retain the variables. The rules are arranged as follows,

- The removal of variables depends on insignificant p -values starting with the highest value in the model.
- If the correlation is high but insignificant p -value also high, then the variable is skipped.
- If the removal of variable causes the increment in the standard error, then the variable is retained.
- If more than one variable are highly insignificant together with weak correlations, then the variables are removed as a group.

- The reduced model must pass all diagnostic tests as described in Chapter 3.

In particular, the nine model selection procedures are *SURE-Autometrics*, *SURE-PcGets*, *SURE-Mine*, *Autometrics-SURE*, *Stepwise-SURE*, *Mine-SURE*, *Autometrics*, *Stepwise* and *Mine*. The performance of models selected using these procedures are compared by measuring forecast errors.

5.2 Measures of Forecasting Errors

In this study, the data is divided into in-sample for estimate the model, and out-sample for model validation. A model that fit 'best' in-sample might not be also 'best' when it comes to prediction using the out-sample data (Bartolomei & Sweet, 1989; Pant & Starbuck, 1990). Hence the accuracy of forecasting using the estimated model is assessed by the out-of-sample tests instead of goodness of fit in the in-sample tests. According to Greene (2012), most of forecasting accuracy measures are designed to evaluate out-sample forecasts. These measures are based on the errors from the forecasts.

An out-of-sample evaluation of forecasting accuracy begins with the division of the data series into a fit period (i.e. in-sample) and test period (i.e. out-sample). The fit period is used to identify and estimate a model while the test period is reserved to assess the model's forecasting accuracy. In this study, the forecasts of one until three-steps ahead are performed recursively until all out-sample data points are exhausted.

Subsequently, the forecast errors are determined by subtracting each of these forecasts from the observe data values in the test period as follows,

$$\hat{\varepsilon}_t(l) = y_{T+l} - \hat{y}_{T+l} \quad (3.31)$$

where l denotes number of steps-ahead forecast. At forecasting origin (T), the forecasts are generated for time periods $T+1, T+2, \dots, T+l$.

5.2.1 Error Measures

This study compared the forecast errors from final model selected by *SURE-Autometrics* and other selection procedures discussed in Section 5.1. The errors are calculated by the difference between the actual values and out-sample forecasts. There are various ways of obtaining the summary statistic of the forecast errors. As discussed in Hyndman and Koehler (2006), it depends on the choice of error measures and the use of statistical operator. For instance, the error measures would be an absolute errors, squared errors, percentage errors, or relative errors, whereas the possible statistical operator is median, arithmetic mean or geometric means. Typically, personal taste or certain criteria such as reliability, resistant to outliers, interpretability, validity or consensus, and descriptive of underlying distribution could be as reference in choosing the error measure (Armstrong & Collopy, 1992; Fildes, 1992).

Hence the root mean square error (RMSE) and geometric root mean square (GRMSE) are chosen for this study for the purpose of comparison analysis with *SURE-PcGets* since the measures were employed in Ismail (2005). Despite of this, the RMSE is the most commonly error measure for assessing the performance of forecasting models and the GRMSE is able to deal with outliers and extreme values. An indicator for a

good forecasting performance is through the smaller values of these measures. Both error measures are described as follows,

$$RMSE = \sqrt{\frac{\sum_{t=T}^{T+n-l} \hat{\varepsilon}_t^2(l)}{n+1-l}} \quad (3.32)$$

$$GRMSE = \left[\prod_{t=T}^{T+n-l} \hat{\varepsilon}_t^2(l) \right]^{\frac{1}{2(n+1-l)}} \quad (3.33)$$

where $\hat{\varepsilon}_t(l)$ is forecast error at the l -step-ahead forecast, n is number of observations in test period, T is the forecast origin time, and l is number of steps-ahead.

The median across all equations for both RMSE and GRMSE are used to represent the selected models. The performance of each model selection procedure are assessed by ranking these medians in ascending orders from one to nine where one will represent the best forecasting performance. If the median values are similar, then the procedure receives the lowest rank.

5.2.2 Equality Test

The measures in previous section assesses the forecast accuracy of an estimated final models based on the forecast error statistics. According to recent study, (Chen, Wan, & Wang, 2014), these measures have some limitations where the values resulted due to chance and the difference between competing models might be stochastically generated. Thus, the judgment about the forecasting performances based on these error measures only, might be inefficient. Hence, this study also included statistical tests as a formal way to determine whether forecasting using final models selected by

SURE-Autometrics predicts more accurately than others. A statistical test was employed to test the equality of forecasts from the competing models. This test also being applied in *SURE-PcGets* (Ismail, 2005) evaluation.

The test was based on the small sample modification of the Diebold-Mariano test proposed by Harvey, Leybourne and Newbold (1997). It focused on the loss differentials denoted by,

$$d_t = \hat{\varepsilon}_{1t}^2 - \hat{\varepsilon}_{2t}^2 \quad (3.34)$$

where $\hat{\varepsilon}_{1t}$ is error measure from models selected by *SURE-Autometrics*, and $\hat{\varepsilon}_{2t}$ is error measure from the competing models selected by other selection procedure. The two procedures have an equal performance if and only if the error measures have zero expectation for all equations in the model. Hence, the null hypothesis is given as,

$$H_0 : E(d_t) = 0 \quad \text{for all } t \quad (3.35)$$

Then, the sample mean loss differential is,

$$\bar{d} = \frac{1}{n} \sum_{t=1}^n d_t \quad (3.36)$$

and the estimated variance of the mean loss differential is,

$$\hat{V}(\bar{d}) = \frac{1}{n} \left[\hat{\gamma}^*(0) + \frac{2}{n} \sum_{k=1}^{h-1} (n-k) \hat{\gamma}^*(k) \right] \quad (3.37)$$

The statistic has asymptotically t distribution with $(n-1)$ degrees of freedom, is formulated as follows,

$$S_1^* = \left[\frac{n+1-2h+n^{-1}h(h-1)}{n} \right]^{\frac{1}{2}} S_1 \quad (3.38)$$

$$S_1 = [\hat{v}(\bar{d})]^{-\frac{1}{2}} \bar{d} \quad (3.39)$$

$$\hat{\gamma}^*(k) = \frac{1}{n-k} \sum_{t=k+1}^n (d_t - \bar{d})(d_{t-k} - \bar{d}) \quad (3.40)$$

where h is the number of steps-ahead.

5.3 Air Passenger Flows Data

In 2005, Ismail conducted a study on demand for air travel between UK and six countries using a seemingly unrelated regression equations (SURE) model. The annual data from year 1961 to year 2002 was used in the empirical study of *SURE-PcGets*. The dependent variable (Y_{it}) was the total annual of international passenger traffic (in thousands) from and to UK (two-way) reported by airports country, which known as a route onwards. The values represent the total passengers carried by all airlines flying between the UK and each route including all passengers carried on scheduled and chartered services, excluding those carried on aircraft chartered by Government Departments, regardless the economy or business classes.

The countries included are Germany, Sweden, Italy, Japan, USA and Canada. These routes were chosen to evade the routes of high tourist intensity such as Spain. Figure 5.1 displays the number of air passengers for the six routes from 1961 to 2002.

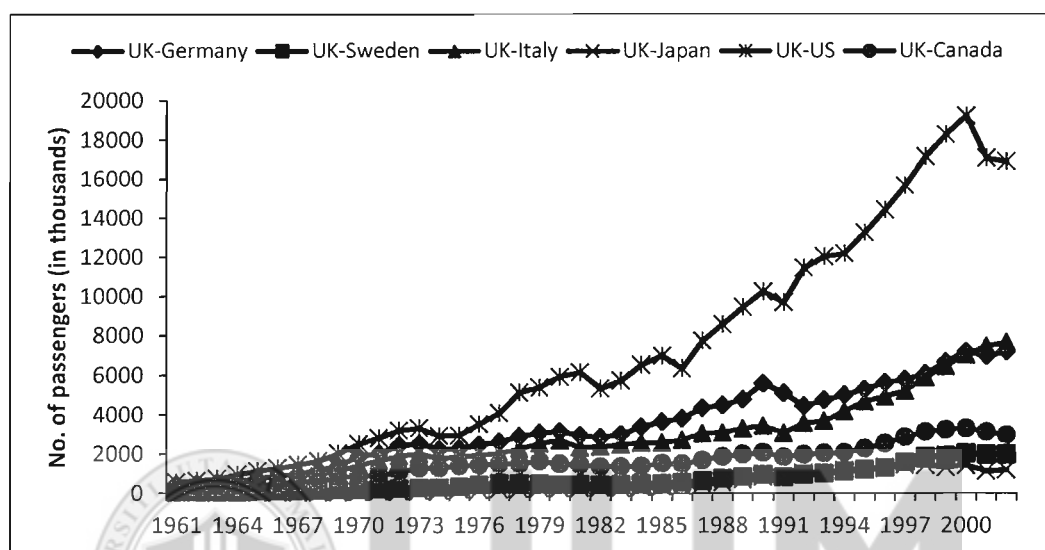


Figure 5.1. Number of Air Passengers for Six Routes

The figure showed the trends of passengers' numbers flying on the six routes are increasing over the years. The increment is noticeably for UK-US route starting from year 1976, but the numbers of passengers have decreased since 2001. This might be due to terrorist incident in September 11, 2001. The trends of flight passengers using UK-Sweden and UK-Japan are the lowest compared to other routes and they are quite similar with the latter has slightly decreased after 1997. Meanwhile, trend of passengers using UK-Italy route arose over the trend of UK-Germany air passengers.

The independent variables are income (in USD billion), trade (in USD million), price ticket (in USD) and 'world' trade (in USD billion). The income variable is a personal disposable income (x_{1t}), whereas the trade (x_{2t}) was included to capture the effects of passengers who travel for business purposes which reflect the strength of economic

relationship between UK and the respective country. The price ticket (x_{i3t}) is the return fare between UK and country using the airports that are selected on the basis of its importance as a gateway to enter or leave a country as shown in Table 5.1. Both income and price were common variables used in many travel demand models (Ismail & Fildes, 2007; Lazim, 1995). Since they are in local currency, it was converted to USD currency after being deflated using consumer price index (CPI). Meanwhile, the ‘world’ trade (x_{i4t}) is a total trade of all industrial countries that is included to improve forecast as well as to consider the effects of transit passengers (Garcia-Ferrer, Highfield, Palm, & Zellner, 1987).

Table 5.1

The Selected Airports to Determine Price Tickets between Countries

| Country | Airports |
|---------|-----------|
| UK | London |
| Germany | Frankfurt |
| Sweden | Stockholm |
| Italy | Rome |
| Japan | Tokyo |
| US | New York |
| Canada | Vancouver |

In this application study, all the variables including the number of passengers have been transformed to achieve stationarity where each variable was log transformed and differenced one time. During the estimation, 36 observations starting from year 1962 to 1997 are used to fit the models and the remaining observations (1998 – 2002) are used to evaluate the models. The model has six equations correspond to the six routes. As guidelines, this study referred to Ismail (2005) in formulating the initial GUMS. It

comprises of three lags of dependent variables (Δy_{it}) and four independent variables (Δx_{ikt}) including their first lag ($\Delta x_{ik(t-1)}$). Hence, the initial GUMS has a total of 11 variables for each equation as follows,

$$\Delta y_{it} = \alpha_{i0} + \sum_{j=1}^3 \alpha_{ij} \Delta y_{i(t-j)} + \sum_{k=1}^4 \sum_{j=0}^1 \phi_{ikj} \Delta x_{ik(t-j)} + \varepsilon_{it}, \quad \begin{array}{l} i = 1, 2, \dots, 6 \\ t = 1, 2, \dots, T \end{array} \quad (3.41)$$

where j is the lag length, Δy_{it} is the growth rate of the number of passengers in year t for route i , Δx_{ikt} is the growth rate of the k^{th} independent variable in year t for route i , ε_{it} are identically independent normally distributed disturbances with mean zero and variance σ^2 , α and ϕ are unknown parameter vectors to be estimated.

Table 5.2 shows the initial GUMS estimated using FGLS by *SURE-Autometrics*, *SURE-PcGets* and *SURE-Mine*, including OLS estimates by *Autometrics*, *Stepwise*, *Mine*, *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE*. Each estimated equation (route) in the GUMS passed all the diagnostic tests except the heteroscedasticity test which is not computed due to insufficient observations. This should not be a matter of concern since heteroscedasticity is often a problem associated with cross-sectional data, but not time-series data. The p -value of MC-QLR test is 0.082 which is significant at 10% level of significance indicating that the seemingly unrelated regression equations (SURE) model is appropriately specified. Additionally, the adjusted R squares (\bar{R}^2) and standard errors for each route also shown in the table.

Table 5.2

Estimated GUMS of Air Passengers' Flows using FGLS and OLS

| Variables | UK-Germany | | UK-Sweden | | UK-Italy | | UK-Japan | | UK-US | | UK-Canada | |
|----------------------------------|---------------------|--------------------|----------------------|--------------------|-----------------------|-----------------------|-----------------------|--------------------|---------------------|--------------------|----------------------|----------------------|
| | FGLS | OLS | FGLS | OLS | FGLS | OLS | FGLS | OLS | FGLS | OLS | FGLS | OLS |
| Constant | -0.023 (-0.930) | -0.028 (-0.858) | 0.029 (0.845) | 0.037 (0.845) | 0.041* (1.889) | 0.047 (1.684) | 0.016 (0.331) | 0.023 (0.387) | 0.011 (0.396) | 0.005 (0.144) | -0.012 (-0.683) | -0.007 (-0.331) |
| Δy_{it-1} | -0.051 (-0.299) | 0.036 (0.145) | 0.020 (0.126) | -0.018 (-0.085) | 0.110 (0.751) | 0.240 (1.208) | 0.047 (0.311) | -0.074 (-0.371) | 0.101 (0.665) | 0.069 (0.312) | 0.152 (0.959) | 0.130 (0.625) |
| Δy_{it-2} | 0.160 (1.079) | 0.181 (0.836) | 0.252 (1.582) | 0.194 (0.915) | 0.022 (0.208) | -0.025 (-0.172) | 0.393*** (2.921) | 0.278 (1.569) | -0.053 (-0.403) | -0.046 (-0.249) | 0.276** (2.526) | 0.318** (2.244) |
| Δy_{it-3} | 0.091 (0.680) | 0.078 (0.382) | 0.094 (0.588) | 0.118 (0.553) | 0.043 (0.414) | 0.095 (0.673) | 0.177 (1.397) | 0.183 (1.093) | 0.086 (0.736) | 0.096 (0.583) | 0.103 (0.903) | 0.137 (0.929) |
| Δx_{it} (Income) | -0.189 (-1.247) | -0.216 (-0.995) | 0.194 (0.800) | 0.234 (0.737) | 0.496*** (3.157) | 0.608** (2.841) | -0.339 (-1.144) | -0.232 (-0.593) | 0.427 (1.278) | 0.513 (0.996) | 0.303* (1.769) | 0.317 (1.430) |
| $\Delta x_{i1(t-1)}$ | -0.153 (-1.030) | -0.017 (-0.078) | -0.393 (-1.565) | -0.419 (-1.268) | 0.323* (1.988) | 0.287 (1.294) | 0.761** (2.340) | 0.536 (1.261) | -0.471 (-1.319) | -0.234 (-0.426) | 0.192 (1.246) | 0.121 (0.608) |
| Δx_{i2t} (Trade) | 0.439*** (3.185) | 0.532** (2.736) | 0.195 (1.540) | 0.211 (1.267) | 0.030 (0.247) | -0.019 (-0.113) | 0.542** (2.814) | 0.471* (1.864) | 0.026 (0.208) | 0.064 (0.348) | 0.204** (2.270) | 0.203* (1.751) |
| $\Delta x_{i2(t-1)}$ | -0.230 (-1.546) | -0.232 (-1.040) | 0.320** (2.357) | 0.360* (2.046) | -0.529*** (-4.747) | -0.525*** (-3.503) | -0.612*** (-3.018) | -0.425 (-1.585) | -0.136 (-1.162) | -0.142 (-0.802) | -0.240** (-2.567) | -0.184 (-1.513) |
| Δx_{i3t} (Price) | 0.008 (0.058) | -0.178 (-0.868) | -0.322** (-2.228) | -0.306 (-1.598) | -0.122** (-2.107) | -0.148* (-1.872) | -0.169 (-0.815) | -0.215 (-0.789) | 0.009 (0.158) | -0.086 (-1.003) | -0.126** (-2.536) | -0.154** (-2.407) |
| $\Delta x_{i3(t-1)}$ | 0.068 (0.487) | 0.004 (0.017) | 0.141 (1.046) | 0.102 (0.567) | -0.033 (-0.687) | -0.039 (-0.600) | 0.166 (0.786) | 0.134 (0.483) | 0.136** (2.151) | 0.090 (0.967) | 0.113** (2.101) | 0.110 (1.587) |
| Δx_{i4t} ('World Trade') | 0.798*** (2.851) | 0.603 (1.629) | 0.528 (1.461) | 0.506 (1.098) | 0.384 (1.646) | 0.281 (0.941) | 0.023 (0.044) | 0.226 (0.336) | 0.972*** (3.338) | 0.965** (2.533) | 0.576*** (3.018) | 0.559** (2.301) |
| $\Delta x_{i4(t-1)}$ | 0.519* (1.757) | 0.492 (1.253) | -0.138 (-0.357) | -0.159 (-0.319) | -0.015 (-0.055) | -0.156 (-0.438) | 0.037 (0.063) | 0.360 (0.483) | 0.188 (0.573) | 0.108 (0.244) | -0.002 (-0.009) | -0.109 (-0.354) |
| \bar{R}^2 | 0.309 | 0.356 | 0.090 | 0.105 | 0.483 | 0.503 | 0.181 | 0.214 | 0.090 | 0.171 | 0.643 | 0.653 |
| Standard errors | 0.060 | 0.072 | 0.078 | 0.096 | 0.051 | 0.062 | 0.118 | 0.145 | 0.070 | 0.083 | 0.038 | 0.047 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () t-value

In the table, values of \bar{R}^2 from OLS estimate are much larger compared to FGLS estimates for all the routes, which in contrast to the standard errors. This is expected since FGLS is more efficient in SURE model. At least 63.6% of the variables in the GUMS are insignificant except for UK-Canada route estimated by FGLS. The route has the highest \bar{R}^2 and the lowest standard error which showing that 7 out of 11 variables are already significant at 10%. The GUMS also reveals that UK-Sweden and UK-US are initially have the lowest \bar{R}^2 for both estimation method, but the standard errors are quite smaller as compared to large value in UK-Japan.

Then, nine model selection procedures as explained in Section 5.1 are used to find the 'best' model from the initial GUMS. Subsequently, the selected models are employed on the one until three-steps ahead forecast. In order to determine which procedure yield the smaller forecast error, two error measures and two types of statistical tests as mentioned in Section 5.2.

5.3.1 Estimated Models of the Six Routes

This section represents the estimated model of air passengers selected using the nine model selection procedures. *SURE-Autometrics*, *SURE-PcGets* and *SURE-Mine* are implemented on the initial GUMS based on FGLS estimation, whereas *Autometrics*, *Stepwise*, *Mine*, *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE* are applied on initial GUMS that is estimated using OLS. All these procedures used 5% as the significance level. The estimated models are separately shown in Table 5.3 to Table 5.8 according to UK-Germany, UK-Sweden, UK-Italy, UK-Japan, UK-US and UK-Canada route, respectively. The models estimated by *SURE-Autometrics*, *SURE-PcGets*, *SURE-Mine*, *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE* displayed

the FGLS coefficients' estimates. While the OLS coefficients presented by *Autometrics*, *Stepwise* and *Mine*. The models selected by *SURE-PcGets* were re-estimated using FGLS based on selected variables from Ismail (2005) since the study allowed for the exclusion of constant term in the model.

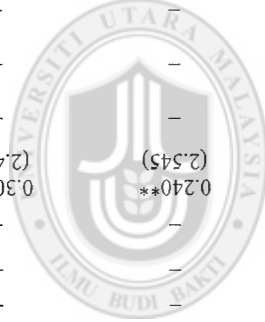
The model selected by *SURE-Autometrics*, *SURE-PcGets* and *SURE-Mine* showed that the p -value of the correlation disturbances amongst the equations using the MC-QLR test are 0.001, 0.007 and <0.001 , respectively. As for both *Autometrics-SURE* and *Stepwise-SURE*, the p -value is 0.010 while *Mine-SURE* resulted in 0.001. These results indicated that these models are more efficient with FGSL estimation. Furthermore, each equation within these models has passed all the diagnostic tests.

The tables also consist of adjusted R^2 and standard error for each of the estimated models. Based on the adjusted R^2 values, models for UK-Canada have the highest values with an average of 0.6, followed by 0.5 in UK-Italy. Amongst the model selection procedures within the UK-Italy though, showed that model estimated by *SURE-PcGets* produced substantially lower than average (0.23). This could be due to number of variables retained in the model since other procedures choose at least three variables, except only one in the model selected by *SURE-PcGets*. Moreover, this procedure is the only one has insignificant variable in the model which is the price (Δx_3) in UK-Sweden and lag one of trade ($\Delta x_{2(t-1)}$) in UK-Japan. It might be due to re-estimation. By focusing on the standard errors of the estimated model, UK-Japan has the highest where all the values were above 10%. While the standard errors for UK-Sweden was close to 10%. Other routes obtained about 5% to 8% for all the nine model selection procedures.

Table 5.3
 Estimated Models for UK-Germany Route

| Model selection procedures | | | | | | | | | |
|----------------------------|------------------|-------------|-----------|------------------|---------------|-----------|-------------|----------|---------|
| Variable | SURE-Autometrics | SURE-PcGets | SURE-Mine | Autometrics-SURE | Stepwise-SURE | Mine-SURE | Autometrics | Stepwise | Mine |
| Δy_{t-1} | - | - | - | - | - | - | - | - | - |
| Δy_{t-2} | - | - | - | - | - | - | - | - | - |
| Δy_{t-3} | - | - | - | - | - | - | - | - | - |
| Δx_{it} (Income) | -0.222** | - | - | - | - | - | - | - | - |
| $\Delta x_{it}(t-1)$ | (-2.683) | - | - | - | - | - | - | - | - |
| Δx_{it} (Trade) | 0.287*** | 0.330*** | 0.304*** | 0.254** | 0.254** | 0.240** | 0.307** | 0.307** | 0.307** |
| Δx_{it} (Price) | (3.298) | (3.047) | (3.251) | (2.589) | (2.589) | (2.545) | (2.444) | (2.444) | (2.444) |
| $\Delta x_{it}^{(2)}(t-1)$ | - | -0.112 | (-2.514) | - | - | - | - | - | - |
| $\Delta x_{it}^{(3)}(t-1)$ | - | (-0.882) | - | - | - | - | - | - | - |
| $\Delta x_{it}^{(World)}$ | 0.732*** | 0.503* | 0.856*** | 0.812*** | 0.812*** | 0.848*** | 0.781** | 0.781** | 0.781** |
| $\Delta x_{it}^{(Trade)}$ | (2.981) | (1.960) | (3.374) | (2.899) | (2.899) | (3.144) | (2.570) | (2.570) | (2.570) |
| $\Delta x_{it}^{(4)}(t-1)$ | - | - | 0.409* | - | - | - | - | - | - |
| \bar{R}^2 | 0.376 | 0.349 | 0.403 | 0.367 | 0.367 | 0.365 | 0.370 | 0.370 | 0.370 |
| Standard errors | 0.067 | 0.068 | 0.064 | 0.068 | 0.068 | 0.068 | 0.071 | 0.071 | 0.071 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () t-value



UUM
 Universiti Utara Malaysia

Estimated Models for UK-Sweden Route

Table 5.4

| Model selection procedures | | | | | | | | | |
|--------------------------------|------------------|-------------|-----------|------------------|---------------|-----------|-------------|----------|---------|
| Variable | SURE-Autometrics | SURE-PCGets | SURE-Mine | Autometrics-SURE | Stepwise-SURE | Mine-SURE | Autometrics | Stepwise | Mine |
| Δy^{t-1} | 0.058*** | 0.055** | 0.041 | 0.041 | 0.041 | 0.041 | 0.041 | 0.041 | 0.041 |
| Δy^{t-2} | 0.271* | 0.334** | - | - | - | - | - | - | - |
| Δy^{t-3} | (1.993) | (2.336) | - | - | - | - | - | - | - |
| Δx^{t-1} (Income) | - | - | - | - | - | - | - | - | - |
| Δx^{t-1} (Trade) | 0.223** | 0.223** | - | - | - | - | - | - | - |
| Δx^{t-1} (Price) | 0.270** | 0.270** | - | - | - | - | - | - | - |
| Δx^{t-1} (Trade) | (2.070) | (2.070) | - | - | - | - | - | - | - |
| Δx^{t-1} (Price) | 0.253** | 0.253** | - | - | - | - | - | - | - |
| Δx^{t-1} (Trade) | (2.373) | (2.373) | - | - | - | - | - | - | - |
| Δx^{t-1} (Price) | (-2.487) | (-2.487) | - | - | - | - | - | - | - |
| Δx^{t-1} (Trade) | - | - | - | - | - | - | - | - | - |
| Δx^{t-1} (World Trade) | - | - | - | - | - | - | - | - | - |
| Δx^{t-1} (World Trade) | 0.901** | 0.901** | 0.901** | 0.901** | 0.901** | 0.901** | 0.901** | 0.901** | 0.901** |
| Δx^{t-1} (World Trade) | (2.561) | (2.561) | (2.532) | (2.532) | (2.532) | (2.522) | (2.423) | (2.423) | (2.423) |
| Δx^{t-1} | - | - | - | - | - | - | - | - | - |
| R^2 | 0.198 | 0.102 | 0.132 | 0.132 | 0.132 | 0.132 | 0.132 | 0.132 | 0.132 |
| Standard errors | 0.084 | 0.091 | 0.092 | 0.092 | 0.092 | 0.092 | 0.092 | 0.095 | 0.095 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () t-value



Table 5.5

Estimated Models for UK-Italy Route

| Variable | Model selection procedures | | | | | | | | |
|---------------------------------|----------------------------|---------------------|-----------------------|-------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | <i>SURE-Autometrics</i> | <i>SURE-PcGets</i> | <i>SURE-Mine</i> | <i>Autometrics-SURE</i> | <i>Stepwise-SURE</i> | <i>Mine-SURE</i> | <i>Autometrics</i> | <i>Stepwise</i> | <i>Mine</i> |
| Constant | 0.035** (2.083) | 0.063*** (4.833) | 0.047*** (2.867) | 0.034* (2.009) | 0.034* (2.009) | 0.034* (2.018) | 0.035* (1.935) | 0.035* (1.935) | 0.035* (1.935) |
| Δy_{it-1} | - | - | - | - | - | - | - | - | - |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - |
| Δy_{it-3} | - | - | - | - | - | - | - | - | - |
| Δx_{1t} (Income) | 0.448*** (5.363) | 0.329*** (3.275) | 0.517*** (6.494) | 0.452*** (5.427) | 0.452*** (5.427) | 0.478*** (5.691) | 0.504*** (4.756) | 0.504*** (4.756) | 0.504*** (4.756) |
| $\Delta x_{1(t-1)}$ | - | - | 0.289** (2.405) | - | - | - | - | - | - |
| Δx_{2t} (Trade) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2(t-1)}$ | -0.298*** (-4.316) | - | -0.496*** (-5.224) | -0.288*** (-4.199) | -0.288*** (-4.199) | -0.295*** (-4.328) | -0.318*** (-3.622) | -0.318*** (-3.622) | -0.318*** (-3.622) |
| Δx_{3t} (Price) | - | - | -0.084* (-2.012) | - | - | - | - | - | - |
| $\Delta x_{3(t-1)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{4t} ('World Trade') | 0.657*** (2.889) | - | 0.514** (2.343) | 0.671*** (2.946) | 0.671*** (2.946) | 0.670*** (2.942) | 0.667** (2.731) | 0.667** (2.731) | 0.667** (2.731) |
| $\Delta x_{4(t-1)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.512 | 0.230 | 0.552 | 0.512 | 0.512 | 0.515 | 0.516 | 0.516 | 0.516 |
| Standard errors | 0.058 | 0.075 | 0.054 | 0.058 | 0.058 | 0.058 | 0.061 | 0.061 | 0.061 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () *t*-value

Table 5.6

Estimated Models for UK-Japan Route

| Variable | Model selection procedures | | | | | | | | |
|----------------------------------|----------------------------|--------------------|-----------------------|-------------------------|----------------------|-----------------------|--------------------|--------------------|----------------------|
| | <i>SURE-Autometrics</i> | <i>SURE-PcGets</i> | <i>SURE-Mine</i> | <i>Autometrics-SURE</i> | <i>Stepwise-SURE</i> | <i>Mine-SURE</i> | <i>Autometrics</i> | <i>Stepwise</i> | <i>Mine</i> |
| Constant | 0.045 (1.374) | 0.100** (2.749) | 0.044 (1.379) | 0.074* (1.785) | 0.074* (1.785) | 0.024 (0.645) | 0.066 (1.508) | 0.066 (1.508) | 0.029 (0.655) |
| Δy_{it-1} | - | - | - | - | - | - | - | - | - |
| Δy_{it-2} | 0.379*** (3.435) | 0.328** (2.325) | 0.383*** (3.560) | - | - | 0.335*** (2.793) | - | - | 0.234 (1.563) |
| Δy_{it-3} | - | - | - | - | - | - | - | - | - |
| Δx_{i1t} (Income) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-1)}$ | 1.168*** (4.361) | - | 0.921*** (3.504) | - | - | 1.012*** (3.736) | - | - | 0.824** (2.414) |
| Δx_{i2t} (Trade) | - | - | 0.303** (2.219) | - | - | - | - | - | - |
| $\Delta x_{i2(t-1)}$ | -0.706*** (-4.153) | -0.113 (-0.701) | -0.657*** (-3.861) | - | - | -0.553*** (-3.232) | - | - | -0.494** (-2.299) |
| Δx_{i3t} (Price) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i3(t-1)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{i4t} ('World Trade') | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i4(t-1)}$ | - | - | - | 1.313** (2.382) | 1.313** (2.382) | 0.533 (1.097) | 1.441** (2.417) | 1.441** (2.417) | 0.909 (1.495) |
| \bar{R}^2 | 0.206 | 0.068 | 0.277 | 0.130 | 0.130 | 0.243 | 0.131 | 0.131 | 0.264 |
| Standard errors | 0.137 | 0.150 | 0.128 | 0.148 | 0.148 | 0.131 | 0.152 | 0.152 | 0.140 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () *t*-value

Table 5.7

Estimated Models for UK-US Route

| Variable | Model selection procedures | | | | | | | | |
|---------------------------------|----------------------------|--------------------|---------------------|-------------------------|----------------------|---------------------|---------------------|---------------------|---------------------|
| | <i>SURE-Autometrics</i> | <i>SURE-PcGets</i> | <i>SURE-Mine</i> | <i>Autometrics-SURE</i> | <i>Stepwise-SURE</i> | <i>Mine-SURE</i> | <i>Autometrics</i> | <i>Stepwise</i> | <i>Mine</i> |
| Constant | 0.031 (1.498) | 0.048** (2.245) | 0.016 (0.782) | 0.031 (1.349) | 0.031 (1.349) | 0.016 (0.770) | 0.029 (1.215) | 0.029 (1.215) | 0.014 (0.622) |
| Δy_{i-1} | - | - | - | - | - | - | - | - | - |
| Δy_{i-2} | - | - | - | - | - | - | - | - | - |
| Δy_{i-3} | - | - | - | - | - | - | - | - | - |
| Δx_{1t} (Income) | - | - | 0.587** (2.124) | - | - | 0.609** (2.073) | - | - | 0.717* (1.786) |
| $\Delta x_{1(t-1)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{2t} (Trade) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2(t-1)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{3t} (Price) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3(t-1)}$ | 0.104** (2.156) | - | 0.149*** (3.354) | - | - | 0.135*** (2.906) | - | - | 0.135** (2.101) |
| Δx_{4t} ('World Trade') | 0.938*** (3.254) | 0.622** (2.268) | 0.960*** (3.424) | 0.911*** (2.968) | 0.911*** (2.968) | 0.948*** (3.411) | 0.943*** (2.927) | 0.943*** (2.927) | 0.929*** (3.007) |
| $\Delta x_{4(t-1)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.247 | 0.165 | 0.298 | 0.191 | 0.191 | 0.301 | 0.191 | 0.191 | 0.302 |
| Standard errors | 0.076 | 0.081 | 0.072 | 0.080 | 0.080 | 0.072 | 0.082 | 0.082 | 0.076 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () *t*-value

Table 5.8

Estimated Models for UK-Canada Route

| Variable | Model selection procedures | | | | | | | | |
|----------------------------------|----------------------------|-----------------------|-----------------------|-------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | <i>SURE-Autometrics</i> | <i>SURE-PcGets</i> | <i>SURE-Mine</i> | <i>Autometrics-SURE</i> | <i>Stepwise-SURE</i> | <i>Mine-SURE</i> | <i>Autometrics</i> | <i>Stepwise</i> | <i>Mine</i> |
| Constant | 0.009 (0.605) | 0.008 (0.549) | 0.008 (0.543) | 0.003 (0.229) | 0.003 (0.229) | 0.001 (0.107) | -0.003 (-0.199) | -0.003 (-0.199) | -0.003 (-0.199) |
| Δy_{it-1} | - | - | - | - | - | - | - | - | - |
| Δy_{it-2} | 0.347*** (3.689) | 0.369*** (3.927) | 0.353*** (4.170) | 0.238** (2.533) | 0.238** (2.533) | 0.247** (2.633) | 0.303** (2.654) | 0.303** (2.654) | 0.303** (2.654) |
| Δy_{it-3} | - | - | - | 0.224** (2.349) | 0.224** (2.349) | 0.214** (2.280) | 0.268** (2.288) | 0.268** (2.288) | 0.268** (2.288) |
| Δx_{it} (Income) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{it(t-1)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{i2t} (Trade) | 0.173*** (2.768) | 0.156** (2.492) | 0.309*** (4.793) | 0.220*** (3.809) | 0.220*** (3.809) | 0.216*** (3.767) | 0.235*** (3.343) | 0.235*** (3.343) | 0.235*** (3.343) |
| $\Delta x_{i2(t-1)}$ | - | - | -0.176*** (-2.897) | - | - | - | - | - | - |
| Δx_{i3t} (Price) | -0.120*** (-2.945) | -0.113*** (-2.769) | - | -0.122*** (-3.316) | -0.122*** (-3.316) | -0.124*** (-3.381) | -0.144*** (-3.197) | -0.144*** (-3.197) | -0.144*** (-3.197) |
| $\Delta x_{i3(t-1)}$ | - | - | 0.112*** (3.005) | - | - | - | - | - | - |
| Δx_{i4t} ('World Trade') | 0.561*** (3.211) | 0.545*** (3.057) | 0.710*** (3.810) | 0.542*** (3.108) | 0.542*** (3.108) | 0.569*** (3.251) | 0.516** (2.575) | 0.516** (2.575) | 0.516** (2.575) |
| $\Delta x_{i4(t-1)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.581 | 0.576 | 0.553 | 0.636 | 0.636 | 0.637 | 0.650 | 0.650 | 0.650 |
| Standard errors | 0.047 | 0.048 | 0.048 | 0.043 | 0.043 | 0.043 | 0.047 | 0.047 | 0.047 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () *t*-value

Basically, *Autometrics* and *Stepwise* have selected similar variables for all routes. The similarities also shared by manual selection which is *Mine* but only for UK-Germany, UK-Sweden, UK-Italy, and UK-Canada. Since *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE* are based on the same principle of single selection, thus the resulted variables in the selected models also possess the similarities. However, the coefficients are different due to estimation are done using the FGLS. Meanwhile, *SURE-Autometrics* and *SURE-PcGets* are able to obtain similar variables only for UK-Canada. There were also variables that were removed from the GUMS by all the model selection procedures. Amongst them are three periods of lags dependent variables in UK-Germany, UK-Italy and UK-US, besides the other three routes removed first period lag.

Afterwards, these models are used to forecast up to three-steps-ahead using the 5 remaining observations in the data (1998 – 2002). For one-step-ahead process, the forecast of 1998 is equivalent to the forecast from the final selected models using data for 1997. Then the data for 1998 is added to the 1962 to 1997 series and the model is re-estimated to obtain the forecast for 1999. The whole process of adding, re-estimating, and forecasting is performed repeatedly until 2002. As for the two-steps ahead, the process started from year 1999 until 2002 while the three-steps-ahead calculated recursively from 2000.

The forecasting accuracies from all procedures are measured by the root mean squares error (RMSE) and the geometric root mean squares error (GRMSE). Both measures were calculated for each route. This study used the median of these measures to represent each model selection procedure. The one up to three-steps-ahead forecasting accuracy

represented by median of RMSE and GRMSE for all model selection procedures are summarised in Table 5.9 and Table 5.10, respectively. The performances are ranked from 1 (smallest error measure) to 9 (largest error measure). Procedures with similar errors received the lowest rank.

Based on RMSE, the first classification of model selection procedure which comprises of *Autometrics*, *Stepwise* and *Mine* outperformed other procedures for one and two-step ahead forecasts. While the best procedure for three-step-ahead was *SURE-PcGets* and *SURE-Autometrics* was rank second placed indicating that simultaneously selection with FGLS estimates were performed better for long forecast horizon. The one and two-step-ahead forecast for both procedures were ranked at the bottom including *SURE-Mine* which using manual selection.

Table 5.9
Air Passengers' Forecasting Performances based on RMSE

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|----------|-----------|----------|-------------|----------|
| | RMSE | Rank | RMSE | Rank | RMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 8.77 | 8 | 9.65 | 9 | 9.35 | 2 |
| 2. <i>SURE-PcGets</i> | 8.71 | 7 | 9.42 | 7 | 9.06 | 1 |
| 3. <i>SURE-Mine</i> | 8.84 | 9 | 9.47 | 8 | 10.55 | 9 |
| 4. <i>Autometrics-SURE</i> | 8.63 | 5 | 9.38 | 5 | 10.21 | 7 |
| 5. <i>Stepwise-SURE</i> | 8.63 | 5 | 9.38 | 5 | 10.21 | 7 |
| 6. <i>Mine-SURE</i> | 8.61 | 4 | 9.37 | 1 | 10.20 | 6 |
| 7. <i>Autometrics</i> | 8.60 | 1 | 9.37 | 1 | 10.19 | 3 |
| 8. <i>Stepwise</i> | 8.60 | 1 | 9.37 | 1 | 10.19 | 3 |
| 9. <i>Mine</i> | 8.60 | 1 | 9.37 | 1 | 10.19 | 3 |

GRMSE values in Table 5.10 however showed the performances of model selection procedures were ranked differently. In general, all the values were smaller than RMSE values suggesting the existence of large or small forecast error amongst the routes. It turns out UK-Japan obtained the largest errors as compared to others. Therefore, performances based on GRMSE are more reliable than RMSE. The best performance obtained by *SURE-PcGets* for short forecast horizon (one and two-steps-ahead) whereas *Autometrics* and *Stepwise* was the best at three-steps-ahead forecast. These findings are contradicted to performances based on RMSE. Moreover, model selection through all non-algorithm procedures were underperformed at all forecast horizon except for two-steps-ahead forecasts where *SURE-Mine* remarkably ranked at second places with only 0.11 differences. *SURE-Autometrics* however was unable to perform well using the GRMSE values.

Table 5.10
Air Passengers' Forecasting Performances based on GRMSE

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|------|-----------|------|-------------|------|
| | GRMSE | Rank | GRMSE | Rank | GRMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 5.15 | 6 | 7.18 | 7 | 8.25 | 6 |
| 2. <i>SURE-PcGets</i> | 4.30 | 1 | 6.87 | 1 | 6.69 | 5 |
| 3. <i>SURE-Mine</i> | 5.64 | 7 | 6.98 | 2 | 8.70 | 8 |
| 4. <i>Autometrics-SURE</i> | 4.92 | 2 | 7.09 | 3 | 6.67 | 3 |
| 5. <i>Stepwise-SURE</i> | 4.92 | 2 | 7.09 | 3 | 6.67 | 3 |
| 6. <i>Mine-SURE</i> | 5.48 | 8 | 7.48 | 8 | 8.53 | 7 |
| 7. <i>Autometrics</i> | 4.99 | 4 | 7.13 | 5 | 6.33 | 1 |
| 8. <i>Stepwise</i> | 4.99 | 4 | 7.13 | 5 | 6.33 | 1 |
| 9. <i>Mine</i> | 5.78 | 9 | 7.50 | 9 | 8.58 | 9 |

Table 5.11

Estimated Models for UK-Germany without UK-Japan Route

| Variable | Model selection procedures | | | | | | | | | |
|-------------------------------------|----------------------------|----------------------|---------------------|---------------------|-----------------------|----------------------|------------------------------------|---------------------|---------------------|---------------------|
| | SURE-Autometrics | | SURE-PcGets | | SURE-Mine | | Autometrics-SURE/ Stepwise-SURE | | Mine-SURE | |
| | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes |
| Constant | 0.012 (0.604) | 0.017 (0.969) | 0.001 (0.028) | 0.020 (1.095) | -0.003 (-0.148) | -0.011 (-0.507) | 0.006 (0.283) | 0.007 (0.342) | 0.005 (0.279) | 0.005 (0.275) |
| Δy_{it-1} | - | - | 0.219 (1.528) | - | - | - | - | - | - | - |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - | - |
| Δy_{it-3} | - | - | - | - | 0.264** (2.754) | - | - | - | - | - |
| Δx_{11t} (Income) | - | - | - | - | - | - | - | - | - | - |
| $\Delta x_{11(t-1)}$ | - | -0.222** (-2.683) | - | - | -0.307*** (-3.458) | - | - | - | - | - |
| Δx_{12t} (Trade) | 0.303*** (2.987) | 0.287*** (3.298) | 0.255** (2.152) | 0.330*** (3.047) | 0.291*** (3.410) | 0.304*** (3.251) | 0.259** (2.627) | 0.254** (2.589) | 0.249** (2.536) | 0.240** (2.545) |
| $\Delta x_{12(t-1)}$ | -0.184** (-2.055) | - | - | - | - | -0.226** (-2.514) | - | - | - | - |
| Δx_{13t} (Price) | - | - | -0.128 (-0.917) | -0.112 (-0.882) | - | - | - | - | - | - |
| $\Delta x_{13(t-1)}$ | - | - | - | - | - | - | - | - | - | - |
| Δx_{14t} ('World Trade') | 0.855*** (3.092) | 0.732*** (2.981) | 0.630** (2.102) | 0.503* (1.960) | 0.797*** (3.070) | 0.856*** (3.374) | 0.827*** (2.929) | 0.812*** (2.899) | 0.837*** (2.964) | 0.848*** (3.144) |
| $\Delta x_{14(t-1)}$ | - | - | - | - | - | 0.409* (2.035) | - | - | - | - |
| \bar{R}^2 | 0.356 | 0.376 | 0.357 | 0.349 | 0.408 | 0.403 | 0.367 | 0.367 | 0.366 | 0.365 |
| Standard errors | 0.068 | 0.067 | 0.066 | 0.068 | 0.064 | 0.064 | 0.068 | 0.068 | 0.068 | 0.068 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () *t*-value

Table 5.12

Estimated Models for UK-Sweden without UK-Japan Route

| Variable | Model selection procedures | | | | | | | | | | | |
|---------------------------------|----------------------------|----------------------|--------------------|--------------------|--------------------|--------------------|------------------------------------|--------------------|--------------------|--------------------|----------------|-----------------|
| | SURE-Autometrics | | SURE-Pc Gets | | SURE-Mine | | Autometrics-SURE/ Stepwise-SURE | | Mine-SURE | | R ² | Standard errors |
| | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | | |
| Constant | 0.041 (1.554) | 0.058*** (2.831) | 0.026 (0.910) | 0.055** (2.572) | 0.041 (1.554) | 0.041 (1.577) | 0.041 (1.532) | 0.041 (1.554) | 0.041 (1.554) | 0.041 (1.568) | 0.132 | 0.092 |
| Δy_{it-1} | - | - | -0.092 (-0.598) | - | - | - | - | - | - | - | 0.132 | 0.092 |
| Δy_{it-2} | - | 0.271* (1.993) | 0.235 (1.607) | 0.334** (2.336) | - | - | - | - | - | - | 0.132 | 0.092 |
| Δy_{it-3} | - | - | - | - | - | - | - | - | - | - | 0.132 | 0.092 |
| Δx_{1t} (Income) | - | - | - | - | - | - | - | - | - | - | 0.132 | 0.092 |
| $\Delta x_{1t}(-1)$ | - | - | - | - | - | - | - | - | - | - | 0.132 | 0.092 |
| Δx_{2t} (Trade) | - | 0.223** (2.070) | - | - | - | - | - | - | - | - | 0.132 | 0.092 |
| $\Delta x_{2t}(-1)$ | - | 0.270** (2.373) | 0.244* (1.952) | 0.253** (2.170) | - | - | - | - | - | - | 0.132 | 0.092 |
| Δx_{3t} (Price) | - | -0.249** (-2.487) | -0.162 (-1.556) | -0.165 (-1.597) | - | - | - | - | - | - | 0.132 | 0.092 |
| $\Delta x_{3t}(-1)$ | - | - | - | - | - | - | - | - | - | - | 0.132 | 0.092 |
| Δx_{4t} ('World Trade') | 0.901** (2.500) | - | 0.805** (2.286) | - | 0.901** (2.500) | 0.901** (2.561) | 0.911** (2.532) | 0.901** (2.500) | 0.901** (2.500) | 0.899** (2.522) | 0.132 | 0.092 |
| $\Delta x_{4t}(-1)$ | - | - | - | - | - | - | - | - | - | - | 0.132 | 0.092 |
| R ² | 0.132 | 0.198 | 0.167 | 0.102 | 0.132 | 0.132 | 0.132 | 0.132 | 0.132 | 0.132 | 0.132 | 0.092 |
| Standard errors | 0.092 | 0.084 | 0.084 | 0.091 | 0.092 | 0.092 | 0.092 | 0.092 | 0.092 | 0.092 | 0.132 | 0.092 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () t-value

Table 5.13

Estimated Models for UK-Italy without UK-Japan Route

| Variable | Model selection procedures | | | | | | | | | | | | | |
|---------------------------------|----------------------------|-----------------------|-----------------------|---------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------------------|-----------------------|-----------------------|------------|--|
| | SURE-Autometrics | | | SURE-PcGets | | | SURE-Mine | | | Autometrics-SURE/ Stepwise-SURE | | | Mine-SURE | |
| | Without UK-Japan | All routes | Without UK-Japan | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | |
| Constant | 0.047*** (2.871) | 0.035** (2.083) | 0.041** (2.535) | 0.063*** (4.833) | 0.047*** (2.867) | 0.046*** (2.823) | 0.034* (1.967) | 0.034* (2.009) | 0.034* (1.996) | 0.034* (2.018) | 0.462*** (5.441) | 0.478*** (5.691) | | |
| Δy_{it-1} | - | - | - | - | - | - | - | - | - | - | - | - | | |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - | - | - | - | | |
| Δy_{it-3} | - | - | - | - | - | - | - | - | - | - | - | - | | |
| Δx_{1t} (Income) | 0.493*** (6.319) | 0.448*** (5.363) | 0.509*** (5.883) | 0.329*** (3.275) | 0.517*** (6.494) | 0.486*** (6.184) | 0.446*** (5.339) | 0.452*** (5.427) | 0.462*** (5.441) | 0.478*** (5.691) | 0.462*** (5.441) | 0.478*** (5.691) | | |
| Δx_{2t} (Trade) | 0.298** (2.472) | - | 0.234 (1.679) | - | 0.289** (2.405) | 0.297** (2.434) | - | - | - | - | - | - | | |
| Δx_{3t} (Price) | -0.513*** (-5.308) | -0.298*** (-4.316) | -0.451*** (-4.165) | - | -0.496*** (-5.224) | -0.484*** (-5.062) | -0.282*** (-4.114) | -0.288*** (-4.199) | -0.292*** (-4.243) | -0.295*** (-4.328) | -0.292*** (-4.243) | -0.295*** (-4.328) | | |
| Δx_{4t} ('World Trade') | -0.102*** (-2.440) | - | -0.086* (-1.802) | - | -0.084* (-2.012) | -0.101** (-2.383) | - | - | - | - | - | - | | |
| Δx_{5t} (World Trade) | 0.515** (2.348) | 0.657*** (2.889) | 0.422* (1.871) | - | 0.514** (2.343) | 0.511** (2.326) | 0.679*** (2.978) | 0.671*** (2.946) | 0.676*** (2.963) | 0.670*** (2.942) | 0.676*** (2.963) | 0.670*** (2.942) | | |
| Δx_{6t} (World Trade) | - | - | - | - | - | - | - | - | - | - | - | - | | |
| \bar{R}^2 | 0.546 | 0.512 | 0.569 | 0.230 | 0.552 | 0.549 | 0.510 | 0.512 | 0.513 | 0.515 | 0.513 | 0.515 | | |
| Standard errors | 0.054 | 0.058 | 0.052 | 0.075 | 0.054 | 0.054 | 0.058 | 0.058 | 0.058 | 0.058 | 0.058 | 0.058 | | |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () t-value

Table 5.14

Estimated Models for UK-US without UK-Japan Route

| Variable | Model selection procedures | | | | | | | | | |
|------------------------------------|----------------------------|---------------------|---------------------|--------------------|---------------------|---------------------|------------------------------------|---------------------|---------------------|---------------------|
| | SURE-Autometrics | | SURE-PcGets | | SURE-Mine | | Autometrics-SURE/ Stepwise-SURE | | Mine-SURE | |
| | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes |
| Constant | 0.030 (1.318) | 0.031 (1.498) | 0.027 (1.226) | 0.048** (2.245) | 0.027 (1.245) | 0.016 (0.782) | 0.029 (1.253) | 0.031 (1.349) | 0.018 (0.824) | 0.016 (0.770) |
| Δy_{it-1} | - | - | - | - | - | - | - | - | - | - |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - | - |
| Δy_{it-3} | - | - | - | - | - | - | - | - | - | - |
| Δx_{1t} (Income) | - | - | - | - | - | 0.587** (2.124) | - | - | 0.537* (1.797) | 0.609** (2.073) |
| $\Delta x_{1(t-1)}$ | - | - | - | - | - | - | - | - | - | - |
| Δx_{2t} (Trade) | - | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2(t-1)}$ | -0.213** (-2.297) | - | - | - | - | - | - | - | - | - |
| Δx_{3t} (Price) | - | - | -0.025 (-0.434) | - | - | - | - | - | - | - |
| $\Delta x_{3(t-1)}$ | - | 0.104** (2.156) | - | - | 0.103** (2.201) | 0.149*** (3.354) | - | - | 0.128** (2.723) | 0.135*** (2.906) |
| Δx_{4t} ('World Trade') | 1.019*** (3.318) | 0.938*** (3.254) | 0.963*** (3.218) | 0.622** (2.268) | 1.020*** (3.427) | 0.960*** (3.424) | 0.943*** (3.020) | 0.911*** (2.968) | 0.953*** (3.327) | 0.948*** (3.411) |
| $\Delta x_{4(t-1)}$ | - | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.193 | 0.247 | 0.195 | 0.165 | 0.249 | 0.298 | 0.191 | 0.191 | 0.297 | 0.301 |
| Standard errors | 0.078 | 0.076 | 0.078 | 0.081 | 0.076 | 0.072 | 0.080 | 0.080 | 0.072 | 0.072 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () *t*-value

Table 5.15

Estimated Models for UK-Canada without UK-Japan Route

| Variable | Model selection procedures | | | | | | | | | |
|----------------------------------|----------------------------|-----------------------|----------------------|-----------------------|-----------------------|-----------------------|------------------------------------|-----------------------|-----------------------|-----------------------|
| | SURE-Autometrics | | SURE-PcGets | | SURE-Mine | | Autometrics-SURE/ Stepwise-SURE | | Mine-SURE | |
| | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes | Without UK-Japan | All routes |
| Constant | 0.002 (0.136) | 0.009 (0.605) | 0.001 (0.098) | 0.008 (0.549) | 0.002 (0.162) | 0.008 (0.543) | -0.001 (-0.098) | 0.003 (0.229) | -0.001 (-0.053) | 0.001 (0.107) |
| Δy_{it-1} | 0.289** (2.374) | - | 0.052 (0.422) | - | 0.306** (2.531) | - | - | - | - | - |
| Δy_{it-2} | 0.320*** (3.369) | 0.347*** (3.689) | 0.340*** (3.064) | 0.369*** (3.927) | 0.321*** (3.404) | 0.353*** (4.170) | 0.271*** (2.799) | 0.238** (2.533) | 0.270** (2.772) | 0.247** (2.633) |
| Δy_{it-3} | - | - | - | - | - | - | 0.248** (2.507) | 0.224** (2.349) | 0.254** (2.548) | 0.214** (2.280) |
| Δx_{i1t} (Income) | - | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-1)}$ | - | - | 0.076 (0.697) | - | - | - | - | - | - | - |
| Δx_{i2t} (Trade) | 0.307*** (4.780) | 0.173*** (2.768) | 0.165** (2.533) | 0.156** (2.492) | 0.316*** (4.945) | 0.309*** (4.793) | 0.219*** (3.673) | 0.220*** (3.809) | 0.226*** (3.775) | 0.216*** (3.767) |
| $\Delta x_{i2(t-1)}$ | -0.221*** (-3.372) | - | - | - | -0.222*** (-3.398) | -0.176*** (-2.897) | - | - | - | - |
| Δx_{i3t} (Price) | - | -0.120*** (-2.945) | -0.116** (-2.768) | -0.113*** (-2.769) | - | - | -0.132*** (-3.449) | -0.122*** (-3.316) | -0.124*** (-3.223) | -0.124*** (-3.381) |
| $\Delta x_{i3(t-1)}$ | 0.172*** (4.161) | - | - | - | 0.172*** (4.200) | 0.112*** (3.005) | - | - | - | - |
| Δx_{i4t} ('World Trade') | 0.541** (2.738) | 0.561*** (3.211) | 0.616*** (2.953) | 0.545*** (3.057) | 0.519** (2.631) | 0.710*** (3.810) | 0.546*** (3.040) | 0.542*** (3.108) | 0.538*** (2.993) | 0.569*** (3.251) |
| $\Delta x_{i4(t-1)}$ | - | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.598 | 0.581 | 0.553 | 0.576 | 0.601 | 0.553 | 0.646 | 0.636 | 0.645 | 0.637 |
| Standard errors | 0.045 | 0.047 | 0.047 | 0.048 | 0.045 | 0.048 | 0.043 | 0.043 | 0.043 | 0.043 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%, () *t*-value

Based on the findings, this study tried to further the analysis by reducing the number of equations. Previous simulation study implied that *SURE-Autometrics* performed well on a small number of multiple equations. Therefore, this data were re-analysed by excluding the UK-Japan route since it obtained the largest forecast errors. The route also has large standard error amongst the estimated models. Besides, it is the only Asian country considered in the data. The new estimated models and forecasting accuracy are explained in the following section.

5.3.2 Air Passengers Flows excluding the UK-Japan Route

Generally, the UK-Japan has the largest standard error for all types of procedures. Therefore this study also assessed the performances of all the nine model selection procedures when the route is removed from the model which leads to specification of five equations model. In this part, model selection using *Autometrics*, *Stepwise* and *Mine* were not re-applied because their selection principle is on the individual equation using OLS estimates. Therefore, the estimated equations for other routes are not affected by the removal of UK-Japan route. On the other hand, *SURE-Autometrics*, *SURE-PcGets*, *SURE-Mine*, *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE* are required to simplify the initial GUMS without the UK-Japan route since they are using FGLS estimates.

The estimated models with or without the route are shown in Table 5.11 until Table 5.15. As expected, *SURE-Autometrics*, *SURE-PcGets* and *SURE-Mine* retained different variables within the equation as changing the number of equations. These procedures excluded or included the variables simultaneously for all the equations based on p-values obtained through FGLS estimation. Thus the selected models are

dissimilar. *SURE-PcGets* attained insignificant variables of lagged one dependent variable for UK-Germany, UK-Italy and UK-Canada. The variable was originally significant in Ismail (2005) study with UK-US and UK-Canada excluded the constant terms. The development of *SURE-Autometrics* followed *Autometrics* where the constant term is not included the reduction processes. Therefore, this study has re-estimated the models selected by *SURE-PcGets* with constant term be in all the equations. Other than that, all the selected models have significant variables at 5% and 1% level of significance. Regrettably, the MC-QLR test for model selected by *SURE-PcGets* is insignificant since the *p*-value equals to 0.145, whereas *SURE-Autometrics*, *SURE-Mine*, *Mine-SURE* and *Autometrics-SURE* obtained significant test with the *p*-value for the latter is 0.003 and the others are 0.001.

Table 5.16
Forecasting Performances (without UK-Japan) based on RMSE

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|-------------------------------|----------|----------|-----------|----------|-------------|----------|
| | RMSE | Rank | RMSE | Rank | RMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 7.90 | 3 | 8.72 | 3 | 9.58 | 2 |
| 2. <i>SURE-PcGets</i> | 7.14 | 1 | 8.11 | 1 | 9.02 | 1 |
| 3. <i>SURE-Mine</i> | 7.79 | 2 | 8.61 | 2 | 9.58 | 2 |
| 4. <i>Autometrics-SURE</i> | 8.33 | 7 | 9.34 | 4 | 9.58 | 2 |
| 5. <i>Stepwise-SURE</i> | 8.33 | 8 | 9.34 | 4 | 9.58 | 2 |
| 6. <i>Mine-SURE</i> | 8.35 | 9 | 9.34 | 4 | 9.58 | 2 |
| 7. <i>Autometrics</i> | 8.32 | 4 | 9.34 | 4 | 9.58 | 2 |
| 8. <i>Stepwise</i> | 8.32 | 4 | 9.34 | 4 | 9.58 | 2 |
| 9. <i>Mine</i> | 8.32 | 4 | 9.34 | 4 | 9.58 | 2 |

Next, the one up to three-steps-ahead forecast were done by the estimated models.

The forecast errors for each route are measured using RMSE and GRMSE. Table 5.16

shows the median of RMSE values across five routes. The model selection procedures are ranked 1 to 9 from smallest to largest value. The lowest rank is given to procedures with similar values. *SURE-PcGets* consistently ranked at 1 for all the forecast horizons. By the exclusion of UK-Japan route, simultaneous model selection with FGLS method using either algorithm or non-algorithm is outperformed other model selection principles. Starting from one up to three-steps-ahead, all the procedures have equivalently performed.

As shown in Table 5.17, the performances of *SURE-Autometrics* has improved by outperformed *SURE-PcGets* on the two and three-steps-ahead forecasts based on GRMSE. Although *SURE-PcGets* is consistently performed well in short term forecast horizon, the procedures ranked at last in the three-steps-ahead forecast which in reverse place of performance using RMSE. Meanwhile, other procedures have equivalent performance since they received similar rank as in RMSE.

Table 5.17

Forecasting Performances (without UK-Japan) based on GRMSE

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|------|-----------|------|-------------|------|
| | GRMSE | Rank | GRMSE | Rank | GRMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 3.65 | 6 | 4.01 | 1 | 6.34 | 1 |
| 2. <i>SURE-PcGets</i> | 2.87 | 1 | 6.65 | 3 | 8.04 | 9 |
| 3. <i>SURE-Mine</i> | 4.54 | 7 | 5.90 | 2 | 7.12 | 2 |
| 4. <i>Autometrics-SURE</i> | 3.50 | 2 | 6.66 | 4 | 7.12 | 2 |
| 5. <i>Stepwise-SURE</i> | 3.50 | 2 | 6.66 | 4 | 7.12 | 2 |
| 6. <i>Mine-SURE</i> | 5.14 | 8 | 7.41 | 9 | 7.12 | 2 |
| 7. <i>Autometrics</i> | 3.60 | 4 | 6.66 | 4 | 7.12 | 2 |
| 8. <i>Stepwise</i> | 3.60 | 4 | 6.66 | 4 | 7.12 | 2 |
| 9. <i>Mine</i> | 5.19 | 9 | 7.40 | 8 | 7.12 | 2 |

The findings in this section demonstrate that forecasting performances of *SURE-Autometrics* has been improved by reducing the number of equations or specifically removing the equation that has poor data quality indicated by the standard error. Further verification is required to validate this by applying the algorithm on four equations since simulation study considered the number on the assessment. Thus, this application will support the conclusion that *SURE-Autometrics* is performing well when number of equations are minimal.

5.3.3 Air Passengers Flows excluding the UK-Japan and UK-Sweden Routes

In this part, *SURE-PcGets* is not involved in the comparison amongst model selection procedures since it was developed for five and six equations only. This time, the UK-Sweden route was excluded from the multiple equations model due to large standard error in the previous estimated model. Similarly, only *SURE-Autometrics*, *SURE-Mine*, *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE* will be affected by the route exclusion.

Table 5.18 and 5.19 indicate the forecasting performances for the eight model selection procedures based on RMSE and GRMSE in the forecast of one up to three-step-ahead. Generally, the GRMSE values are still lower than RMSE. This could be due to UK-US route since the evaluation contains year 2001 where September 11th has happened. This incident really affects the number of passengers' in this route where it has been dropped tremendously according to Figure 5.1 (pg. 91).

Table 5.18

Forecasting Performances (without UK-Japan and UK-Sweden) based on RMSE

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|------|-----------|------|-------------|------|
| | RMSE | Rank | RMSE | Rank | RMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 6.62 | 2 | 7.31 | 1 | 8.67 | 1 |
| 2. <i>SURE-Mine</i> | 6.52 | 1 | 7.47 | 2 | 8.83 | 8 |
| 3. <i>Autometrics-SURE</i> | 6.88 | 4 | 7.63 | 4 | 8.73 | 2 |
| 4. <i>Stepwise-SURE</i> | 6.88 | 4 | 7.63 | 4 | 8.73 | 2 |
| 5. <i>Mine-SURE</i> | 6.87 | 3 | 7.62 | 3 | 8.74 | 4 |
| 6. <i>Autometrics</i> | 6.91 | 6 | 7.66 | 6 | 8.79 | 5 |
| 7. <i>Stepwise</i> | 6.91 | 6 | 7.66 | 6 | 8.79 | 5 |
| 8. <i>Mine</i> | 6.91 | 6 | 7.66 | 6 | 8.79 | 5 |

Table 5.19

Forecasting Performances (without UK-Japan and UK-Sweden) based on GRMSE

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|------|-----------|------|-------------|------|
| | GRMSE | Rank | GRMSE | Rank | GRMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 3.20 | 1 | 3.14 | 1 | 5.73 | 1 |
| 2. <i>SURE-Mine</i> | 4.23 | 6 | 4.75 | 2 | 8.19 | 8 |
| 3. <i>Autometrics-SURE</i> | 3.49 | 2 | 5.03 | 3 | 6.73 | 2 |
| 4. <i>Stepwise-SURE</i> | 3.49 | 2 | 5.03 | 3 | 6.73 | 2 |
| 5. <i>Mine-SURE</i> | 4.33 | 7 | 5.43 | 7 | 7.21 | 6 |
| 6. <i>Autometrics</i> | 3.55 | 4 | 5.10 | 5 | 6.88 | 4 |
| 7. <i>Stepwise</i> | 3.55 | 4 | 5.10 | 5 | 6.88 | 4 |
| 8. <i>Mine</i> | 4.40 | 8 | 5.47 | 8 | 7.35 | 7 |

Based on RMSE, the forecasting performances of *SURE-Autometrics* in four equations model is getting better as compared to five and six equations model. *SURE-Mine* also well performed except for the three-step-ahead forecast. The ranking

indicate that UK-Germany, UK-Italy, UK-US and UK-Canada have better forecasts using FGLS estimation because *Autometrics*, *Stepwise* and *Mine* were unable to outperform other model selection procedures.

Through GRMSE, *SURE-Autometrics* shows an outstanding performance by being at the first rank for all forecast horizons. Meanwhile, *SURE-Mine*, *Mine-SURE* and *Mine* which representing the non-algorithms procedures have not performed as compared to *Autometrics-SURE*, *Stepwise-SURE*, *Autometrics* and *Stepwise*.

Overall, application using the air passengers' flows data has verified the findings in simulation experiments study. *SURE-Autometrics* was able to simplify the GUMS by selected different models from the other model selection procedures. Forecasting using the selected models uncover the outstanding performances of *SURE-Autometrics* in four equations model.

5.4 National Growth Rates Data

In addition to air passengers flows data analysed in previous section, this study also evaluate the forecasting performance of *SURE-Autometrics* using national growth rates data based on the study by Garcia-Ferrer, Highfield, Palm and Zellner (1987). It contains of the annual gross domestic product (GDP, Y_t) from 1951 until 1981 for nine countries. Garcia-Ferrer has provided the data of GDP, real stock return (x_{1t}), 'world' stock return (x_{2t}), and money (M1, x_{3t}). Subsequently, updated data until 2003 was obtained online from International Financial Statistics (IFS) database (www.imf.org/en/Data) only for six countries due to the availability of the data. The countries are Denmark, Ireland, Italy, Netherlands, UK and US. Based on the study, the stock return is defined as the industrial share prices and 'world' stock return is the

median of stock return of the six countries. The GDP measured in constant year was deflated using GDP deflator, whereas the real stock return and the money were deflated using consumer price index (CPI). Similar to the data in Section 5.3, these data too were log transformed and differenced one time to avoid the issues with non-stationarity. The observations from 1952 to 1998 ($T = 47$) are used to fit the estimated models and the remaining 5 observations (1999 – 2003) are used for the validation purpose.

According to Garcia-Ferrer et al. (1987), the estimated models for national growth rates which formulated within the specific-to-general approach, consisted of three lags of the dependent variable, two lags of stock return, one lag of ‘world’ stock return and one lag of money growth rate. Taking this information into consideration including the sample size of 47, the maximum lag length of four for each of the independent variables are chosen in the formulation of the initial model so that it will be as general as possible. However there is a possibility of endogeneity biases exist in the model, but this issue are not discussed in this study since the aim is on the forecasting accuracy amongst different procedures of model selection which is not affected by the endogeneity problem.

Hence, the initial GUMS consists of four lags of dependent variable, three predictors, each with four period of lags with a total of 19 variables in each equation which defined as follows,

$$\Delta y_{it} = \alpha_{i0} + \sum_{j=1}^4 \alpha_{ij} \Delta y_{i(t-j)} + \sum_{k=1}^3 \sum_{j=0}^4 \varphi_{ikj} \Delta x_{ik(t-j)} + \varepsilon_{it} \quad (3.42)$$

where j is the lag length, $i = 1, 2, \dots, 6$ (countries), $t = 1, 2, \dots, T$ (time periods) and Δy_{it} is the growth rate of the GDP in year t for country i . Δx_{ikt} is the growth rate of the k th predictor in year t for country i , ε_{it} are identically independently distributed random errors with mean zero and variance σ^2 . α and ϕ are unknown parameter vectors to be estimated. The initial GUMS is estimated using FGLS whenever *SURE-Autometrics*, *SURE-PcGets* and *SURE-Mine* are applied, whereas other model selection procedures employed OLS method of estimation. Each estimated equation for national growth rates has passes all the diagnostic tests except the heteroscedasticity test since it is unable to compute due to insufficient observations. The p -value of MC-QLR test of contemporaneous correlation disturbances is 0.091 which is significant at 10% level of significance indicating that the seemingly unrelated regression equations (SURE) model is appropriately specified.

Table 5.20 presents the estimated GUMS based on FGLS and OLS method, including the adjusted R square (\bar{R}^2) and standard errors for each equation. In general, the equations within the GUMS estimated by FGLS have smaller values of \bar{R}^2 and standard error as compared to GUMS with OLS estimates. Regardless of whether it is FGLS or OLS method, Denmark has the highest value of \bar{R}^2 whereas US obtained the lowest values for both \bar{R}^2 and standard error. Moreover, both methods showed that only one out of 19 variables is significant at 10% for US model. Generally, the percentages of significant variables in the GUMS with FGLS are much higher than OLS. Both estimation methods indicated that money is significant at 1% for all countries except for the US where only lag one of money is significant.

Table 5.20

Estimated GUMS of National Growth Rates using FGLS and OLS

| Variables | Denmark | | Ireland | | Italy | | Netherland | | UK | | US | |
|----------------------------|-----------|----------|----------|----------|----------|---------|------------|----------|-----------|-----------|----------|---------|
| | FGLS | OLS | FGLS | OLS | FGLS | OLS | FGLS | OLS | FGLS | OLS | FGLS | OLS |
| Constant | -0.015** | -0.016 | 0.006 | 0.006 | 0.039** | 0.034 | 0.005 | 0.004 | 0.004 | 0.003 | 0.033*** | 0.030** |
| Δy_{it-1} | -0.420*** | -0.463** | 0.333** | 0.338 | 0.460*** | 0.430** | -0.060 | 0.047 | 0.108 | 0.073 | -0.206 | -0.114 |
| Δy_{it-2} | -0.045 | -0.126 | -0.051 | -0.109 | -0.114 | -0.066 | -0.187 | -0.131 | 0.151 | 0.218 | -0.084 | 0.027 |
| Δy_{it-3} | 0.164 | 0.274 | -0.041 | 0.032 | 0.040 | 0.009 | 0.330* | 0.302 | -0.234* | -0.167 | 0.005 | -0.019 |
| Δy_{it-4} | 0.180 | 0.235 | 0.244 | 0.221 | -0.243 | -0.105 | 0.007 | 0.071 | -0.437*** | -0.415** | 0.028 | -0.045 |
| Δx_{it} (Stock) | -0.077** | -0.123** | 0.019 | 0.007 | 0.099 | 0.115 | 0.229** | 0.183 | -0.348* | -0.363 | 0.046 | 0.060 |
| $\Delta x_{i1(t-1)}$ | 0.022 | -0.034 | -0.001 | -0.010 | 0.182** | 0.116 | 0.015 | -0.035 | 0.010 | 0.107 | 0.081 | 0.114 |
| $\Delta x_{i1(t-2)}$ | -0.097** | -0.113 | 0.199*** | 0.172** | -0.212** | -0.149 | 0.208* | 0.136 | -0.265 | -0.369 | -0.053 | -0.028 |
| $\Delta x_{i1(t-3)}$ | -0.021 | -0.021 | -0.149** | -0.114 | 0.131 | 0.018 | -0.077 | -0.072 | 0.336* | 0.342 | 0.027 | 0.017 |
| $\Delta x_{i1(t-4)}$ | -0.048 | -0.082 | -0.021 | -0.026 | 0.139 | 0.146 | -0.243* | -0.274 | -0.702*** | -0.676*** | -0.032 | -0.058 |
| Δx_{i2t} ('World') | 0.050 | 0.101 | -0.057 | -0.046 | 0.198 | 0.178 | -0.388* | -0.293 | 0.334* | 0.356 | 0.024 | -0.005 |
| $\Delta x_{i2(t-1)}$ | 0.180* | 0.248* | 0.059 | 0.062 | -0.450** | -0.322 | -0.026 | 0.052 | -0.215 | -0.215 | -0.075 | -0.087 |
| $\Delta x_{i2(t-2)}$ | 0.046 | 0.054 | -0.189* | -0.163 | 0.132 | 0.075 | -0.382* | -0.267 | 0.456** | 0.533** | -0.027 | -0.020 |
| $\Delta x_{i2(t-3)}$ | -0.024 | -0.040 | 0.192* | 0.141 | -0.338* | -0.201 | 0.130 | 0.143 | -0.445** | -0.452* | 0.037 | 0.041 |
| $\Delta x_{i2(t-4)}$ | -0.231** | -0.268* | 0.020 | 0.037 | -0.014 | -0.034 | 0.404* | 0.450 | 0.617*** | 0.544** | -0.001 | -0.003 |
| Δx_{i3t} (Money) | 0.850*** | 0.911*** | 0.694*** | 0.707*** | 0.350*** | 0.399** | 0.895*** | 0.929*** | 0.390*** | 0.429*** | -0.101 | -0.078 |
| $\Delta x_{i3(t-1)}$ | 0.565*** | 0.583*** | -0.314** | -0.293 | 0.113 | 0.060 | 0.168 | 0.061 | -0.121 | -0.124 | 0.348*** | 0.416** |
| $\Delta x_{i3(t-2)}$ | 0.164 | 0.243 | 0.153 | 0.181 | -0.002 | 0.053 | 0.244 | 0.163 | -0.059 | -0.090 | 0.190 | 0.030 |
| $\Delta x_{i3(t-3)}$ | -0.132 | -0.169 | -0.011 | -0.055 | -0.123 | -0.132 | -0.312* | -0.297 | 0.119 | 0.087 | -0.032 | -0.086 |
| $\Delta x_{i3(t-4)}$ | 0.048 | 0.044 | -0.194 | -0.176 | 0.238* | 0.193 | 0.028 | -0.058 | 0.138* | 0.145 | -0.101 | -0.032 |
| \bar{R}^2 | 0.802 | 0.825 | 0.788 | 0.794 | 0.431 | 0.479 | 0.689 | 0.703 | 0.751 | 0.760 | 0.126 | 0.178 |
| Standard errors | 0.034 | 0.044 | 0.031 | 0.042 | 0.063 | 0.082 | 0.045 | 0.060 | 0.040 | 0.054 | 0.017 | 0.022 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%

5.4.1 Estimated Models of National Growth Rates

All the nine procedures were applied to select the 'best' model by simplifying the initial GUMS stated in previous section. The simplified models are shown in Table 5.21 until 5.26 according to Denmark, Ireland, Italy, Netherland, UK, and US country, respectively. Each table corresponds to each country displays estimated equation by *SURE-Autometrics*, *SURE-PcGets*, *SURE-Mine*, *Autometrics-SURE*, *Stepwise-SURE*, *Autometrics*, *Stepwise* and *Mine*, as well as \bar{R}^2 and standard error values. The variables in the model selected by *SURE-PcGets* were obtained from Ismail (2005) and re-estimated because there were equations without constant term within the model. The test of independence amongst the estimated models indicated that FGLS is more appropriate than OLS where the *p*-values of MC-QLR test are 0.037, 0.004, <0.001, 0.081, 0.041 and 0.054, respectively selected by *SURE-Autometrics*, *SURE-PcGets*, *SURE-Mine*, *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE*. However, estimated model of Netherland (0.006) and US (0.002) growth countries selected by *Stepwise* and *Stepwise-SURE* have failed normality assumption on disturbances term.

The equation of Denmark originally has six to seven significant variables estimated either by OLS or FLGS. After the simplification, *SURE-Autometrics* was able to increase the numbers by getting nine variables significant at 5%. Thus, the equation has the highest \bar{R}^2 with the smallest standard error compared to other procedures. *Autometrics*, *Stepwise* and *Mine* selected similar variables, so did *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE* where the estimated values were slightly different since FGLS was used instead of OLS. *SURE-PcGets* however retained only five variables. In this equation, only *SURE-Autometrics* and *SURE-Mine* retained lag one the growth rates.

Table 5.21

Estimated Models of Denmark Growth Rate

| Variable | Model selection procedures | | | | | | | | |
|----------------------------|----------------------------|--------------------|------------------|-------------------------|----------------------|------------------|--------------------|-----------------|-------------|
| | <i>SURE-Autometrics</i> | <i>SURE-PcGets</i> | <i>SURE-Mine</i> | <i>Autometrics-SURE</i> | <i>Stepwise-SURE</i> | <i>Mine-SURE</i> | <i>Autometrics</i> | <i>Stepwise</i> | <i>Mine</i> |
| Constant | -0.012* | -0.003 | -0.015** | -0.010 | -0.010 | -0.010 | -0.015* | -0.015* | -0.015* |
| Δy_{it-1} | -0.367*** | - | -0.362*** | - | - | - | - | - | - |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - |
| Δy_{it-3} | - | - | - | - | - | - | - | - | - |
| Δy_{it-4} | 0.246*** | - | 0.221*** | 0.208*** | 0.202*** | 0.214*** | 0.243*** | 0.243*** | 0.243*** |
| Δx_{it} (Stock) | -0.062** | - | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-1)}$ | - | 0.111*** | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-2)}$ | -0.073*** | -0.083*** | - | -0.094*** | -0.088*** | -0.089*** | -0.115*** | -0.115*** | -0.115*** |
| $\Delta x_{i1(t-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-4)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{i2t} ('World') | - | -0.041 | - | - | - | - | - | - | - |
| $\Delta x_{i2(t-1)}$ | 0.259*** | - | 0.256*** | 0.254*** | 0.272*** | 0.255*** | 0.245*** | 0.245*** | 0.245*** |
| $\Delta x_{i2(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i2(t-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i2(t-4)}$ | -0.190*** | - | -0.193*** | -0.199*** | -0.224*** | -0.205*** | -0.249*** | -0.249*** | -0.249*** |
| Δx_{i3t} (Money) | 0.752*** | 0.711*** | 0.731*** | 0.763*** | 0.765*** | 0.753 | 0.858*** | 0.858*** | 0.858*** |
| $\Delta x_{i3(t-1)}$ | 0.467*** | 0.195*** | 0.459*** | 0.210*** | 0.209*** | 0.218*** | 0.246*** | 0.246*** | 0.246*** |
| $\Delta x_{i3(t-2)}$ | 0.130*** | - | 0.136*** | - | - | - | - | - | - |
| $\Delta x_{i3(t-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i3(t-4)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.828 | 0.745 | 0.791 | 0.814 | 0.814 | 0.812 | 0.826 | 0.826 | 0.826 |
| Standard errors | 0.038 | 0.049 | 0.043 | 0.041 | 0.041 | 0.042 | 0.044 | 0.044 | 0.044 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%

Table 5.22

Estimated Models of Ireland Growth Rate

| Variable | Model selection procedures | | | | | | | | |
|---------------------------|----------------------------|-------------|---------------|----------------------|---------------|-----------|-------------|----------|----------|
| | SURE- Autometrics | SURE-PcGets | SURE- Mine | Autometrics- SURE | Stepwise-SURE | Mine-SURE | Autometrics | Stepwise | Mine |
| Constant | 0.010* | 0.007 | 0.007 | 0.009 | 0.009** | 0.009 | 0.009 | 0.009 | 0.009 |
| Δy_{it-1} | - | 0.227* | 0.304** | - | - | - | - | - | - |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - |
| Δy_{it-3} | - | - | - | - | - | - | - | - | - |
| Δy_{it-4} | - | - | - | - | - | - | - | - | - |
| Δx_{1t} (Stock) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{1t(r-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{1t(r-2)}$ | 0.102*** | 0.109*** | 0.116*** | 0.113*** | 0.110*** | 0.113*** | 0.108*** | 0.108*** | 0.108*** |
| $\Delta x_{1t(r-3)}$ | - | - | -0.058* | - | - | - | - | - | - |
| $\Delta x_{1t(r-4)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{2t} ('World') | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(r-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(r-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(r-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(r-4)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{3t} (Money) | 0.693*** | 0.736*** | 0.705*** | 0.703*** | 0.711*** | 0.698*** | 0.711*** | 0.711*** | 0.711*** |
| $\Delta x_{3t(r-1)}$ | - | -0.205* | -0.234** | - | - | - | - | - | - |
| $\Delta x_{3t(r-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3t(r-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3t(r-4)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.834 | 0.834 | 0.841 | 0.834 | 0.835 | 0.834 | 0.835 | 0.835 | 0.835 |
| Standard errors | 0.036 | 0.035 | 0.034 | 0.036 | 0.036 | 0.036 | 0.037 | 0.037 | 0.037 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%

Table 5.23

Estimated Models of Italy Growth Rate

| Variable | Model selection procedures | | | | | | | | |
|---------------------------|----------------------------|-------------|-----------|------------------|---------------|-----------|-------------|----------|----------|
| | SURE-Autometrics | SURE-PcGets | SURE-Mine | Autometrics-SURE | Stepwise-SURE | Mine-SURE | Autometrics | Stepwise | Mine |
| Constant | 0.028** | 0.027** | 0.028** | 0.028** | 0.028** | 0.027 | 0.030** | 0.030** | 0.030** |
| Δy_{t-1} | 0.232** | 0.245*** | 0.239** | 0.284*** | 0.264*** | 0.265*** | 0.254** | 0.254** | 0.254** |
| Δy_{t-2} | - | - | - | - | - | - | - | - | - |
| Δy_{t-3} | - | - | - | - | - | - | - | - | - |
| Δy_{t-4} | - | - | - | - | - | - | - | - | - |
| Δx_{1t} (Stock) | 0.167*** | 0.166*** | 0.169*** | 0.171*** | 0.181*** | 0.175*** | 0.158*** | 0.158*** | 0.158*** |
| $\Delta x_{1t(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{1t(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{1t(t-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{1t(t-4)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{2t} ('World') | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-4)}$ | - | - | - | - | - | - | - | - | - |
| Δx_{3t} (Money) | 0.456*** | 0.439*** | 0.457*** | 0.468*** | 0.462*** | 0.450*** | 0.549*** | 0.549*** | 0.549*** |
| $\Delta x_{3t(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3t(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3t(t-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3t(t-4)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.588 | 0.584 | 0.589 | 0.593 | 0.590 | 0.588 | 0.603 | 0.603 | 0.603 |
| Standard errors | 0.070 | 0.070 | 0.070 | 0.069 | 0.070 | 0.070 | 0.072 | 0.072 | 0.072 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%

Table 5.24

Estimated Models of Netherland Growth Rate

| Variable | Model selection procedures | | | | | | | | |
|---------------------------|----------------------------|--------------------|------------------|-------------------------|----------------------|------------------|--------------------|-----------------|-------------|
| | <i>SURE-Autometrics</i> | <i>SURE-PcGets</i> | <i>SURE-Mine</i> | <i>Autometrics-SURE</i> | <i>Stepwise-SURE</i> | <i>Mine-SURE</i> | <i>Autometrics</i> | <i>Stepwise</i> | <i>Mine</i> |
| Constant | 0.014 | 0.016* | 0.016* | 0.018** | 0.019** | 0.015* | 0.016* | 0.016* | 0.012 |
| Δy_{it-1} | - | - | - | - | - | - | - | - | - |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - |
| Δy_{it-3} | 0.300** | - | 0.303** | - | - | 0.296** | - | - | 0.296* |
| Δy_{it-4} | - | - | - | - | - | - | - | - | - |
| Δx_{1t} (Stock) | - | 0.199** | - | - | - | - | - | - | - |
| $\Delta x_{1t(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{1t(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{1t(t-3)}$ | - | -0.016 | - | - | - | - | - | - | - |
| $\Delta x_{1t(t-4)}$ | -0.243*** | - | -0.237*** | -0.189** | - | -0.231*** | -0.220** | - | -0.258*** |
| Δx_{2t} ('World') | - | -0.294* | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-3)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{2t(t-4)}$ | 0.415** | - | 0.405** | 0.325** | - | 0.395** | 0.368** | - | 0.434** |
| Δx_{3t} (Money) | 0.851*** | 0.842*** | 0.829*** | 0.849*** | 0.768*** | 0.826*** | 0.906*** | 0.852*** | 0.909*** |
| $\Delta x_{3t(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3t(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{3t(t-3)}$ | -0.312** | - | -0.333** | - | - | -0.306** | - | - | -0.301** |
| $\Delta x_{3t(t-4)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.776 | 0.743 | 0.772 | 0.761 | 0.737 | 0.772 | 0.765 | 0.745 | 0.779 |
| Standard errors | 0.048 | 0.053 | 0.049 | 0.051 | 0.055 | 0.049 | 0.054 | 0.056 | 0.052 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%

Table 5.25

Estimated Models of UK Growth Rate

| Variable | Model selection procedures | | | | | | | | |
|----------------------------|----------------------------|--------------------|------------------|-------------------------|----------------------|------------------|--------------------|-----------------|-------------|
| | <i>SURE-Autometrics</i> | <i>SURE-PcGets</i> | <i>SURE-Mine</i> | <i>Autometrics-SURE</i> | <i>Stepwise-SURE</i> | <i>Mine-SURE</i> | <i>Autometrics</i> | <i>Stepwise</i> | <i>Mine</i> |
| Constant | 0.002 | -0.001 | 0.001 | 0.002 | 0.003 | 0.002 | 0.001 | 0.001 | 0.001 |
| Δy_{it-1} | - | - | - | - | - | - | - | - | - |
| Δy_{it-2} | - | - | - | - | - | - | - | - | - |
| Δy_{it-3} | -0.245** | - | -0.238** | - | - | - | - | - | - |
| Δy_{it-4} | -0.302*** | - | -0.260*** | -0.242*** | -0.255*** | -0.238*** | -0.202** | -0.202** | -0.202** |
| Δx_{i1t} (Stock) | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-3)}$ | - | -0.038 | - | - | - | - | - | - | - |
| $\Delta x_{i1(t-4)}$ | -0.615*** | - | -0.658*** | -0.616*** | -0.595*** | -0.599*** | -0.640*** | -0.640*** | -0.640*** |
| Δx_{i2t} ('World') | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i2(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i2(t-2)}$ | 0.229*** | - | 0.232*** | 0.212** | 0.202** | 0.201** | 0.207** | 0.207** | 0.207** |
| $\Delta x_{i2(t-3)}$ | -0.098 | - | - | - | - | - | - | - | - |
| $\Delta x_{i2(t-4)}$ | 0.485*** | - | 0.487*** | 0.504*** | 0.482*** | 0.484*** | 0.512*** | 0.512*** | 0.512*** |
| Δx_{i3t} (Money) | 0.339*** | 0.458*** | 0.349*** | 0.356*** | 0.352*** | 0.361*** | 0.394*** | 0.394*** | 0.394*** |
| $\Delta x_{i3(t-1)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i3(t-2)}$ | - | - | - | - | - | - | - | - | - |
| $\Delta x_{i3(t-3)}$ | 0.203*** | - | 0.197*** | - | - | - | - | - | - |
| $\Delta x_{i3(t-4)}$ | - | - | - | - | - | - | - | - | - |
| \bar{R}^2 | 0.767 | 0.633 | 0.766 | 0.765 | 0.762 | 0.765 | 0.769 | 0.769 | 0.769 |
| Standard errors | 0.047 | 0.064 | 0.048 | 0.049 | 0.050 | 0.049 | 0.053 | 0.053 | 0.053 |

*** Significant at 1%, ** Significant at 5%, * Significant at 10%

As for Ireland, all the procedures selected lagged two of stock return and money except that *SURE-PcGets* and *SURE-Mine* have retained additional variables which are lagged one of dependent variable and lagged one of money. *SURE-Mine* also included lagged three of stock return which only significant at 10%. The values of \bar{R}^2 and standard errors in this equation were quite similar for all procedures.

Regardless differences in all procedures, the estimated growth rates for Italy revealed an exact variables where \bar{R}^2 is 60.3% and standard errors is 0.072 for *Autometrics*, *Stepwise* and *Mine*. These values are somewhat higher than others because OLS estimates. The selected variables are lagged one of growth rates, the stock return and the money.

Stepwise and *Stepwise-SURE* only has retained the money when estimating the growth rates for Netherland with high \bar{R}^2 which is 74.5% and 73.7%, respectively. This indicated that the money was an important variable. Other procedures also retained this variable with additional variables such as lagged four of the stock return and lagged four of the 'world' stock return. Both *SURE-Autometrics* and *SURE-PcGets* included the basic of these variables in the equation. Amongst procedures that employed FGLS estimates, *SURE-Autometrics* obtained largest \bar{R}^2 (77.6%) and smallest standard error (0.048).

The equation in estimating of UK growth rate were equivalent when using *Autometrics*, *Stepwise* and *Mine* since they selected lagged four of dependent variable, lagged four of stock return, lagged two and four of 'world' stock return and the money. *SURE-Autometrics* and *SURE-Mine* also choose these variables with the

addition of lagged three of dependent variable and lagged three of the money. The estimated growth rate for US country failed the Chow test at 1% level of significance. Therefore *Autometrics* included the insignificant lagged one of dependent variable to maintain the congruency which is the reason of *Autometrics-SURE* also has this variable. However, *SURE-Mine* included the stock return which significant at 10% and passed all the diagnostic test. *SURE-Autometrics* did not retain any of these variables because the algorithm has tolerance for the Chow test at 0.5%.

Overall, the estimated equation for US growth rate obtained not only lowest standard error but also smallest \bar{R}^2 value. The most reliable variables in estimating the growth rates for any countries are the stock return and the money in spite of model selection procedures implemented.

5.4.2 Forecast Accuracy of National Growth Rates

Subsequently, the estimated models in Section 5.4.1 were employed for the forecast of one up to three-steps-ahead. The performances of all the model selection procedures were compared using the median of RMSE and GRMSE across equations to represent each procedure. Table 5.27 shows the percentages of forecasts error measured by RMSE for one until three-step-ahead forecasts. The values are ranked from 1 (the smallest) to 9 (the largest) to indicate the forecasting performances for the procedures where similar values received the lowest rank.

The RMSE for one-step-ahead forecasts shows that the model selected simultaneously using FGLS estimation such as *SURE-PcGets*, *SURE-Autometrics* and *SURE-Mine* outperformed others where the former is the best procedure. The performances are

similar for two-steps-ahead forecasts with *SURE-Mine* shifted the rank with *SURE-PcGets* to be the best procedure and sustained it up to three-steps-ahead forecasts. Regardless the forecasts horizon, *SURE-Autometrics* maintained the performances at the second placed. *SURE-PcGets* however were distinctively underperformed when forecasted at three-steps-ahead since it is ranked at 9.

Table 5.27

National Growth Rates' Forecasting Performances based on RMSE

| Model Selection Procedures | One-Step | | Two-Step | | Three-Step | |
|----------------------------|----------|------|----------|------|------------|------|
| | RMSE | Rank | RMSE | Rank | RMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 5.11 | 2 | 5.42 | 2 | 4.95 | 2 |
| 2. <i>SURE-PcGets</i> | 4.67 | 1 | 5.59 | 3 | 6.24 | 9 |
| 3. <i>SURE-Mine</i> | 5.24 | 3 | 5.04 | 1 | 4.85 | 1 |
| 4. <i>Autometrics-SURE</i> | 5.89 | 7 | 6.35 | 7 | 6.02 | 8 |
| 5. <i>Stepwise-SURE</i> | 6.04 | 9 | 6.42 | 9 | 5.95 | 7 |
| 6. <i>Mine-SURE</i> | 5.28 | 4 | 5.79 | 4 | 5.39 | 4 |
| 7. <i>Autometrics</i> | 5.83 | 6 | 6.30 | 6 | 5.93 | 6 |
| 8. <i>Stepwise</i> | 6.00 | 8 | 6.41 | 8 | 5.92 | 5 |
| 9. <i>Mine</i> | 5.29 | 5 | 5.79 | 4 | 5.36 | 3 |

As described in Section 5.1, all these procedures are classified into four categories where the last belongs to manually selection procedures. By disregarding the best three ranked forecasting performances based on RMSE, the manual selection procedures was better than the algorithm procedures. Amongst individually selection procedure where OLS is the estimation method, *Mine* outperformed *Autometrics* and *Stepwise* for all forecasting horizon where the rank is 5, 4 and 3, respectively. This likeness also occurs within the second model selection classification which

comprises of *Autometrics-SURE* and *Stepwise-SURE*. Forecast of one up to three-step-ahead reveals that *Mine-SURE* (rank 4) consistently outperformed these procedures.

Table 5.28

National Growth Rates' Forecasting Performances based on GRMSE

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|------|-----------|------|-------------|------|
| | GRMSE | Rank | GRMSE | Rank | GRMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 4.33 | 7 | 4.20 | 3 | 3.80 | 4 |
| 2. <i>SURE-PcGets</i> | 3.52 | 2 | 4.41 | 5 | 5.39 | 9 |
| 3. <i>SURE-Mine</i> | 3.81 | 3 | 3.57 | 1 | 3.23 | 1 |
| 4. <i>Autometrics-SURE</i> | 4.13 | 5 | 5.07 | 9 | 4.86 | 8 |
| 5. <i>Stepwise-SURE</i> | 4.27 | 6 | 5.03 | 8 | 4.53 | 6 |
| 6. <i>Mine-SURE</i> | 3.50 | 1 | 4.14 | 2 | 3.68 | 2 |
| 7. <i>Autometrics</i> | 4.44 | 8 | 4.97 | 6 | 4.79 | 7 |
| 8. <i>Stepwise</i> | 4.44 | 8 | 4.97 | 6 | 4.48 | 5 |
| 9. <i>Mine</i> | 3.95 | 4 | 4.23 | 4 | 3.75 | 3 |

Table 5.28 shows the summaries of one up to three-step-ahead forecasting performances measured by GRMSE. Unlike RMSE, the best procedure measured using GRMSE is the *Mine-SURE* for one-step forecast, and *SURE-Mine* for two and three-step-ahead forecast. Although the values are slightly lower, the *SURE-Mine* obtained similar ranks as in RMSE. The *SURE-Autometrics* however has performed very poorly (rank 7) at one-step-ahead forecasts. By chance, the difference with *SURE-Mine* is insignificant using the modified Diebold-Mariano test ($S_1^* = 0.6895$). The performance of *SURE-Autometrics* for two and three-step-ahead also decline but it is not as tremendously bad as one-step-ahead forecast. The procedure is ranked at 3

and 4, respectively as compared ranked 2 using RMSE. The three-steps-ahead forecasts also reveals that *SURE-PcGets*, *Autometrics-SURE*, *Stepwise* and *Mine* has equally ranked for both error measures where the former at the bottom.

Again, the manual selection procedures outperformed algorithm procedures on the individual selection using OLS and initially individual selection with FGLS estimates for the final model. Particularly, *Mine* outperformed *Autometrics* and *Stepwise*, whereas *Mine-SURE* outperformed *Autometrics-SURE* and *Stepwise-SURE* for all forecasting horizon.

Table 5.29

Measures of Forecast Error for Italy Growth Rate

| Model Selection Procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | RMSE | GRMSE | RMSE | GRMSE | RMSE | GRMSE |
| 1. <i>SURE-Autometrics</i> | 2.52 | 1.02 | 0.89 | 0.53 | 0.60 | 0.40 |
| 2. <i>SURE-PcGets</i> | 2.56 | 0.98 | 0.89 | 0.50 | 0.58 | 0.38 |
| 3. <i>SURE-Mine</i> | 2.56 | 1.03 | 0.89 | 0.53 | 0.60 | 0.40 |
| 4. <i>Autometrics-SURE</i> | 2.63 | 0.98 | 0.88 | 0.47 | 0.58 | 0.36 |
| 5. <i>Stepwise-SURE</i> | 2.62 | 1.00 | 0.87 | 0.51 | 0.57 | 0.39 |
| 6. <i>Mine-SURE</i> | 2.59 | 1.01 | 0.88 | 0.52 | 0.58 | 0.39 |
| 7. <i>Autometrics</i> | 2.69 | 0.83 | 1.01 | 0.58 | 0.72 | 0.22 |
| 8. <i>Stepwise</i> | 2.69 | 0.83 | 1.01 | 0.58 | 0.72 | 0.22 |
| 9. <i>Mine</i> | 2.69 | 0.83 | 1.01 | 0.58 | 0.72 | 0.22 |

By using this data, all model selection procedures were able to select the stock return and money to estimate the growth rate of Italy. Hence Table 5.29 shows the error measures of one up to three-step-ahead forecast specifically for this country. On the

basis of RMSE, *SURE-Autometrics* (2.52) has the lowest error only for short horizon, while *Stepwise-SURE* performed on the other horizons. Meanwhile, GRMSE indicated that all the single model selection procedures with OLS estimation which are *Autometrics*, *Stepwise* and *Mine* are the best procedures for one and three-step-ahead forecasts.

Based on the forecasting analysis, *SURE-Autometrics* was able to perform quite well for all forecasts horizon on the basis of RMSE. However, GRMSE indicated that it has underperformed. Previous results also implied that the performances of algorithm become better as the number of multiple equations decreased. Additionally, growth rate for Italy suggested that it has the best forecast accuracy when the model is selected by single equation procedure with OLS estimation method. Therefore, the data are re-analysed using five equations where Italy has been chosen to be excluded from the group.

However, only eight procedures of model selection were involved since *SURE-PcGets* only applied on six equations model. Subsequently, the selected models were used to forecast the countries' growth rate for one up to three-step-ahead. The forecast errors from each country are measured using RMSE and GRMSE. The median of these values are shown in Table 5.30 and 5.31, respectively. Similarly, the forecasting performances of the procedures are ranked according smallest to highest value.

Table 5.30 indicated that *Mine-SURE* is the best model selection procedure for one-step-ahead forecast, whereas both two and three-step-ahead are best forecasted by model selected by *SURE-Mine*. *SURE-Autometrics* received the second rank for all

forecast horizons. This implied that modelling with or without Italy does not affect the forecasting performances of model selected by *SURE-Autometrics* since the ranked remain unchanged for both multiple equations models. RMSE also showed that procedures using FGLS is outperformed model with OLS estimates.

Table 5.30

Forecasting Performances (without Italy) based on RMSE

| Model selection procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|----------|-----------|----------|-------------|----------|
| | RMSE | Rank | RMSE | Rank | RMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 5.76 | 2 | 6.36 | 2 | 5.74 | 2 |
| 2. <i>SURE-Mine</i> | 5.76 | 2 | 4.92 | 1 | 4.71 | 1 |
| 3. <i>Autometrics-SURE</i> | 5.82 | 5 | 6.38 | 3 | 6.16 | 8 |
| 4. <i>Stepwise-SURE</i> | 6.12 | 7 | 6.41 | 5 | 6.04 | 6 |
| 5. <i>Mine-SURE</i> | 5.68 | 1 | 6.38 | 3 | 5.81 | 4 |
| 6. <i>Autometrics</i> | 5.86 | 6 | 6.42 | 6 | 6.11 | 7 |
| 7. <i>Stepwise</i> | 6.20 | 8 | 6.42 | 6 | 6.08 | 5 |
| 8. <i>Mine</i> | 5.80 | 4 | 6.42 | 6 | 5.75 | 3 |

Even with GRMSE as shown in Table 5.31, *Mine-SURE* once more outperformed other model selection procedures for one-step-ahead including the three-step-ahead forecast. While the best performance of forecast at two-step-ahead still retained by *SURE-Mine* and *SURE-Autometrics* maintained at the second place. By using the GRMSE, the ranked of *SURE-Autometrics* has declined from the second to fourth place for one-step-ahead and third place for three-step-ahead forecast. These results revealed that manual selection procedures outperformed algorithm procedures where *Mine* has the third rank for one and two-step-ahead forecasts.

Table 5.31

Forecasting Performances (without Italy) based on GRMSE

| Model selection procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|----------|-----------|----------|-------------|----------|
| | GRMSE | Rank | GRMSE | Rank | GRMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 4.50 | 4 | 4.26 | 2 | 3.86 | 3 |
| 2. <i>SURE-Mine</i> | 3.88 | 2 | 3.96 | 1 | 3.79 | 2 |
| 3. <i>Autometrics-SURE</i> | 5.02 | 7 | 5.62 | 5 | 5.96 | 8 |
| 4. <i>Stepwise-SURE</i> | 5.06 | 8 | 5.63 | 6 | 5.28 | 5 |
| 5. <i>Mine-SURE</i> | 3.72 | 1 | 4.35 | 4 | 3.71 | 1 |
| 6. <i>Autometrics</i> | 4.95 | 5 | 5.65 | 7 | 5.93 | 7 |
| 7. <i>Stepwise</i> | 4.95 | 5 | 5.65 | 7 | 5.32 | 6 |
| 8. <i>Mine</i> | 3.98 | 3 | 4.30 | 3 | 3.87 | 4 |

Since *SURE-Autometrics* was unable to show an outstanding performance, the number of equations has to be reduced. As revealed by air passengers' data and simulation results, the algorithm performed well on four equations model. Hence, another country is excluded from the group of countries. Initially, Denmark was removed due to large forecast errors amongst other countries. However the MC-QLR test indicated that the contemporaneous correlation disturbances amongst equations are insignificant which lead to inappropriate of SURE specification. Thus Denmark is added back to the model. Subsequently, US country is removed since the estimated model has the lowest \bar{R}^2 value and the forecast errors are extremely lower as compared to other countries. Moreover the MC-QLR test showed significant result of dependencies between disturbances term.

Eight model selection procedures were employed in finding the best model representing the growth rate of Denmark, Ireland, Netherland and UK. The

procedures are listed in Table 5.32 as well as the error measure of one until three-step-ahead forecasted from the selected models. The values represent the median of RMSE between the four countries. The performance of each procedure is ranked for 1 (smallest) to 8 (largest) where ties received the lowest rank.

Table 5.32

Forecasting Performances (without Italy and US) based on RMSE

| Model selection procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|------|-----------|------|-------------|------|
| | RMSE | Rank | RMSE | Rank | RMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 6.50 | 6 | 6.60 | 8 | 6.24 | 2 |
| 2. <i>SURE-Mine</i> | 5.94 | 1 | 5.67 | 1 | 5.71 | 1 |
| 3. <i>Autometrics-SURE</i> | 6.41 | 3 | 6.39 | 3 | 6.43 | 5 |
| 4. <i>Stepwise-SURE</i> | 6.60 | 7 | 6.40 | 4 | 6.43 | 5 |
| 5. <i>Mine-SURE</i> | 6.33 | 2 | 6.37 | 2 | 6.29 | 3 |
| 6. <i>Autometrics</i> | 6.48 | 5 | 6.53 | 5 | 6.54 | 8 |
| 7. <i>Stepwise</i> | 6.65 | 8 | 6.53 | 5 | 6.53 | 7 |
| 8. <i>Mine</i> | 6.45 | 4 | 6.53 | 5 | 6.36 | 4 |

Forecasting performances for multiple equations model without Italy and US showed *SURE-Mine* has outperformed other model selection procedures for all forecast horizons. Besides that, *Mine-SURE* keep up the good performances by receiving second rank for both one and two-step-ahead forecasts. *SURE-Autometrics* though was only able to sustain the good performance for forecast at three-step-ahead. The algorithm obtained sixth place for the short forecast horizon and ranked at last for two-step-ahead forecast indicating poor performance so far as compared to previous results.

Table 5.33

Forecasting Performances (without Italy and US) based on GRMSE

| Model selection procedures | One-Step | | Two-Steps | | Three-Steps | |
|----------------------------|----------|------|-----------|------|-------------|------|
| | GRMSE | Rank | GRMSE | Rank | GRMSE | Rank |
| 1. <i>SURE-Autometrics</i> | 3.85 | 1 | 5.11 | 4 | 4.71 | 1 |
| 2. <i>SURE-Mine</i> | 3.88 | 2 | 4.85 | 1 | 4.94 | 2 |
| 3. <i>Autometrics-SURE</i> | 5.29 | 8 | 5.74 | 7 | 6.08 | 7 |
| 4. <i>Stepwise-SURE</i> | 5.28 | 7 | 5.59 | 5 | 5.78 | 5 |
| 5. <i>Mine-SURE</i> | 4.29 | 3 | 4.94 | 2 | 4.97 | 3 |
| 6. <i>Autometrics</i> | 5.04 | 5 | 5.81 | 8 | 6.14 | 8 |
| 7. <i>Stepwise</i> | 5.04 | 5 | 5.66 | 6 | 5.84 | 6 |
| 8. <i>Mine</i> | 4.46 | 4 | 4.97 | 3 | 5.11 | 4 |

Table 5.33 shows the forecasting performances of four countries estimated by model selected by the eight procedures using median of GRMSE. Ultimately, *SURE-Autometrics* was able to be the best procedure in one and three-steps-ahead forecasts but the performance was ranked at fourth for two-steps-ahead forecast. Nonetheless, *SURE-Mine* again indicates a good performance since it received only first or second ranked.

5.5 Summary of Findings

Both data used in this study have six equations represented by routes in the air passengers' flows data and countries in national growth rates data. The routes were UK-Germany, UK-Sweden, UK-Italy, UK-Japan, UK-US and UK-Canada. Meanwhile the countries were Denmark, Ireland, Italy, Netherlands, UK and US. The initial GUMS of air passengers has 11 variables with 33 observations while GUMS of growth rates has 19 variables with 47 observations. Nine model selection procedures

were applied to simplify the GUMS. The procedures are *Autometrics*, *Stepwise*, *Autometrics-SURE*, *Stepwise-SURE*, *SURE-Autometrics*, *SURE-PcGets*, *Mine*, *Mine-SURE*, and *SURE-Mine*.

For air passengers' model selection, *Autometrics* and *Stepwise* has similarly selected variables in the equation for all the routes. Therefore, *Autometrics-SURE* and *Stepwise-SURE* also has similar variables with different estimated coefficients. They shared similar principle in the selection procedures but the final model is estimated using different method. In the same way, *Mine* and *Mine-SURE* has similar variables except the estimates. Meanwhile, *SURE-Autometrics*, *SURE-PcGets* and *SURE-Mine* has selected different variables for each route. Only model selected by *SURE-PcGets* contained insignificant variables. The forecasts using the selected models revealed RMSE values were larger than GRMSE. The reason could be due to the presence of year 2001 in the out-sample data where September 11th incident had happened. As concurred with Armstrong and Fildes (1995), RMSE does greatly affected by extreme values. Consequently, GRMSE is preferred in indicating the forecast accuracy. However, other procedures outperformed *SURE-Autometrics*. It was presumed since the simulation study already showed poor performance in six equations model.

Following Ismail (2005), the number of equations reduced to five where UK-Japan was removed due to highest standard error. The estimated models using *Autometrics*, *Stepwise* and *Mine* are still similar because they were selected individually. Thus the removal of equation does not affect the estimated model, whereas the estimated values were affected in models selected by *Autometrics-SURE*, *Stepwise-SURE* and *Mine-SURE* since it used FGLS method. Other procedures have selected models

differently as the number of equations changed. The forecasting accuracy for *SURE-Autometrics* did improved a lot by obtained the smallest GRMSE values in two and three-step-ahead forecast. In conjunction with simulation study, forecasting accuracy for four equations model was measured by eliminate the UK-Sweden route. As verifying the study, *SURE-Autometrics* showed an outstanding performance in the all forecasts horizon based on GRMSE. Without both routes, RMSE also conformed *SURE-Autometrics* forecasting accuracy.

The application on national growth rates data is to validate the forecasting performances for model selected from large set of initial GUMS. The findings were not comparable as air passengers' data. Even though the procedures are different in the selection principle, surprisingly they were able to obtain similar variables for Italy growth rates. *Autometrics* and *Autometrics-SURE* retained one insignificant variable to maintain the congruency of in the US model. Unlike previous data, *Autometrics* and *Stepwise* included different variables in the growth rates of Netherland and US. As for the forecasting accuracy, the values of RMSE and GRMSE were different only slightly. *SURE-Autometrics* was performed well by ranked at the second place for all forecast horizon based on RMSE values. However, these findings were not supported by GRMSE values which showed severe performance in one-step-ahead forecast. The accuracy increased for two and three-step-ahead forecasts.

As expected, *SURE-Autometrics* was performed well when the number of equations is reduced to five and four by the exclusion of Italy followed by US. Even though the algorithm has not produced the lowest RMSE value, it still in the top three ranked. By using GRMSE indicator, *SURE-Autometrics* has outstanding performances when

there are four equations in the model. However both error measures showed that the forecasts are more accurate with model estimated by FGLS method and selected using manual procedures where *SURE-Mine* and *Mine-SURE* maintained the top three performances' ranked.

The national growth rates data revealed that manual procedures outperformed algorithm procedures could be due to failure in several diagnostic tests whenever a certain variable is excluded from the equations. The tests that failed are either normality test or heteroscedasticity test or Chow test. By using the manual procedures, these problems can be monitored closely.

From the above discussion, it can be summarised that *SURE-Autometrics* appears to perform well on model with four equations. Nevertheless, the performances depend on whether a relatively good quality data are used. This suggests that the formulation of initial GUMS also plays an important role in the success of this procedure, confirming Hendry and Doornik (2009) conclusion about *Autometrics*.

CHAPTER SIX

CONCLUSION AND FUTURE RESEARCH

6.1 Conclusion

Common practices in the model formulation involved human intervention. The specification of model such as inclusion of variables is decided through theories and personal knowledge. This situation indicates the existence of tacit knowledge which can be gained through research experiences and cannot be articulated. Consequently the model building becomes a challenging practice especially for non-experts due to inexperience in research studies and lack of statistical knowledge. After years of research on the area of automatic modelling, Hendry and Doornik (2014) has concluded that the automatic modeller using model selection technique would be able to obtain a better model as compared to human modeller. This automatic modelling tool relies on algorithm that provides a step by step guidance in the model selection processes which will lessen the role of tacit knowledge. Regardless the experience and expertise, modellers are able to practice in a wide range of research contexts by employing the algorithms.

Most of the existing algorithms focus on the modelling of single equation. Thus, this study has explored the possibility of model selection for multiple equations using automatic approach by algorithm. Hence, the development of an algorithm known as *SURE-Autometrics* has been established. The selection strategy is on the basis of general-to-specific (GETS) modelling approach. Particularly, the algorithm is developed to select the 'best' model for seemingly unrelated regression equations (SURE).

The selection procedure begins with the formulation of general model consists of all the possible predictors which then is reduced to a parsimonious model. Since there are many types of multiple equations, this study highlighted the seemingly unrelated regression equations (SURE). This type of model has one of the properties in model selection algorithm within GETS approach where each equation should be congruent. In other words, there is no violation in any assumptions regarding the disturbances for each equation. Thus the disturbances are identically independent normally distributed (i.i.d) with homoscedasticity variance. However, these multiple equations are correlated to each other through the disturbances where the feasible generalised least squares (FGLS) method of estimation is more efficient rather than ordinary least squares (OLS). Model selection using OLS estimates require the equations to be selected individually.

Basically, the new algorithm was developed using the search strategy in *Autometrics* algorithm with additional dependence test of correlation disturbances in the series of diagnostic tests for model adequacy checking analysis. The inclusion of dependence test is to ensure that the model is appropriate for SURE specification. Otherwise, the equations are better using single equation model selection such as *Autometrics*. Therefore, while maintaining the search method in *Autometrics*, the new algorithm replaced the single equation estimation (OLS) with multiple equations estimation (FGLS), besides the inclusion of independence test in the diagnostic checking procedure. Hence, the name of the new algorithm is *SURE-Autometrics*.

The *SURE-Autometrics* has five development phases. The first phase deals with the formulation of an initial specification of the general unrestricted model (GUMS), and

then followed by the second phase which can be considered as pre-search reduction process. At this phase, the highest insignificant variables are deleted to reduce the complexities of variables in the initial GUMS. Third phase is the procedure of finding the all possible simplified models from the GUMS after pre-search. The tree search strategy is used to minimize the computational efficiency so that only terminal models are considered instead of all the possible models. The terminal is defined as a model with significant variables, valid reduction from the GUMS and each equation is congruent. The fourth phase is similar to the third by reiterated the tree search strategy to produce more terminal models. Finally, the last phase is dedicated to select the 'best' final simplified model amongst all the terminal models resulted from the third and fourth phase based in information criteria. The selected final model is denoted as the specific unrestricted model (SUMS). The fully developed algorithm is then transformed into a computer programming code via *GAUSS* (version 9.0) language.

Afterwards, the performance of *SURE-Autometrics* was assessed through simulation experiments. The algorithm was replicated 100 times for each of experimental conditions using artificial data that were generated based on true specification models and requires the usage of real and simulated data. Since the data-generating process is known, the performances are measured by the percentages of similarities in the inclusion of variables between selected model and true model. This measure also equivalent with the probability of *SURE-Autometrics* finds the correct specification. There are five correct specification models which are described as follows,

- S1: Each equation consists of constant term and disturbances where the constant is estimated by the mean of dependent variable and the coefficient for disturbances represent by the standard deviation of dependent variable.
- S2: Each equation has constant term, one lag of dependent variable and disturbances where the coefficients' are estimated by FGLS and disturbances' coefficient is the standard error of each estimated model.
- S3: Each equation has constant term, one lag dependent variable, one independent variable including the lag one variable and disturbances. The coefficients' associated with the variables are estimated by FGLS and coefficient's for disturbances is represented by the standard error of model.
- S4: Each equation is specified similarly to S3 except that different independent variable is considered. The coefficients' associated with the variables are estimated by FGLS and coefficient's for disturbances is represented by the standard error of model.
- S5: Each equation is specified by the union of S3 and S4 for each equation. The coefficients' associated with the variables are estimated by FGLS and coefficient's for disturbances is represented by the standard error of model.

The values of disturbances in the true specifications were simulated using standard normal distribution and were allowed to correlate amongst the equations from weakest to strongest strength.

During the simulation of *SURE-Autometrics*, initial GUMS is formulated according to two sets where small set contains 18 variables and large set comprises of 39 variables. Subsequently, *SURE-Autometrics* simplified the GUMS at 5% and 1% significance level for sample of 146 observations and half of it which is 73 observations.

The algorithm assessment involved the specification searches on 120 experiment conditions that were designed from various characteristics that rise during the data-generating process (DGP) and variation during *SURE-Autometrics* simplification. Since the algorithm was developed for multiple equations model, the performances were not only performed on six equations model as in Ismail (2005), but also on two and four equations model. Hence, the total conditions were 360. In assessing the performances, each condition was replicated 100 times where the outcomes were classified into four categories. Category 1 indicates that the true specification has been selected, whereas Category 2 considers the true specification nested in the selected model. Failure of selecting model without the right specification is classified to Category 3, and Category 4 belongs to model with equations resulted in different category. The outcomes in Category 1 were similar as the probability of *SURE-Autometrics* finds the correct specification.

The algorithm has well performed by achieving high probabilities which is at least 80% (Doornik, 2009; Hoover & Perez, 1999; Ismail, 2005). Table 6.1 shows the conditions where *SURE-Autometrics* able to obtain the high probability. It is clearly showed that the algorithm is not performed in the model with large number of equations since none of the conditions obtained high probability in correctly specified all five true models. However, the outcomes were actually at least 70% which

signifies the possibility that the algorithm also is likely to perform well for a model with large number of equations. Low probabilities might be due to unable to consider the model resulted with different variables amongst the equations. As the number of multiple equations increase, the different between equations become larger. Hence, simple model such as two equations is able to get higher probabilities.

Table 6.1

Experiment Conditions with Probabilities above 80%

| True specification | Number of equations | | |
|--------------------|--|---|---|
| | 2 | 4 | 6 |
| S1 | $\rho = 0.9, 0.6, 0.2$ $n = 146, 73$ $k = 39, 18$ $\alpha = 0.05, 0.01$ | $\rho = 0.9, 0.6, 0.2$ $n = 73$ $k = 18, 39$ $\alpha = 0.05, 0.01$ | $\rho = 0.6, 0.2$ $n = 146$ $k = 39, 18$ $\alpha = 0.05$ |
| S2 | $\rho = 0.9, 0.6, 0.2$ $n = 146, 73$ $k = 39, 18$ $\alpha = 0.05, 0.01$ | $\rho = 0.9, 0.6,$ $n = 73$ $k = 18, 38$ $\alpha = 0.05, 0.01$ | - |
| S3 | $\rho = 0.9, 0.6, 0.2$ $n = 146, 73$ $k = 39, 18$ $\alpha = 0.05, 0.01$ | $\rho = 0.9, 0.6, 0.2$ $n = 146, 73$ $k = 39$ $\alpha = 0.05$ | - |
| S4 | $\rho = 0.9, 0.6, 0.2$ $n = 146, 73$ $k = 39, 18$ $\alpha = 0.05, 0.01$ | $\rho = 0.9, 0.6, 0.2$ $n = 146, 73$ $k = 39$ $\alpha = 0.05$ | - |
| S5 | $\rho = 0.9, 0.6, 0.2$ $n = 146, 73$ $k = 18$ $\alpha = 0.05$ | - | - |

Based on these simulation findings, it can be concluded that the *SURE-Autometrics* has an outstanding performances for all conditions in the model of two equations.

Amongst the five specification models, S5 received lower probability as number of equations increased. This is because S5 contains five relevant variables and added with 13 to 36 irrelevant variables in the initial phase of GUMS. During the evaluation, the final selected model retained irrelevant variables and removed relevant variables. These variables are varied across the equations. Details inspection revealed that within the model, there is equation able to remove all the irrelevant variables but failed to retain at least one of the relevant variables. This situation is actually occurs to S3 and S4 which originally has three variables and added up with irrelevant variables from 15 to 36. As the number of equations increased, the selected variables across the equations become difficult to assess.

Overall findings also suggested that the performance of *SURE-Autometrics* indicated by the probability of correct specification decreased as the number of equations increased. This can be seen starting from the model of four equations where not all conditions were able to achieve high percentages and none of the conditions in S5 are found to be the best. This is expected since the true model of S5 contains five predictors and all four equations must be able to reduce all the irrelevant variables in order to be counted. However, the probabilities were not very bad because it was at least 70%.

To conclude, the algorithm performed well when the number of equations and number of predictors in the true specification models were as minimal as possible. The target size or significance level setting also affects the outcomes of simulation results where probabilities are higher for 5% level as compared to 1% level of significance.

Afterwards, two sets of real data were used for validation of *SURE-Autometrics*. The algorithm is applied in modelling air passengers' flows data, and national growth rates data. These data were based on Ismail (2005). In the application, several model selection using algorithm and non-algorithm procedures were considered. The non-algorithm signifies the selection through manual process where the removal and inclusion of variables are using trial and error process based on author's judgment. All these procedures are classified according to type of selection and method of estimation. The classification and the corresponding procedures are as follows,

1. *Autometrics* and *Stepwise* are algorithm procedures that select the equations individually using OLS estimation method.
2. *Autometrics-SURE* and *Stepwise-SURE* are algorithm procedures that select the equations individually using OLS and the final selected multiple equations will be estimated by FGLS method.
3. *SURE-Autometrics* and *SURE-PcGets* are algorithm procedures that simultaneously select the multiple equations using FGLS method.
4. *Mine*, *Mine-SURE*, and *SURE-Mine* are non-algorithm procedures where the selection form and estimation method for each procedure corresponds to each classification above, respectively.

Specifically these nine model selection procedures were classified based on the selection process (algorithm vs. manual), type of model (single vs. multiple equations) and method of estimation (OLS vs. FGLS). For each of data set, the model can be

specified as SURE model or several equations that are independent of each other's. Then, the selection were done automatically (i.e. algorithm) and manually where each can be estimated using FGLS and OLS.

Subsequently, the forecast of one until three-steps ahead were carried out by the selected models. The forecasting performances amongst these nine model selection procedures were compared to determine which procedure yields the smallest error in forecasts. The forecast errors were measured using the root mean square error (RMSE) and geometric root mean square error (GRMSE). Since there were multiple equations, median of these measures was chosen to represent the procedures. The best procedure is identified according to the ranked of these medians.

Throughout the real data studies, it was revealed that application of model selection using algorithm procedures where it is automatically select a model are more easier than using non-algorithm procedures. The modellers require spending extra time in deciding the removal or inclusion of variables since it involved trial and error processes while maintaining the congruency of each equation. Both real data revealed that there was more than one possible model that is 'best' in representing data. Although the equations within initial model have similar number of variables, the final selected model might be different in terms of number of variables as well as the variables itself.

As concluded in the simulation study, *SURE-Autometrics* demonstrated excellent performances for one until three-steps ahead forecast when the equations number is not large. However, it also depends on the data because the forecast of growth rates

for six countries showed *SURE-Autometrics* is the second best procedure. Therefore, it can be assumed that the algorithm has the possibility to give high accuracy in the forecasting.

Finally, this study concludes that the model selection algorithm for multiple equations has successfully developed. The algorithm is recognised as *SURE-Autometrics* where performances from both simulation experiments and real data application indicated that it has surpassed the algorithm with similar principles which is *SURE-PcGets*. The latter has been developed only for five and six equations model while *SURE-Autometrics* can be applied to at least two equations model. The severe results in experiments suggested the need in improving the search strategies in the algorithm.

Realising there are others model selection procedures can be implemented, the application studies illustrated that *SURE-Autometrics* selects different model from others. Based on the forecasting accuracy, the selected models sometimes outperformed other procedures and sometimes are underperformed. However, the differences are insignificant. Amongst the two sets of data, the national growth rates has problem with the congruency of estimated equations within the model as compared to air passengers' data. The estimated national growth rates suffered from failure in normal distribution, heteroscedasticity variance and parameter constancy. Thus, the performances of *SURE-Autometrics* based on forecasting accuracy also rely on the quality of the data.

Even though non-algorithm procedures which manually selecting the model sometimes have outstanding performances, this type of procedure are time

consuming. During the real data analyses, it took about more than 10 minutes to employ the non-algorithm procedures as compared to less than 3 minutes using the algorithm. However, these average times are insignificant for other researchers since it will be varied according to the software used in the execution of the procedures, the complexities of the multiple equations, the quality of data, and the tacit knowledge acquired by the researchers. The time could be longer especially for young researchers due to difficulties in deciding which variables should be removed or retained in the model. It involves the process of trial and error in removing variables which could lead to the problem of mass significance as discussed in Chapter 2. The *SURE-Autometrics*, *SURE-PcGets* and *Autometrics* are model selection algorithms that were developed within the GETS approach. These algorithms actually have more than one possible model that is best represents the data. Since it aims to help the practitioners in modelling, automatically there is only one model presented for the purpose of forecasting. The selected model is based on the information criterion. Therefore, *SURE-Autometrics* could possibly achieve a similar model as selected by other procedures, but it was neglected during the selection procedure.

This study also concludes that the search strategy in *SURE-Autometrics* needs an improvement. The drawback is identified during the manual selection procedures for modelling growth rates of six countries. Throughout the selection process where variables are removed or included in the model, the data showed the problem of maintaining the congruency for each country. Additionally, the removal of variables does not only affects the significance of other variables within the equation, but also influence the significance of variables in other equations by using FGLS estimation. The modellers will not realize this behaviour by using algorithm procedures. The

search strategy implemented on *SURE-Autometrics* is based on sequential procedure where the selections are one by one but the estimation involved all the equations. The problem could be arise if the data lack of quality such as failure in any diagnostic tests or the dependent variable also correlated with independent variables in other equations. This could be the reason of *SURE-Autometrics* has inconsistent performances based on RMSE and GRMSE for one up to three-step-ahead forecasts.

Nevertheless, the findings from real data applications also demonstrated that modellers need to pay more concentration during the formulation of initial GUMS because it is the most crucial part in any model selection algorithm within the GETS approach. Furthermore, model selection using non-algorithm procedures is time consuming and the modellers require knowledge and experience to make judgment about the removal and inclusion of the variables. This problem is overcome by the algorithm offered by this study. It is suitable to be implemented by students who still in the learning process or novice in research areas.

6.2 Suggestions for Further Research

Based on the conclusions, there are three aspects that can be considered for future study which are algorithm development, assessment of the simulation experiments and application using real data. At first, the development of *SURE-Autometrics* needs to be advanced in two ways. One way is by providing an alternative method for model estimation. Rather than FGLS method, the full information maximum likelihood (FIML) or iterative feasible generalised least squares (IFGLS) can be used to estimate the SURE model. Second way is to improve the search strategy in model selection

procedure through parallel computing. This technique will gain the computational efficiency for model selection of multiple equations.

As for assessment of simulation experiments, additional measures could be used to evaluate the ability of *SURE-Autometrics* in finding the true model when the data-generating process (DGP) is known. Specifically, if the algorithm is unable to obtain similar specification as the true model, the performance could be assessed by the average proportion of irrelevant variables that are retained in the final model which indicate that it survives the reduction process. This statistic may reflect the size of hypothesis testing. Hence, the average proportion of relevant variables will indicate the success of the reduction which also signifies the power of hypothesis testing. Moreover, the experimental condition could be varied in terms of unequal number of observation within the model.

Lastly, the application of algorithm on real data could be extended to cross-sectional type of data. In this situation, the second phase of *SURE-Autometrics* which focused on the reduction of lag variable should be turn off. At the same time, analysis of variance (ANOVA) can be considered for the comparison of forecast errors amongst the model selection procedures to determine whether the differences are significant.

Sharing the same belief as Hendry and Doornik (2014), Ismail (2005) and other researchers (Castle et al., 2013), model selection algorithm could advance the automatic modelling approach. This approach however still needs human intervention in the first phase, which is very crucial in determining the success of the model selection procedure. It is hoped that the existence and exploration of automatic

approaches will bridge the gap between applied and theoretical modelling and lessen the role of tacit knowledge.



REFERENCES

- Armstrong, J. S., & Collopy, F. (1992). Error measures for generalising about forecasting methods: Empirical comparison. *International Journal of Forecasting*.
- Armstrong, J. S., & Fildes, R. (1995). On the Selection of Error Measures for Comparison among Forecasting Methods. *Journal of Forecasting*, 14, 67–71.
- Bartolomei, S. M., & Sweet, A. L. (1989). A note on a comparison of exponential smoothing methods for forecasting seasonal series. *International Journal of Forecasting*, 5, 111–116.
- Beasley, T. M. (2008). Seemingly unrelated regression (SUR) models as a solution to path analytic models with correlated errors. *Multiple Linear Regression Viewpoints*, 34(1), 1–7.
- Bhatti, M. I., Al-Shanfari, H., & Hossain, M. Z. (2006). *Econometric Analysis of Model Selection and Model Testing*. Aldershot: Ashgate Publishing Limited.
- Breiman, L. (1995). Better subset regression using the nonnegative garrote. *Technometrics*, 37, 373–384.
- Castle, J. L., Doornik, J. A., & Hendry, D. F. (2011). Evaluating automatic model selection. *Journal of Time Series Econometrics*, 3(1), 33.
- Castle, J. L., Qin, X., & Reed, W. R. (2013). Using model selection algorithms to obtain reliable coefficient estimates. *Journal of Economic Surveys*, 27(2), 269–296. <http://doi.org/10.1111/j.1467-6419.2011.00704.x>
- Chen, H., Wan, Q., & Wang, Y. (2014). Refined Diebold-Mariano Test Methods for the Evaluation of Wind Power Forecasting Models. *Energies*, 7(7), 4185–4198. <http://doi.org/10.3390/en7074185>
- Denton, F. T. (1985). Data Mining as an Industry. *The Review of Economics and Statistics*, 67(1), 124–127.
- Derksen, S., & Keselman, H. J. (1992). Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology*, 45, 265–282.
- Doornik, J. A. (2008). Encompassing and automatic model selection. *Oxford Bulletin of Economics and Statistics*, 70, 915–925. <http://doi.org/10.1111/j.1468-0084.2008.00536.x>
- Doornik, J. A. (2009). Autometrics. In J. L. Castle & N. Shephard (Eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry* (pp. 88–121). New York: Oxford University Press.
- Doornik, J. A., & Hendry, D. F. (2007). *Empirical Econometric Modelling using PcGive 12: Volume 1*. London: Timberlake Consultants Ltd.
- Dufour, J.-M., & Khalaf, L. (2002). Exact tests for contemporaneous correlation of

- disturbances in seemingly unrelated regressions. *Journal of Econometrics*, 106, 143–170.
- Efroymson, M. A. (1960). Multiple Regression Analysis. *Mathematical Methods for Digital Computers*. New York: Wiley.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.
- Ericsson, N. R., & Kamin, S. B. (2009). Constructive Data Mining: Modeling Argentine Broad Money Demand. In J. L. Castle & N. Shephard (Eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry* (pp. 412–439). New York: Oxford University Press.
- Fernandez, S., Smith, C. R., & Wenger, J. B. (2007). Employment, privatization, and managerial choice: Does contracting out reduce public sector employment? *Journal of Policy Analysis and Management*, 26, 57–77.
- Fildes, R. (1992). The evaluation of extrapolative forecasting methods. *International Journal of Forecasting*.
- Fildes, R., Wei, Y., & Ismail, S. (2011). Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. *International Journal of Forecasting*, 27(3), 902–922.
<http://doi.org/10.1016/j.ijforecast.2009.06.002>
- Fisher, S. (1993). The role of macroeconomics factors in growth. *Journal of Monetary Economics*, 32, 485–512.
- Foster, D. P., & Stine, R. A. (2004). Variable selection in data mining: Building a predictive model for bankruptcy. *Journal of the American Statistical Association*, 99(466), 303–313.
- Garcia-Ferrer, A., Highfield, R. A., Palm, F., & Zellner, A. (1987). Macroeconomic Forecasting Using International Data. *Journal of Business & Economic Statistics*, 8(1), 53–67. Retrieved from
<http://core.kmi.open.ac.uk/download/pdf/6750786.pdf>
- Granger, C. W. J. (1999). *Empirical Modelling in Economics: Specification and Evaluation*. New York: Cambridge University Press.
- Granger, C. W. J., & Hendry, D. F. (2005). A Dialogue Concerning A New Instrument for Econometric Modeling. *Econometric Theory*, 21, 278–297.
<http://doi.org/10.1017/S0266466605050164>
- Granger, C. W. J., Hendry, D. F., & Hansen, B. E. (2005). Challenges for econometric model selection. *Econometric Theory*, 21, 60–68.
<http://doi.org/10.1017/S0266466605050048>
- Greene, W. H. (2012). *Econometric Analysis* (7th ed.). Edinburgh Gate: Pearson Education Limited.
- Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23, 461–465.
- Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction

- mean squared errors. *International Journal of Forecasting*, 13, 281–291.
- Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*. *Springer Series in Statistics*. New York: Springer.
- Hendry, D. F. (1980). Econometrics-Alchemy or Science? *Economica*, 47(188), 387–406.
- Hendry, D. F. (1995). *Dynamic Econometrics*. Oxford University Press.
- Hendry, D. F. (2001). Achievements and challenges in econometric methodology. *Journal of Econometrics*, 100(1), 7–10. [http://doi.org/10.1016/S0304-4076\(00\)00045-2](http://doi.org/10.1016/S0304-4076(00)00045-2)
- Hendry, D. F., & Doornik, J. A. (2014). *Empirical Model Discovery and Theory Evaluation: Automatic Selection Methods in Econometrics*. MIT Press.
- Hendry, D. F., & Krolzig, H.-M. (1999). Improving on “Data mining reconsidered” by K.D. Hoover and S.J. Perez. *Econometrics Journal*, 2, 202–219.
- Hendry, D. F., & Krolzig, H.-M. (2001). *Automatic Econometric Model Selection Using PcGets 1.0*. London: Timberlake Consultants Press.
- Hendry, D. F., & Krolzig, H.-M. (2003). New Developments in Automatic General-to-Specific Modeling. In *Econometrics and the Philosophy of Economics: Theory-data confrontations in economics* (pp. 379–419). Princeton: Princeton University Press.
- Hendry, D. F., & Krolzig, H.-M. (2004). We ran one regression. *Oxford Bulletin of Economics and Statistics*, 66, 799–810.
- Hendry, D. F., & Krolzig, H.-M. (2005). The properties of automatic Gets modelling. *Economic Journal*, 115(502), C32–C61. <http://doi.org/10.1111/j.0013-0133.2005.00979.x>
- Hendry, D. F., & Reade, J. J. (2008). *Modelling and forecasting using model averaging*. Working paper.
- Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (1999). Bayesian Model Averaging: A Tutorial. *Statistical Science*, 14(4), 382–401.
- Hoover, K. D., & Perez, S. J. (1999). Data mining reconsidered: Encompassing and the general-to-specific approach to specification search. *Econometrics Journal*, 2, 167–191.
- Hoover, K. D., & Perez, S. J. (2000). Three attitudes towards data mining. *Journal of Economic Methodology*, 7(2), 195–210. <http://doi.org/10.1080/13501780050045083>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688. <http://doi.org/10.1016/j.ijforecast.2006.03.001>
- Ismail, S. (2005). *Algorithmic approaches to multiple time series forecasting*. University of Lancaster, Lancaster.

- Ismail, S., & Fildes, R. (2007). Algorithmic approaches to multiple time series forecasting. In *The 27th Annual International Symposium on Forecasting*. New York.
- Ismail, S., Yusof, N., & T-Muda, T.-Z. (2015). Algorithmic approaches in model selection of the air passengers flows data. *Proceedings of the 5th International Conference on Computing and Informatics*, 32–37.
- Judge, G. G., Hill, R. C., Griffiths, W. E., Lütkepohl, H., & Lee, T.-C. (1988). *Introduction to the Theory and Practice of Econometrics* (2nd ed.). New York: Wiley.
- Kontoghiorghes, E. J. (2004). Computational methods for modifying seemingly unrelated regressions models. *Journal of Computational and Applied Mathematics*, 162(1), 247–261. <http://doi.org/10.1016/j.cam.2003.08.024>
- Krolzig, H.-M. (2001). General-to-specific reductions of vector autoregressive process. *Econometrics Studies: A Festschrift in Honour of Joachim Frohn*. Münster: LIT.
- Krolzig, H.-M., & Hendry, D. F. (2001). Computer automation of general-to-specific model selection procedures. *Journal of Economic Dynamics and Control*, 25, 831–866.
- Kurata, H. (2004). One-sided tests for independence of seemingly unrelated regression equations. *Journal of Multivariate Analysis*, 90(2), 393–406. <http://doi.org/10.1016/j.jmva.2003.09.003>
- Lazim, M. A. (1995). *Econometric forecasting model and model evaluation: A case study of air passenger traffic flow*. Lancaster University. Retrieved from http://eprints.uitm.edu.my/3170/1/MOHAMAD_ALIAS_LAZIM_95.pdf
- Leamer, E. E. (1978). Specification searches: Ad hoc inference with non-experimental data. New York: Wiley.
- Leeb, H., & Pötscher, B. M. (2005). Model selection and inference: Facts and fiction. *Econometric Theory*.
- Leeb, H., & Pötscher, B. M. (2009). Model Selection. In T. G. Andersen, R. A. Davis, J.-P. Kreiss, & T. V. Mikosch (Eds.), *Handbook of Financial Time Series* (pp. 889–925). New York: Springer-Verlag Berlin Heidelberg.
- Lovell, M. C. (1983). Data mining. *The Review of Economics and Statistics*, 65(1), 1–12.
- Magnus, J. R. (1999). The success of econometrics. *De Economist*, 147(1), 55–77.
- Magnus, J. R., & Morgan, M. S. (1997). Design of the Experiment. *Journal of Applied Econometrics*, 12(5), 459–465.
- Magnus, J. R., & Morgan, M. S. (1999). *Methodology and Tacit Knowledge: Two Experiments in Econometrics*. New York: John Wiley.
- Mariano, R. S. (2002). Testing forecast accuracy. *A Companion to Economic Forecasting*, (July), 284–298. <http://doi.org/10.1007/s10614-008-9144-4>

- Miller, A. J. (1984). Selection of Subsets of Regression Variables. *Journal of Royal Statistical Society*, 147(3), 389–425.
- Miller, A. J. (2002). Subset selection in regression. In *Monographs on Statistics and Applied Probability* (Vol. 95). Florida: Chapman & Hall/ CRC.
- Mizon, G. E. (1995). Progressive modelling of macroeconomic time series: The LSE methodology. *Macroeconomics: Developments, Tensions, and Prospects*. Dordrecht: Kluwer Academic Press.
- Mycielski, J., & Kurcewicz, M. (2004). *A specification search algorithm for cointegrated systems*. *Computing in Economics and Finance*. Retrieved from <http://econpapers.repec.org/RePEc:sce:scecf4:321>
- Oksanen, E. H. (1987). A Note On Seemingly Unrelated Regression Equations with Residual Vectors as Explanatory Variables. *Statistics & Probability Letters*, 6, 103–105.
- Pagan, A. (1987). Three Econometric Methodologies: A Critical Appraisal. *Journal of Economic Surveys*, 1(1), 3–24.
- Pant, P. N., & Starbuck, W. H. (1990). Innocents in the forecast: Forecasting and research methods. *Journal of Management*, 16, 433–460.
- Pérez-Amaral, T., Gallo, G. M., & White, H. (2003). A flexible tool for model building: The relevant transformation of the inputs network approach (RETINA). *Oxford Bulletin of Economics and Statistics*, (65), 821–838.
- Philips, P. C. B. (2003). Laws and limits of econometrics. *The Economic Journal*, 113, 26–52.
- Pindyck, R. S., & Rubinfeld, D. L. (1998). *Econometric Models and Economic Forecasts*. Boston, Massachusetts: Irwin/McGraw-Hill.
- Reade, J. J. (2007). Modelling and forecasting football attendances. *Oxonomics*, 2, 27–32. <http://doi.org/10.1111/j.1752-5209.2007.0015.x>
- Romano, J. P., Shaikh, A., & Wolf, M. (2008). Formalized data snooping based on generalized error rates. *Econometric Theory*, 24, 404–447.
- Santos, C., Hendry, D. F., & Johansen, S. (2007). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, 23(2), 317–335. <http://doi.org/10.1007/s00180-007-0054-z>
- Schwartz, J. (2006). Family structure as a source of female and male homicide in the United States. *Homicide Studies*, 10(4), 253–278.
- Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1), 1–48.
- Srivastava, V. K., & Dwivedi, T. D. (1979). Estimation of seemingly unrelated regression equations: A brief survey. *Journal of Econometrics*, 10, 15–32.
- Srivastava, V. K., & Giles, D. E. A. (1987). *Seemingly Unrelated Regression Equations Models: Estimation and Inference*. *Statistics: Textbooks and Monographs*. New York: Marcel Dekker, Inc.
- Srivastava, V. K., & Maekawa, K. (1995). Efficiency properties of feasible

- generalized least squares estimators in SURE models under non-normal disturbances. *Journal of Econometrics*, 66(03630015), 99–121.
- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 58(1), 267–288.
- Timm, N. H. (2002). Seemingly Unrelated Regression Models. In *Applied Multivariate Analysis* (pp. 311–349). New York: Springer.
- Timm, N. H., & Al-Subaihi, A. A. (2001). Testing model specification in seemingly unrelated regression models. *Communications in Statistics: Theory and Methods*, 30(4), 579–590.
- Tsay, W.-J. (2004). Testing for contemporaneous correlation of disturbances in seemingly unrelated regressions with serial dependence. *Economic Letters*, 83, 69–76.
- Verzilli, C. J., Stallard, N., & Whittaker, J. C. (2005). Bayesian modelling of multivariate quantitative traits using seemingly unrelated regressions. *Genetic Epidemiology*, 38, 313–325. <http://doi.org/10.1002/gepi.20072>
- White, H. (2000). A Reality Check for Data Snooping. *Econometrica*, 68(5), 1097–1126.
- Whittingham, M. J., Stephens, P. A., Bradbury, R. B., & Freckleton, R. P. (2006). Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, 75(5), 1182–1189. <http://doi.org/10.1111/j.1365-2656.2006.01141.x>
- Yusof, N., & Ismail, S. (2011). Independence test in SURE-Autometrics algorithm. *Proceedings of the International Symposium on Forecasting*.
- Yusof, N., & Ismail, S. (2014). Lag variables reduction in multiple models selection algorithm. *Proceedings of the International Conference on the Analysis and Mathematical Applications in Engineering and Science*, 169–173.
- Yusof, N., Ismail, S., & T-Muda, T.-Z. (2015). Assessing the simulation performances of multiple model selection algorithm. *Proceedings of the 5th International Conference on Computing and Informatics*, 25–31.
- Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association*, 57(298), 348–368.
- Zellner, A. (1963). Estimators for seemingly unrelated regressions: Some exact finite sample results. *Journal of the American Statistical Association*, 58, 977–992.