A Multivariate Analysis of Household Debts in South Africa and the USA during the 2007-2009 Financial Crisis

N.D. Moroke

Thesis submitted for the degree Doctor of Philosophy in Statistics at the Mafikeng Campus of the North-West University

Promoter: Prof. J. Mukuddem-Petersen

Co-Promoter: Prof. M.A. Petersen

November 2013

TABLE OF CONTENTS

DECLARATION I

ACKNOWLEDGEMENTS II

ABSTRACT III

GLOSSARY OF TERMS V

LIST OF ACRONYMS VII

LIST OF TABLES VIII

LIST OF FIGURES IX

CHAPTER 1 1

ORIENTATION OF THE STUDY 1

1.1 Introduction 1

1.2 Problem statement 5

1.3 Contribution of this study 6

1.4 Research aims and objectives 7

  1.4.1 Research aims 7

  1.4.2 Research objectives 9

1.5 Research questions / hypothesis 9

  1.5.1 Research questions 9

  1.5.2 Research hypothesis 10

1.6 Research methodology 10

1.7 Possible results and conclusions 11

1.8 Scope limitations and delimitations 11

  1.8.1 Limitations 11

  1.8.2 Delimitations 12

1.9 Ethical Considerations 12

1.10 Thesis Structure 12

CHAPTER 2 14

THEORY AND LITERATURE REVIEW 14

2.1 Introduction 14

2.2 Theoretical framework 14

  2.2.1 The Keynesian theory 15

2.2.2 The Life-Cycle and Permanent Income Hypotheses 15
2.2.2.1 Household Debt Determination 16
2.3 Empirical framework 18
2.3.1 Procedure for multivariate time series analysis (MTSA) 18
2.3.2 Nature of data 21
2.3.3 Testing for a Unit Root 21
2.3.4 Cointegration analysis 22
2.3.4.1 Testing Order of Cointegration 23
2.3.4.2 Specifiying the VAR(p) Order 23
2.3.4.3 Testing for cointegration 24
2.3.4.4 Normalisation 24
2.4 The VECM approach 25
2.5 The VAR model 25
2.6 The VEC causality test 26
2.7 Studies that incorporated the LCH, PIH theories and Johansen's procedure 27
2.8 Fisher's LDA approach 32
2.8.1 Introduction 32
2.8.2 Studies that used DA 33
2.8.3 Studies that used FA 39
2.9 Summary 40

CHAPTER 3 42
DATA AND METHODOLOGY 42
3.1 Introduction 42
3.2 Why household enter into debts? 43
3.3 The research approach 43
3.3.1 Information gathering 43
3.3.2 Methods of data analysis 44
3.4 Procedure for data analysis 44
3.4.1 Variables and data processing 46
3.5 Preliminary data analysis 51
3.6 Model specification 53
3.7 Model estimation 54
3.7.1 Unit root tests 56

3.7.1.1 Augmented Dickey-Fuller Test 56
3.7.1.2 Phillips Perron Test 57
3.7.1.3 Kwiatkowski, Phillips, Schmidt and Shin 58
3.7.2 Convergence of Household Debts 59
3.7.3 Cointegration analysis 61
3.7.3.1 The VAR model and optimal lag selection 61
3.7.3.2 The Johansen Multivariate Cointegration test 61
3.7.4 Causality analysis 66
3.8 Parameter and model diagnostic testing 69
3.8.1 Normality test 71
3.8.2 Serial correlation among residuals 68
3.8.3 Heteroscedasicity assumption 72
3.9 The Generalised Impulse Response Function (GIRF) analysis 73
3.10 Overview of Discriminant Analysis (DA) 74
3.10.1 Discriminant Analysis framework 76
3.10.2 Testing the homogeneity of the variance-covariance matrix 77
3.10.3 Significance of independent variables 78
3.11 Evaluating group differences 78
3.12 Comparison of SA and the USA 80
3.13 Summary 80

CHAPTER 4 81
DATA ANALYSIS AND RESULTS 81
4.1 Introduction 81
4.2 Empirical findings 82
4.2.1 Exploratory data analysis 82
4.3.2 Association between household debts determinants 84
4.3.3 Sampling and variable adequacy analysis 86
4.4 Trend analysis of Household debt determinants 87
4.5 Stationarity results 100
4.6 Convergence results 105
4.7 Cointegration results 106
4.8 Estimated model 108
4.9 Causality test results 115

4.10 Model Diagnostics Results 120
4.10 Generalised Impulse Response Analysis Results 120
4.11 Forecasting 131
4.12 Discriminant Analysis results 133
4.12.1 Confirmation of assumptions 133
4.12.2 Variable selection results 134
4.12.3 Fisher's Linear Discriminant Model 137
4.12.4 Comparison between SA and the USA 139
4.13 Summary 140

CHAPTER 5 142
CONCLUSION AND RECOMMENDATIONS 142
5.1 Introduction 142
5.2 Research questions and conclusions 142
5.3 Recommendations 153
5.4 Summary 156

CHAPTER 6 157
APPENDICES 157
Appendix A: VAR (1) Model 157
Appendix B: GIRF 159
Appendix C: Forecasts of HHD 160
Appendix D: Stepwise Discriminant Analysis Results 160
Appendix E: Convergence analysis results 163

REFERENCES 165

LIST OF TABLES

Table 3.1: Summary of variables 51
Table 3.2: Variable names and descriptions 54
Table 3.3: Descriptive Statistics 83
Table 3.4: Sample and variable adequacy results 86
Table 3.5: ADF and PP tests 104
Table 3.6: KPSS test 104

Table 3.7: Default Minimum Information Criterion 107
Table 3.8: Cointegration tests 108
Table 3.9: SA VAR(1) model parameter estimates 109
Table 3.10: USA VAR(1) model parameter estimates 110
Table 3.11: SA HHD VEC Model 111
Table 3.12: USA HHD VEC Model 112
Table 3.13: SA Error correction model 114
Table 3.14: USA Error correction model 115
Table 3.15: SA Toda-Yamamoto Wald Causality Test 116
Table 3.16: SA Toda-Yamamoto Wald Causality Test 116
Table 3.17: SA model diagnostics 118
Table 3.18: USA model diagnostics 119
Table 3.19: Test of homogeneity of within covariance matrices 134
Table 3.20: Forward selection 134
Table 3.21: Communalities 136
Table 3.22: Linear Discriminant Function 138
Table 3.23: Standardised Discriminant Function 138
Table 3.24: Mahalanobis distance statistic 139

LIST OF FIGURES

Figure 2.1: Model building framework for MTSA 21
Figure 3.1: Stages in the preliminary data analysis 46
Figure 3.2: Steps for estimation, diagnostic checking and forecasting 56
Figure 3.3: Stages in Discriminant Analysis 76
Figure 4.1: Description of household debt determinants 85
Figure 4.2: SA and USA HHD plot 88
Figure 4.3: SA and USA HP plot 90
Figure 4.4: SA and USA CPI plot 92
Figure 4.5: SA and USA INC plot 93
Figure 4.6: SA and USA HC plot 94
Figure 4.7: SA and USA GDP plot 95
Figure 4.8: SA and USA IR plot 96
Figure 4.9: SA and USA HS plot 97

Figure 4.10: SA and USA ER plot 98
Figure 4.11: SA and USA UR plot 99
Figure 4.12: SA and USA First Dif. HHD, HP and CPI 101
Figure 4.13: SA and USA First Dif. INC, IR and GDP 102
Figure 4.14: SA and USA First Dif. HC, HS, ER and UR 103
Figure 4.15: Time series plot of household debt differential 105
Figure 4.16: Response to impulse in SA and USA HP 121
Figure 4.17: Response to impulse in SA and USA CPI 122
Figure 4.18: Response to impulse in SA and USA INC 123
Figure 4.19: Response to impulse in SA and USA IR 124
Figure 4.20: Response to impulse in SA and USA GDP 125
Figure 4.21: Response to impulse in SA and USA HC 126
Figure 4.22: Response to impulse in SA and USA HS 127
Figure 4.23: Response to impulse in SA and USA ER 128
Figure 4.24: Response to impulse in SA and USA UR 129
Figure 4.25: Forecasts of SA HHD 130
Figure 4.26: Forecasts of USA HHD 131

DECLARATION BY STUDENT

I declare that the submitted work "A Multivariate Analysis of Household Debt in South Africa and the USA during the Financial Crisis" was completed by me, the undersigned, and that I have only used the permitted reference sources or materials and have not engaged in any plagiarism. All references and other sources I have used have aptly been acknowledged in the work. I also declare that this work is the result of my own investigations except where otherwise identified by references and that I have not plagiarised the work of others.

Full Names: ....

Signature: Date:
Student

Signature: Date:
Supervisor

Signature: Date:
Co-supervisor

ACKNOWLEDGEMENTS

First and foremost, I would like to express my gratitude the 'Most High God' for the life, strength and mercy provided. With Him, nothing is impossible.

I am deeply indebted to my promoters Professors. Janine Mukuddem-Petersen and Mark Petersen who made stimulating suggestions and encouraged me throughout the study and the writing of this thesis. Your valued contribution and guidance are greatly cherished.

To the management of the faculty, colleagues and my friends, thank you for your interest and support.

I owe thanks to my children Othusitse and Koketso and my brothers Otshepeng and Tshireletso for their understanding and also for giving me space during my studies.

I greatly thank my parents, who constantly reminded me that good things come to those who wait patiently and those who work as if they will not see tomorrow anymore.

In conclusion, all those who directly or indirectly contributed to the success of this thesis are acknowledged.

ABSTRACT

This study takes place against the backdrop of the 2007-2009 financial crisis which had impacted heavily on the wellbeing of households worldwide. This crisis was triggered by the SMC which emerged in the USA. The crisis resulted in increased household debt (HHD) levels in many countries of differing financial sophistication. The crisis also had a deleterious effect on unemployment and interest rates, investment and savings as well as household credit. Recent financial reports show that the effects of this crisis are still being felt in many countries. In the main, this study aims to analyse trends in household debts in SA and the USA and determine the effects of related determinants on household debts. Furthermore, the study identifies differences and/or similarities between the economies of the aforementioned countries.

In our study, time-series data collected for the period 90Q1 to 13Q1 were obtained from the South African Reserve Bank (SARB), Federal Reserve Bank (the Fed) and Statistics South Africa (SSA). The data consisted of determinants of HHD such as house prices (HP), consumer price index (CPI), disposable income (INC), interest rates (IR), gross domestic product (GDP), household consumption expenditure (HC), household savings (HS), exchange rates (ER), unemployment rates (UR). As far as the methodology is concerned, we mainly employed Johansen's multivariate and discriminant methods which are executed by using the statistical package, SAS. In particular, our techniques enable us to consider variables contributing to HHD and the relationships between them.

The findings of this study provide a clear picture of the extent of the impact of the 2007-2009 financial crisis on household debt in SA and the USA. As reflected in the dynamics of the household debt determinants, these countries were both significantly affected during this period. The Error Correction Model (ECM) confirmed that in the long run, increases in household debt in SA can be attributed, in order of importance, to household consumption expenditure, gross domestic product and house prices. The findings for the USA indicated that household consumption and expenditure as well as household income may be associated with increased debts in the long run. The error correction terms were obtained as -13.78 % and -1.54 % for SA and the USA respectively. The Toda-Yamamoto causality test revealed several important causal relationships. GDP, exchange rates and unemployment rates were found to have unidirectional relationships with other variables in the case of SA with the former two having this relationship in the USA. The generalised impulse response function (GIRF) analysis showed that HHD responds negatively to shocks from changes in house prices. Moreover, forecasts for both countries report increased debts going forward. Fisher's linear discriminant analysis (LDA) identified five main determinants to be significantly associated with debts in these countries. Moreover, the Mahalanobis distance and Hoteling's test confirm that as much as SA household debt is different from the USA, there are also significant similarities.

In conclusion, the Johansen tests highlighted the main causes of HHD in SA and the USA both in the short and long -run. Toda-Yamamoto tests revealed relationships among the ten variables in his study. The GIRF forecasts confirmed using the 95 % level of confidence that in the next five years, household indebtedness will be even higher in SA and the USA. With Fisher's LDA reporting consumer price index, household income, gross domestic product, household consumption and expenditure and household savings were identified as the main contributors to debt. Finally, our study recommends the formulation and implementation of policies that will effectively help in dealing with household indebtedness in SA and the USA.

KEYWORDS: Multivariate Analysis, Vector Error Correction Model, Discriminant Analysis Household Debts, South Africa, the USA.

JEL Classification: C32; E21; G01

GLOSSARY OF TERMS

Multivariate Analysis refers to an instance where two or more statistical outcome variables are observed and analysed concurrently.

Household Debts is the amount of money that all adults in the household owe financial institutions including consumer debt as well as mortgage loans.

Stationary process is a stochastic process whose joint probability distribution remains unchanged when shifted in time. Subsequently, the mean and variance also do not change over time and do not follow any trend.

House prices refer to the price of residential units.

Consumer price index is an inclusive measure used for estimating the changes of price in a basket of goods and services representative of consumption expenditure in an economy.

Household disposable income refers to the sum of household final consumption expenditure and savings less the change in net equity of households in pension funds. This amount corresponds to the sum of wages and salaries, mixed income, net property income, net current transfers and social benefits.

Interest rate is a portion of percentage charged or paid for the use of money. It is expressed as an annual percentage of the principal and is subject to change as a result of inflation and Reserve Banks' policies.

Gross domestic product refers to the total market value of all final goods and services produced in a country in a given year. This product equals to total consumer, investment and government spending, including the value of exports, less the value of imports.

Household consumption expenditure is the market value of all goods and services, including durable products such as cars, home appliances and electronics purchased by households.

Household savings refers to the difference between a household's disposable income and its consumption. The household saving rate is calculated by dividing household saving by household disposable income.

Exchange rate is the price for which the currency of a country is exchanged for another country's currency.

Unemployment rate is a percentage of total workforce who does not have a job, have actively looked for work in the past four weeks, and are currently available for work. This includes people who were temporarily laid off and are waiting to be called back to that job.

LIST OF ACRONYMS

SA	South Africa
US	United States
LDA	Linear Discriminant Analysis
FA	Factor Analysis
ADF	Augmented Dickey Fuller
PP	Phillips Perron
KPSS	Kwiatkowski, Phillips, Schmidt and Shin
AIC	Akaike Information Creteria
SAS	Systems Analysis Software
ECM	Error Correction Model
VECM	Vector Error Correction Model
VAR	Vector Autoregression
GIRF	Generalised Impulse Response Function
LCH	Life-Cycle Hypothesis
PIH	Permanent-Cycle Hypothesis
OECD	Organization for Economic Corporation and Development
SARB	South African Reserve Bank
FRB	Federal Reserve Bank
SMC	Subprime Mortgage Crisis
ESDC	European Sovereign Crisis

LIST OF TABLES

Table 3.1: Summary of variables
Table 3.2: Variable names and descriptions
Table 4.1: Descriptive Statistics
Table 4.2: Sample and variable adequacy results
Table 4.3: ADF and PP tests
Table 4.4: KPSS stationarity test
Table 4.5: Default Minimum Information Criterion
Table 4.6: Cointegration tests
Table 4.7: SA VAR(1) model parameter estimates
Table 4.8: US VAR(1) model parameter estimates
Table 4.9: SA HHD VEC Model
Table 4.10: US HHD VEC Model
Table 4.11: SA Error Correction Model
Table 4.12: US Error Correction Model
Table 4.13: SA Toda-Yamamoto Wald Causality Test
Table 4.14: US Toda-Yamamoto Wald Causality Test
Table 4.15: SA model diagnostics
Table 4.16: US model diagnostics
Table 4.17: Test of homogeneity of within covariance matrices
Table 4.18: Forward selection
Table 4.19: Communalities
Table 4.20: Linear Discriminant Function
Table 4.21: Standardised Discriminant Function
Table 4.22: Mahalanobis distance statistic

LIST OF FIGURES

Figure 2.1: Model building framework for MTSA
Figure 3.1: Stages in the preliminary data analysis
Figure 3.2: Steps for estimation, diagnostic checking and forecasting
Figure 3.3: Stages in Discriminant Analysis
Figure 4.1: Description of household debt determinants
Figure 4.2: SA and USA HHD plot
Figure 4.3: SA and USA HP plot
Figure 4.4: SA and USA CPI plot
Figure 4.5: SA and USA INC plot
Figure 4.6: SA and USA HC plot
Figure 4.7: SA and USA GDP plot
Figure 4.8: SA and USA IR plot
Figure 4.9: SA and USA HS plot
Figure 4.10: SA and USA ER plot
Figure 4.11: SA and USA UR plot
Figure 4.12: SA and USA First Dif. HHD, HP and CPI
Figure 4.13: SA and USA First Dif. INC, IR and GDP
Figure 4.14: SA and USA First Dif. HC, HS, ER and UR
Figure 4.15: Time series plot of household debt differential
Figure 4.16: Response to impulse in SA and USA HP
Figure 4.17: Response to impulse in SA and USA CPI
Figure 4.18: Response to impulse in SA and USA INC
Figure 4.19: Response to impulse in SA and USA IR
Figure 4.20: Response to impulse in SA and USA GDP
Figure 4.21: Response to impulse in SA and USA HC
Figure 4.22: Response to impulse in SA and USA HS
Figure 4.23: Response to impulse in SA and USA ER
Figure 4.24: Response to impulse in SA and USA UR
Figure 4.25: Forecasts of SA HHD
Figure 4.26: Forecasts of USA HHD

CHAPTER 1

ORIENTATION OF THE STUDY

Statistics are like bikinis. What they reveal is suggestive, but what they conceal is vital.
~ Aaron Levenstein

1.1 INTRODUCTION / BACKGROUND
1.2 PROBLEM STATEMENT
1.3 CONTRIBUTION OF THIS STUDY
1.4 RESEARCH AIMS AND OBJECTIVES
1.5 RESEARCH QUESTIONS / HYPOTHESIS
1.6 RESEARCH METHODOLOGY
1.7 POSSIBLE RESULTS AND CONCLUSIONS
1.8 SCOPE LIMITATIONS AND DELIMITATIONS
1.9 ETHICAL CONSIDERATIONS
1.10 THESIS STRUCTURE

1.1 Introduction

Several years have elapsed since the start of the 2007-2009 financial crisis that has given rise to the largest wave of banking crises since the great depression. Laeven & Valencia (2012) reported that unlike previous crises over this period, the recent wave of crises has affected most of the advanced economies. The effects of these crises are still lingering and in many cases the crisis is still on-going. Like many countries worldwide, South Africa (SA) also suffered the effects of this crisis.

This study identifies the three most recent crises which affected many economies in the world, i.e. the subprime mortgage crisis (SMC) of 2007-2008, the financial crisis and the recent European sovereign crisis (ESDC). The Wall Street Journal (2013) reported that the fiscal cliff in the United States of America (USA) was the sharp decline in the budget deficit. This could have occurred at the beginning of 2013 due to increased taxes and reduced spending as required by previously enacted laws. The deficit was projected to be reduced by roughly half in 2013. The congregational budget office had estimated that the fiscal cliff would have likely led to a mild recession with higher unemployment in 2013, followed by strengthening in the labour market with increased economic growth.

According to the United States (US) *Wall Street Journal*, the American Taxpayer Relief Act 2012 largely eliminated the fiscal cliff with a higher deficit and a smaller revenue increase as compared to the previously enacted laws. The expectation was that adjustments to spending would be resolved at the beginning of 2013. The 2007-2009 financial crisis forms the basis of this study, in view of the fact that it has had an enormous impact on economies, with stock markets falling, financial institutions collapsing and governments forced to intervene with bailouts, while refocusing on regulatory reform.

The U.S. subprime mortgage crisis (SMC) is considered to be a set of events and conditions that led to the late 2000s financial crisis. As reports indicate, this crisis was characterised by a rise in subprime mortgage delinquencies and foreclosures, and the resulting decline of securities backed by the said mortgages. As explained by The State of the Nation's Housing (2008), the percentage of new lower-quality subprime mortgages rose from the historical 8% or lower range to approximately 20% from 2004 to 2006, with much higher ratios in some parts of the U.S. Mark (2009) emphasised that these two changes were part of a broader trend of lowered lending standards and higher-risk mortgage products. As economists' reports show, the U.S. households became increasingly indebted, with the ratio of debt to disposable personal income rising from 77% in 1990 to 127% at the end of 2007 and much of this increase was mortgage-related.

Several factors contributed to the rapid growth in subprime lending - low interest rates, rapid house price appreciation, easier credit, and new mortgage instruments (Kojucharov, Martin, & Martin 2008). The results of these authors' study revealed that between 2001 and 2005, interest rates on both fixed and adjustable rate subprime mortgages declined by over 2% points. This considerably reduced the borrowing costs for prospective homebuyers. In addition, the results revealed that hybrid products such as 2/28 and 3/27 adjustable rate mortgages (ARMs) attracted a larger pool of borrowers by offering lower initial costs. Furthermore, the high house price appreciation also allowed many borrowers to refinance out of the subprime pool. This brought about the temptation to mortgage originators to loosen their underwriting standards, and to lend to borrowers with increasingly questionable credit histories and income documentation (Demyanyk & Van Hemert 2008).

A recent study by Petersen *et al.* (2012) reported on other effects of the SMC which directly relates to the foundations of the financial industry instability. These, as the report indicates, have caused numerous iconic Wall Street investment banks and prominent depository institutions to collapse. Furthermore, the crisis thwarted credit extension to households and businesses and as a result created credit crunches and, ultimately recessions. Based on these problems, Petersen *et al.* (2012) formulated the IDIOM hypothesis highlighting the main causes of the SMC as being largely caused by the intricacy and design of subprime mortgage origination. The study further highlighted factors such as securitization and agents as being the causes of information problems and valuation opaqueness.

Due to the 2007-2008 SMC discussed in the previous paragraphs, the 2007-2009 financial crisis in the U.S emerged. Like many crises, this crisis had many effects of which some spilled over to other continents including Africa. Naudé (2009) warned that the overall effects of the financial crisis on developing countries such as Africa, especially those dependent on trade with USA like SA would surely be negative. Though it was reported that SA was not severely affected by the financial crisis, the spill-over effects of this crisis began to hit the industries in the country. These effects resulted in assets being devalued, tightening of financial conditions, closing down of companies, escalation in the rates of unemployment and economic growth decrease.

Furthermore, due to these effects, several losses were registered in almost all the companies in the financial industry, causing serious problems in their business. Taking out loans was brought to a minimum and the severity of the whole situation seriously shook people's confidence in banks and other financial institutions. This also forced investors to withdraw funds from the financial markets, thus exacerbating the situation even more. The increase in unemployment and poverty in the whole world contributed to the decline in the aggregate demand. Furthermore, this resulted in a drop of the GDP in almost all the countries. Amongst other countries, the USA recorded a decline of 2.3%. The South African Reserve Bank's 2010 quarterly report reveals that SA's GDP decline was 15.3% in 2009. The director of the Federal Reserve's Alan Greenspan advocated significant reductions in Central Bank interest rates. Economic reporters are of the view that this might have served as an invitation to companies and the population to increase their investments and consumption.

Allen & Giovannetti (2010) postulated that the seeds of this crisis can be traced to the low interest rate policies adopted by the Federal Reserve Bank and other central banks after the collapse of the technology stock bubble. The interest rates cuts affected the investors as well as households. Households were as a result compelled to lower their savings rates and this escalated to household indebtedness. As indicated by Aron & Muellbauer (2000), low savings rates make a contribution to persistent structural weakness in SA. This in turn causes households to borrow rather than to save. This situation was worsened during the crisis and the result was a high percentage of household debt in the country.

The report by the Central Bank shows that the South African household debts vis-à-vis disposable income was about 82% in 2009 and this dropped to 75.6% during the first quarter of 2013. The US Federal Reserve reported that the household debts as a percentage of disposable income dropped from 128% in 2007 to 112% in 2011. The reports further indicated that experts were worried as this could get worse as banks push into unsecured loans. The high debt level incurred by the household sector has created economic and political concern in developed and developing countries. Despite the fact that household debt levels in SA are lower than that of the USA, within the context of SA, this figure remains at historically at high levels.

Among other researchers on household indebtedness in SA, Meniago *et al.* (2013) studied household debts and the financial crisis in SA using Econometric Methods; Hassan & Strazicich (2011) studied the financial crisis in Africa and Madubeko (2010) did research on the financial crisis and its impact on the South African economy. The following authors also provided the descriptive overview of household indebtedness in SA, *viz.* Mashigo (2006), Prinsloo (2002) and Daniels (2001). In the American context, Gartner (2013), Brown, Haughwout, Lee, & van der Klaauw (2013) and Cynamon & Fazzari (2008) delivered reports on the subject. The former study focused on the USA alone and employed Regression method and charts for analysis. The second study only focused on the exploratory analysis using the bar charts and the last used univariate Econometric Methods. None of these studies used extensive Multivariate Methods in their studies and none did a comparative analysis between SA and the USA.

1.2 Problem statement

A country is said to be in a good state of finance if it has low debt levels. This applies to both the government and the households in that country. When providing loans to customers, banks must be very prescriptive and discriminative. The 2007-2009 USA financial crises have had a significant impact on both the wealth and economic behaviour of many households in the world.

The USA household debt was recorded as being 127% during 2007-2008. According to International Monetary Fund report of 2012, mortgages, credit card debts, and consumer loans increased household debts in the USA from 99% to 138% of income. The Canadian economic outlook report 2012, confirmed that the newly revised debt levels are close to the peak witnessed in the USA, at the height of the subprime mortgage crisis. He further reported that the USA debt to disposable income is in the range of 140% subsequent to American's deleveraging after the recession.

Household indebtedness has increased rapidly across countries at different levels of economic development. In countries like SA, the rapid growth of microcredit ramped up the debts of many of the poorest households in the world (Chen, Rasmussen & Reille, 2010). The crisis has had a huge impact on unemployment and interest rates, low or lack of investment and saving by members of household, hunger for borrowing, inability to repay creditors, *etcetera*. As discussed in section 1.1, the level of household debts to disposable income in both the American and the South African context has decreased as compared to the figures during the financial crisis. Despite this decrease, these figures are still high and such high levels of household debts may have adverse effects on the economy. With the current state of household indebtedness, the household sector may run the risk of being too exposed to several adverse shocks such as unemployment rates, price shocks and shocks from income, *etcetera*. As literature indicates, financial crises emitted devastating shocks to vulnerable economies, especially SA where the level of corruption in terms of finances is at its height. This study extensively investigates the significant causes of high levels of household indebtedness and the shocks that the related determinants might create to the economy.

In view of the above, this study mainly focuses on the 2007-2009 financial crisis and its impact on the household indebtedness in SA and the USA. A comprehensive multivariate analysis is used to help ascertain which variables are good predictors of the rising indebtedness and how long these shocks will be present in the economies of SA and the USA. Furthermore, the study also investigates which of these two countries will be more vulnerable to these shocks. Based on the findings of this study, measures may be proposed to help in the alleviation of credit conditions thereby encouraging a stable financial household sector in the two countries.

1.3 Contribution of this study

Lindhe (2000) and Hansen (2003) highlighted that there are a number of studies which demonstrate a positive relationship between financial cycles (i.e. macroeconomic variables) and household debts. These studies used econometric methods. The focus of these studies was on a particular country and as a result, no comparison with other countries was made. Unlike previous studies, this study employs multivariate techniques such as Fisher's Linear Discriminant Analyses (Fisher's LDA), the Vector Autoregression (VAR) and the Vector Error Correction Model (VECM). Fisher's LDA allows the researcher to study the differences between two or more groups of objects with respect to several indicator variables simultaneously (Klecka 1980). In this study the two groups are SA and the USA and household debts determinants to disposable income are the discriminating variables for this technique. Moreover, this technique is also used in this study to determine the differences and/or similarities between these groups in relation to the average household debts determinants. The condition for constructing and continuing the analysis with the VAR or VEC models depends on whether the results reveal features of non stationarity and no cointegration between the variables under study. The multivariate cointegration and causality analysis precede the analysis with Fisher's LDA.

This study identifies and uses groups of variables that define household debts according to Modigliani (1966)'s theory and information from past studies. The results of this analysis may be used by other researchers to study differences between other countries relative to their household debts. These results may also be interpreted to understand the influence each determinant has on the economic wellbeing of households in each country. Suggestions on how to deal with household debts given the relationships anticipated may be formulated and submitted to the responsible authorities. The outcomes may also help the researcher understand the household debts during the financial crises and may also be used to avoid future crises.

This study is of importance not only to advance the researcher's knowledge of the determinants of household debts in SA and the USA. It is also meant to support the current debate on the effectiveness of saving enticements, on the role that institutions can play and on comprehensive financial systems for all. Apart from making a contribution to existing literature, this study may give researchers in the field of economics an idea about the interrelations between household debts determinants. It may be easier for them to formulate policies that may effectively help in reducing the problem on the basis of the contribution of each determinant. The recommendations made for this study may be used by policy makers to advocate policies that may encourage household savings and in turn curtail household debts. The findings could also help in understanding the extent to which the USA financial crisis has affected SA and how long the effects will remain effective. Again, these results could contribute to the existing literature and encourage researchers to use Fisher's LDA rather than descriptive statistical analysis to compare other countries' economies. The results from this empirical analysis may help us understand the differences and/similarities between SA and the USA relative to the related household debts.

Since no studies on this topic have been undertaken before, this study forms a basis for an application of multivariate analysis to household debt determinants for other countries during the financial crises. In essence, to the best of the researcher's knowledge, this study is the first of its kind to use a multivariate approach to study household debts in SA and the USA. No study has tried to construct the VEC model using household debts determinants. Additional novelty includes the comparison of household debts determinants by using Fisher's LDA. Another contribution includes investigating the discriminatory power of the determinants of household debts in these countries. The two measures such as the Cronbach's alpha and the Keiser-Meyer-Olkin have not been used in the past in studies similar to one being undertaken and therefore this serves as another innovation about this study. These measures have only been used in FA where questionnaires are used as data collection tools. The former is used in this study to assess if the selected independent variables are true measures of household debts and the latter is used as a measure of sample adequacy.

1.4 Research aims and objectives

The goal of this study is to investigate the significant difference between SA and the USA in terms of household debts in lieu of the financial crises. To achieve this goal, multivariate techniques such as the VECM/VAR and Fisher's LDA are used to address the following aims and objectives.

1.4.1 Research aims

The following are the aims of this study.

Research Aim 1.4.1.1: *To identify the constraints associated with the 2007-2009 financial crisis and to discover how SA and the USA household debt levels have been affected before, during and after this crisis.*

Research Aim 1.4.1.2: *To identify potential factors contributing to the rise of household indebtedness.*

Research Aim 1.4.1.3: *To investigate whether the chosen household debt determinants measure the said construct.*

Research Aim 1.4.1.4: *To construct models for household debt in SA and the USA for the period 1990 Q1- 2013 Q1. This will help to ascertain the relationship between household and other macroeconomic determinants.*

Research Aim 1.4.1.5: *To determine the forecasts of household debts in SA and the USA.*

Research Aim 1.4.1.6: *To identify the variables with more discriminatory power in the systems of household debts in these countries.*

Research Aim 1.4.1.7: *To investigate the differences and/or similarities between SA and the USA in terms of the average score profiles on household debt determinants.*

Research Aim 1.4.1.8: *To formulate some suggestions pertaining to how the two countries' household debts can be reduced to a reasonable level.*

1.4.2 Research objectives

The objectives of this study include:

Research objective 1.4.2.1: *To critically screen literature related to the 2007-2009 financial crisis and to identify those parameters related to household indebtedness in SA and the USA. To graphically analyse* *possible trends in the South African and American level of household indebtedness in relation to the periods "before", "during" and "after" the aforementioned crisis.*

Research Objective 1.4.2.2: *To consider related studies and economic theory to help identify factors contributing to the rise of household indebtedness.*

Research Objective 1.4.2.3: *To use the Cronbach's Alpha to assess whether the selected determinants are real measures of household debt.*

Research Objective 1.4.2.4: *To use time series data and the application of the VECM to analyse the relationships between household debt and its main determinants during the period Jan 1990 Q1- 2013 Q1.*

Research Objective 1.4.2.5: *To use the GIRF to forecast household debts in SA and the USA.*

Research Objective 1.4.2.6: *To use Fisher's LDA approach to determine linear combinations of household debt determinants in SA and the USA.*

Research Objective 1.4.2.7: *To use the Mahalanobis distance measure and Hoteling's test to identify the variable(s) that account for most differences in these two countries.*

Research Objective 1.4.1.8.8: *To utilize the results of the study to formulate some recommendations to reduce household debt levels of the two countries in question.*

1.5 Research questions / Hypothesis

In this section, specific research questions and a hypothesis are provided.

1.5.1 Research questions

Research Question 1.5.1.1: *What parameters are associated with the 2007-2009 financial crisis? What is the level of South African and American household debts before the inception, during and after the period of the aforementioned crisis?*

Research Question 1.5.1.2: *What are the main explanatory factors regarding the relationship between household debt and other macroeconomic determinants?*

Research Question 1.5.1.3: *Are the chosen household debt determinants measuring the same construct?*

Research Question 1.5.1.4: *What is the relationship between household debt to disposable income and related determinants?*

Research Question 1.5.1.5: *What are the forecasts of household debts in SA and the USA?*

Research Question 1.5.1.6: *Which of the determinants of household debt discriminate between the two countries?*

Research Question 1.5.1.7: *Do SA and USA households differ significantly in relation to their debt levels?*

Research Question 1.5.1.8: *What measures can be taken in SA and the USA to reduce household debt to a reasonable level?*

1.5.2 Research hypothesis

This study evaluates the following hypotheses:

Research Hypothesis 1.5.2.1: *This study hypothesises that the financial crisis of 2007-2009 has a negative impact on the level of household indebtedness in SA and the USA.*

This hypothesis is presented as follows:

H_0: The financial crisis has a negative effect on household debts in SA and the USA.
H_1: The financial crisis has a positive effect on household debts in SA and the USA.

Research hypothesis 1.5.2.2: *This study also hypothesises that on average SA and the USA are not different on the basis of the determinants of household debts.*

$$H_0 : \mu_{RSA} = \mu_{USA}$$
$$H_1 : \mu_{RSA} \neq \mu_{USA}$$

1.6 Research methodology

The research method that was followed to compile this report is of a descriptive and inferential nature. The method includes a quantitative overview of the existing literature. The main motivation for undertaking this study, apart from using theory to investigate the determinants of household in SA and the USA, is to use Fisher's LDA in investigating the difference between these countries. The long-run and short-run relationships between household debts and the related determinants are also investigated with multivariate techniques such as the Johansen Multivariate cointegration and Toda-Yamamoto Granger causality tests. Data used in this study spans the period *1990 Q1- 2013 Q1* and is sourced from the OECD, SARB, STATSSA, US Federal Housing Finance Agency, the USA Department of the Treasury and other relevant sources. Information related to this study is obtainable from amongst others the following sources; JSTORE, Springer, Wiley Interscience, online and net library, Sage journals, Social Science Research Network (SSRN), Science Direct, Blackwell Synergy, search engines such as Yahoo and Google, *etcetera* also helps in this research. Data is captured on Statistical package for Social Sciences (SPSS) version 21 and analysed with Statistical Software Analysis (SAS) version 9.

1.7 Possible results and conclusions

In this study, the following outcomes are expected:

- Obtaining information concerning the determinants of household debt and financial crisis in SA and the USA.

- Determining that the results for VECM will reveal causal relationships between household debts in SA and the USA.
- Determining the independent variables used in the analysis measure household debts in SA and the USA.
- Finding that the results highlight the difference and/similarities between SA and the USA.
- The findings confirming the generalisability of LDA results.

1.8 Scope limitations and delimitations

1.8.1 Limitations

A few limitations are noted as follows:

- The unavailability of data of some of the variables could compel us to exclude these variables or use the proxies if available. Not all the variables are measured on the same scale for SA and the USA, therefore the researcher may have to use a smaller number of variables than anticipated, for instance, availability of tax rates in the USA databases and quarterly unemployment rates in SA. The shortcoming of such measures is that the model may not accurately capture the phenomenon under review. Moreover, the use of proxies may not correctly exemplify the characteristics and behaviour of the actual variable, which may result in misleading findings.
- The study is limited to only the households of SA and the USA. No other country may be considered. There are very limited studies that used Fisher's LDA to compare two countries on the basis of their household debts. This may serve as a limitation on the number of sources consulted.

1.8.2 Delimitations

- Since some of the data are expressed in different measures, such as millions, ratios, billions, indices, monthly, quarterly, *etcetera*, transformations are used to make the variables similar.
- This study uses data collected from 1990 Q1-2013 Q1 giving a total of 93 observations. This caters for the methods of data analysis used in this study and furthermore helps in obtaining best linear unbiased estimates (BLUE), robust and

The results from the multivariate cointegration and causality analyses are estimated and discussed in this chapter. These results are based on the different tests described in the previous chapter. This chapter also provides and discusses the results obtained from the discriminant analysis. It is expected that these results help in deciding whether there is/are difference/similarities between SA and the USA.

Chapter 5

This chapter presents the summary and conclusion of the study. Recommendations for further research are also given in this chapter.

References and Appendices are given after the last chapter.

CHAPTER 2

THEORY AND LITERATURE REVIEW

2.1 INTRODUCTION

2.2 THEORETICAL FRAMEWORK

2.2 EMPIRICAL FRAMEWORK

2.3 STUDIES THAT ADOPTED THE LCH, PIH AND JOHANSEN'S PROCEDURE

2.4 FISHER'S LDA APPROACH

2.5 SUMMARY

2.1 Introduction

In order to provide a framework for the investigation carried out in this study, this chapter presents an overview of the conceptual theories explaining household debts. Also reviewed in this chapter is the literature regarding the classification of the related determinants. This serves as a guide in selecting the variables to be measured. This framework also facilitates this study in estimating the statistical relationship of the variables in order to construct the model. Empirical studies on the determinants of household debt particularly in the USA and SA are also reviewed.

2.2 Theoretical framework

There are a number of theories formulated which seek to explain household indebtedness. One of these theories was developed by Keynes (1936, in Modigliani 1986). Keynes thought of the subject in the form of the absolute income hypothesis and Modigliani followed it up with the life cycle hypothesis (LCH) in 1954. Later Friedman (1957) devised a theory with the permanent income hypothesis. This section provides a review of these theories and identifies the connection between them. Also discussed is the model's implication for household borrowing. The consequences of relaxing some of the assumptions of the framework are also examined.

2.2.1 The Keynesian Theory

Keynes' (1936) theory is more focused on the savings. This theory assumes that consumption is determined by the current level of income. Keynes was worried that as a country and its people became richer, they would not spend as much and would save more. Consequently, this might threaten future growth. The author views savings in the context of theory of demands and the consumption function. Income, he regards it as the systematic determinant of individual and national saving. He emphasised that the average propensity of the households to save would increase as these households reached a higher income level. Keynes also postulated that as consumption of households increased, their income would also increase. This increase in consumption, however; would not be as great as the increase in income. Keynes' consumption function is linear with a constant marginal propensity to consume and is described as follows:

$$C = a + bY \qquad [2.1]$$

Where ***C*** represents the consumption spending by households

a = represents autonomous consumption

b = is the marginal propensity to consume (MPC)

Y = is the disposable income

Though Modigliani and Friedman accepted the Keynesian theory, not enough evidence could be gathered to accentuate it. Hence the two theories, the LCH and PIH, were developed in order to explain the saving and borrowing behaviour of households.

2.2.2 The Life-Cycle and Permanent Income Hypotheses

From the economic perspective, the life-cycle and permanent income hypotheses are useful points of departure for examining trends in household borrowing. The LCH of savings is an initiative of Ando & Modigliani (1963) and the permanent income hypothesis was devised by Friedman (1957). Both these hypotheses were derived from the Keynesian theory. The proposition for the LCH is that household savings and consumption are a reflection of the life cycle stage of the household. Modigliani (1986) further proposed that consumption is a linear function of available cash and the discounted value of future income. This means that a household can choose a path of consumption to maximise utility over its life time subject to an intertemporal budget constraint (Rinaldi & Sanchis-Arellano 2006). This limits a household to consume no more than the sum of the present discounted value of its labour income and its current net worth (*i.e.* its asset holdings less its liabilities). During times when their income is low relative to average income over the life-time, households may borrow to fund current consumption. They may alternatively run down their asset holdings. Debelle (2004) advises that by then the loan would be repaid (assets would be accumulated) in periods where income is high relative to average life-time income.

Ando & Modigliani (1963) based their proposal on the assumption that as income is expected to increase during working years and decline at retirement, households will be inclined to borrow when they are still young, save during middle age, and start spending during retirement. As the permanent income hypothesis (PIH) suggests, the debt position of an individual household can therefore be determined by the path of future income and the interest rate (relative to the discount rate) (Debelle 2004). As suggested by Yilmazer & DeVaney (2005), household borrowing can be explained by focusing on demand-side determinants. According to the life-cycle model, aggregate household debt depends on other factors such as the households' demographics, the expected path of future income and real interest rates. This study does not consider the demographic variable as one of the determinants of household debt, but income and interest rates are included in the list of determinants. The life-cycle and permanent income hypotheses consider consumer spending as a function of expected lifetime earnings, consisting of wage earnings and income from assets (Rinaldi & Sanchis-Arellano 2006). These two theories identify a number of variables that influence household choice of consumption and level of borrowing (Ando & Modigliani 1963, Friedman 1957 and Debelle 2004).

2.2.2.1 Household Debt Determinants

Future income: An increase in future income expectations drives consumption up, boosting borrowing.

Demographics: Individual income profiles vary substantially with age. Shifts in the age composition of the population can help explain changes in household debt.

Personal saving: The reduction of household saving causes sizable increase in household net worth associated with increased house prices and stock prices (Crawford & Faruqui 2011) future income. Emphasised by Debelle (2002) the higher house prices are relative to household income, and the more debt households have to incur to buy housing.

Uncertainty: The households' attitude towards risk such as income uncertainty and time-varying interest rates is also an important factor in life-cycle consumption and borrowing decisions. With uncertainty, households choose to build up precautionary savings/wealth. If uncertainty is reduced, the households' rational response is to reduce precautionary saving (Ando & Modigliani 1963). To emphasise this assumption as pointed out by Dynan & Kohn (2007), having this framework in mind, historically low real and nominal interest rates, a substantial decrease in macroeconomic volatility, changes in taxation, or the ageing of the baby boom generation, have all been pointed out as potential factors behind the observed increase in indebtedness.

Interest rates: Another macro-factor that may influence household debt is the interest rate. This is because a long period of low interest rate may be an incentive for households to borrow (Debelle 2004). For example, when the interest rate is halved, households can double their take-out loan and still face the same servicing cost (Ewell 2013).

2.3 Empirical framework

2.3.1 Procedure for Multivariate Time Series Analysis (MTSA)

This study intends building an econometric multivariate model using the VEC framework. The model involves one dependent (household debt) and more than one independent variable (related household debt determinants). The variables are selected by reviewing household debt-related literature and the three theories discussed in previous sections. After it has been decided on the relevant variables to analyse, they are transformed using the natural log transformation as suggested by Box, Jenkins & Reinsel (2008). This is done to address irregularities in the data. Sims (1980) suggested that all the variables should be included firstly endogenously in the VAR model. He further warned that to test for exogeneity using the same model is often not plausible due to limitations of insufficient number of observations for some variables or nature of the data analysed. As the data for this study are collected for two different countries, it is anticipated that not all the variables are measured on the same scale (e.g. millions, indexes, rands, dollars, *etcetera*) and that there might be inequality in the number of observations collected. To help address this problem, the data are transformed prior the analysis and depending on the level where variables reveal shortages, extrapolation or interpolation methods are used to maximise the number of observations.

The nature of data used in this study requires that the chosen variables be tested for stationarity. In the economic practice, two tests have been widely used to determine the order of integration (Heilmann 2010). The two traditional tests of unit root, the ADF and the PP, are used in this study together with the empirical stationarity test KPSS. The application of KPSS is intended to confirm the power of ADF and PP tests. Researchers in the field of econometrics cautioned that if the series do not display evidence of stationarity, then there is persistence. This study takes this advice into consideration. It is also noted in this study that if more than one variable proves to have unit roots, the Johansen (1988) framework is adopted to firstly test for long-run equilibrium relationships among these non-stationary variables. This approach takes into consideration multiple variables and is chosen as one of the data analysis techniques in this study. It has also been widely used by amongst others Chowdhury (1994) and Nijs, Dekimpe, Steenkamp, & Hanssens (2001).

The decision to build a VAR model in this study is based upon the condition that the series contains unit root, but no cointegration between the series exist. A VEC model is built if there are cointegrating series. This study anticipates both stationary and non-stationary series, as a result considers a mixed model. A Bayesian approach for estimating VAR model as suggested by Holden (1995) is considered. Theory suggests that this step be followed up with choosing the optimal lag structure and identification of the model. Once the estimated model has been chosen, diagnostic checks in relation to the assumptions of this model are performed to assess if the model is valid. This study uses the two information criteria AIC and SBC to select the appropriate lag. Lastly, the study uses the estimated model to do causality and impulse response analyses. Figure 2.1 summarises the framework as discussed here.

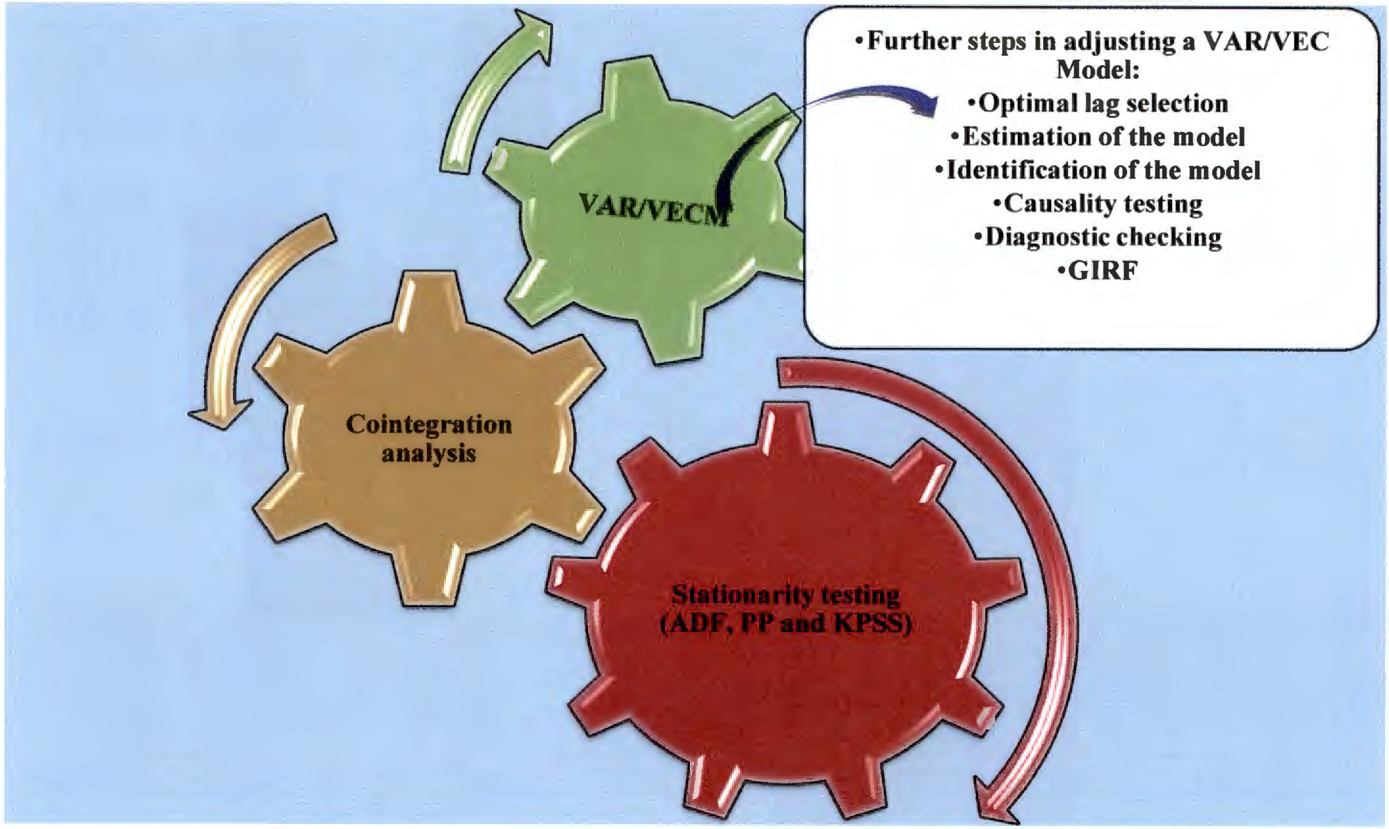


Figure 2.1: Model-building framework for MTSA

2.3.2 Nature of data

This study uses quarterly time series data selected with the guidance of the LCH and PIH theories discussed in this chapter. Information on the determinants of household debt in SA and the USA is analysed. The data used in this study data are based on economics and consist of observations on a variable or several variables gathered over time. Given that past events can influence future events and lags in behaviour are predominant in the social sciences, time is an important measurement in a time series data set. Unlike the arrangement of cross-sectional data, the consecutive ordering of observations in a time series conveys potentially important information. One important key feature making time series data more difficult to analyse than cross-sectional data is the fact that economic observations can rarely, if ever, be assumed to be independent of time. The independent variables selected for this study are thought to be related to one another, either positively or negatively. Once variables have been decided on and data on these variables collected, the next step is to test it for the presence of unit root as Figure 2.1 suggests.

2.3.3 Testing for a Unit Root

One basic requirement for the presence of cointegration is that all the time series that are integrated of the same order. Therefore the researcher needs tests to figure out whether the series are stationary or not (Heilmann 2010). As discussed in literature, non-stationarity is a property common to many macroeconomic and financial time series. This means that a variable has no clear tendency to return to a constant value or a linear trend (Nobel 2003). In time series analysis, non-stationarity occurs when the mean or the variance of a series are not constant over time. 'A process is said to be stationary if it is in a state of "statistical equilibrium" (Woodward, Gray & Elliott 2012:5)'. As Nelson & Plosser (1982) established, many macroeconomic time series contain unit roots dominated by stochastic trends. Unit roots are useful in examining if the series display properties of stationarity. This is so because a non-stationary regressor has a tendency of invalidating may practical results. If a series proves to be non-stationary, it has to be transformed to some stationary time series and then be analysed. Two types of stationarity exist; a strong and weak stationarity. A weak stationarity has the following properties;

for all *t* **[2.2]**

$$E(X_t) = \mu$$

$E(X_t^2) = \sigma^2$ for all t [2.3]

$\text{cov}(X_t, X_k) = \text{cov}(X_{t+s}, X_{k+s})$ *for all, t,k,s* [2.4]

These expressions mean flat looking series, without trend, with constant variance over time and no periodic fluctuations (seasonality). All the variables analysed in this study are collected on a quarterly basis and are therefore assumed to be non-stationary. Two commonly used stationarity tests proposed by Dickey & Fuller (1976; 1981) and Phillips & Perron (1988) are reviewed in Chapter 3. Once the assumption of stationarity has been validated, the next step assesses the series for cointegration as discussed in the next section.

2.3.4 Cointegration analysis

The concept of cointegration in this study rests on the ground-breaking works of Johansen's (1991) framework. This framework deals with the relationship between non-stationary time series sharing common stochastic trend, (*i.e.* two time series processes that both contain a unit root). Even though these series can wander arbitrarily, there might be some economic forces which tie them together to establish some kind of long-run equilibrium. This relationship is common for economic data. Modelling this long-run relationship by econometric methods and providing a unified framework for estimations and testing is the great methodological contribution of Johansen which is presented in this study.

In the Johansen cointegration test, the estimation of the maximum likelihood is based on the multivariate framework, such as the vector autoregressive (VAR) system. Brooks (2008) suggests this method as it gives advantage of identifying all cointegrating vectors within a given set of variables (*i.e.* multivariate cointegration). This method is also known for testing of hypotheses on the cointegrating relationships themselves. As literature commends, the Johansen approach has better asymptotic properties which yield more robust results. Once the cointegrating relations have been ascertained, the VECM is estimated to test for the long-run dynamics and correct for errors accounted for in the analyses.

The following are the summary points of the Johansen approach adopted in this study;

- Testing order of cointegration

- Specifying the VAR *(k)* order
- Testing for cointegration
- Normalisation
- Testing the hypothesis/diagnostic

2.3.4.1 Testing Order of Cointegration

This study performs the multivariate cointegration analysis with the help of the Johansen methodology prior to using the multivariate Granger causality test also developed from the multivariate framework. Later Fisher's LDA is used to compare SA and the USA based on the household debt determinants chosen. All these procedures are in the form of regression analysis. According to Granger & Newbold (1974) and Phillips (1986), regression results may be spurious if the variables are non-stationary. To avoid spurious estimation results, it is essential to determine the order of integration for each series. To affirm the order of integration for each series, the study applies the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests. The idea is to ascertain that the variables or series are integrated of order one I(1) so that the combination of these variables will produce residuals that will be integrated of order zero I(0) indicating existence of a cointegrating relationship.

However, it is still possible to include variables of different orders such as I(0) or I(1), as argued by Harris (1995). The argument by Harris is that stationary variables, I(0) might play a key role in establishing a long-run relationship between non-stationary variables, especially if theory suggests that the variables should be included. In this study, two procedures for testing stationarity among the variables before estimating the parameters are considered. The variables are visually inspected and the unit root testing using the ADF and PP are used. KPSS is used to confirm these two tests. The unit root testing approaches are discussed below.

2.3.4.2 Specifying the VAR(*k*) Order

As explained earlier, the procedure for the Johansen test uses the maximum likelihood system. Suppose the VAR of lag *k* is assumed, the following representation is expected.

Let $y_t = (y_{1t}, \ldots, y_{kt})'$, $t = 0,1, \ldots$, denote a *k*-dimensional time series vector of random variables of interest. The *p*th-order VAR process is written as

$$y_t = \delta + \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + \varepsilon_t \qquad [2.5]$$

where the ε_t is a vector white noise process with $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$ such that $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Sigma$, and $E(\varepsilon_t \varepsilon_s') = 0$ for $t \neq s$,. y_t is *n by 1* vector of I(1) endogenous variables (household debt and its determinants) in the VAR system.

Analysing and modelling the series jointly enable this study to understand the dynamic relationships over time among the series. This also helps to improve the accuracy of forecasts for individual series by utilising the additional information available from the related series and their forecasts.

2.3.4.3 Testing for cointegration

To confirm that the cointegrating vectors estimated are truly stationary, a graph of the residuals is plotted. This graph plots the residuals from the linear combination of variables. Visual inspection of the graph is expected to show that the residuals are stationary confirming the cointegrating relation. Having established the deterministic trend assumptions, the actual Johansen cointegration test is performed. Johansen (1988; 1991) provides two different likelihood ratio tests to determine the value of *r* such as;

- The trace test
- The maximum eigenvalue test

2.3.4.4 Normalisation

This study examines whether all the estimated coefficients are error corrected. The focus is on household debt as related to disposable income. The coefficient of the error correction term corresponding to this variable in the cointegrating equation is examined. To check whether a differenced household debt to disposable income variable is a cointegrating vector, the sign of the error correction parameter is expected to be negative and significant. If this error correction parameter meets these conditions, it is regarded as well behaved according to Aziakpono (2006). The true cointegrating vector becomes a candidate for normalisation.

2.4 The VECM Approach

This method developed by Johansen (1988; 1991), is the most widely used method in econometrics. Models using this approach are known as vector error correction (VECM) or cointegrating VAR (CIVAR) models. VECM can be seen as a scaled-down VAR model in which the structural coefficients are identified. The VECM is also a multivariate generalisation of the ECM model. This method is a commonly used tool for analysing and testing for common stochastic trends between non-stationary macroeconomic and financial time series data. Recent contributions to time series literature propose a time varying coefficients of the VECM. The justification for the VECM approach is that identification and testing for the significance of the structural coefficients, underlying the theoretical relationships, is important (Rao 2006). This study estimates the VECM on condition that there is a set of cointegrating vectors as suggested by figure 2.1. The procedure for estimation of the VECM is reviewed in chapter 3.

2.5 The VAR Model

As recommended by many researchers in the field of econometrics, the VAR model is one of the most successful, flexible and easy to use models for the analysis of multivariate time series. This model extends the univariate autoregressive model to dynamic multivariate time series. The model has been proven by researchers to be useful for describing the dynamic behaviour of economic and financial time series and also for forecasting. The VAR model was made popular by Sims (1980) who demonstrated that VARs provide a flexible and tractable framework for analysing economic time series. As illustrated in Figure 2.1, the VAR model in this study is estimated only if the variables proof not to be cointegrated.

A considerable number of studies considered the estimation of VAR. Amongst other researchers, Jacobsen and Vatne (2011) analysed the effect of house price changes on debt secured on dwellings in Norway. The study used both the macro-time series and micro-panel data to estimate a structural vector auto regression. Norwegian monthly data from December 1997 to April 2010 were analysed. All variables entered the VAR in levels, and none of them could (clearly) be rejected as non-stationary. Nevertheless, evidence of three cointegrating vectors was found and the results confirmed that the estimated VAR satisfied the stability condition.

The results of the macro analysis showed that positive house price innovations have positive and persistent effects on households debt secured on dwellings. Results from the micro-data analysis suggested that the effect of house price changes on the borrowing decision differs from the effect on the instalment decision among existing home owners. The VAR model predicted that the size of both loans and instalments increase with income. As loan sizes increase, the instalments fell with increasing collateral value. The results were in support of the existence of a wealth channel but did not provide support for a collateral channel.

One other study by Laina (2011) investigated the dynamic effects of total debt to GDP. The methodology used in the said study was based on time-series regression analysis, in which a structural VAR model was estimated. Granger causality tests were used to study the dynamic interactions, impulse response functions and forecast error variance decompositions. Quarterly USA GDP and real total debt growth data was collected for period 1959 to 2010 and this data was used in the analysis. The main finding of this study was that real total debt growth affects real GDP growth, but no feedback from real GDP growth to real total debt growth was found. The results showed that the response of real GDP growth to a shock in real total debt growth seemed to be transitory, but the level effect being persistent. In both cases the effect was found to be in the same direction.

2.6 The VEC Causality Test

The concept of causality in this study is preceded by the cointegration analysis of variables. As affirmed by Granger (1988), once the variables are found to be cointegrated, then there must be Granger cause in at least one direction. This means that a long-run equilibrium relationship holds.

For this reason, this study estimates the multivariate p*th* order of VECM for testing a modified Toda & Yamamoto (1995) Granger causality generalising all the household debts determinants selected. According to the concept of Granger causality, 'an independent variable causes a dependent variable' if and only if the past values of the former help to predict the changes of the latter, while 'the latter causes the former' if and only if the past values of the latter help to predict the changes of the former. However, Granger (1988) noted that if a set of variables are cointegrated, there must be short- and long-run causality which cannot be captured by the standard first difference in VAR model. In this case, this study implements the Granger causality test with the VECM framework as illustrated in chapter 3.

2.7 Studies that incorporated the LCH, PIH Theories and the Johansen's Procedure

In order to study the financial situation of households, this study takes a general look at the savings behaviour of households. Aron, Muellbauer & Prinsloo (1999) are three of the few who reviewed literature on household savings and debt. The study findings by Prinsloo (1999) showed that there was a decline in personal savings expressed as a percentage of disposable income over the past 20 years. His study also reported a rise in the household sector's indebtedness at a rate that almost corresponded with the rate of increase in the disposable income. The current study concurs with this claim in reference to the (South African Reserve Bank 2009) report which shows that 13.65 million households in SA had at their disposal a combined income of R1,631 billion. This figure characterises a nominal growth of 6.5% compared with 2008 when the household income was R1.1531 billion. Prinsloo's study was solely based on estimating the composition of household balance sheet, the determinants of savings and the impact of financial liberalisation in one country, SA. The study did not critically examine the indebtedness of households. This study highlights household debt and saving as the main concepts. According to the LCH, if households do not save, they will be forced to go into debt in order to satisfy their consumption needs.

Findings from few of the studies report a relationship between debt and savings. For instance, Kotzé & Smit (2008) amongst other authors are of the view that high levels of household debt are due to a lack of a comprehensive savings culture among people. Mainly this was believed to have been caused by financial illiteracy on the part of the consumers who spend almost all of their income leaving little if nothing for their savings and investments (Lorgat 2003). This claim is verified with the VECM.

There are several studies which have examined the impact of a financial crisis on household debt using econometric methods. For instance, Meniago *et al.* (2013) adopted the LCH and PIH to study the financial crisis and household debt in SA. The study used the long-run cointegration and short run ECM. Independent variables used in this study were CP, interest rates, GDP, prime rates and consumption expenditure. The findings confirmed that these variables moved together in the long-run. The study also confirmed that in the short run about 3.8% of the disequilibrium is corrected each quarter. The Engle-Granger cointegration results proved no connection between household debt and unemployment on the basis of the chosen time line.

Another study on the subject was done by Philbrick & Gustafsson (2010). This study explored the determinants of household debt to disposable income ratio in the Australian context using both the long-run cointegration analysis and a short run ECM. Considered in this study is the LCH. The results revealed a positive long-run cointegration between the change in the debt ratio and house prices. A negative cointegration was shown on interest rates. The results of the ECM affirmed a positive short run relation between debt ratio and the change in house prices and the consumer sentiment index. The opposite was reported for inflation and the error term.

Studies by Meniago *et al.* (2013) and Philbrick & Gustafsson (2010) serve as a base for the current study with regards to the theoretical foundations of LCH and PIH. These studies differ from the study in progress in view of the fact that the latter uses the multivariate approach and studies two countries simultaneously. Mulivariate cointegration and causality precede the Fisher's LDA in this study. Cussen, O'Leary & Smith (2012) reported on the impact of the financial turmoil on the quarterly bulletin of April 2012. Ólafsson & Vignisdóttir (2012) published a paper on 'households' position in the financial crisis in Iceland'. These studies used descriptive statistics for data analysis. Cussen *et al.* did a cross country comparison with descriptive statistics. The current study solely uses descriptive statistics to describe the distribution of variables, but not for comparison purposes.

Meng, Hoang & Siriwardana (2011) used the cointegrated vector autoregression (CVAR) model to explore the possible causes of Australian household debt. Seven variables (both demographic and economic) explaining household debt were utilised in the analysis. The results reported showed GDP to be the most substantial cause of household debt followed by house prices and the number of new dwellings. Interest, unemployment and inflation rates were reported to have a negative effect on Australian household debt with interest rates being the most significant. The study by Meng *et al.* is similar to the current one in that it similar variables save for the demographics coupled with others as suggested by the theoretical frameworks adopted are used. However, this study does not use the CVAR. Instead, the VECM is used in the analysis of data as it is anticipated that the results may show the cointegrating series following the nature of data collected.

The study by Anundsen & Jansen (2011) focused on self-reinforcing effects between housing prices and credit in Norway. This study estimated the interaction between housing prices and household borrowing in this country. A multivariate long-run and short-run analysis was performed where real household prices, real disposable household income and real household debt, conditioning on the real after tax interest rate, the number of house transactions and the volume of housing capital were used. Two cointegrating equations which determine equilibrium housing prices and household debt were identified. The study showed the long-run equations embedded in a system of two error correction equations which were estimated simultaneously. The equations were estimated using quarterly data spanning the period 1986 to 2008. Impulse responses demonstrated self-reinforcing feedback effects between real disposable household income and household debt. In the current study, only quarterly time series variables defining household debt for SA and the USA and multivariate approaches as used in Anundsen & Jansen (2011) study are considered. However, cross-sectional data are not taken into account.

The study by Kim (2011) used multi-equation econometric frameworks to investigate the impact of household debt on aggregate performance in the USA. In this study the VAR analysis capturing the transitory feedback effects observed a bidirectional positive feedback process between aggregate income and debt. The VECM estimation showed negative long-run relationships between household debt and output. This empirical model was also extended to include investment and corporate debt as independent variables. The results of this study were found to be in contrast with the results of empirical model without corporate sector variables. Also found was the negative long-run relationship between household debt and GDP. This was proven by the positive cointegrating coefficients in the cointegrating equations. Impulse response functions from these extended empirical models also indicated that investment may be an important channel through which household debt affects output. The current study takes note of some of the variables used in the previous study and the approaches used for data analysis.

Loganathan, Sukemi & Sanusi (2010) analysed the long-run and short-run relationship between external debt and macroeconomics performance of Malaysia. Government's revenue, balance of payment and government reserve in this study were referred to as macroeconomics performance. Annual time-series data spanning the period 1988-2008 and econometric techniques were used. The cointegration approach was employed to investigate the long-run relationship; and VECM was used to investigate the long term dynamics. The findings of this study suggested a significant long-rung and short-run relationship between external debt and macroeconomics variables performance. Also reported is a 13% of speed of adjustment to restore equilibrium condition in the long-run. Overall, the findings of this study showed that Malaysia's external debt was 'sustainable' with its macroeconomics performance, although have faced several unstable economic scenario for the last two decades. This study differs from the current in that only one country was considered and national annual data were used, otherwise the approach used and household related variables are a key.

Several studies and economic theories are in support of the fact that unemployment is one of the most significant determinants of household borrowing. This is also affirmed by Debelle (2004) who emphasises that this variable is one of the most significant negative shocks to household income. This author highlights that the majority of households enter into debt due the fact that they are unemployed. The study findings by Nieto (2007) who explored the determinants of household credit in Spain reported unemployment to be the most significant determinant. Further findings of this study through an econometric analysis also confirm that unemployment influence household credit in the short-run and the long-run. This confirmed that households are inclined to borrow more when unemployed. Unemployment is included as one of the determinants in the analysis.

Hatzinikolaou & Polasek (2005) adopted a multivariate cointegration approach to study the commodity price view of the Australian dollar. This study used Australian quarterly data for the period 1984 to 2003. The analysis yielded two cointegrating relations, one for the interest rate differential and the other for the nominal exchange rate. This study also uses multivariate cointegrating approaches, quarterly data, interest and exchange rates as some of the variables to be analysed.

The Hong Kong Monetary Authority (2002) reported the findings of the nexus of credit, household debt service and consumption. The study used quarterly data of 1982 to 2001 and the LCH and PIH were adopted to select the variables. Variables used were real private consumption, GDP, CP, real interest rate and best lending rate. The estimated parameters of the error correction model indicated that real GDP and housing prices sum to unity in the long-run relationship and the statistical testing could not reject the hypothesis. The results showed that consumer credit has been recording strong growth over the past decade attributable to both demand and supply side factors. In addition, the debt service burden has declined substantially over the past two years following successive interest rate cuts. It was also reported that households with a mortgage were likely to benefit from an increase in expendable income. Also shown was the fact that despite such signs of easing liquidity constraints, private consumption remains subdued.

The empirical findings from quantitative analyses suggested that real income and net wealth, as captured by property prices, were the major determinants of consumer spending. Furthermore, changes in the real interest rate and the debt-service burden were also found to be significant and negatively associated with consumption in the short run. Additionally, the study revealed that credit conditions, proxied by the spread between lending and borrowing rates, also appeared to play a significant role. Nevertheless, the quantity of consumer credit has little explanatory power for consumption in addition to that of income and property prices. This study takes note of the variables used in the one being discussed and the theory used to select them. However, financial variables are not considered in the current study. The focus is on macroeconomic variables.

As literature indicates, recent studies on OECD countries (for example by Antzoulatos 1996 and Ludvigson 1999) reveal that expansion in consumer credit is strongly associated with consumption growth. In addition as reported by Bacchetta & Gerlach (1997), credit tightness, as proxied by the difference between banks' lending and borrowing rates, appears to be negatively related to the growth of consumption. It is also found that financial liberalisation, particularly the deregulation of interest rates and the credit market, may reduce liquidity constraints faced by households, and in turn help support consumption (Boone, Girouard & Wanner 2001). These study findings suggest that consumer credit plays a significant role in determining consumption, and the supply of credit is influenced by the monetary conditions as well as the competitive environment in which financial institutions operate.

Maki (2000) confirmed that little evidence has been found that household debt-service burdens are negatively related to future consumption. Some theoretical models suggest that high debt-service burdens could make household consumption more sensitive to a drop in income or expected future income. As McCarthy (1997) hinted, there may also be an indirect causal relationship between debt service and consumption, implying that a rise in the debt-service burden may cause a rise in loan defaults, leading to a tightening in credit availability, and in turn a negative impact on consumption.

2.8 Fisher's LDA Approach

2.8.1 Introduction

In this study, the theoretical foundation that supports Fisher's Linear DA (LDA) is adopted. This helps in exploring the differences between the dependent and independent variables in the study. The model showing a linear combination of discriminant scores is described prior exploring these differences. The description of LDA is summarised below.

Afifi, May & Clark (2012) and Cramer (2003) define DA as a parametric technique to determine which weightings of quantitative variables or predictors best discriminate between two or more than two groups of cases and do so better than chance. Holmes's (2010) definition is that DA is a multivariate technique that distinguishes statistically between two or more groups of cases on the basis of measure of characteristics on which the groups are expected to differ. This technique works by finding one or more linear combinations of the predictors that yield maximal difference among the groups. As (Betz 1987) hinted, researchers use DA to characterise the nature of groups difference by interpreting the contributions of the individual predictors to this linear combination and also to identify which variables best capture group differences.

This method creates a discriminant function which is a linear combination of the weightings and scores on these variables. The maximum number of functions is either the number of predictors or the number of groups minus one, whichever of these two values is the smaller. As explained by Cramer, the best-known property for LDA is that it is a Bayes rule under a normality condition about the predictor distribution. More precisely, the condition requires that for the i_{th} class, $i = 1,..., K$, the p-dimensional predictor variable $x = (x_1,..., x_p)$ follows a multivariate normal distribution with mean μ_i and a common covariance Σ. Together with the prior probability $\pi_i, 1,..., K$, about the relative occurrence frequency for each class, this basic normality assumption leads to a Bayes discriminant rule which coincides with the rule of LDA.

Another way of deriving LDA originates from the consideration of group separation when there are only two groups, $K = 2$. This method was invented by Fisher (1936). The idea is to find a linear combination of the predictors and then use the model to investigate the differences between the groups. This method is reviewed in chapter 3.

2.8.2 Studies that used DA

As announced before, the DA method was first introduced by Fisher in 1936. He proposed that the evaluation of the countries' development level should be based on either using a system of indicators or one synthetic indicator. Jaba, Jemna, Viorică & Balan (2007) consider this approach to have disadvantages as it takes into account the level of the indicators and not the relationships between the variables that explain the economic development. Jaba *et al.* (2007) applied the DA for the identification of a typology for the Romanian counties by the economic development level. Literature suggests that the DA presents the advantage to synthesise a set of variables using the discriminant function (DF). Moreover, if used, it will express the relationship between the variables from the set used to characterise the development level and the DF *Score*. The authors used a set of variables of the development recorded at the level of the 42 counties of Romania, grouped into eight development regions. Their study used the independent variables that characterise the economic and social development level of the Romanian counties. The variables used are; number of physicians per 1000 inhabitants urbanization level, nominal net mean wages, higher education school population, number of people that have had accidents at work place, research and development expenses, number of inhabitants and number of libraries. In the present study, social variables are not used. Only economic variables determining household debt are included in the model as the independent variables.

In the Awh & Waters (2012) research, a case study involving economic, demographic and attitudinal characteristics of bank charge card holders was studied. The objective of the study was to establish the extent to which differences exist between active and inactive bank charge card holders. This study compared two groups of customers and the independent variables used were collected from the three characteristics mentioned. The DA results proved that the economic characteristics differentiate the active bank charge card holders from the inactive card holders. This study has few similarities with the present in that economic variables are used in the analysis and SA and the USA are compared based on these variables. Fisher's LDA is used to do these comparative analyses. Demographic variables are not considered in this study.

Turkan, Polat & Gunay (2011) used Linear DA, Logistic Regression (LR) and Artificial Neural Network Models (ANNM) to compare domestic and foreign commercial banks in accordance with their financial performances. The data used in this study involved 14 domestic and 17 foreign commercial banks operating in Turkey in 2009. The study analysed 19 financial ratios as a basis to compare the banks. FA was used to group the variables according to their correlations. Five factors explaining 84% of variance collected all 19 variables. DA, LR and ANN were built in SPSS Clementine 11, and were used to develop the model for predicting the group membership of commercial banks. In this case, DA, stepwise DA, LR, forward LR and four types ANN models were developed using five factors as independent variables and the status of the banks (one for domestic banks and two for foreign banks) as dependent variable. Stepwise LDA and forward LR analyses were used to check the most important variables.

The results of the study above showed ANN-Prune model to have the best predictive ability with a percentage of correct classification of 87.1. In addition, DA, forward LR and ANN-Quick models were the second best models giving similar results in terms of correct classification rate. This study is similar to one undertaken on the basis of the methods used for data analysis such as DA and FA. The latter is used in this study to determine the relationship between the independent variables and to assess the reliability and sampling adequacy of the variables and data. DA is used to investigate the significant differences between SA and the US in terms of the household debt determinants. The difference is that the current study does not consider using methods such as ANN and LR.

Singh, Sandhu & Kundu (2010) studied a case of investors' adoption of internet stock trading. The purpose of their paper was to examine whether investors who adopted internet stock trading perceived differently from those of non-adopters. Primary data based on 299 investors (149 adopters and 150 non-adopters) were analysed using multivariate techniques like FA and DA. Results indicated that attitude dimensions and demographic variables contributed significantly in classifying investors as adopters or non-adopters in Internet trading. As regards attitude dimensions, 'variety of financial products and safety' contributed significantly in discriminating between adopters and non-adopters of Internet trading followed by the factor such as 'convenience and transparency'. As far as the demographics were concerned, the mature/older, experienced, and businessmen investors were less likely to use internet stock trading as compared to young, inexperienced, and non-businessmen investors. The only difference between the present study and one being discussed is that the former uses secondary data on economic variables with a country as grouping variable. Otherwise, a combination of FA and DA is adopted in the analysis of data.

Narco-Burds (2010) used a DA to compare urban and rural nursing students that could account for an increased percentage of rural NCLEX-RN failures. A convenience sample comprised 398 graduates of an associate degree nursing programme in the Midwest. Variables included used both the demographic and academic characteristics of students. The study findings showed that there were no significant differences between urban and rural nursing students in terms of their demographic and academic personalities. This study is partially similar to the present in terms of the objective. As already explained present study is comparative and uses the DA to compare the two countries in relation to the household debt. Demographic variables are not used in the present study to cater for the methods chosen for data analysis.

Afolabi (2008) conducted a study which examined loan repayments among small-scale farmers in South-Western Nigeria. The study specifically identified the socio-economic characteristics that discriminate between loan defaulters and non-defaulters. A multistage sampling technique was used to select the respondents and a structured questionnaire was administered to them. Fisher's LDA was used and a model considering the socio-economic characteristics postulated for the loan defaulters and non-defaulters showed that six variables, *i.e.* age of farmers, gross farm income, net farm income, interest rates charged and farming experience were significant in differentiating between defaulters and non-defaulters. Although the current study has same intentions as Afolabi's, the researcher does not use a combination of social and economic variables. A grouping variable in the present study is country and is categorised as SA and the USA. We used time series data on economic variables only.

Another study was conducted where DA was used to identify characteristics that distinguish between the three groups of farmers. Kasiska (2007) designed an analysis that involved fully certified organic, partially certified organic and non-organic farmers in the Umbumbulu District, KwaZulu-Natal in SA. A questionnaire was administered to 200 farmers and a sample of 151 was selected for the analysis. Socio-economic variables were used as independent variables. The results of this study yielded two discriminant functions which suggested that farmers with higher household sizes, incomes, input costs per hectare and number of chickens owned, locations further from innovators and less risk aversion were more likely to be certified as organic as compared to other farmers. The current study is limited to only two groups and only economic variables are used as independent variables in the analysis unlike Kasiska's study. Questionnaires are not considered in this study.

Sherry, Henson & Lewis (2003) used DA to examine the differences between college students and adolescents. Five personality factors were used as variables in this study. Costa & McCrae (1992) furthered Sherry *et al.'s* study but instead compared three groups, with adults as the third group. These three groups had ten observations each. In this study, only two groups are used and the researcher intends to use more than ten observations for each group to cater for the statistical methods used. This study is similar to that of Sherry *et al.* because these authors used a grouping variable of two and the other study used a grouping variable of three. The researcher intends not to introduce the third country in the analysis.

The study by Kayaranni (2003) investigated differences between web-shoppers and non-shoppers, in terms of compatibility, relative advantage and demographics. Stepwise discriminant analysis was applied on a sample of 165 personally interviewed consumers, which showed that compatibility and relative advantage were overall successful, whereas, demographics were unsuccessful, in distinguishing web-shoppers from non-shoppers. Significant variables included three factors of compatibility (use of direct shopping; use of web browsing activities at home; and use of web browsing activities at the office), and two factors of relative advantage (motives and impediments). This study is similar to the current in part. The study under discussion uses the same method of data analysis to compare SA and the USA. Secondary economic data are used.

Uslay, Teach & Shwartz (2002) did a cross-cultural analysis of students' attitudinal differences. The study collected data on 343 students from the USA, Spain and Turkey. A stepwise DA was used to explore attitudes regarding entrepreneurship and compare them across cultures. The second goal was to examine the role of cross-cultural differences in the home environment and the role of parents' work experience on students' interest in entrepreneurship. The third goal was to understand gender differences regarding entrepreneurial interests. Others goals included exploring possible differences and factors that may have affected attitudes about entrepreneurship. According to literature, the USA is regarded as a melting pot of cultures at its economic height, Spain is a European Union member with a vibrant economic height, and Turkey does not only have chronically high inflation rates, but also extraordinary economic growth.

The results of the study showed that men and women had different attitudes about running one's own business. The females appeared to be less certain in doing so. The gender differences were also significant for the Turkish sample when analysed separately. This study is not completely different from the present. The study undertaken intends to use the same method of data analysis specifically for comparison purposes just like Aslay *et al.* The current study is also more into economic issues as opposed to the one under discussion.

Feigner & Sverdlov (2002) applied the multivariate DA (MDA) to evaluate the differences between three groups of patients (dependent variable) suffering from depression. A separate drug (independent variables) was administered to each group and the significant difference between these drugs was evaluated at a later stage. The results of this study showed that DA was successful in identifying the differences between these groups. Unlike this study, the current study only compares two countries, SA and the USA, and more than three independent variables are included in the analysis.

Galbraith & Jiaqing (1999) used both DA and Cluster Analysis (CA) to study industrial performance and related issues such as changes in the wage structure. The procedure was applied to time series data to firstly explore the group pattern, and then the forces that promote the formation of that group pattern. The independent variables used in their study are GNP and interest rates. Wage was used as a discriminant variable with levels 1 and 0 for high and low wages respectively. The present study also uses GNP and interest rates and other related independent variables. Though cluster and factor analyses have common characteristics of data reduction and determining relationships between variables, the former is not used in this study. However, multivariate time series data, FA and Fisher's LDA are employed in the analysis of data.

Galbraith and Jiaqing recommended the application of CA and LDA to fields that have time series data available. The use of CA and DA on multivariate time series data was first done by Galbraith and Calmon on industrial wage rate in 1990, 1994 and 1996. Further work on this area includes Ferguson & Galbraith (1998) on American industrial performance and wage structures, Galbraith & Kim (1998) on Korean industrial policy and Calistri (in progress) on industrial structure and wage change in the OECD. The present study also utilises multivariate economic time series data which are analysed following LDA and FA amongst other techniques just as the authors recommend.

Venable & Thompson (1998) used the DA to investigate what the psychological variables distinguish between care takers who prematurely take their children out of counselling and those who do not. This study used six psychological constructs. Using this technique, the authors obtained results distinguishing parents who prematurely terminated their children's counselling from other parents. The present study also uses two groups (SA and the USA) and several number of economic variables determining household debts. The researcher expects to get the results that help in distinguish one country from another on the basis of these variables.

Walsh *et al.* (1996) effectively used DA to compare 9th graders and 12th graders on a scale measuring values. Their results indicated that values such as physical prowess and prestige were more discriminative of 9th graders, whereas authority, working conditions and variety were discriminative of 12th graders. In this study, the researher expect to see the results describing the difference between the countries and also to see some of the variables showing discriminative power on one country and other variables discriminating on the other.

Doyle, Saunders, & Wong (1992) used MDA in their studies to compare success factors of foreign firms operating in Britain. Ghymn (1983) used this technique to compare different criteria for evaluating suppliers in different countries. Ramakrishnam (1986) tested USA bank performance in the international marketplace; whilst Materna (1988) tested for differences among decision makers global orientation. None of these studies did a comparative analysis of countries based on their household debt determinants.

Mahoney (1979) used a DA to compare the characteristics of pro and anti-individuals relative to factors related to attitude toward sex education in the public schools. Based on previous research on the subject, it was predicted that pro and anti-sex-education individuals would differ in some of the characteristics. The ability of these factors to distinguish between the groups subsamples was examined through LDA and the results suggested significant differences. Variables that showed discriminating characteristics on the groups were also identified. There is similarity between the two studies in that the present study compares two countries using DA though different variables are used. It is expected that the results of the current study show significant differences between these countries in some if variables. Also expected are the results revealing that some variables have more discriminating power than others.

Altman (1968) used a two-group DA for bankruptcy analysis of industrial firms. For each firm, 22 variables were initially chosen for analysis on the basis of their popularity in the literature and their potential relevance to the study. These ratios were then allocated to five categories: liquidity, profitability, leverage, solvency, and activity. One ratio was then chosen from each category for inclusion in the discriminant model. Altman's results showed that the DF had done well in predicting non-bankrupt firms. He calculated a *Z* score and used this as the basis to classify the firms as non-bankrupt or bankrupt. Firms having a Z-score above a cut-off point fall into the non-bankrupt sector, while those firms having a Z-score below the cut-off are bankrupt firms. Though the present study is similar to Altman's in that metric dependent variable is chosen as a variable of interest and a two group DA is used, classification analysis is not performed.

2.8.3 Studies that used FA

FA provides a method of reducing the number of variables to be treated by discarding the linear combinations that have small variances. Only the variables with large variances are studied, hence simplifying the financial evaluation a group. Below is literature on studies that applied FA in the analysis.

Vogel (1976) used FA to identify groups of countries in which similar marketing practices could be used. Arkaah & Yaprak (1988) used this method to identify factors perceived to be important in foreign direct investment decisions. As literature suggests, there must be a relationship between independent variables. In the current study, FA is employed to confirm the interrelationships between household debt determinants and to check household debts. The sample used in the analysis is alsochecked for adequacy with the FA procedure.

Pinches & Mingo (1973) used FA to screen the data in a manner similar to the previous study. From the original group of 35 variables, FA enabled these authors to screen out 6 key variables to be used in the final DFs. In total, 180 bonds were analysed and randomly assigned to one of two sample groups, 132 forming the original sample for model development and 48 forming a holdout sample for testing purposes. Using Bartlett's V-statistic, or eigenvalue determination, three DFs of the six variables were determined. This MDA was used to correctly rate 92 of the 132 bonds, corresponding to a correct prediction rate of 69.7%. This encourages us to use that FA and DA as their combination have provide reliable results.

In Johnson's (1979) study, the original numbers of variables were 306 manufacturing firms and 159 retailing firms. Factor scores were determined simply by taking a linear combination (or composite) of the original variables, such as $F_i = b_1 y_1 + b_2 y_2 + ... + b_j y_j + ... + b_p y_p$, where F_i is the i_{th} factor score, y_j $(j = 1,2,..., p\_)$ were the original variables, and $b_1, b_2,..., b_p$ were the coefficients (weight) to be estimated. For manufacturing firms Johnson obtained 306 factor scores for each factor; for the retailers he obtained 159 factor scores. By using the simple correlation coefficient formula, he then calculated 61 factor loadings for each industry. Johnson study used the confirmatory factor analysis approach while the present study uses the exploratory FA approach instead.

2.9 Summary

In this chapter, the three theoretical frameworks, Keynesian, LCH and PIH developed by the Keynes, Ando and Modigliani and Friedman respectively are discussed and linkages between them established. These frameworks are used to identify the variables for this study and for the identification of an econometric model. Also reviewed are the MTSA, VEC and the Fisher's LDA empirical frameworks which are used in building the models and assessing their significance. Previous studies used univariate measures such as ANOVA and the variance to make comparisons. DA has not been used before in similar studies. Fisher's LDA is not only used in constructing a multivariate model, but also in investigating the differences between SA and the USA. Related literature to these frameworks is also reviewed to help identify the variables to be used in the analysis and also to emphasise the significance of the methods of data analysis for this study.

CHAPTER 3

DATA AND METHODOLOGY

"A statistical analysis, properly conducted, is a delicate dissection of uncertainties, a surgery of suppositions"

~ M.J. Moroney

"... a hypothesis test tells us whether the observed data are consistent with the null hypothesis, and a confidence interval tells us which hypotheses are consistent with the data."
~ William C. Blackwelder

3.1 INTRODUCTION

3.2 DEFINITION OF HOUSEHOLD DEBT

3.3 THE RESEARCH APPROACH

3.4 PRE-DATA ANALYSIS

3.5 MODEL SPECIFICATION

3.6 MODEL ESTIMATION PROCEDURE

3.7 PARAMETER AND MODEL DIAGNOSTIC TESTING

3.8 THE IMPULSE RESPONSE FUNCTION (IRF) ANALYSIS

3.9 OVERVIEW OF DISCRIMINANT ANALYSIS

3.10 DISCRIMINANT FRAMEWORK

3.11 ASSESSING VALIDITY OF THE DA RESULTS

3.12 SUMMARY

PART A: ECONOMETRIC ANALYSIS

3.1 Introduction

This chapter thoroughly discusses the methodological process as illustrated in Figure 2.1 of the previous chapter. The first stage identifies the variables and secondly the nature of the data used in the analysis which are then assessed. In this stage of the methodology, the data are pre-processed and the characteristics of each variable are identified for subsequent analysis. The procedure in this stage begins with normalisation of each variable to achieve comparability among variables. Since same order of integration is essential for all endogenous variables in the VECM model, the stationary condition and integration order of each variable is identified using unit root test.

In the third stage, multivariate cointegration and causality frameworks are used to evaluate the long and short run relationship between variables. Based on the outcomes in the second step, further steps such as lag selection, model estimation, identification and diagnostic checking are carried out. The VAR or VECM is chosen based on cointegration test results. The final stage of methodology is forecasting using the GIRF. This chapter also deliberates on Fisher's LDA framework. This framework is used as a basis for constructing a linear discriminant function which is used in this study to help identify differences between SA and the USA.

3.2 Why households enter into debts?

There are various reasons why households enter into credit transactions. In essence, the use of consumer debt is related to the eagerness of consumer sacrifice to future consumption in order to obtain greater satisfaction from current consumption (van der Walt & Prinsloo 1995). The authors explained that households surrender future consumption and use income earned at a later stage to settle debts and meet interest commitments. They further said that for this reason, households will have less spendable funds at their disposal at that stage if they do not increase repayments of their credit commitments. Prinsloo (2000) defines household debt as an obligation or liability arising from borrowing money or taking goods or services on credit. As Prinsloo clarifies, debt a contract states the terms of borrowing, the interest and redemption payments that the borrower must make and what collateral the borrower has to provide. He explains secured debt as an instance where the borrower provides collateral, which the lender is entitled to take over if the borrower does not make the promised payments. Generally, mortgage debt is secured on houses, other buildings or land.

3.3 The research approach

This study uses quantitative methods. Secondary data is used and the literature is employed as a source of information regarding which variables to analyse. The research approach of this study is motivated by the methods that are used to analyse data, the objectives stated and the type of data envisaged.

3.3.1 Information gathering

Relevant information for this study was gathered from different available sources. The data used in this study are secondary and are available on several websites. Multivariate methods are used for data analysis. The following readily available sources are consulted; JSTORE, Springer, Wiley Interscience, online and net library, Sage journals, Social Science Research Network (SSRN), Science Direct, Blackwell Synergy, search engines such as Yahoo and Google, *etcetera* are also used to help obtain information needed for this research.

3.3.2 Methods of data analysis

The methods employed in this study are the Multivariate time series and Fisher's LDA. The intention of this study is to use Fisher's LDA to compare SA and the USA in relation to the countries' average household debt determinants. No study has been undertaken before where this technique was used to compare countries using the determinants of household debt. The study rearranges the collected data in a systematic manner for interpretation purposes. The analytic tools such as the descriptive, inferential, graphs, *etcetera* are used.

3.4 Procedure for data analysis

This study analyses multiple time series variables using multivariate methods. It is essential that preliminary data analysis prior the actual investigation be undertaken. It is expected as theory dictates that economic data are not stationary by nature. These data are normalised to induce stationarity. Three separate procedures are followed since the study employs three different techniques for data analyses. The stages are shown on Figures 3.1, 3.2 and 3.3 for the preliminary, econometric and discriminant analyses. Figure 3.1 summarises the stages for preliminary data analyses and is general for the three methods. The procedure below combines all the three stages including the pre-data analysis.

1) The first step of this analysis involves gathering data that have been modified into the same frequency and same base year.
2) This step performs univariate data analysis and checks the adequacy of the sample and reliability of the variables. This helps to understand the distribution and the nature of the data to be used. Grubbs test is used to check if the data contain outliers or not.
3) The third step tests the variables for the Unit root. The Augmented Dickey-Fuller (ADF),

Phillips-Perron's and the KPSS tests are used. The convergence of household debts is also looked at in this step.

4) The model estimation procedure begins in this step. The study chooses the optimal lag length that is best suitable for the model in this step. This is done using the Akaike information criterion (AIC) and Schwartz Bayesian criterion. The variables are cointegrated using the Johansen test. This involves testing for the number of cointegrating relationships (rank of the matrix) using both the trace and maximum eigenvalue statistics. If the results of the Johansen test supports that the variables are found to be cointegrated, then the analysis is continued with the VECM, otherwise the VAR approach is used. A modified Granger causality test by Toda & Yamamoto (1995) is used to determine the relationships between variables.
5) Diagnostic tests such as the autocorrelation test, the normality, heteroskedasticity and the stability tests are performed to check if the estimated model is good. If the model is found good, then it can be used for policy analysis.
6) This step uses the GIRF to do the forecasting.
7) This step performs the Fisher's LDA to investigate the differences and/or similarities between SA and the USA.

This procedure is summarised on Figure 3.1.

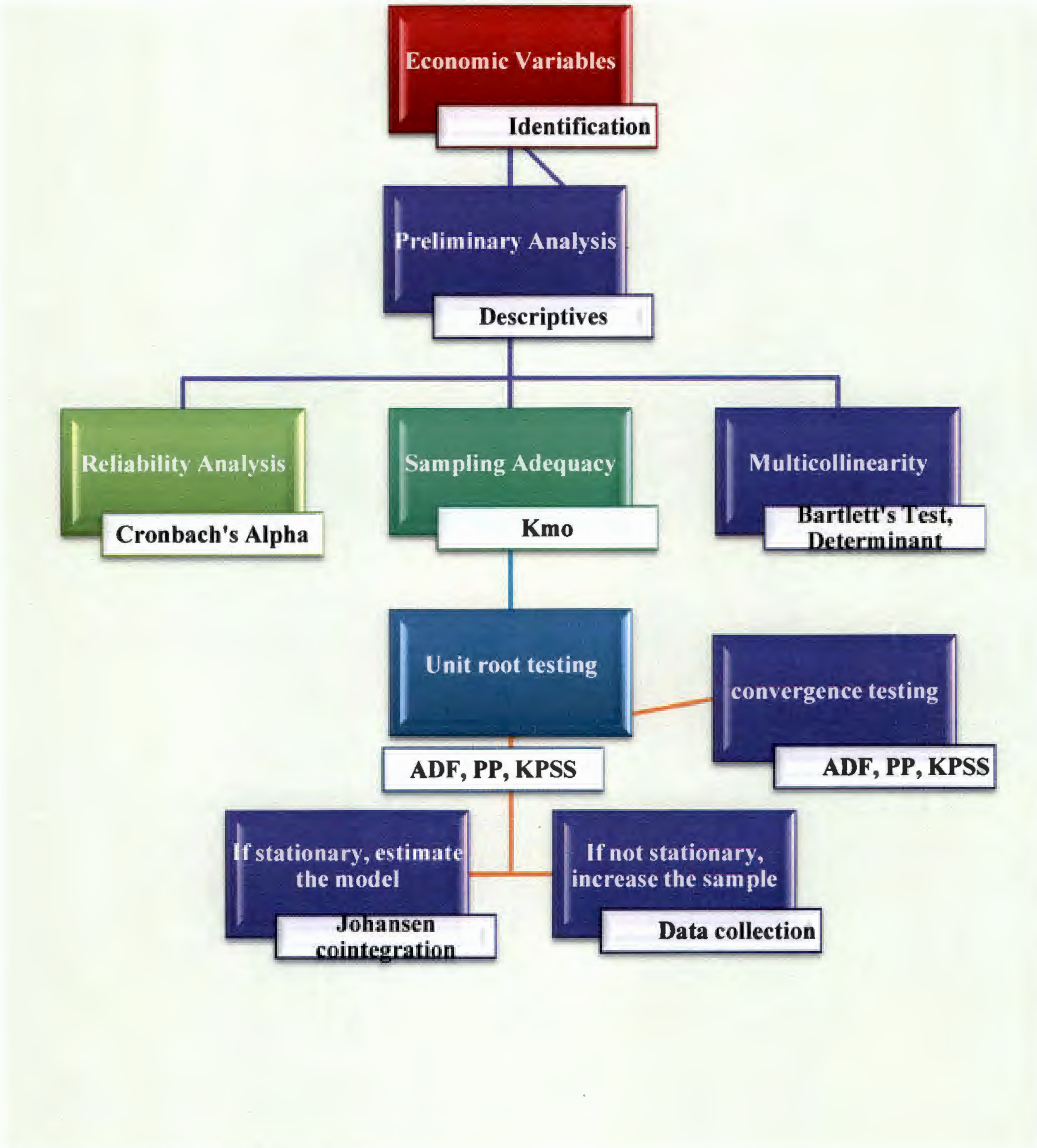


Figure 3.1: Stages in the preliminary data analysis

3.4.1 Variables and data processing

This section gives a brief definition of the variables selected for analysis in this study. The LCH and PIH together with the reviewed literature were used in selecting these variables. These theories are discussed in Chapter 2. The target population for this study is households in SA and the USA. Time series data on household debts determinants in these countries are used in the analysis. These data are collected on a quarterly basis and cover the period 1990 Q1 to 2013 Q1 resulting in 88 observations. The data for SA are sourced from the SARB, STATSSA, OECD and ABSA Bank amongst others. The USA data source is the USA Department of Commerce, USA Federal Housing Finance Agency, the USA Department of the Treasury and OECD. The capturing of data is done on the Statistical Package for Social Sciences (SPSS) version 21. The actual analysis of data is done using the Statistical Software Analysis (SAS) version 9.3.

The level of household indebtedness in SA and the USA is influenced by many factors. Based on theory and literature, this study uses house prices, consumer prices, household income, prime rate, gross domestic product, household expenditure, household savings, personal taxes, unemployment rates, mortgage debt, household investment, interest rates and personal taxes to help us portray the behaviour of households towards debt taking. The analysis involves both the dependent and independent variables. Below is a summary of the data available for this research.

Dependent variable

The dependent variable might be influenced or modified by some exposure and is also referred to it as an outcome variable. This study employs two dependent variables on separate occasions. The VECM approach is adopted to analyse the long and short run relationship between the variables with household debt to disposable income as one variable of interest and household debt to disposable income is used in this regard as a dependent variable. The definition of this variable is given below.

Household debt to disposable income

As Clark & Daniel (2006) expounds, the ratio of household debt to disposable income measures the percentage of the average households' disposable income that is used to repay debt. The motivation for this according to these authors is that this ratio will increase if the interest rate increases, since, for example, mortgage repayments will increase, or as a result of the acceptance of additional debt. Additionally, this ratio will also increase if the amount of disposable income decreases. A further example given is that an increase in the domestic fuel price consequently increases the price of most goods. Henceforth, this ratio is a direct measure of the average households' affordability and thus indicates their attitude towards spending.

The second dependent variable is the country (SA and the USA). This variable is non-metric and is also mutually exclusive. The discriminant analysis requires that a dependent variable of this nature and the independent variables must be measured in a continuous scale as those used in this study.

Independent variables

An independent variable is a hypothesised cause of influence on a dependent variable. A discrepancy in the independent variable is determined by the performance of a dependent variable. This study analyses independent variables derived from the literature review and also on the basis of the LCH and PIH. It is expected as literature suggests that these variables are a good measure of household debt. FA is used in this study to ascertain this assumption. It is also anticipated that some of these variables have a discriminatory power on the countries. The theoretical discussion in the previous chapter suggests a number of possible variables to include in the econometric analysis of the determinants of household debt-to-income ratio. The following is a short discussion of the variables chosen to include in the model.

House price

In general, most of the people obtain household loans for the purpose of financing their mortgages. Therefore, house price plays a significant role in determining the rise and the fall of household indebtedness. According to a study by Subhanij (2007), rise in house prices could encourage households to consume more and build up debt by betting on higher expected future incomes. Moreover, the higher house prices are relative to household income, the more debt households have to incur to buy housing (Debelle 2002). This variable is included to capture both the wealth effect, from a change in asset prices, and also the fact that when house prices increase households must take on more debt in order to purchase housing.

Interest rate

Interest rate is another macro-factor that influences household debt. This is because a long period of low interest rate may be an incentive for households to borrow. For example, when the interest rate is halved, households can double their take-out loan and still face the same servicing cost. Kearns & Woods (2006) employed household level data to explore why *households fell into mortgage arrears during the 1990s in Ireland. His study suggests that a modest rise of interest rates would result in a substantial repayment burdens for significant number of newly mortgaged households. Kearns and Woods further advised that the continued mortgage growth lending may lead to a higher rate of mortgage arrears among households, especially if the lending criteria are relaxed and households accept the higher repayment burdens. The researcher includes the interest rate (IR) in the analysis, i.e. the gap between lending and deposit rates, to capture the effect of changes in financial regulations and innovations.*

Inflation rate

The decline in inflation has two effects on household borrowing. First, the reduction in borrowing costs allows a greater number of households to borrow and/or increase the average level of debt per borrowing household. Second, with lower inflation, the real value of the debt (which is fixed in nominal terms) is not eroded as fast as in the past. Households may be surprised in later years by the proportion of income still required to service their debt, and hence have lower than desired consumption. The higher aggregate debt/income ratio means that households will be more exposed to shocks, and will also remain exposed for a longer period of time than in the past (Debelle 2004). The consumer prices (CP) is used as a proxy of the inflation rate and this variable is included to capture the effect of the speed at which the debt-to-income ratio erodes. This also helps in investigating if households are inclined to go in for debt when the prices of goods and services fluctuate.

Unemployment rate and household income

As Debelle (2004) cautions, the largest and most significant negative shock to household income is unemployment. According to his report, greater household indebtedness and higher debt service levels will heighten the sensitivity of households to a rise in unemployment, amplifying the effect of a negative shock to the economy. Households with debt will find it more difficult to maintain their mortgage payments through a period of unemployment, and hence will be more likely to default (Debelle 2004). An increase in the unemployment rate implies a tighter budget for households and curbs household consumption from all perspectives. The results of his study shows that personal saving rate increases during recessions, indicating that households consumed less and thus borrowed less. Therefore, the unemployment rate is expected to have a negative effect on the debt service ratio. The *inclusion of income in the model helps to ascertain if an increase in the debt-to-income-ratio is due to the lack of money in the households.*

Gross Domestic Product, household consumption expenditure and unemployment rate

Rising GDP growth coupled with a period of low unemployment may lead to higher consumer confidence and rising consumer spending. This may cause the private consumption expenditure (PCE) to expand accordingly. In other words, households could increasingly take out loans to finance their consumption. A measure of GDP variance is included in the analysis in order to capture the effect of output volatility. This also measures how the increase in the economic growth of the country could possibly affect household debt. As the LCH theory affirms, the more households spend, the more their income will decrease and the more they will opt for more debt.

Household savings

*Clark & Daniel (2006) define personal saving as the amount by which the current income of households exceeds their current expenditures. Prinsloo (2002) views this factor as that part of current income, after the payment of direct taxes that is not consumed or transferred as part of household current expenditure. According to Prinsloo, dissaving occurs when current expenditures exceed current income. Reports reveal that the saving of a household or of any other organisation is equivalent to the increase in the net asset value of the household or organisation, in terms of accounting practices. Moreover, as credit commitments of households increase, a decline in their saving occurs. Real household savings are integrated into the model to investigate whether an increase in household debt is due to a lack of saving behaviour among households as theory suggests. Based on the Keynesian theory, the interest rate is inversely related to household consumption and positively related to household saving. As the interest rate decreases, household tend to consume more and save less (as banks offer less interest rates), implying more debt accumulation (*Mason & Jayadev 2012*).*

Table 3.1: Summary of variables

Variable	Measure	*Source*
Dependent		
1. Country; SA (0) or the USA (1)	Nominal	
2. Real household debt to disposable income	Quarterly data at seasonally adjusted annual data measured in millions , 1990-Q1 to 2013-Q1	SARB and US Department of Commerce (Bureau of Economic Analysis)
Independent		
1. House Price (HP)	Quarterly data at seasonally adjusted annual rates, 1990-Q1 to 2013-Q1	ABSA Bank and US Federal Housing Finance Agency
2. CP	Yearly indices on the same period of the previous year, 1990-Q1 to 2013-Q1	Stats SA and OECD.STAT
3. Real household income	Seasonally adjusted Quarterly data at seasonally adjusted annual rates, 1990-Q1 to 2013-Q1	SARB and US Department of Commerce (Bureau of Economic Analysis)
4. Real GDP at market price	Quarterly data at seasonally adjusted annual rates, 1990-Q1 to 2013-Q1	SARB OECD.STAT US Department of Commerce (Bureau of Economic Analysis)
5. Real household consumption expenditure	Quarterly data at seasonally adjusted annual data, 1990-Q1 to 2013-Q1	SARB US Department of Commerce (Bureau of Economic Analysis)
6. Real household savings	Quarterly at seasonally adjusted annual rates, 1990-Q1 to 2013-Q1	SARB US Department of Commerce (Bureau of Economic Analysis)
7. Exchange rates	Quarterly data at seasonally adjusted annual data, 1990-Q1 to 2013-Q1	US Department of Commerce (Bureau of Economic Analysis)
8. Unemployment rates	Quarterly data seasonally adjusted, 1990-Q1 to 2013-Q1	STATSSA and US Department of Commerce (Bureau of Economic Analysis)
9. Interest rates	Quarterly data at seasonally adjusted annual data, 1990-Q1 to 2013-Q1	SARB US Department of Commerce (Bureau of Economic Analysis)

This study analyses quarterly data that are seasonally adjusted. It offers several advantages over studies that employ annual data and also allows the researcher to capture the inflationary trend experienced in SA and the US thereby enhancing the strength and reliability of the results.

3.5 Preliminary data analysis

The intention of this section is to assess the variables and sample chosen for data analysis. This is achieved by calculating the descriptive statistics such as the mean and the variance of individual variables. Time series plots are used here to describe the movement of each variable over time. Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures (Trochim 2006). This will help in understanding the description of and differences between household debt in each country and also to univariately assess the differences between each pair of household debt determinant per country.

Furthermore, the sample used in this study is checked for adequacy. Since this study uses multivariate techniques, it is required that the sample and variables chosen for analysis be as adequate and reliable as possible. One of the objectives of this study is to investigate the long and short run relationships between the variables. This implies that household debt determinants chosen are expected to be correlated to a certain degree and should multicollinearity be present, it has to be addressed. It is therefore imperative to make sure prior the analysis that these requirements are satisfied. The Keiser Meyer Olkin (KMO) is used in this study to determine whether the chosen sample is acceptable or not. This measure is calculated using the following formula;

$$KMO = \frac{\sum (correlations)^2}{\sum (correlations)^2 + \sum (partial\ correlations)^2} = \frac{\underbrace{\sum\sum r_{ij}^2}_{i \neq j}}{\underbrace{\sum\sum r_{ij}^2}_{i \neq j} + \underbrace{\sum\sum a_{ij}^2}_{i \neq j}} \qquad [3.1]$$

Where

r_{ij} = Pearson correlation between items i and j

a_{ij} = partial correlation coefficient between items i and j

As literature dictates, the KMO value must be greater than 0.5 but less than or equal to 1 to conclude that the sample is sufficient. Kaiser (1974) suggested measures in the ranges; 1 to 0.9 be characterised as marvellous, 0.8 and 0.89 as meritorious, 0.7 to 0.79 as middling, 0.6 to 0.69 as mediocre, 0.5 to 0.59 as miserable and 0 to 0.49s as unacceptable. To evaluate the problem of multicollinearity, this study uses the Bartlett's test for sphericity. This is a test statistic used to examine the hypothesis that though the variables are related, the correlation between them is not significant. The formula below describes the Bartlett's test:

$$\chi^2 = -\left[(n-1) - \frac{1}{6}\left(2p + 1 + \frac{2}{p}\right)\right]\left[\ln S + p\ln\left(\frac{1}{p}\right)\sum \lambda_i\right] \qquad [3.2]$$

where

p= number of variables, k = number of factors, λ_i= *ith* eigenvalue of the sample covariance matrix and the $df = (p-1)(p-2)/2$.

Churchill (1995) proposed the rejection of the associated hypothesis at least 5% level of significance. The variables are checked to determine whether they are truly measures of household debt with Cronbach's alpha. This is a measure of how well a set of items measure a single unidimensional latent construct. Byrne, Shavelson, and Muthen (1989) and Chin (2000) suggest the variables chosen as predictors must be highly associated with each other and represent a single concept. This suggestion is supported by amongst others Blaha, Merydith, Wallbrown and Dowd (2001) and Diamantopoulos and Siguaw (2006). Cronbach's coefficient alpha ranges from 0 to 1 with values closer to 0 implying that the items do not measure the same construct and values closer to 1 providing a different implication.

The following formula is used to describe Cronbach's (α):

$$\alpha = \frac{kr}{1 + (k-1)} \qquad [3.3]$$

Where k is the average correlations between the variables and r is the number of variables.

Kline (1999) and Cronbach & Shavelson (2004) use the following rules of thumb to describe Cronbach's alpha α:

≥ 0.9	Excellent
$0.8 \leq \alpha < 0.9$	Good
$0.7 \leq \alpha < 0.8$	Acceptable
$0.6 \leq \alpha < 0.7$	questionable
$0.5 \leq \alpha < 0.6$	Poor
$\alpha < 0.5$	unacceptable

3.6 Model specification

Guided by both the theoretical and empirical literature in the previous sections, a number of possible determinants of household debts have been suggested. Amongst others, the following list constitutes factors affecting households' indebtedness; interest rates, house prices, the tax system, consumer prices, unemployment rate, household income, gross domestic product, household consumption expenditure, household savings and discount rates. All the variables are expressed in natural log form. This study specifies the following multivariate VAR model both for the aforementioned countries:

$$\ln HHD = f(\ln HP, \ln CPI, \ln INC, \ln IR, \ln GDP, \ln HC, \ln HS, \ln ER, \ln UR) \qquad [3.4]$$

Expressed in regression form, the model becomes

$$\ln HHD_t = \beta_0 + \underbrace{\beta_1 \ln HP_t}_{+} + \underbrace{\beta_2 \ln CPI_t}_{-} + \underbrace{\beta_3 \ln INC_t}_{-} + \underbrace{\beta_4 \ln IR_t}_{-} + \underbrace{\beta_5 \ln GDP_t}_{+} + \underbrace{\beta_6 \ln HC_t}_{+} + \underbrace{\beta_7 \ln HS_t}_{-} + \underbrace{\beta_8 \ln ER_t}_{-} + \underbrace{\beta_9 \ln UR_t}_{-} + \mu_t \qquad [3.5]$$

Where

β_0 = intercept

ε = error term

$\beta's$ = coefficients of the independent variables as described in Table 3.1 below:

Table 3.2 Variable names and descriptions

Name	Description	Sign
*ln*HP	Natural log of house prices	+
*ln*CP	Natural log of consumer prices	-
*ln*INC	Natural log of income	-
*ln*IR	Natural log of interest rate	-
*ln*GDP	Natural log of gross domestic product	+
lnHC	Natural log of household consumption	+
lnHS	Natural log of household savings	-
lnER	Natural log of exchange rates	-
lnUR	Natural log of unemployment	-

Based on the results of the causal hypotheses as stated in Chapter 1, the corresponding bidirectional hypothesis has to be examined. The causal relation in a VAR model is investigated and the vector of variables is formulated in equation [3.1]. This model type relates each variable to its own past values and to past values of all other variables included in the model (Dritsakis 2007).

3.7 Model estimation

The following figure summarises the econometric analytic procedure used in this study.

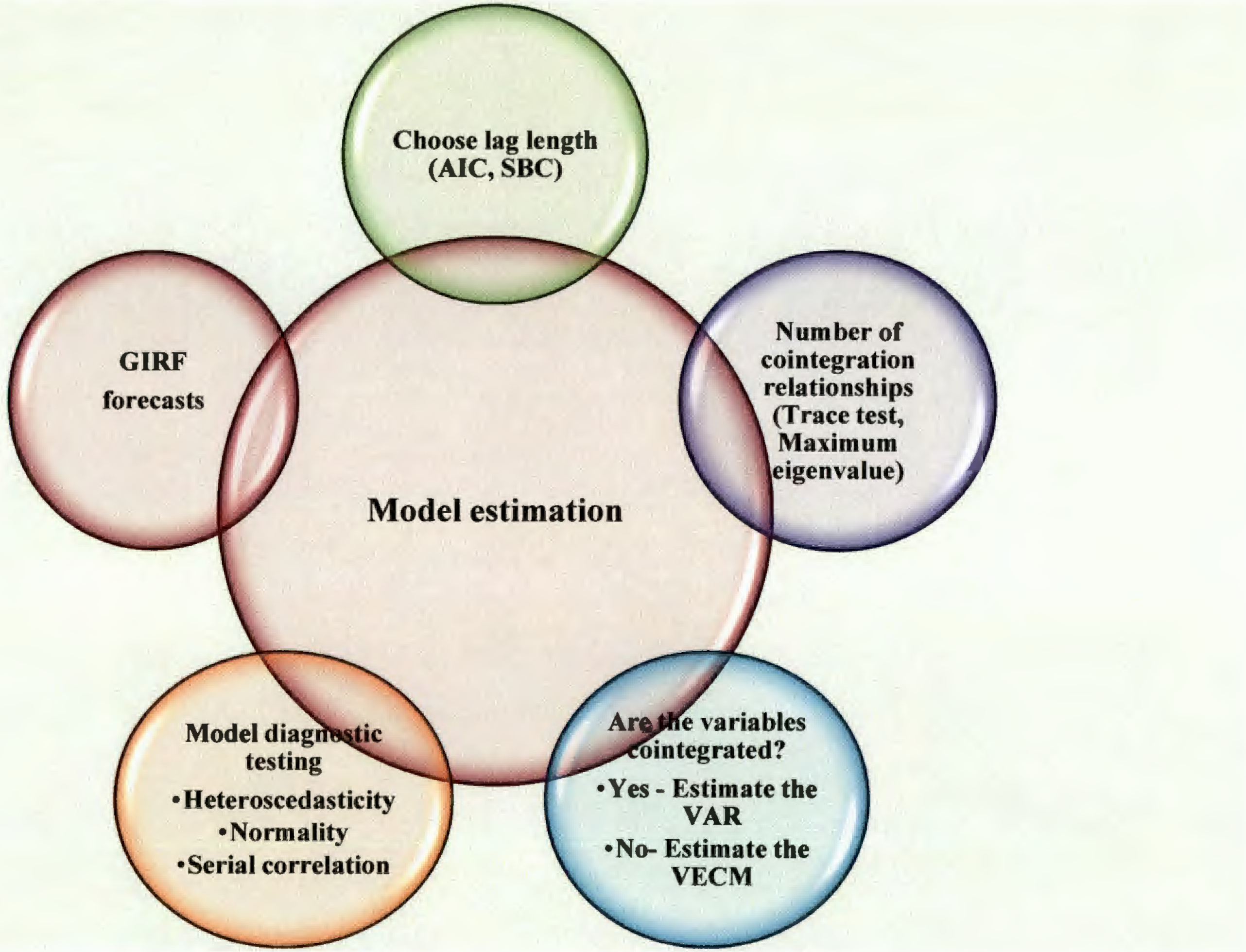


Figure 3.2: Steps for estimation, diagnostic checking and forecasting

3.7.1 Unit Root Tests

This study follows a three-stage procedure to test the direction of causality between household debt and related determinants $(HP, CPI, INC, IR, GDP, HC, HS, ER, UR)$ of the two countries. The first stage tests the order of integration using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests. The model is estimated using VEC frameworks as discussed in the next section.

3.7.1.1 Augmented Dickey-Fuller Test (ADF)

The study uses the ADF test to assess the variables for & as discussed in Chapter 2. This testing procedure was first introduced by Dickey & Fuller (1976; 1981). The basis for ADF test is upon estimating the regression

$$\Delta Y_t = \alpha_0 + \beta_0 Y_{t-1} + \sum_{i=1}^{k} \beta_i \Delta Y_{t-i} + \varepsilon_t \qquad [3.6]$$

Where

Δ is the first difference operator; t is the time drift; k represents the number of lags used and ε is the error term; α's and β's are the model bounds.

The correct value for k is determined using the Akaike and the Schwarz-Bayesian information criteria. The ADF test includes a constant and time trend. For the decision rule, assuming that the series $\{ Y_t t^T{}_{=1} \}$ follows the AR *(p)* process, Hamilton (1990) shows that the rejection or acceptance of the null hypothesis of a unit root,

$H_0 : \phi_1 = 1$ *(the series has unit root)*

$H_A : \phi_1 < 1$ *(the series is stationary)*

is based on running the regression

$$Z_t = \mu + (\phi_1 - 1) Y_{t-1} + \sum_{j=1}^{p-1} C_j Z_{t-j} + \varepsilon_t \qquad [3.7]$$

Where

$Z_{t-j} = Y_{t-j} - Y_{t-j-1}$ for j = 0, 1, 2,..., *p-1* and ε_t is a white noise process. H_o: $\phi_1 = 1$ is rejected if the test statistic

$$\hat{\tau}_{ADF} = \frac{\hat{\phi}_1 - 1}{se(\hat{\phi}_1)} \qquad [3.8]$$

is less than an appropriate critical value of t for the Dickey-Fuller Test. $se(\hat{\phi}_1)$ is the standard error of $\hat{\phi}_1$. Alternatively, the null hypothesis is rejected if the probability value of the test statistic is less than some convectional level of significance. The ADF tests use a parametric auto-regression to approximate the autoregressive moving average (ARIMA) structure of the errors in the test regression.

3.7.1.2 The Phillips-Perron Test

The Phillips-Perron (PP) test was suggested by Phillips & Perron (1988) as a test for stationarity. A PP unit root test differs from the ADF test mainly in how they deal with serial correlation and heteroskedasticity in the errors. In particular, the ADF tests use a parametric autoregression to approximate the ARIMA structure of the errors in the test regression and the PP ignore any serial correlation in the test regression. The PP test is estimated by equation

$$\Delta Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t \qquad [3.9]$$

Where

ε_t is $I(0)$ and may be heteroskedastic.

The test statistic is calculated with the equation:

$$\tau_{pp} = \left(\frac{t_{\phi-1}}{\xi}\right)\Gamma_0^{\frac{1}{2}} - \frac{N}{2}\left(\frac{\xi^2 - \Gamma_0}{\xi \hat{\sigma}}\right) se(\phi - 1) \qquad [3.10]$$

Where

$t_{\phi-1}$ is the test statistic of $\phi-1$, $se(\phi-1)$ is the standard error of $\phi-1$, $\hat{\sigma}$ is the standard error of the test regression and Γ is the truncation lag. The asymptotic distributions of the PP test statistics are the same as those of the ADF test statistics. Here again, the null hypothesis of a unit root H_o: $\phi = 1$ is rejected if $\hat{\tau}_{pp}$ is less than the appropriate critical value at some level of significance.

3.7.1.3 Kwiatkowski, Phillips, Schmidt and Shin (KPSS) Test

This section reviews an empirical test for stationarity KPSS invented by (Kwiatkowski *et al.* (1992). This test enables the researcher to test whether the series have a deterministic trend versus the stochastic trend. It is used in this study to confirm the decision made by the ADF and PP tests. This statistic tests the following hypothesis;

$H_0 : Y_t \sim I(0)$ (level(or trend) stationary)

$H_1 : Y_t \sim I(1)$ (difference stationary)

The KPSS test statistic is given as;

$$KPSS = n^{-2}\sum_{t=1}^{n}\frac{S_t}{\hat{\sigma}^2} \qquad [3.11]$$

Where $S_t = \sum_{i=1}^{t} e_t$ and $\hat{\sigma}^2$ is the estimate of the long-run variance of the residuals. The null hypothesis is rejected when the KPSS is larger than the critical value, because that is the evidence that the series wander from its mean.

3.7.2 Convergence of Household Debts

The economic justification of the convergence hypothesis arises within the standard Solow neoclassical growth model, but has been applied by several authors in other economic areas. Given the household debts of SA and USA, Φ_t and Θ_t, let Ω_t denotes the spread between the two debts, that is:

$$\{\Omega_t\} = \Phi_t - \Theta_t,.$$ [3.12]

Assuming that the time series $\{\Omega_t\}$ (*i.e.*, Ω_t is a sequence of real numbers for t=1,2, ...) has a limit of A, then:

$$\lim_{t\to\infty}\{\Omega_t\} = A.$$ [3.13]

Then if the power series in x, given by, $\sum_{t=1}^{\infty}\Omega_t x^t$, converges for a value $x = x_0$, it also converges absolutely for all x such that $|x|<|x_0|$ and diverges for all x such that $|x|>|x_0|$. Then following Busetti, Forni, Harvey & Venditti (2007), convergence can be modelled as an AR(*p*) form as:

$$\Delta\Omega_t = \alpha + (\rho - 1)\Omega_{t-1} + \sum_{j=1}^{p-1}\delta_j \Delta\Omega_{t-j} + \eta_t,$$ [3.14]

where $0<\rho<1$, $\alpha = A(1-\rho)$ is the mean of $\Delta\Omega_t$, η_t is white noise with mean 0 and variance, σ_η^2. Equation [3.14] is equivalent to the autoregressive model for testing for a unit root in Ω_t. If Ω_t is stationary, then $E(\Omega_t) = E(\Omega_{t-1}) = \alpha$. Following Harvey & Carvalho (2002) and Havery & Bates (2003), the household debts of USA and RSA will only converge if the differential $\Delta\Omega_t$ is stationary and bound bounded by a long-run variance. Testing equation [3.14] for the presence of a unit root gives rise to two definitions of the convergence hypothesis – one linked to long-run convergence and another linked to the concept of catching up in economics.

The absence of a unit root in Ω_t supports the convergence between the two debts. On the other hand, the nature of the deterministic trend should suggest a kind of catching-up or lagging-behind between the two debts. In order to examine the two issues simultaneously, equation [3.14] is modified to include a deterministic trend:

$$\Delta\Omega_t = \alpha + \beta t + (\rho - 1)\Omega_{t-1} + \sum_{j=1}^{p-1}\delta_j \Delta\Omega_{t-j} + \eta_t.$$ [3.15]

According to Gómez-Zaldívar & Gomez & Ventosa-Santaularia (2012), if the coefficient of the deterministic trend, β, is equal to zero, there is an indication of a divergence process. On the other hand, if the coefficient of the deterministic trend, β, is negative, then there is a loose lagging-behind process between the two debts, while if the coefficient is positive, then there is a catching-up process between the two debts.

3.7.3 Cointegration analysis

3.7.3.1 The VAR Model and Optimal Lag Selection

In order to choose appropriate lag length for the VAR order, this study uses the information criteria AIC and SBIC. These criteria are used the same way as in unit root testing. These information criteria ensure that residuals are Gaussian. The AIC was introduced by Akaike (1974) and the SBIC by Schwartz (1978) as measures used to choose the appropriate lag. Brooks (2008) however, suggested that if these information criteria provide conflicting results, the one which produce white noise residual and most economically interpretable results should be chosen. These criteria are estimated using the following equations:

$$AIC-2\left(\frac{l}{n}\right)+\left(\frac{k}{n}\right) \qquad [3.16]$$

$$SBC-2\left(\frac{l}{n}\right)+k\frac{\log(n)}{n} \qquad [3.17]$$

Where: l is the log of the likelihood function, k is the number of parameters in the model, n is the number of observations.

These criteria choose the model with small lags. The Johansen procedure is designed to statistically determine the number of cointegrating vectors r in the VAR (Gujarati 2003).

3.7.3.2 The Johansen Multivariate Cointegration Test

This section reviews a multivariate cointegration framework used for testing the existence of a long-run equilibrium relationship between the variables. The specification of VECM contains the cointegration relations, so it assumes that the economy converges to the long-run

ECT_{t-1} represents the deviation from equilibrium in period t and the coefficient δ characterises the response of the dependent variable (HHD) in each period to departures from equilibrium. The VECM leads to a better understanding of the nature of any nonstationarity among the different component series and can also improve longer term forecasting over an unconstrained model. The following are representations of unconstrained Johansen multivariate VECM model for HHD and the eight related determinants:

$$\begin{aligned}\Delta\ln HHD_t = \alpha_{0HHD} + \sum_{i=1}^{p}\eta_{iHHD}\Delta\ln HHD_{t-i} + \ldots + \sum_{i=1}^{p}\varpi_{iHHD}\Delta\ln UR_{t-i} \\ + \sigma_{1HHD}\ln HHD_{t-i} + \sigma_{2HHD}\ln HP_{t-i} + \ldots + \sigma_{10HHD}\ln UR_{t-i} + \upsilon_{1t}\end{aligned} \qquad [3.24]$$

$$\begin{aligned}\Delta\ln HP_t = \alpha_{0HP} + \sum_{i=1}^{p}\beta_{iHP}\Delta\ln HP_{t-i} + \ldots + \sum_{i=1}^{p}\varpi_{iHP}\Delta\ln UR_{t-i} \\ + \sigma_{1HP}\ln HP_{t-i} + \sigma_{2HP}\ln HHD_{t-i} + \ldots + \sigma_{10HP}\ln UR_{t-i} + \upsilon_{1t}\end{aligned} \qquad [3.25]$$

$$\begin{aligned}\Delta\ln CPI_t = \alpha_{CPI} + \sum_{i=1}^{p}\phi_{iCPI}\Delta\ln CPI_{t-i} + \ldots + \sum_{i=1}^{p}\varpi_{CPI}\Delta\ln UR_{t-i} \\ + \sigma_{1CPI}\ln CPI_{t-i} + \sigma_{2CPI}\ln HP_{t-i} + \ldots + \sigma_{10CPI}\ln UR_{t-i} + \upsilon_{1t}\end{aligned} \qquad [3.26]$$

$$\begin{aligned}\Delta\ln INC_t = \alpha_{0INC} + \sum_{i=1}^{p}\theta_{iINC}\Delta\ln INC_{t-i} + \ldots + \sum_{i=1}^{p}\varpi_{iINC}\Delta\ln UR_{t-i} \\ + \sigma_{1INC}\ln INC_{t-i} + \sigma_{2INC}\ln CPI_{t-i} + \ldots + \sigma_{10INC}\ln UR_{t-i} + \upsilon_{1t}\end{aligned} \qquad [3.27]$$

$$\begin{aligned}\Delta\ln IR_t = \alpha_{0IR} + \sum_{i=1}^{p}\vartheta_{iIR}\Delta\ln IR_{t-i} + \ldots + \sum_{i=1}^{p}\varpi_{iIR}\Delta\ln UR_{t-i} \\ + \sigma_{1IR}\ln IR_{t-i} + \sigma_{2IR}\ln HC_{t-i} + \ldots + \sigma_{10IR}\ln UR_{t-i} + \upsilon_{1t}\end{aligned} \qquad [3.28]$$

$$\begin{aligned}\Delta\ln GDP_t = \alpha_{0GDP} + \sum_{i=1}^{p}\rho_{iGDP}\Delta\ln GDP_{t-i} + \ldots + \sum_{i=1}^{p}\varpi_{iHHD}\Delta\ln UR_{t-i} \\ + \sigma_{1GDP}\ln GDP_{t-i} + \sigma_{2GDP}\ln IR_{t-i} + \ldots + \sigma_{10GDP}\ln UR_{t-i} + \upsilon_{1t}\end{aligned} \qquad [3.29]$$

$$\begin{aligned}\Delta\ln HC_t = \alpha_{0HC} + \sum_{i=1}^{p}\vartheta_{iHC}\Delta\ln HC_{t-i} + \ldots + \sum_{i=1}^{p}\varpi_{iHHD}\Delta\ln UR_{t-i} \\ + \sigma_{1HC}\ln HC_{t-i} + \sigma_{2HC}\ln GDP_{t-i} + \ldots + \sigma_{10HC}\ln UR_{t-i} + \upsilon_{1t}\end{aligned} \qquad [3.30]$$

$$\Delta \ln HS_t = \alpha_{0HS} + \sum_{i=1}^{p} \gamma_{iHS} \Delta \ln HS_{t-i} + ... + \sum_{i=1}^{p} \varpi_{iHHD} \Delta \ln UR_{t-i}$$
$$+ \sigma_{1HS} \ln HS_{t-i} + \sigma_{2HS} \ln HC_{t-i} + ... + \sigma_{10HS} \ln UR_{t-i} + \upsilon_{1t} \quad [3.31]$$

$$\Delta \ln ER_t = \alpha_{0ER} + \sum_{i=1}^{p} \Gamma_{ER} \Delta \ln ER_{t-i} + ... + \sum_{i=1}^{p} \eta_{ER} \Delta \ln HHD_{t-1}$$
$$+ \sigma_{1ER} \ln ER_{t-i} + \sigma_{2ER} \ln UR_{t-i} + ... + \sigma_{10ER} \ln HHD_{t-i} + \upsilon_{1t} \quad [3.32]$$

$$\Delta \ln UR_t = \alpha_{0UR} + \sum_{i=1}^{p} \varpi_{iUR} \Delta \ln UR_{t-i} + ... + \sum_{i=1}^{p} \eta_{iUR} \Delta \ln HHD_{t-1}$$
$$+ \sigma_{1UR} \ln UR_{t-i} + \sigma_{2UR} \ln ER_{t-i} + ... + \sigma_{10UR} \ln HHD_{t-i} + \upsilon_{1t} \quad [3.33]$$

Equations [3.24] to [3.33] are summarised in a reduced form of the VAR as follows;

$$\Delta y_t = \alpha + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Phi_i^* \Delta y_{t-1} + \varepsilon \quad [3.34]$$

Where Δ is the differencing operator and all the variables are I (0), $\Pi = \alpha\beta'$ and α and β are vector of matrices. $\ln HP$ is the natural log of house prices, *etcetera*. The existence of the long-run relations between the variables is determined through testing the significance of the lagged levels of the variables. When a long-run relationship exists between the variables, the *F* test indicates which variable should be normalised. In equation [3.25], where HP is the dependent variable, the null hypothesis that there are at most *r* cointegrating vectors;

$$H_0 : \lambda_i = 0 \quad \textit{for i=r+1,...,k}$$

The opposing hypothesis is $H_0 : \lambda_i \neq 0$.

Thus, the F test for this equation is denoted as $F_{HPI}(HPI \mid HHD, INF, INC, IR, GDP, HC, HS, TR)$.

To confirm the correctness of the results, this study uses the Johansen cointegration test. Johansen (1988) suggests two test statistics to test the associated hypothesis above. In addition, Fountis & Dickey (1989) suggested an examination of eigenvalues following the multivariate setting. Steps for this methodology are as follows;

Step 1: Fit linear multivariate time series as

$$
\begin{aligned}
X_{1,t} &= \phi_{1,1}X_{1,t-1} + \phi_{1,2}X_{1,t-2} + \ldots + \phi_{1,p}X_{1,t-p} + \varepsilon_{1,t}, \\
X_{2,t} &= \phi_{2,1}X_{2,t-1} + \phi_{2,2}X_{1,t-2} + \ldots + \phi_{2,p}X_{2,t-p} + \varepsilon_{2,t}, \\
&\quad \cdot \qquad\qquad \cdot \qquad\qquad \cdot \\
&\quad \cdot \qquad\qquad \cdot \qquad\qquad \cdot \\
X_{n,t} &= \phi_{n,1}X_{n,t-1} + \phi_{n,2}X_{n,t-2} + \ldots + \phi_{n,p}X_{n,t-p} + \varepsilon_{n,t},
\end{aligned}
\qquad [3.35]
$$

$$\Rightarrow \qquad X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_p X_{t-p}.$$

Step 2: Compute the largest eigenvalue, $\lambda_{\max}$, based on the characteristic equation

$$\left|\lambda^p \mathrm{I} - \Phi_1 \lambda^{p-1} - \Phi_p \lambda^{p-2} - \ldots - \Phi_p\right| = 0, \qquad [3.36]$$

Where I is the $p \times p$ matrix.

Step 3: Test of the null hypothesis of unit root is based on the following test statistic

$$\hat{\tau}_{mfd} = N(\lambda_{\max} - 1), \qquad [3.37]$$

Where $\lambda_{\max}$ is the largest eigenvalue based on step 2.

Step 4: At the 0.05 significance level, obtain the critical value from the table. Reject H_0 if $\hat{\tau}_{mfd} >$ critical value or alternatively if the observed probability is less than the level of significance.

The following are the Johansen trace and maximum eigenvalues formulae also calculated through these four steps;

$$J_{Trace} = -N \sum_{i=r+1}^{n} \ln(1 - \hat{\lambda}_{\max}) \qquad [3.38]$$

$$\Delta \ln HS_t = \alpha_8 + \sum_{i=1}^{p} \gamma_i \Delta \ln HS_{t-i} + \sum_{i=1}^{p} \vartheta_i \Delta \ln HC_{t-i} + ... + \sum_{i=1}^{p} \varpi_i \Delta \ln UR_{t-i} + \psi_1 ECT_{t-1} + \upsilon_{8t} \quad [3.47]$$

$$\Delta \ln ER_t = \alpha_9 + \sum_{i=1}^{p} \tau_i \Delta \ln ER_{t-i} + \sum_{i=1}^{p} \gamma_i \Delta \ln HS_{t-i} + ... + \sum_{i=1}^{p} \varpi_i \Delta \ln UR_{t-i} + \psi_1 ECT_{t-1} + \upsilon_9 \quad [3.48]$$

$$\Delta \ln UR_t = \alpha_9 + \sum_{i=1}^{p} \varpi_i \Delta \ln UR_{t-i} + \sum_{i=1}^{p} \gamma_i \Delta \ln HS_{t-i} + ... + \sum_{i=1}^{p} \tau_i \Delta \ln ER_{t-i} + \psi_1 ECT_{t-1} + \upsilon_9 \quad [3.49]$$

where

Δ is the first difference operator, ln is the natural logarithm, $i = 1,2,3,...,7$ (depends on the number of household debts determinants), β, ϕ, θ, κ, ρ, ϑ, γ are the parameters in the model. The residuals υ_{it} are assumed to be normally distributed and to follow a white noise process. ECT_{t-1} is the one period lagged error-correction term derived from the cointegration equation. The ECT_{t-1} variable is excluded from that model if the variables are not cointegrated. The optimal lag length p is determined by the AIC and SBC as discussed previously.

In the next step of this section, the Likelihood Ratio (LR) statistic is used to determine the direction of causality between the variables of interest. The expected results are that household debt to disposable income determinants is well behaved, i.e. negative and significant according to the LR. The null hypothesis for this test is that the series is weakly exogenous (Harris 1995). If the observed probability value is greater than a conventional level of significance, this hypothesis is not rejected, implying that that variable is not a true cointegrating vector. Assuming HHD and HP are subjected to causality testing, either one or two of the following possible conclusions is expected;

1. HP Granger causes HHD, but HHD does not Granger cause HP
2. HHD Granger causes HP, but HP does not Granger cause HHD
3. HP Granger causes HHD and HHD Granger causes HP, i.e., there is a feedback system or bidirectional causality.
4. HP does not Granger cause HHD and HHD does not Granger cause X HP

On the basis of these outcomes, this section of the study tests the following hypotheses:

$H_{01}: \phi_1 = \phi_2 = ... = \phi_p = 0$; *HHD does not Granger cause HP*

$H_{02}: \beta_1 = \beta_2 = ... = \beta_p = 0$; *HP does not Granger cause HHD*

The same process can be applied to the rest of the variables. Granger causality uses the *F* statistic to check the statistical significance of the short-run causal effects. The *F* statistic is calculated as follows if it is less than the critical value of *F*:

$$F = \frac{(RSS_R - RSS_U)/k}{RSS_U/(n-2k-1)} \sim F_{(k,n-2k-1)} \qquad [3.50]$$

Where

RSS_U = is the sum of squared residuals from unrestricted equations

RSS_R = is the sum of squared residuals when the restrictions are imposed

n = is the sample size

k = is the lag length of variables.

The null hypothesis that one variable does not cause another is rejected if the observed *F* test exceeds the expected statistic at 5% level of significance. Alternatively, the associated observed probability of this statistic is compared with 5% significance level and the hypothesis is rejected if the former is less than the latter. The *t*-statistic is used to assess the statistical significance of the long-run causal effect on the coefficient of the ECT_{t-1}. The significance of this statistic is tested the same way as the F statistic. Though Granger causality test is widely used, it has its limitations.

First, a two-variable Granger-Causality test without considering the effect of other variables is subject to possible specification bias (Alimi & Ofonyelu 2013). As pointed out by Gujarati (2003), a causality test is sensitive to model specification and the number of lags. It would reveal different results if it was relevant and was not included in the model. Therefore, the empirical evidence of a two-variable Granger-Causality is fragile because of this problem.

Second, time series data are often non-stationary (Maddala 2001) and this could create a situation of spurious regression. According to Gujarati & Porter (2009) when the variables are integrated, the F-test procedure becomes invalid as the test statistics do not have a standard distribution. The test can still be used for significance testing of individual coefficients with t-statistic, but one may not be able to use F-statistic to jointly test the Granger-Causality. Enders (2005) proved that in some specific cases, using F-statistic to jointly test first differential VAR is permissible, only when the two-variable VAR has lagged length of two periods and only one variable is nonstationary. Other deficiencies of these tests are discussed in Toda & Phillips (1994).

This study aims at obtaining optimal results and intends using methods that will help as much as possible to obtain the same. This study intends to assess the causal effects of all the nine variables simultaneously. The suitable test to achieve all these is to use a modified Granger causality testing statistic. This statistic was proposed by Toda & Yamamoto (1995). This is found to be an interesting yet simple procedure which requires the estimation of an augmented VAR which guarantees the asymptotic distribution of the Wald statistic. This statistic follows a χ^2 distribution with m degrees of freedom. This testing procedure is reported to be robust to the integration and cointegration properties of the process.

3.8 Parameter and model diagnostic testing

Norat (2005) is of the view that the fitted model be checked for misspecification as this may give guide to a researcher on how this model can be improved. The author suggested that the checks should include tests for residual serial correlation, normality and heteroscedasticity to help in the validation of the parameter estimation outcomes achieved by the model. This is a very important stage in this study since one of the objectives is to investigate the impact of household debt determinants on the household debt to disposable income in SA and the USA. The estimated parameters are assessed for stability and the estimated models are assessed if they satisfy the assumptions for classical normal linear regression model (CNLRM) mentioned. Diagnostic checks test the stochastic properties of the model, such as residual normality, autocorrelation and heteroscedasticity amongst others.

As Hansen (1992) cautions, the estimated parameters of a time series may vary over time. Parameter tests are important since unstable parameters can result in model misspecification, which have the potential to bias the results (Narayan & Smyth 2009). According to Pesaran & Pesaran (1997), the short-run dynamics are essential in testing for the stability of the long-run coefficients. The Pesaran & Pesaran test involves estimating the ECM for the equation which shows a long-run relationship as reviewed in the previous section. Once the parameters are proven to be stable, the model is tested for significance based on the assumptions as discussed below:

3.8.1 Normality test

This assumption is especially critical when the researcher intends constructing reference intervals for variables (Royston 1991). The violation of normality assumption makes it impossible for a researcher to draw accurate and reliable conclusions as pointed out by Field (2009) and Oztuna & Elhan (2006). Furthermore, ignoring this assumption becomes problematic when the sample is less than 30 observations. Elliott & Woodward (2007) recommend the application of parametric procedures even when the data are not normally distributed. This distribution of the data may only be ignored if the samples consist of hundreds of observations (Altman & Bland 1995). Though this study intends analysing about 93 observations per variable, the assumption of normality is still looked into. The study recommends and uses the D'Agostino (1971) test of normality. The mechanics of performing this test are to use a random sample of size $n \geq 50$. The observations are then ordered to calculate the test statistic. The anticipated sample in this study is greater than 50 because quarterly data set spanning the period 1990 Q1 to 2013 Q1 is analysed. The null and alternative hypotheses about the error terms in the estimated VAR and VECM models are set as follows:

H_0 :*The error terms of the residuals are normally distributed; i.e.* $\hat{\varepsilon_t} \sim N\left(0; {\sigma_{\varepsilon_t}}^2\right)$

H_A : *The error terms of the residuals are not normally distributed;*

Let $X_1, X_2, \ldots, X_p$ represent a random sample of household debts and associated determinants. Also let $X_{,1,n}, X_{2,n}, \ldots, X_{p,n}$ represent the ordered observations derived from this sample. The D'Agostino, D statistic is:

$$D = \frac{T}{n^2 S}, \qquad [3.51]$$

Where

$$T = \sum_{i=1}^{n}\left\{i - \frac{1}{2}(n+1)\right\}X_{i,n}, \qquad [3.52]$$

$$S^2 = \frac{\sum\left(X_i - \bar{X}\right)^2}{n}, \qquad [3.53]$$

$\bar{X}$ being the sample mean. Reject the hypothesis of normality if X is greater than an upper percentile or less than a lower percentile (Rencher 2002).

3.8.2 Serial correlation among residuals

One method for verifying the adequacy of the model is the portmanteau test proposed by Box & Pierce (1970) and Tsay (2005). The test was subsequently modified by Ljung & Box (1978) in response to Davis, Triggs & Newbold (1977) who argued that for moderate sample lengths, the true significance levels were likely to be much lower than predicted by asymptotic theory. Ljung & Box (1978) concluded that the modified test provided an improved approximation that would be adequate for practical purposes. If the model is correct, the Ljung-Box- Q statistic is represented as follows;

$$Q = n(n+2)\sum_{k=1}^{K}(n-k)^{-1}\hat{\rho}_k^2 \qquad [3.54]$$

The statistic is asymptotically distributed as χ^2 with m degrees of freedom (Vogelvang, 2005). Where ρ is called the sample autocorrelation function (SACF) of γ_t (Tsay, 2005), n is the number of observations, k is the highest order for autocorrelation for which to test and ρ_j^2 is the j^{th} autocorrelation. This statistic is best suited for time series variables. The square of this test used to check the presence of autocorrelation in the model as follows;

highlight the presence of heteroscedasticity when *LM*-statistical is higher than the critical value or if the associated *p*-value is less than 5% level of significance.

3.9 The Generalised Impulse Response Function (GIRF) Analysis

The purpose of causality tests as discussed above is only to show the type of causal relationship between the variables. In order to examine how long these impacts will remain effective, the study uses the impulse response. Literature suggests impulse response analysis as a useful tool to examine the effect of a shock over time on various variables in a system. Brooks (2008) and Elder (2003), explain it as a measure used to determine the responsiveness of the dependent variables in the VAR to fluctuations of each of the other variables. Sims (1980) notes that if a variable is truly exogenous with respect to other variables in the system, own innovations will explain all of the variables' forecast error variance. If the errors are correlated as is usually the case according to Sims, this study cannot associate a shock with any one particular variable.

By providing the time path of the impact of a shock on the future values of all the variables in the multivariate dynamic system, the impulse response analysis should give better insights into short-term and long-term linkages among all the variables employed (Hurley 2010). This study also uses the GIRF once the VECM has been estimated to examine the response of debts to shock as a result of changes in related determinants. This helps to show the response of each variable in the system to a shock in any of the other variables. The GIRF is calculated from the moving average representation of the VECM as follows:

$$Z_t = \sum_{i=0}^{\infty} y_t \varepsilon_t \qquad [3.56]$$

Where

$$y_t = c + \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \ldots + \Phi_q y_{t-p} + a_t \qquad [3.57]$$

If the system is stabilised, the following equation is given

$$y_t = \mu + \Psi(L)a_t = \mu + a_t + \Psi_1 a_{t-1} + \ldots + \Psi(L) = [\Phi(L)]^{-1} \qquad [3.58]$$

Redating at time $t+s$:

$$y_{t+s} = \mu + a_{t+s} + \Psi_1 a_{t+s-1} + \ldots + \Psi_{s-1} a_{\cdot t} + \Psi_1 a_{s+1} a_{t-1} + \ldots \qquad [3.59]$$

$\frac{\partial Y_{t+s}}{\partial a'_t} = \Psi_s = \left[\psi_{ij}{}^{(s)}\right]$ are multipliers and $\frac{\partial y_{i,t+s}}{\partial a_{jt}} = \psi_{ij}{}^{(s)}$ is the reaction of the i-variable to a unit change in innovation j.

Following Pesaran & Shin (1998) in Aziakpono (2006), the GIRF of variable $y_{i,t+s}$ with respect to standard error shock in y_{jt} with other variables dated t or held constant is given by equation;

$$\frac{\partial y_{i,t+s}}{\partial a_{jt}} = \psi_{ij} \qquad [3.60]$$

The GIRFs are unique and do not require prior orthogonalisation of the shocks (reordering of variables in the system) (Ben-Kaabia, Gil & Chebbi 2002).

PART B DISCRIMINANT ANALYSIS

3.10 Overview of Discriminant Analysis (DA)

The main intention of this section is to compare SA and the USA in relation to the average household debts using Fisher's LDA. It is not the intention of this study to compare the statistical methods used in comparative research. Only factors underpinning household debts of these countries are subjected to this method. Other factors such as government and corporate debts determinants do not form part of the analysis.

DA is a multivariate analysis technique that relates a categorical dependent variable with metric independent variables (or categorical) (Siqueira, Otuki & da Costa 2012). The PROC STEPDISC is used in this study to obtain the best discrimination possible. This is a procedure in which the independent variables are being successively included in the model. This help in testing, with every inclusion of the independent variable if that particular variable contributes to the improvement of the discriminating power of the model. In the end, the set of variables that best discriminate the groups of the discriminant function predominates (Siqueira *et al.* 2012). The F-test is used to evaluate the significance of the determinants included in the model. The variable with the largest F-statistic is viewed as having the most discriminatory power and is also the most preferred according to the analysis. The number of steps is determined by the number of significant variables identified by PROC STEPDISC. The process stops with the last most significant variable, implying that no further variables are added in the analysis. The main focus of this section is to build the discriminant functions, represented as a linear combination of the determinants of household debt.

PROC DISCRIM is used to produce these functions and statistics that help in evaluating the usefulness of the selected household debts determinants from PROC STEPDISC. These functions may help in identifying those determinants that discriminate best between SA and the USA as categories of the dependent variable. Again, it is the intention of this section to test if there are major differences between the two countries based on their centroids. Once the results are obtained they are interpreted and tested for validation. Figure 3.3 summarises the DA framework adopted to accomplish these objectives.

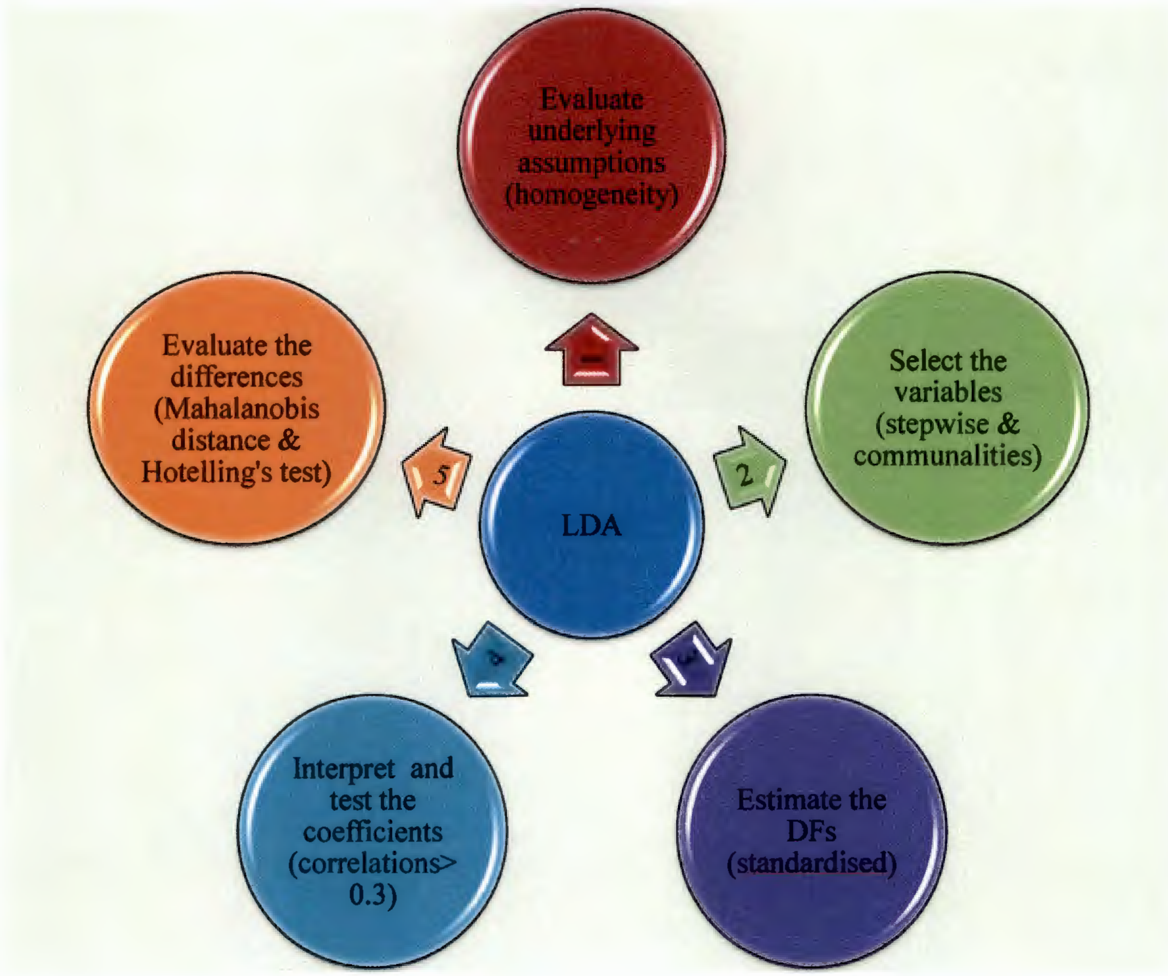


Figure 3.3: Stages in Discriminant Analysis

<u>Adapted from:</u> Hair *et al.* (2010)

3.10.1 Discriminant Analysis framework

This section adopts and reviews the discriminant framework which was first developed by Fisher (1936). Suppose two populations π_1 and π_2 each have p variables, and $X' = (X_1, X_2, ..., X_p)$ then Y is a linear combination of the p X's;

$$Y = a'X = a_1X_1 + a_2X_2 + ... + a_pX_p, . \quad [3.61]$$

The basic idea is to transform X into a single variable Y in such a way that Y allows us to separate π_1 and π_2 as much as possible. This linear combination exhibits the largest difference in the group means relative to the within-group variance. The derived variate z is known as Fisher's discriminant function, or the first canonical variate.

x = the p determinants of household debt

Y = a non-metric dependent variable (SA and the USA)

a = the discriminant weights.

The μ_i and Σ are calculated for each variable and the latter is checked for homogeneity with the statistic created by Box (1954) named Box M's test. Multivariate data are used in this study and the underlying normality assumption in the variables is assessed under section 3.7. The modification of equation [3.61] to suit this study is as follows:

$$HHD = a_1 \ln HPI + a_2 \ln INF + a_3 \ln INC + a_4 \ln IR + a_5 \ln GDP + a_6 \ln HC + a_7 \ln HS + a_8 \ln TR \qquad [3.62]$$

3.10.2 Testing the homogeneity of the variance-covariance matrix

The variance-covariance matrices of equation [3.62] are assumed to be homogeneous and this assumption is validated with the Box'sM test calculated as:

$$M = (n-g)\log\left|S_{pooled}\right| = \sum_{i=1}^{g}(n_i - 1)\log\left|S_i\right| \qquad [3.63]$$

Where S_{pooled} is a pooled sample variance-covariance matrix for the two populations and is calculated as;

$$S_{pooled} = \frac{1}{n_1 + n_2 - 2}\left[(n_1 - 1)s_1 + (n_2 - 1)s_2\right] \qquad [3.64]$$

The hypothesised statements associated with equation [3.63] are:

$H_0 : \Sigma_{RSA} = \Sigma_{USA}$ *(the variance-covariance matrices are homogeneous)*

$H_1 : \Sigma_{RSA} \neq \Sigma_{USA}$ *(the variance-covariance matrices are heterogeneous)*

For this assumption to hold, Box's M must be non-significant. In light of the sensitivity of the test, an alpha level of 0.001 is recommended (Brace, Kemp & Snelgar 2009). The hypothesis is rejected if the observed probability value associated with equation [3.63] is less than alpha.

Morrison (1976) on the other hand recommends the rejection of the null hypothesis at a significance level of 0.1. This implies that the DA ultimately provides more meaningful results. If this assumption is violated, the determinants are log transformed.

3.10.3 Significance of independent variables

To assess the importance of each variable to the discriminant function, this study observes the standardised coefficients of the function. The standardised coefficients are the product of the non-standardised coefficients by the roots of covariance (Siqueira *et al.* 2012). The standardised discriminant function coefficients in this study are used to assess each household debt determinant's unique contribution to discriminant function (George & Fernandez, 2000). A low standardised coefficient means that the groups do not differ much on that variable or that a variable's correlation with the grouping variable is redundant with that of another variable in the model (Hassan 1986). In other words, the larger the standardised coefficient, the greater is the contribution of the respective variable to the discrimination between groups (Klecka 1980). Also as stated by Tabachnick & Fidell (2013) by convention, correlations in excess of 0.33 (10% of variance) may be considered eligible while lower ones are not.

As Hassan (1986) suggests, another way to determine which independent variables define a particular discriminant function is to look at the factor structure. The structure matrix gives coefficients of correlation between independent variables and discriminant function. Generally, any predictor with a loading (a score) of 0.30 or more is considered to be important in defining the discriminant dimension (Wuensch 2001).

3.11 Evaluating group differences

To measure how far apart the two groups are in terms of values of z, the statistic below is computed as:

$$D^2 = \frac{(\mu_1 - \mu_2)^2}{S_Z^2} \qquad [3.65]$$

The coefficients $(a_1, a_2, \dots, a_9)$ are selected so as to maximise D^2. This statistic is a Mahalanobis distance and is interpreted as the squared distance between the means of the standardised values of the *Ys*. A larger value of D^2 indicates that it is easier discriminate

$H_0 : \mu_{RSA} = \mu_{USA}$ *(no difference between the mean vectors)*

$H_1 : \mu_{RSA} \neq \mu_{USA}$ *(at least one of the mean vectors is different)*

Hotelling (1951) invented the test statistic to validate this hypothesis. This test is used in this study to compare SA and the USA relative to the household debt determinants. The test statistic is calculated as:

$$T^2 = \frac{n_i n_j}{n_i + n_j} D^2 \qquad [3.69]$$

Where n_i and n_j represent the sample sizes for SA and the USA respectively. This test is distributed as $T^2{}_{(p,n_1+n_2-2)}$. The null hypothesis is rejected if $T^2 \geq T^2{}_{(\alpha;p,n_1+n_2-2)}$.

3.13 Summary

This chapter outlined the determinants of household debts and also provided their definitions. Justifications for the inclusions of these determinants in the model were also provided. The model was reviewed based on the two frameworks adopted in this study. The processes for the Johansen cointegration and the Toda-Yamamoto Granger causality methods were presented. The statistical tests for the diagnostic testing of the models were also reviewed. Forecasting procedure was also discussed. To guard against the issue of over or underestimation, the selected model is subjected to the ECM and this was also explained in detail. Another multivariate framework by R.A. Fisher was also adopted and modified. This modified function helps in coming up with a model that shows linear combination of household debt and related determinants for both countries. This model will also be used to come up with statistics that compare SA and the USA based on the household debt determinants.

4.2 Empirical findings

This section is divided into two parts. The first presents the results from both univariate and multivariate econometric analysis and the second part focuses on data analysis using LDA. The univariate analysis in the first part covers the preparation of data for analysis, checks data and variables for appropriateness and ensures that the variables are stationary prior the analysis. This also includes cleaning of data and ensuring that the variables are measured using the same scale. Also the variables are transformed with the natural logs to remove the irregularities. The multivariate cointegration approach by Johansen (1988) is adopted in building the final model for HHD using SA and USA data. Test statistics from this approach are also used to assess the long-run relationships. This is precedes the causality analysis between the variables using Toda & Yamamoto (1995) approach. The diagnostic checking of the chosen model is done using the procedure as discussed in Chapter 3 and the results of the forecasts of the output variable are provided. Simultaneously, data on the determinants of household debts for SA and the USA is analysed using the Fisher's LDA developed by Fisher (1936) in the second part where these countries are compared.

4.2.1 Exploratory data analysis

This step performs univariate data analysis and checks the adequacy of the sample and reliability of the variables. This helps to understand the distribution and mark the differences in the variables if any. The results below are obtained from the original data that has not been manipulated. This may help in making significant actual and unbiased differenced and/or similarities between SA and the USA.

Table 4.1: Descriptive statistics

SA										
Variable	**HHD**	**HP**	**CP**	**INC**	**IR**	**GDP**	**HC**	**HS**	**ER**	**UR**
N	93	93	93	93	93	93	93	93	93	93
Mean	63.7	464382.3	87.8	816412.0	11.9	1319819.1	888293.7	15.7	6.0	23.5
Variance	106.3	123280000000.0	1128.31	281343000000.0	11.0	797704000000.0	49936715092.0	2.2	5.0	6.6
US										
N	93	93	93	93	93	93	93	93	93	93
Mean	12.1	221141.9	90.6	31505.6	5.2	12440.8	7275.3	5.50	1.66	5.9
Variance	1.0	3456962679.0	327.43	16062580.5	2.7	4868511.0	5626238.7	2.6	.02	2.2

HHD: Household debts, **HP**: House prices, **CP**: Consumer prices, **INC**: Disposable income, **IR**: Interest rates, **GDP**: Gross domestic product, **HC**: Household consumption expenditure, **HS**: Household savings, **ER**: Exchange rates, **UR**: Unemployment rates

Please note that due to the inaccessibility of percentage of household debts to disposable income in the USA, a proxy (percentage household debts ratio) had to be used instead.

Table 4.1 above gives an overview of the main variables used in the analysis. There are 93 observations for each of the 10 variables. The average housing prices in SA is more than that in the USA. The variances for these countries in terms of house prices confirm a remarkable difference, with SA revealing a larger difference. In terms of this analysis, the average house prices in SA are about R464382.30 and $221141.90 for the USA consumer prices, interest and savings' rates for these countries are nearly the same. The figures clearly reveal that these countries possess some similarities as well as differences. However, this analysis alone cannot be used to make conclusion about these countries hence we use inferential methods to be more convinced of the information the descriptive analysis convey. Next, this study presents the results affirming whether the variables chosen for this study are real measures of HHD or not by looking at the scatter plot matrices.

4.3.2 Association between household debt determinants

This section examines the interrelations between household debt determinants and the results are presented in Figure 4.1. The figures are also helpful in confirming the correct specification of the model to be estimated. This also helps in examining the dynamic effects.

As Figure 5.1 above depicts, there are positive and inverse relations between the chosen variables for the period 1990Q1 to 2013Q1. However, the relationship does not seem to be completely linear at least with CP, INC and GDP. These figures are slightly exponential implying a possibility of a nonlinear relationship. All the variables show the characteristics common with most economic data such as non-normality in the form of fat tails on the histograms. This problem is solved by subjecting all the variables to logarithmic transformation prior to differencing. This will also help eliminate the inconsistencies in the variables. The next section discusses the results used for confirming that the sample and variables used in this study is sufficient.

4.3.3 Sampling and variable adequacy analysis

Table 4.2: Sample and variable adequacy

Measure	SA	USA
Cronbach's Alpha	0.854598	0.854598
KMO	0.832	0.774
Bartlett's test	2098.68 (df=36)	2040.49 (DF=36)
Determinant	4.605E-011	6.826E-011

KMO: Keiser-Meyer-Olkin

Table 4.2 gives a summary of the results confirming that the variables chosen for this study are a true measure of HHD. This is shown by a Cronbach's alpha (0.855) which is good according to Kline (1999) and Cronbach & Shavelson (2004). The variables chosen for analysis are unidimensional and this is agreeable in terms of Byrne *et al.* (1989), Chin (2000), Blaha (2001) and Diamantopoulos & Siguaw (2006). The KMO measure in the context of SA is meritorious and is middling in relation to the USA data. This affirms that the sample selected for this study is adequate and that the results can be generalised (Kaiser 1974). As Bartlett's test and the determinant of the correlation matrix suggest, multicollinearity is not a serious problem. Had the determinant of the correlation matrix been equal to zero, it would be concluded that the said problem would inflate the variance-covariance matrices and cause the false rejection of the null hypothesis (Gujarati 2010). The critical values of chi square associated with Bartlett's statistic at 5 % significance level and 36 degrees of freedom is 55.758. The null hypothesis that the matrix is an identity is rejected confirming that multicollinearity is not a problem (Churchill 1995).

4.4 Trend analysis of household debt determinants

This section provides visual plots of the series and this is usually the first step in the analysis as it paints a picture of the possible nature of time series before pursuing formal tests for stationarity. These plots also give an indication of the behaviour of the variables, whether or not they have a trend. This also allows the detection of any data capturing errors, and structural breaks or drifts that may bias the unit root tests and gives an idea of the trends and stationarity of the data set. The figures below represent plots of all variables used in the model.

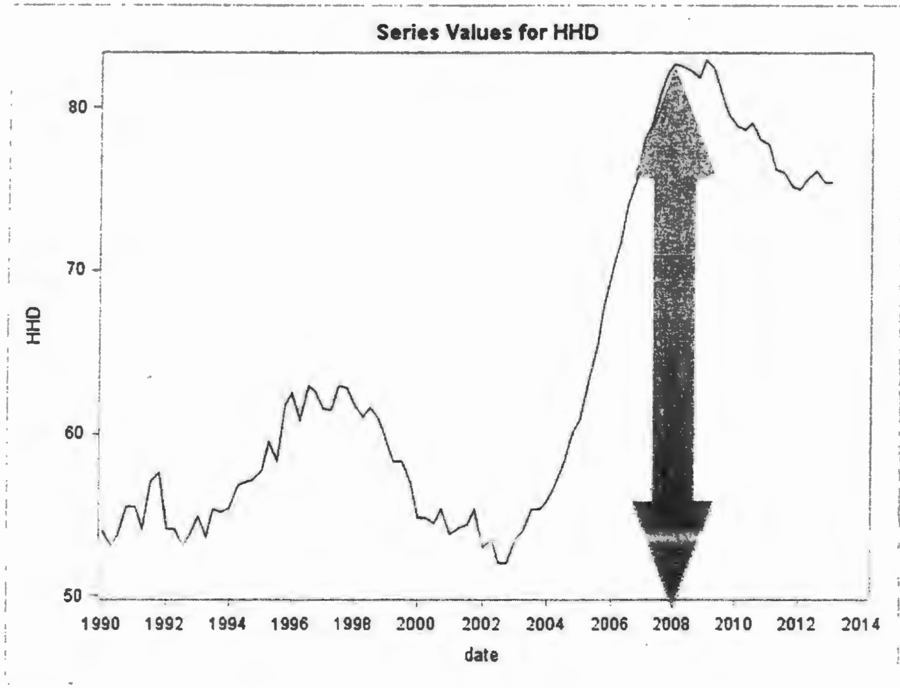


Figure 4.2a: SA HHD plot

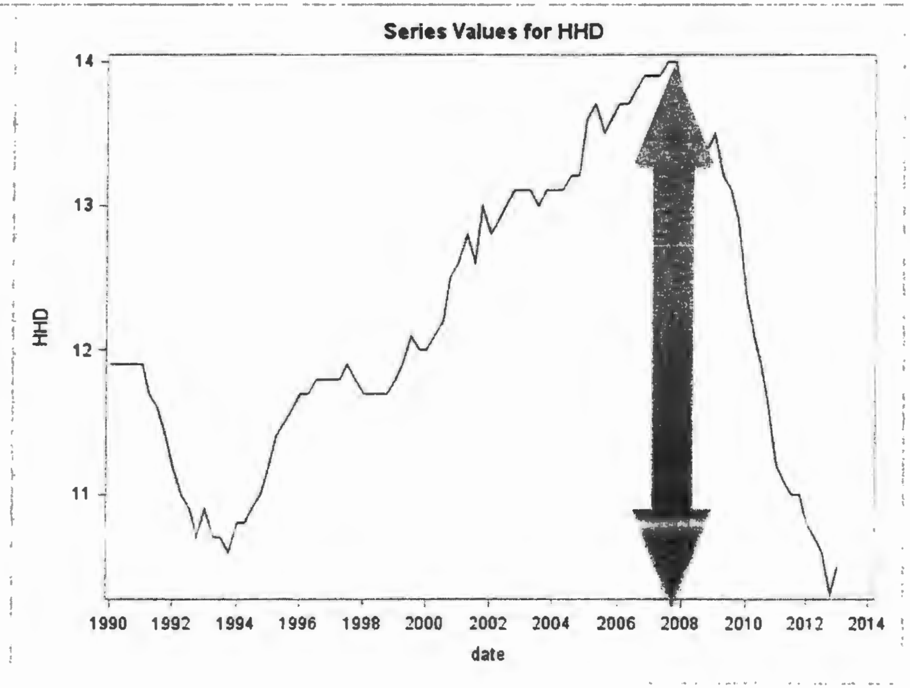
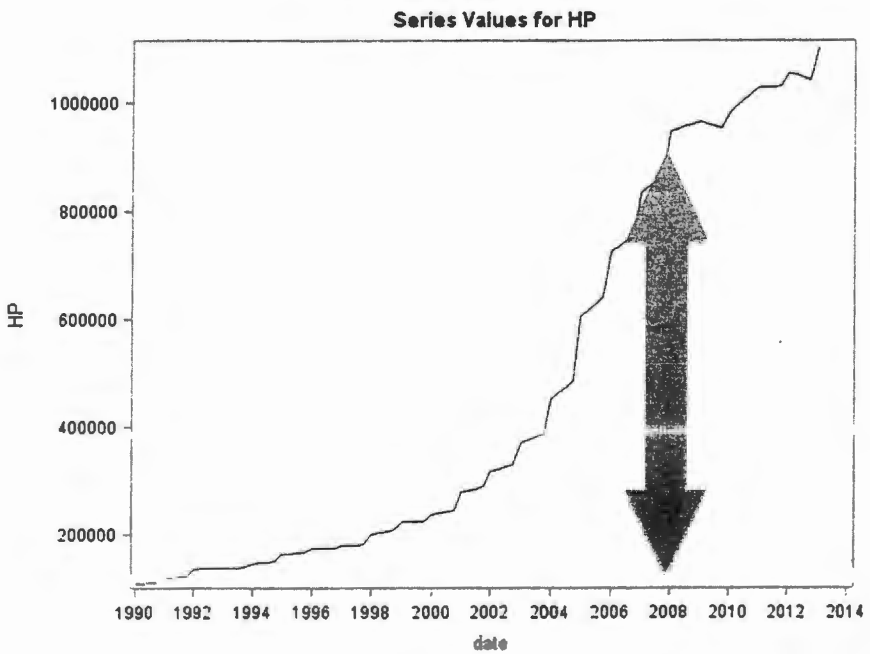


Figure 4.2b: USA HHD plot

It is evident from the figures that household debts in SA and the USA were a bit lower between 1990 and 1994. The first higher peak is visible in 1997. Household debt to disposable income in South Africa shows fluctuations between 52 % and 70 % between 1990 and 2006 confirming the reports by Kotzé & Smit (2008). The period 2007 and 2009 marks the financial crisis which emerged in the USA. This crisis affected the whole world including South Africa as is also visible on the figure. Displayed as well is the recent radical deceleration of household debts in both countries with the USA showing that the growth rate of household debts has recently hit lowest values for the first time.



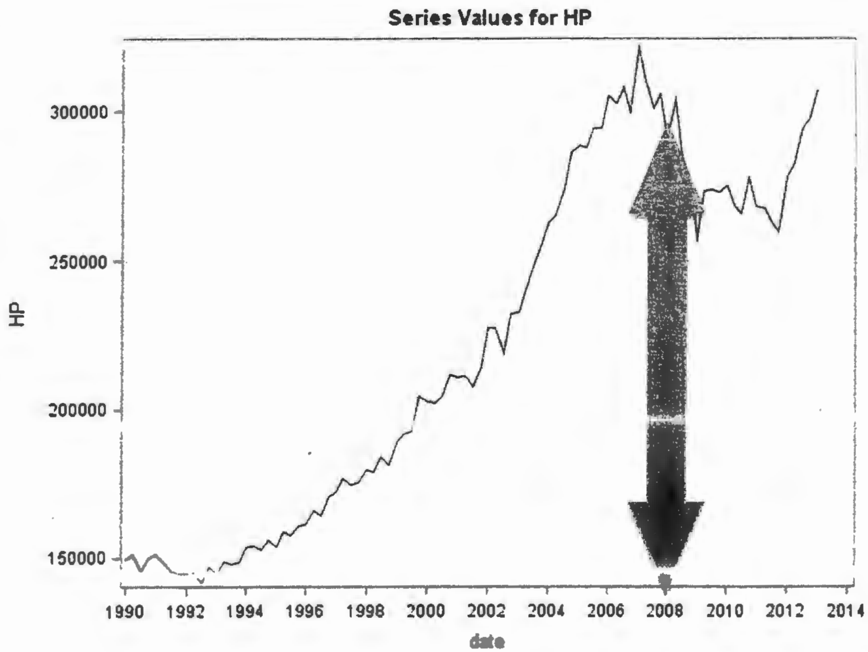


Figure 4.3a: SA HP plot

Figure 4.3b: USA HP plot

Figure 4.3 shows that prior the financial crisis period, *i.e.* (2006-2007), house prices were very high in the USA as compared to South Africa. This could be as a result of the SMC which took place between 2007 and 2008 which also emerged in the USA. The prices started soaring during the GFC period 2007-2009. The spill over effects of the crisis affected the financial markets and caused house prices to increase in an upward sloping trend as the figure displays. America experienced some shocks in 2011 and this caused house prices to increase from 2012 onwards. Although showing signs of revival, the figure reveals a likelihood of continued increase in house prices in these countries. Trebbi, Mian & Sufi (2012) reported that the value of households' financial assets in the USA has bounced back since 2009, but the value of their real estate assets has not. This was reported to be a cause for continued dampening of consumer spending in the country.

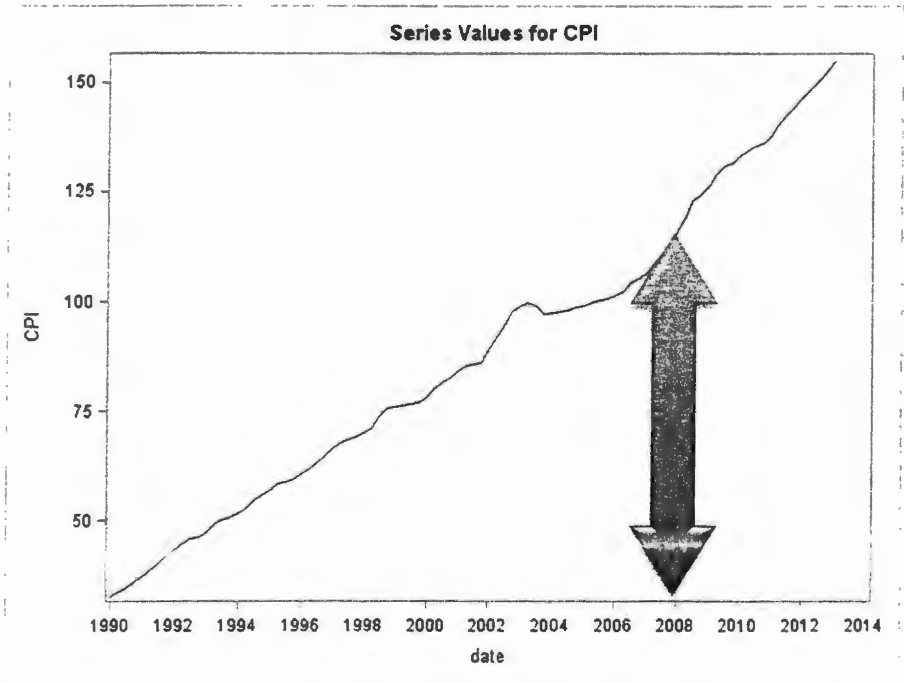


Figure 4.4a: SA CP plot

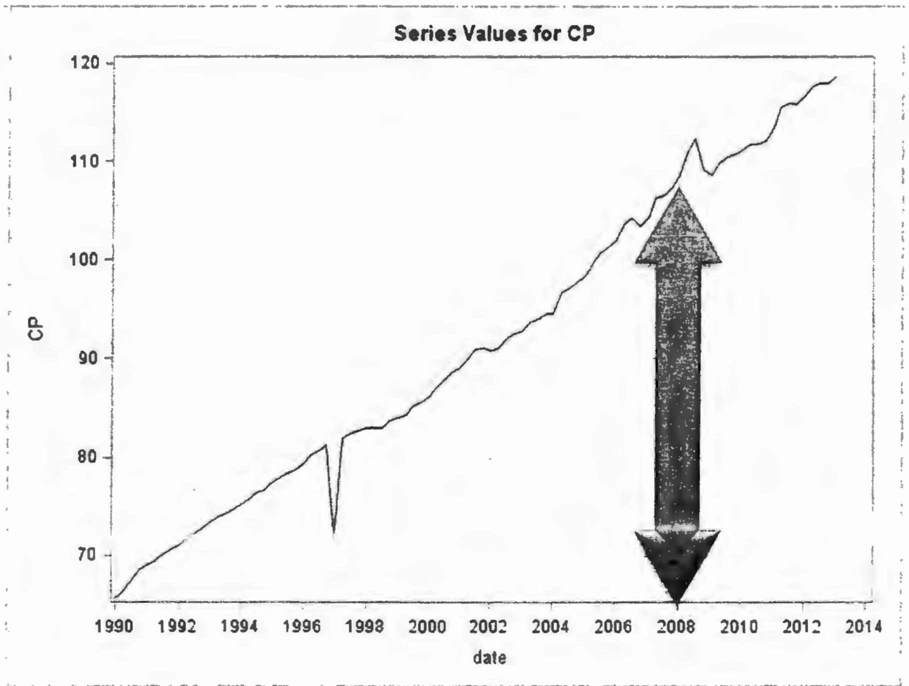


Figure 4.4b: USA CP plot

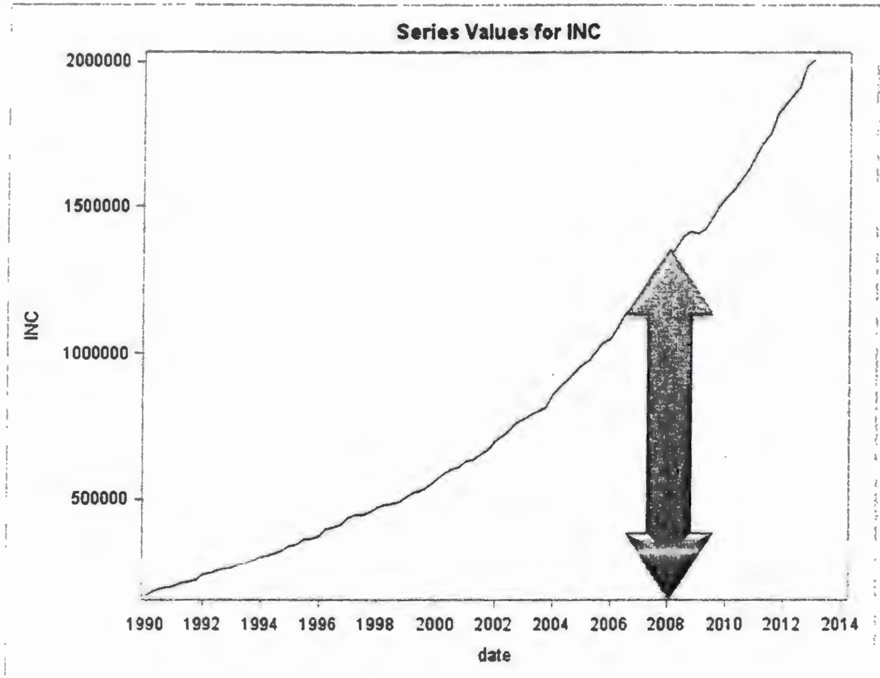


Figure 4.5a: SA INC plot

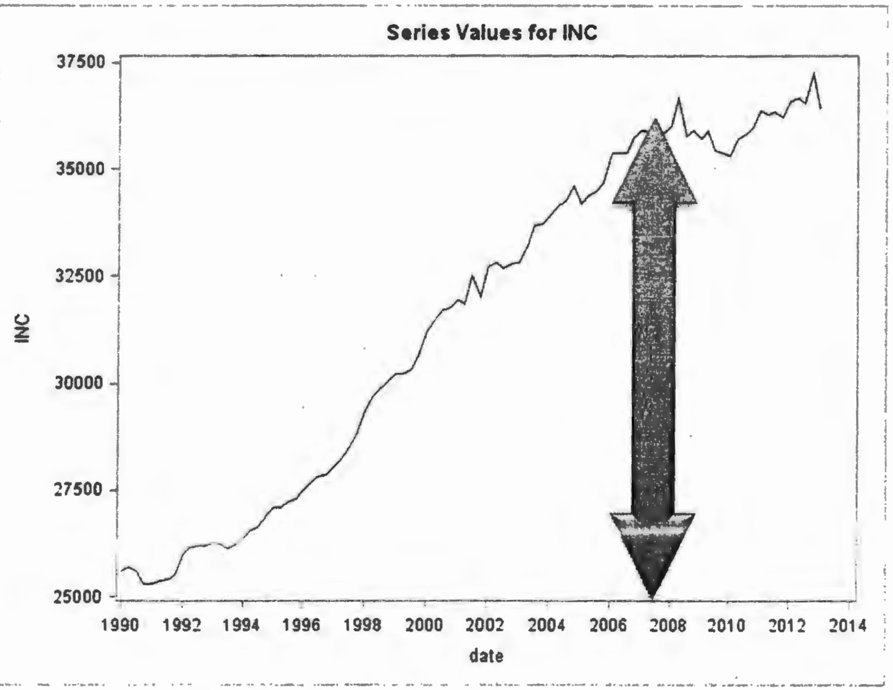


Figure 4.5b: USA INC plot

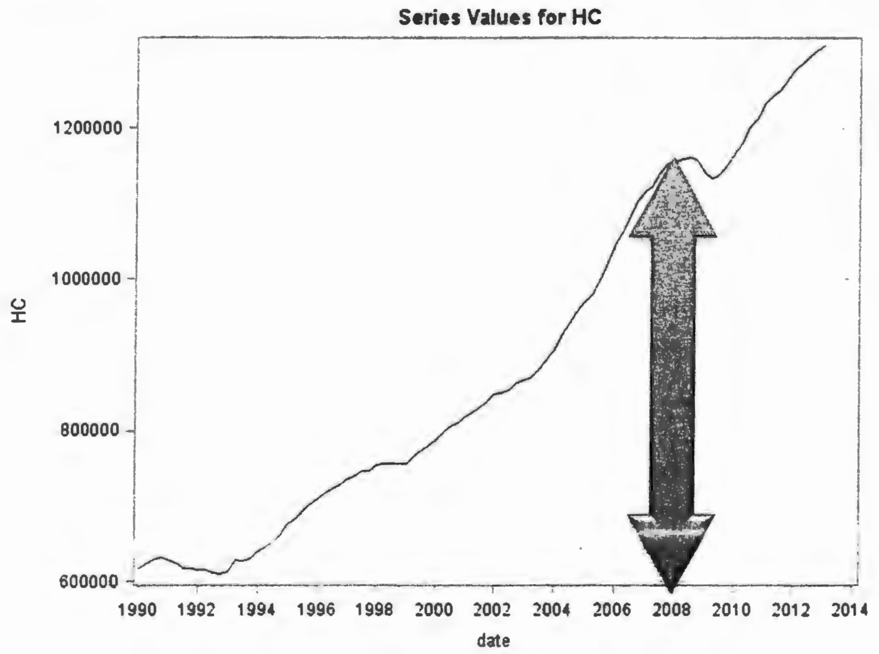


Figure 4.6a: SA HC plot

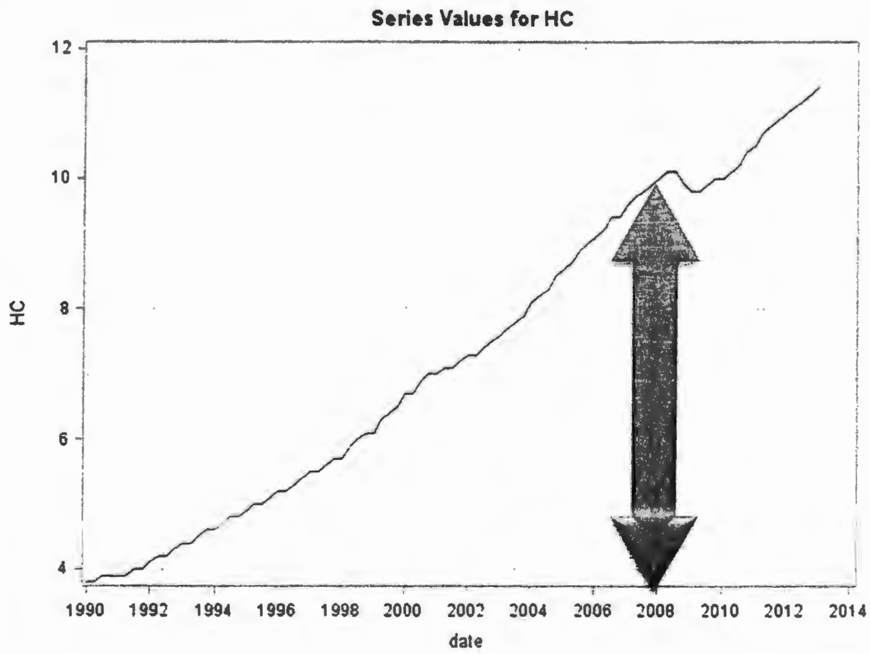
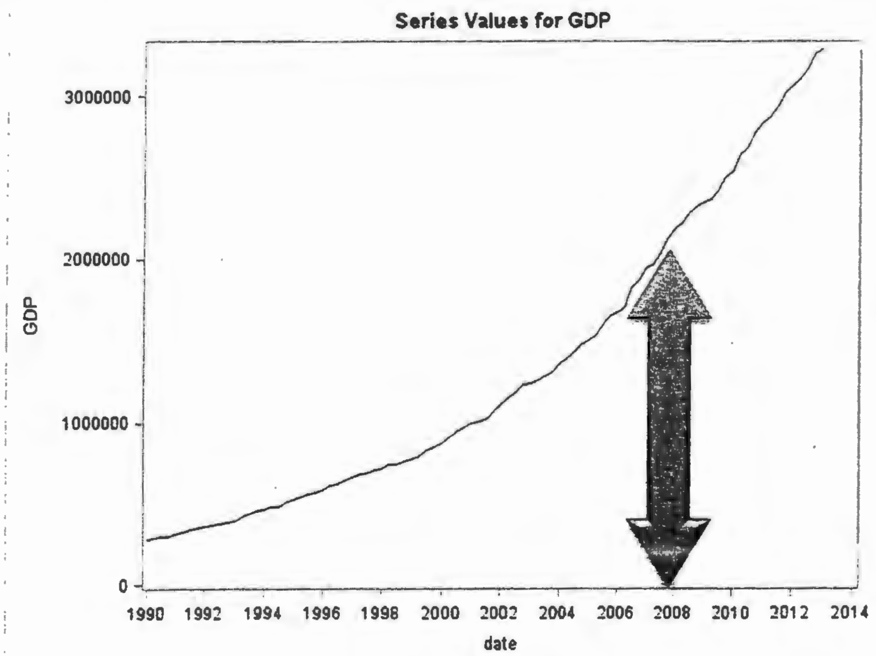


Figure 4.6b: USA HC plot



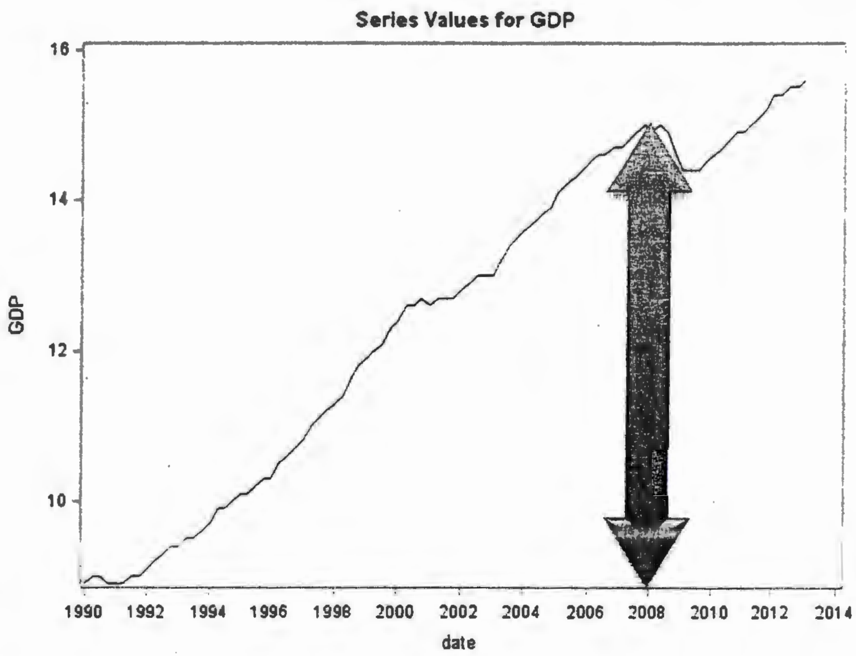


Figure 4.7a: SA GDP plot

Figure 4.7b: USA GDP plot

Both SA and the USA consumer priceses, household disposable incomes, gross domestic prices and household consumption expenditures are explained by upward trends. There is a possibility of further increases in these determinants as shown in Figures 4.4 to 4.7.

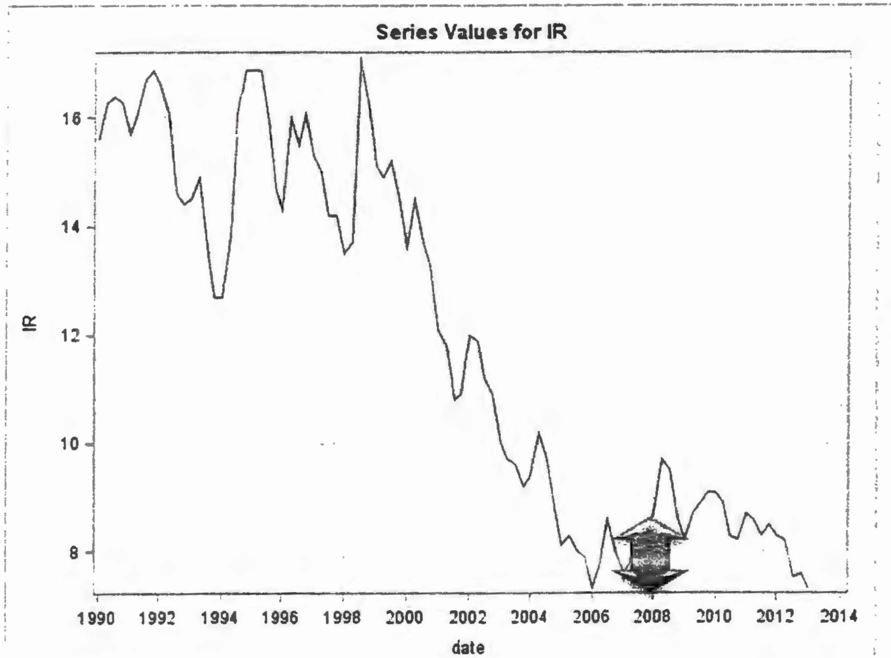


Figure 4.8a: SA IR plot

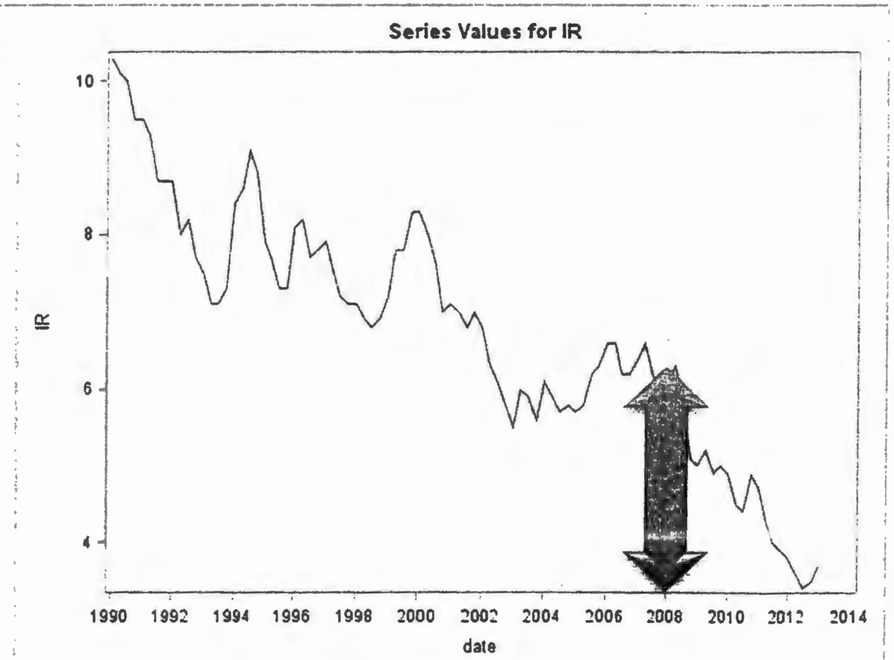


Figure 4.8b: USA IR plot

Interest rates movements in both SA and the USA are explained by downward seasonal graph. During the period of financial crisis, interest rates were about 9 % and less than 7 % for SA and the USA repectively. This shows that SA experienced 2 % higher rates of interest during this period than the USA. In the first quarter of 2013 this determinant was down to about 7 % for SA and 3 % for the USA as the figures reveal.

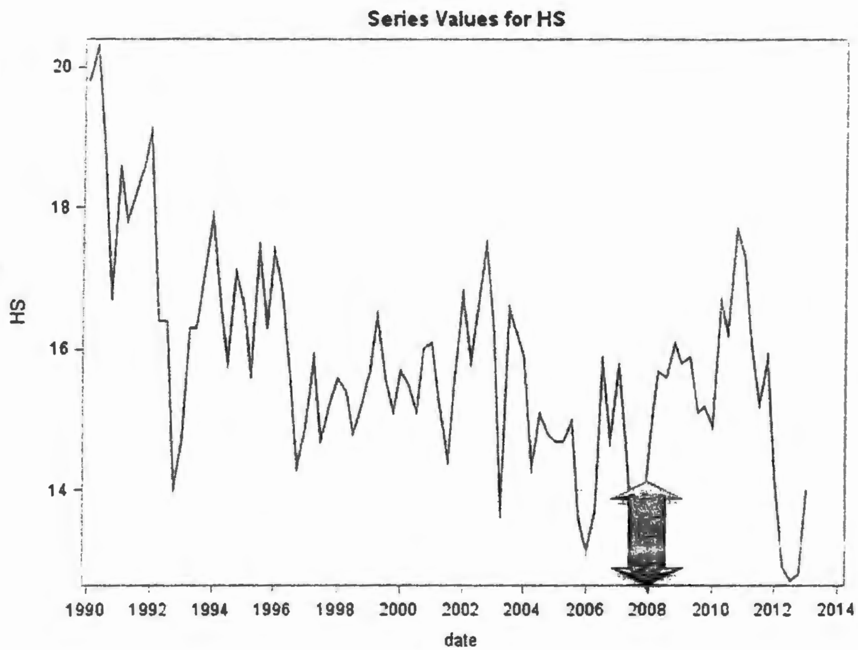


Figure 4.9a: SA HS plot

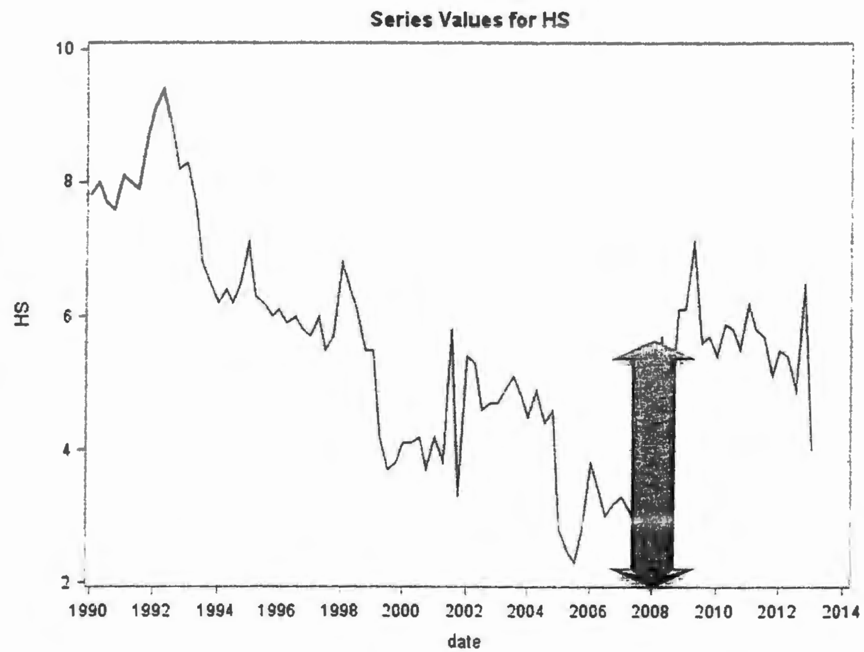


Figure 4.9b: USA HS plot

Household savings seem to be following the same pattern in both countries. As the figures reveal, SA savings rates were about 19 % and USA had almost 8 % in 1990. This determinant seems not to be resistant to shocks, and is indicative of many spikes on the the plots. Savings kept on decreasing for both countries and became even lower during the GFC. SA shows a persistent structural break, which makes it difficult to say whether this determinant is increasing or decreasing just as Aron & Muellbauer (2002) indicated. In the USA perspective, the personal savings declined sharply from 1990s, reaching as low as 1.0 % by 2005. It is likely that the reduction of household saving was in large measure a consequence of the sizable increase in household net worth associated with increased house prices and perhaps consumer prices occurring at that time. As Elwell (2013) reported, a rapid rise in wealth made it less urgent to divert current income to saving hence the figure.

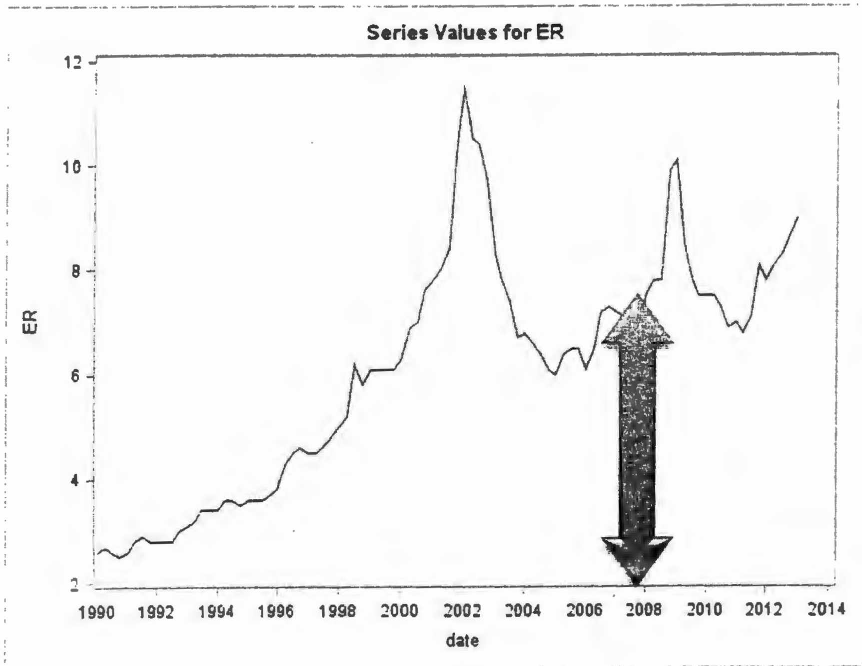


Figure 4.10a: SA ER plot

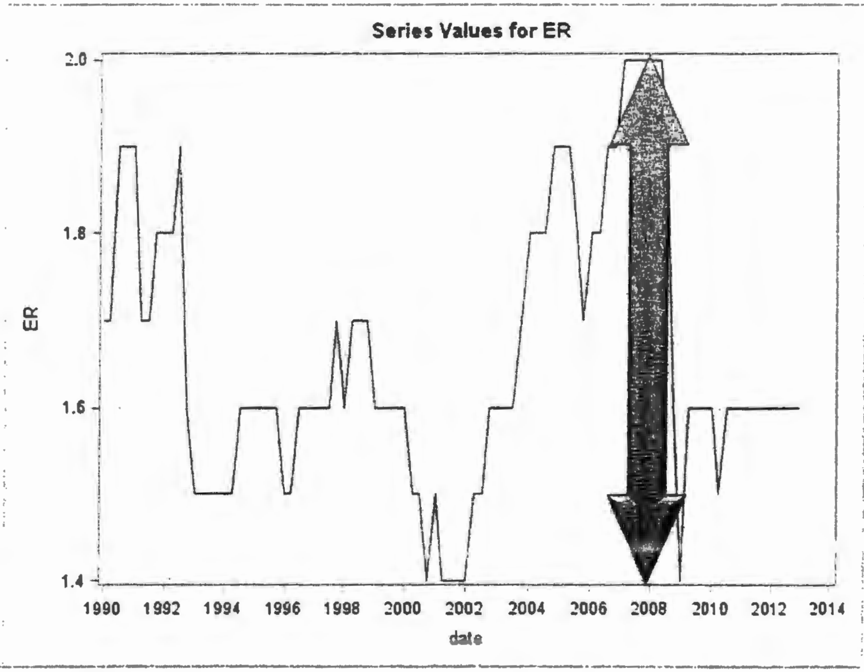


Figure 4.10b: USA ER plot

Figure 4.10 shows the behaviour of nominal quarterly exchange rates from 90Q1 to 13Q1. As literature reports, the volatility of the rand has increased significantly since the Central Bank allowed the exchange rate to be market oriented. The general trend has in the past years been downward, implying a general depreciation in the nominal value of the rand. During the year 2002, the rand reached its highest peak for the first time and the response to this shock was negative until around 2005. The opposite is reported about the USA dollar. Up until the financial crisis, South Africa was all smiles with the USA dollar selling at around R 7.00. There is still hope that this currency might do well in future while the USA dollar shows signs of stability as Figure 4.10b reveals.

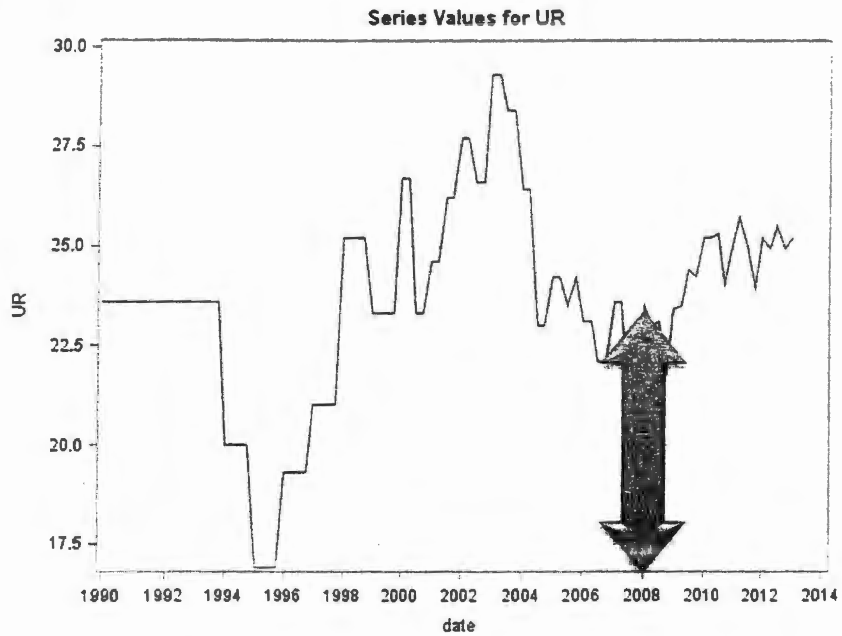


Figure 4.11a: SA UR plot

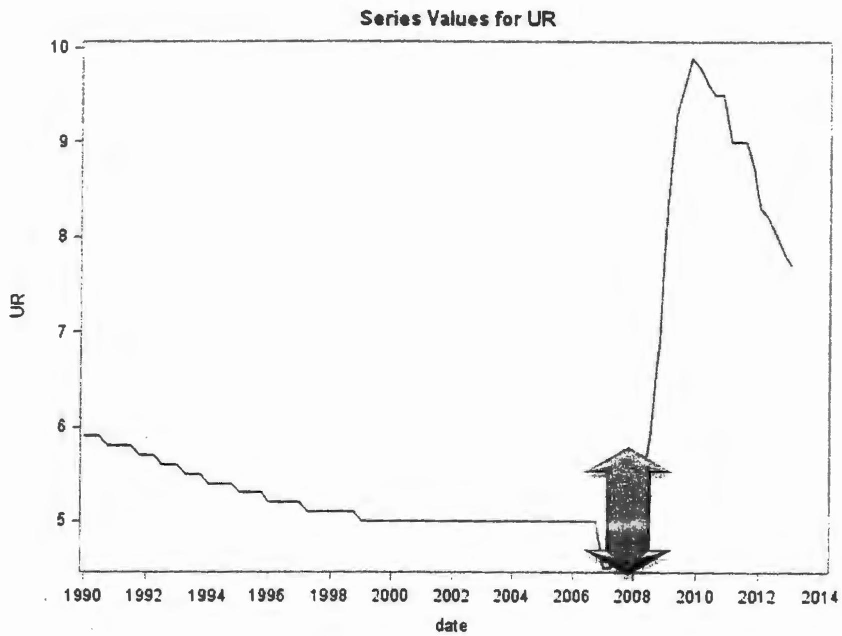


Figure 4.11b: USA UR plot

Prior to the finacial crisis, unemployment rate in South Africa stood at around 23 %. The US rates were much lower at around 6 % at the beginning of the crisis. This determinant peaked to about 10 % in 2009 and edged down to almost 8 % at the beginning of 2013. Elwell (2013) commented that these figures are still high at that stage of the economic recovery. The author highlighted that a considerable share of the improvement in the unemployment rate is not the result of workers finding jobs, but by discouraged workers leaving the ranks of the officially unemployed.

To summarise, household debt determinants are explained by different movements. Some are explained by upward or downward trends while others such as interest, exhange, unemployment rates, *etcetera* have drifts and random walks. This means that the researcher may have to impose restrictions such as the intercept when estimating the model. Nonetheless, all these variables are non-stationary and differencing is applied to them in an attempt to stabilise them. The results of stationarity tests on differenced series are summaried on Figures 4.12 to 4.14 and Table 4.3.

4.5 Stationarity results

This section uses informal and formal tests to assess the variables for stationarity. First order of differencing is applied to normalise these variables. For stationarity tests, the study include a lag of up to four to correct for autocorrelation since the data are collected on a quarterly basis. The normalised results are obtained using the ADF, PP and KPSS tests and the results are summarised in Table 4.3. These results are preceded by the differenced plots.

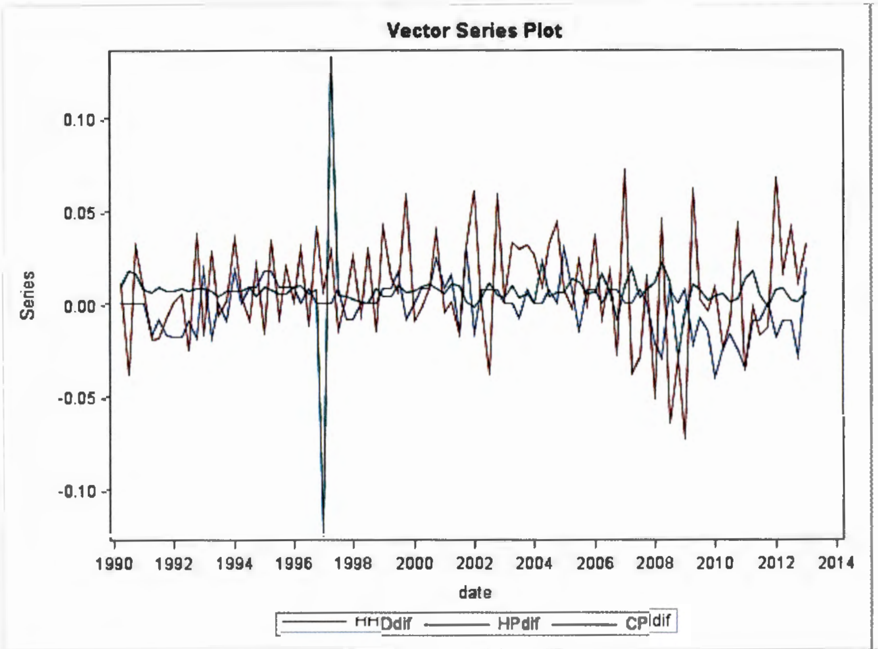


Figure 4.12a: SA First Dif. HHD, HP and CP

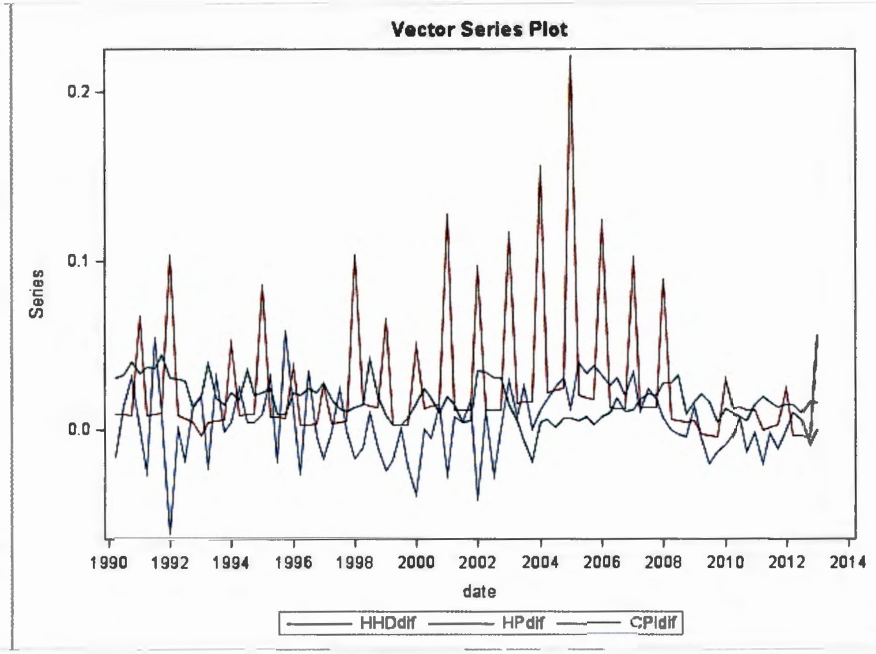


Figure 4.12b: USA First Dif. HHD, HP and CP

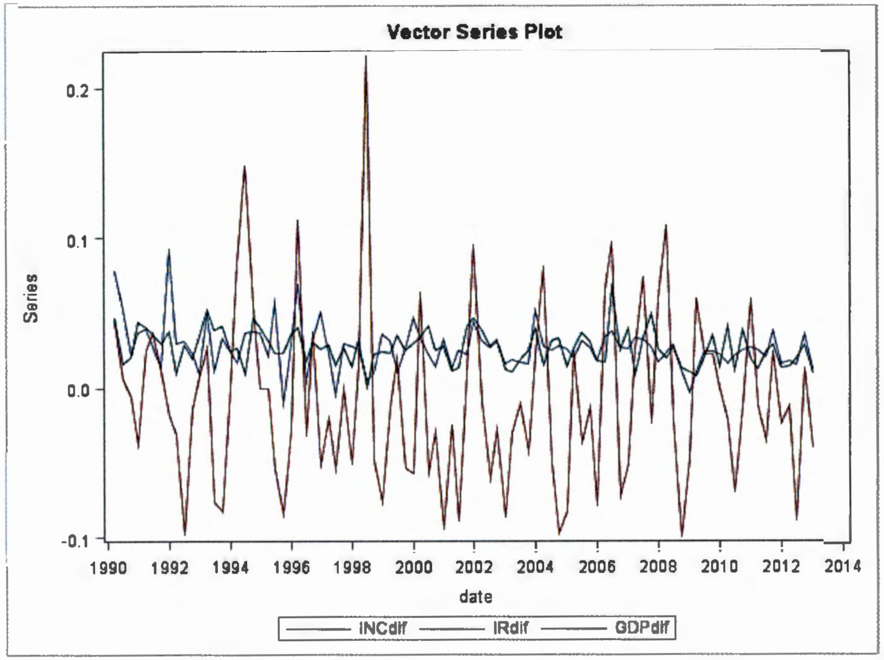


Figure 4.13a: SA First Dif. INC, IR and GDP

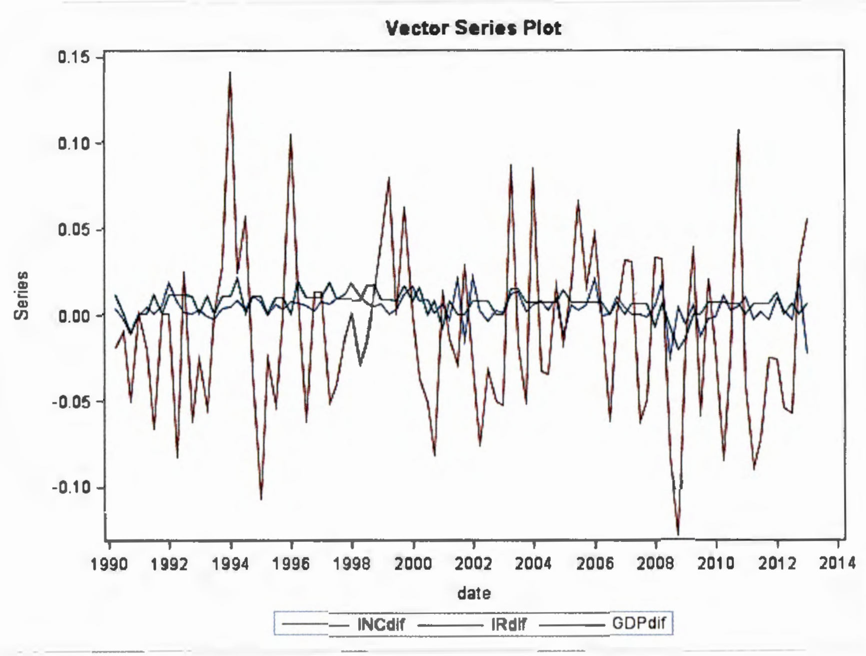


Figure 4.13b: USA First Dif. INC, IR and GDP

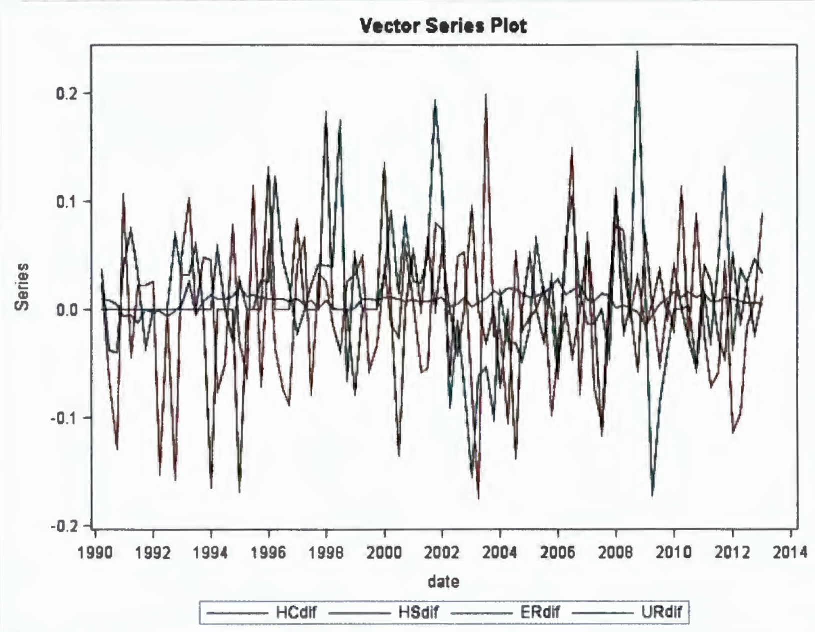


Figure 4.14a: SA First Dif. HC, HS, ER and UR

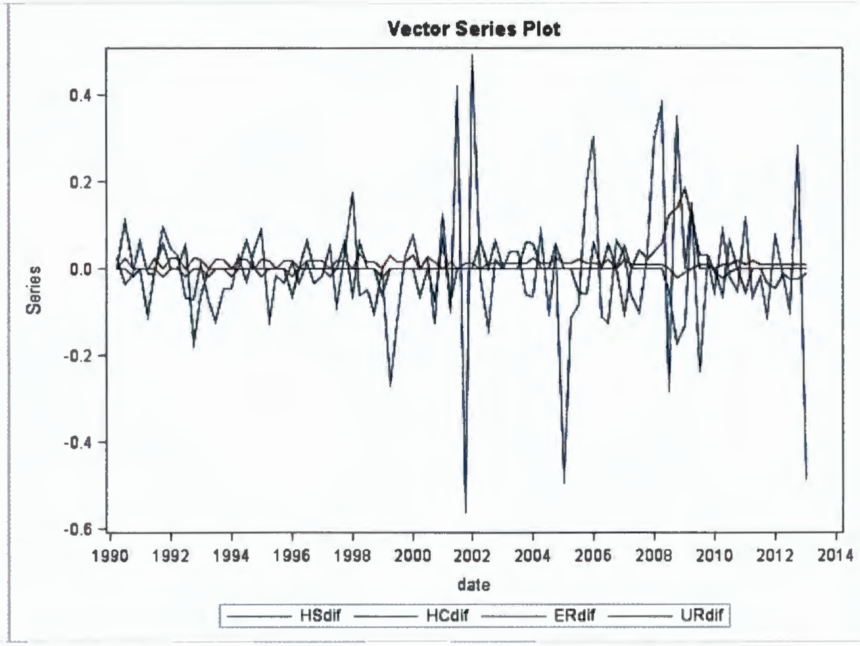


Figure 4.14b: USA First Dif. HC, HS, ER and UR

Figures 4.12 to 4.14 reveal the first differenced results of the ten macroeconomic variables. These figures indicate that all the variables are stationary after first difference.

Table 4.3: ADF and PP tests

SA			
VARIABLE	ADF	PP	DECISION
HHD	-4.83 (0.0009)	-7.75 (0.0001)	Reject the null hypothesis
HP	-6.99 (0.0001)	-9.52 (0.0001)	Reject the null hypothesis
CP	-4.55 (0.0022)	-4.88 (0.0007)	Reject the null hypothesis
INC	-6.76 (0.0001)	-8.71 (0.0001)	Reject the null hypothesis
IR	-7.75 (0.0001)	-8.26 (0.0001)	Reject the null hypothesis
GDP	-6.72 (0.0001)	-10.53 (0.0001)	Reject the null hypothesis
HC	-3.64 (0.0001)	-4.38 (0.0038)	Reject the null hypothesis
HS	-8.72 (0.0001)	-11.56 (0.0001)	Reject the null hypothesis
ER	-6.28 (0.0001)	-7.22 (0.0001)	Reject the null hypothesis
UR	-8.01 (0.0001)	-9.93 (0.0001)	Reject the null hypothesis
USA			
VARIABLE	ADF	PP	DECISION
HHD	-4.07 (0.0098)	-7.17 (0.0001)	Reject the null hypothesis
HP	-6.02 (0.0001)	-11.79 (0.0001)	Reject the null hypothesis
CP	-11.31 (0.0001)	-16.96 (0.0001)	Reject the null hypothesis
INC	-6.39 (0.0001)	-11.68 (0.0001)	Reject the null hypothesis
IR	-7.36 (0.0001)	-9.47 (00001)	Reject the null hypothesis
GDP	-4.17 90.0072)	-5.81 (0.0001)	Reject the null hypothesis
HC	-4.39 (0.0037)	-5.37 (0.0001)	Reject the null hypothesis
HS	-7.21 (0.0001)	-13.96 (0.0001)	Reject the null hypothesis
ER	-8.10 (0.0001)	-7.27 (0.0001)	Reject the null hypothesis
UR	-3.38 (0.0599)	-3.56 (0.0388)	Reject the null hypothesis

Table 4.3 gives a summary of the probability values on parentheses associated with individual variables. All these probabilities are observed for ADF and PP tests and are compared with a 5% level of significance. After differencing, the results show that the observed ADF and PP probabilities are less than the 5% significance level all suggesting no unit root in the series. This implies that the variables are stationary and are integrated of the same order. This confirms what Figure 4.12 to 4.14 depicts. This decision is affirmed with the KPSS test summarised on Table 4.4.

Table 4.4: KPSS Stationarity Test

SA					
Type	Lags	Eta	Prob 10 %	Prob 5 %	Prob 1 %
Single Mean	3	0.0874	0.3470	0.4630	0.7390
Trend	3	0.0862	0.1190	0.1460	0.2160
USA					
Type	Lags	Eta	Prob 10 %	Prob 5 %	Prob 1 %
Single Mean	3	0.0946	0.3470	0.4630	0.7390
Trend	3	0.0961	0.1190	0.1460	0.2160

The KPSS critical value at 0.05 significance level is 0.463. The observed statistics associated with the mean are less than this critical value. Based on this information the null hypothesis cannot be rejected, implying that the series are stationary. This is in accordance with the decision taken using the ADF and PP tests.

Since the variables satisfy stationarity conditions, the next step uses the Johansen approach to estimate the parameters of a VAR (1) model. As explained in chapter 3, the construction of VECM is informed by the results obtained from cointegration analysis.

4.6 Convergence results

Figure 4.15 below is the time series plot of the household debt differential for RSA and USA, Ω=HHD_DIFF. A visual assessment of the plot suggests that HHD_DIFF could be non-stationary. A formal test of unit root using the Phillips-Perron method produced the results in the table below. Appendix E1 presents the Philips-Perron test of unit root in the household debt differential.

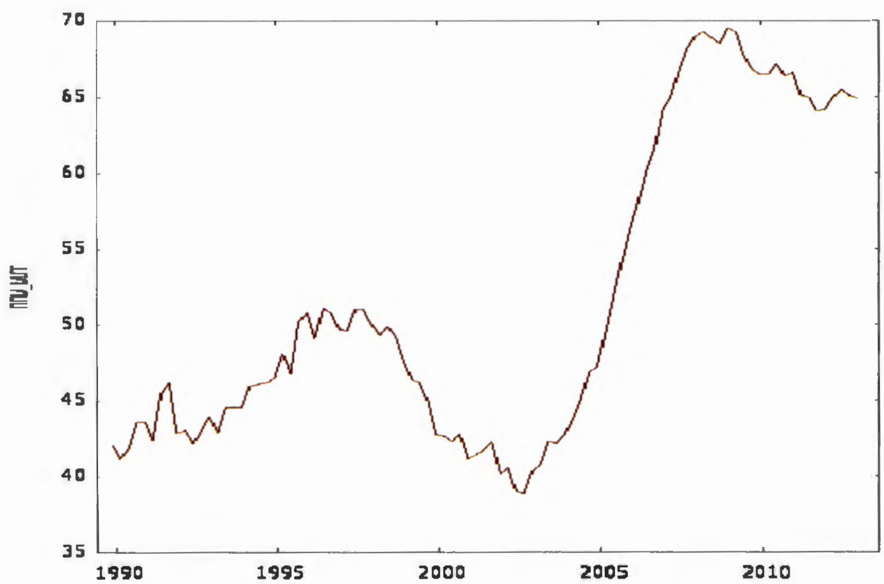


Figure 4.15: Time Series Plot of Household Debt Differential

From the Phillips-Perron test results, the test statistic of -1.4968 is greater than the 5% critical value of -3.459397, suggesting the null hypothesis of a unit root in the household differential cannot be rejected. Hence, household debt differential between South Africa and the USA is non-stationary. This conclusion is confirmed by the KPSS unit root test results also shown in

Appendix E2. The presence of a unit root in the household differential suggests that the two household debt differentials diverge. Also from Appendix E1, the deterministic time trend, @TREND, is not statistically significant at the 0.05 level (*t-value* = 1.1149, *prob*= 0.2679 > 0.05). However, its coefficient, 0.008498, being positive indicates there is a loose lagging-behind process between household debt burdens of South Africa and the USA.

4.7 Cointegration results

Cointegration analysis is conducted in this study using the Johansen (1988) procedure. This helps in determining whether there is a long and short run equilibrium relationship between SA and USA household debts and its theoretical determinants. The approach is advantageous as it allows the integration of the long-run and the short-run relationship between variables within a unified framework to ensure elimination of spurious relations as suggested by Brooks (2008). One of the conditions of cointegration is that two or more variables have to be integrated of the same order and this has been confirmed with stationarity results.

One of the objectives of this study is to analyse the long-run and short-run relationships between household debt and its main determinants during the period 90Q1 to 13Q1. This process is preceded by a lag length determination which is accomplished in this study by using the minimum information criteria (MINIC) AIC and SBC. The detailed results for these two measures are summarised in Table 4.6 for SA and the USA, respectively. These results are obtained using SAS PROC VARMAX. The table showing MINIC chosen by SAS are summarised on Table 4.5.

Table 4.5a: Default Minimum Information Criterion

SA	
AIC	-76.4636
SBC	-73.4484
USA	
AIC	-79.229
SBC	-76.2138

Table 4.5b: Minimum Information Criterion

			SA:AIC			
Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	-53.49873	-53.89992	-52.77407	-51.63181	-50.82874	-51.45387
AR 1	**-76.46362**	-76.46825	-75.98166	-75.61917	-76.6122	-80.62406
AR 2	-76.62648	-76.89497	-76.51184	-76.41954	-78.89583	-88.50296
			SA:SBC			
AR 0	-53.2264	-50.42099	-46.94699	-43.72801	-40.72912	-39.73557
AR 1	**-73.44843**	-70.37145	-68.01489	-66.1474	-65.18232	-66.1887
AR 2	-70.83219	-68.27562	-66.5437	-66.92564	-66.56584	-69.98311
			USA: AIC			
AR 0	-58.57668	-57.37487	-56.09162	-54.80699	-54.2561	-54.74244
AR 1	**-79.22903**	-78.69282	-77.96713	-78.06888	-79.66333	-83.08709
AR 2	-79.69407	-79.78797	-79.3084	-80.81069	-85.12969	-101.7298
			USA: SBC			
AR 0	-58.30435	-54.69419	-50.93913	-47.8864	-44.94374	-42.14654
AR 1	**-76.21385**	-73.35538	-70.5153	-69.19418	-67.82545	-67.4007
AR 2	-73.89978	-71.20147	-68.74987	-67.77122	-67.64267	-69.11102
AR 3	-70.753	-68.86906	-67.51273	-66.84623	-68.39117	

AIC: Akaike information criterion, SBC: Schwartz Bayesian criterion, AR: Autoregression, MA: Moving average

The MINIC chooses lag one as highlighted for both countries. This implies that the analysis is executed with lag one as an optimal lag. Next, the study looks at the cointegration tests for the variables used in the analysis. The Trace test assesses the null hypothesis of r cointegrating vectors against the alternative hypothesis of n cointegrating vectors as set out in Chapter 3. If $r=0$, it means that there is no relationship among the variables that are stationary. Table 4.6 gives the results indicating the order for the Johansen cointegration rank tests and also confirms if there is a long-run relationship between the variables.

Table 4.6: Cointegration tests

		Trace			Max eigenvalue		
H_0: Rank= r	H_1: Rank> r	USA	SA	5 % Critical Value	USA	SA	5 % Critical Value
0	0	326.0672	452.3907	232.60	105.3766	109.7215	62.81
1	1	242.2137	293.7431	192.30	88.5401	95.4438	57.12
2	2	173.6271	194.7045	155.75	63.8624	68.2503	51.42
3	3	**119.2882**	131.1307	**123.04**	**44.9380**	56.6987	**45.28**
4	4	82.7080	**86.6066**	**93.92**	26.7583	**31.0888**	**39.37**
5	5	53.8661	52.7427	68.68	22.8977	25.3193	33.46
6	6	34.6115	29.0491	47.21	14.0998	15.0285	27.07
7	7	20.0463	15.1736	29.38	7.4046	10.2022	20.97
8	8	11.7308	4.0680	15.34	5.2918	6.0080	14.07
9	9	3.7227	0.0702	3.84	0.9877	1.5898	3.76

As Table 4.6 suggests, the Trace and the Maximum eigenvalue tests reject the null hypothesis H_0: Rank=0, i.e. there is no cointegration both for SA and the USA. The test statistics of about 452.39 and 362.07 for SA and the USA respectively are greater than the critical value of approximately 232.60 at 5 % significance level as also shown in the output. The first eigenvalues are also greater than the corresponding critical values. The cointegration test, therefore, concludes that there is cointegrating relationships (vectors) at the 5 % level of significance with orders 4 and 3 for SA and the USA, respectively. The trace test provides a similar conclusion. The maximum eigenvalues 44.938 and 31.088 for the USA and SA in that order are less than the critical values parallel to them confirming this decision.

Since cointegration relationship between variables has been confirmed, the next step constructs the models, firstly the unstandardised and then the standardised household debt models for SA and the USA. The cointegrating equations are firstly built before the ECM. The ECM uses an error term to tie the short run behaviour to the long-run (Gujarati 2010; LeSage 1990). The VAR(1) and VECM(1) are represented as equations [3.14] and [3.15] in Chapter 3.

4.7 Estimated model

The cointegration testing affirmed that there is a possibility of cointegration or equilibrium relationship between the variables. This helps in identifying the variables that may have long and short term effects on household debts in SA and the USA. Also, this may help in building a model that can be used for improvement of longer term forecasting of household indebtedness. The results for VAR(1) and VECM(1) models are summarised on Tables 4.7 and 4.8 for SA and the USA respectively. These models are estimated based on the unrestricted conditions and the information obtained from the cointegration tests.

4.7.1 VAR (1)

Table 4.7 gives parameter estimates for VAR (1) models based on lag one for SA and the USA respectively.

Table 4.7: SA VAR (1) Model Parameter Estimates

Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	PR > |t|	Variable
HHD	CONST1	-1.30312	0.99847	-1.31	0.1955	1
	AR1_1_1	0.86272	0.03817	22.60	0.0001	HHD(t-1)
	AR1_1_2	0.02162	0.03059	0.71	0.4818	HP(t-1)
	AR1_1_3	0.05080	0.14393	0.35	0.7251	CP(t-1)
	AR1_1_4	0.49749	0.13520	3.68	0.0004	INC(t-1)
	AR1_1_5	-0.05997	0.02560	-2.34	0.0216	IR(t-1)
	AR1_1_6	-0.59461	0.15618	-3.81	0.0003	GDP(t-1)
	AR1_1_7	0.23653	0.16024	1.48	0.1438	HC(t-1)
	AR1_1_8	0.06741	0.03185	2.12	0.0374	HS(t-1)
	AR1_1_9	-0.00933	0.01238	-0.75	0.4533	ER(t-1)
	AR1_1_10	-0.09631	0.02207	-4.36	0.0001	UR(t-1)

The VAR (1) obtained from the SA data in mathematical form is as follows;

$$\Delta y_t = [-1.303] + [\ 0.862\ ]HHD_{t-1} + [\ 0.021\ ]HP_{t-1} + \ldots - [\ 0.009\ ]ER_{t-1} - [0.096]UR_{t-10}$$

This model identifies five significant parameters such as INC, IR, GDP, HS and UR except HHD which is a variable of interest in this study. Also revealed are the five signs concurring with theory.

SA have been normalised on household debt as a dependent variable in this study. This helps in expressing the long-run relationship between this variable and the related determinants and their effects thereof.

The outputs for VECM (1) for the said countries respectively are given on Table 4.9 and Table 4.10.

4.7.2 VECM (1)

This table shows the parameter estimates in terms of the constant and lag one coefficients of the first difference (y_{t-1}) contained in the estimates $(\alpha\beta')$ for SA and the USA in that order.

Table 4.9: SA HHD VEC Model

Type of Model	VECM(1)			
Estimation Method	Maximum Likelihood Estimation			
Cointegrated Rank	4			
Long-Run Parameter Beta Estimates When RANK=4				
Variable	1	2	3	4
HHD	1.000	1.000	1.000	1.000
HP	0.048	0.169	-1.756	-3.428
CP	3.441	-0.732	-8.799	-27.937
INC	-0.765	-2.284	16.013	-1.971
IR	-0.229	0.724	0.831	2.591
GDP	-1.282	2.680	-6.415	31.303
HC	-0.021	-1.134	-5.911	-28.992
HS	0.259	-0.419	2.714	-5.099
ER	-0.189	0.216	0.084	0.395
UR	0.638	0.787	0.441	-0.202

The HHD four normalised equations obtained from the analysis of SA data are shown in Table 4.9b. Only the third equation with reversed signs is estimated as follows;

$$\Delta \ln HHD + \beta_1 \Delta \ln HP + \beta_2 \Delta \ln CPI + \ldots - \beta_8 \Delta \ln ER - \beta_9 \Delta \ln UR$$

$$\Delta \ln HHD = 0.302 - 1.756\, \Delta \ln HP - 8.799\, \Delta \ln CPI + 16.013\, \Delta \ln INC + \ldots - 0.441\, \Delta \ln UR$$

This equation is not consistent with theory in terms of CP only. However, the results show that a 1 % decrease in INC and HS will negatively affect HHD by reducing it by 1601 % and 271 % individually in the short run. Also a percentage decrease in UR and IR will also reduce HHD by 44 % and 83 % respectively. The opposite is true about HC and GDP. If these two are increased by 1 %, the result will be an increase in HHD by 591 % and 641.5 % respectively. It is observed from this output that only HC will have a negative effect on HHD in the long-run across the three equations. The results of this analysis confirm the suggestions of most of the authors.

Table 4.10: US HHD VEC Model

Type of Model	VECM(1)		
Estimation Method	Maximum Likelihood Estimation		
Cointegrated Rank	3		
Long-Run Parameter Beta Estimates When RANK=3			
Variable	1	2	3
HHD	1.000	1.000	1.000
HP	-0.325	-13.006	-10.002
CP	3.018	-75.703	24.663
INC	1.991	38.054	29.860
IR	0.135	-1.195	-7.968
GDP	-0.649	-99.590	10.908
HC	-1.453	83.705	-33.462
HS	0.264	-2.891	-8.114
ER	0.283	6.902	7.531
UR	0.224	1.533	6.668

Table 4.10 presents the results of the long-run beta parameter estimates for the three cointegrating equations obtained from the American data. In the long-run, HP will have a positive effect on HHD as the three equations signify. This effect will be negative but for a short while. On the other hand INC, ER and UR will have positive effects in the long-run with the short run reporting opposite effects. Other variables display mixture of effects on HHD. The mathematical representation of the first equation with HHD integrated is shown below;

$$\Delta \ln HHD - \beta_1 \Delta \ln HP + \beta_2 \Delta \ln CPI + \beta_3 \Delta \ln INC + \beta_4 \Delta \ln IR - \beta_5 \Delta \ln GDP - \beta_6 \Delta \ln HC + \beta_7 \Delta \ln HS + \beta_8 \Delta \ln ER + \beta_9 \Delta \ln UR$$

The signs of these coefficients except for the *lnHHD* are reversed and these are checked against those anticipated. The final equation showing the long run relationship between HHD and related variables is as follows;

$$\Delta \ln HHD = 1.856 - 0.325 \Delta \ln HP + 3.019 \Delta \ln CPI + 1.991 \Delta \ln INC + \ldots + 0.283 \Delta \ln ER + 0.225 \Delta \ln UR$$

All the signs of this model are consistent with economic theory. According to this equation, it is observed that in the long-run, a 1 % decrease in CP and INC will induce HHD reduction by 301 % and 199 % respectively. The results are in support of Debelle's (2004) view that people with low income have reduced debts as compared to those with higher earnings. The LCH on the other hand states that when consumer prices are low, HHD will also be reduced. However, these findings do not concur with Prinsloo's (2002) findings who reported a rise in HHD as a result of rising INC. The results also confirm that if UR and IR are reduced by 1 % in the short run, HHD will also go down by 22.5 % and 13.6 % in that respect. These findings are consistent with Hurst & Stafford's (2004) and Debelle's (2004) suggestions that when interest rates are halved, households could double their take-out loan and accumulate more debts.

The model shows that HP will have positive impact on HHD. These findings are in support of Debelle (2004), Subhanji (2007) and Ando & Modigliani (1963) theories who highlighted that rise in HP could encourage households to consume more and build up debts by betting higher expectations for the future. Inversely, if HC and GDP go up by 1 %, HHD will definitely increase by 145.3 % and 65 % separately. The sign of HS confirm the LCH and Kotzé & Smit (2008) theory stating that when household dissave, they tend to go into debt in order to satisfy their consumption and increase their debts. The second equation shows opposite signs for CP, IR, HC and HS and in equation three the different signs are shown for IR, GDP and HS. Based on these findings, this study selects the first equation as the one that best fits the data in the context of the USA.

The coefficients of the VECM for SA and the USA are overvaluing the contribution of the determinants to HHD. However, this is corrected with the ECM. If actual equilibrium value is too high, the error correction term will reduce it while if it is too low, the error correction term will raise it (Nwachukwu & Egwaikhide 2007). The ECM is obtained using equation [3.23] in Chapter 3 and the results are summarised in Tables 4.11 and 4.12 for SA and the USA respectively.

According to the results, also with reference to equation [3.23], $\beta's$ are the short run coefficients of household debt determinants, α is the intercept, ε_t is a white noise error term which corrected the equilibrium. ECT_{t-i} is an error correction term that guides the independent variables of the system to restore equilibrium or correct disequilibrium. The negative sign of the error correction term informs of the rate at which it corrected the previous period disequilibrium of the system.

Table 4.11 SA Error Correction Model

Type of Model	ECMX(1,0)					
Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
D_HHD	CONST1	-0.123	0.882	-0.14	0.889	1
	XL0_1_1	0.075	0.026	2.83	0.005	HP(t)
	XL0_1_2	0.255	0.127	2.00	0.049	CP(t)
	XL0_1_3	-0.581	0.129	-4.50	0.0001	INC(t)
	XL0_1_4	-0.053	0.023	-2.23	0.028	IR(t)
	XL0_1_5	0.244	0.144	1.70	0.093	GDP(t)
	XL0_1_6	0.276	0.141	1.96	0.053	HC(t)
	XL0_1_7	-0.123	0.028	-4.32	0.0001	HS(t)
	XL0_1_8	-0.001	0.011	-0.07	0.946	ER(t)
	XL0_1_9	-0.095	0.018	-5.15	0.0001	UR(t)
	AR1_1_1	-0.138	0.032			HHD(t-1)

The estimated short run model obtained for SA is

$$\Delta \hat{l} nHHD_t = -0.124 + 0.076 \ln \Delta HPI_{t-i} + 0.256 \Delta \ln CPI_{t-i} + \ldots - 0.096 \Delta \ln UR_{t-i} - 0.138 ECT_{t-i}$$

The signs of this model correspond to those dictated by theory of household debts. There is an adjustment in the magnitudes of the long-run coefficients. All the variables in the system appear to be significant at 5 % significance as their allied probabilities confirm. Though ER seem to be less significant in the model, the sign of the adjustment factor is negative and indicates that a long-run equilibrium characterised the relationship among the variables in a system. The coefficient 0.1378 implies that the system corrected its previous disequilibrium period due to its shocks in one period at an adjustments speed of 13.78 % per quarter, hence the signs of the estimated coefficients. The results report an equilibrium HP and HHD confirming the findings by amongst others Philbrick & Gustafsson (2010) and Anundsen & Jansen (2011). The study by Loganathan *et al.* (2010) calculated a speed of adjustment as 13 % from the VECM constructed on macroeconomic variables together with private debt as a dependent variable. This study reported an adjustment speed of 13.78 % which is just 0.78 % higher than the previous study findings.

Table 4.12: USA Error Correction Model

Type of Model	ECMX(1,0)					
Model Parameter Estimates						
Equation	Parameter	Estimate	Standard Error	t Value	Pr > |t|	Variable
D_HHD	CONST1	0.520	1.263	0.41	0.681	1
	XL0_1_1	-0.017	0.033	-0.50	0.619	HP(t)
	XL0_1_2	-0.028	0.069	-0.40	0.689	CP(t)
	XL0_1_3	0.012	0.154	0.08	0.935	INC(t)
	XL0_1_4	-0.011	0.014	-0.79	0.430	IR(t)
	XL0_1_5	-0.100	0.131	-0.77	0.445	GDP(t)
	XL0_1_6	0.050	0.078	0.64	0.525	HC(t)
	XL0_1_7	-0.034	0.009	-3.46	0.001	HS(t)
	XL0_1_8	-0.026	0.016	-1.59	0.116	ER(t)
	XL0_1_9	-0.020	0.011	-1.88	0.063	UR(t)
	AR1_1_1	-0.015	0.028			HHD(t-1)

The short run model for the USA is as follows;

$$l n\hat{HHD}_t = 0.520 - 0.017 \ln \Delta HP_{t-l} - 0.028 \Delta \ln CPI_{t-l} + \ldots - 0.020 \Delta \ln UR_{t-l} - 0.0154 ECT_{t-l}$$

The signs of the short run model are slightly different from those of the long-run relationship. For example HP and GDP will have a negative effect on HHD and INC will have a positive effect in the short run. HS proves to be the most significant determinant as the accompanying probability values show. This implies that in the short run, a 1 % increase in HS will result in 3.43 % increased HHD when reduced UR will consequently cause a decrease of 2.03 %. These coefficients' magnitudes have decreased as compared to those of the long-run relationship. The models display HS as the only most significant determinant and the adjustment factor -0.0154 implying that the system corrected its previous disequilibrium period due to its shocks in one period at an adjustments speed of 1.54 % on quarterly basis.

4.9 Causality test results

From the ADF test results of stationarity, HHD and nine hypothesised variables in the dataset have unit roots at the 5% significance level. The results show that the variables exhibit stationary characteristic after first difference. This section applies a modified Granger Causality test by Toda & Yamamoto (1995) to investigate casual relationships between these variables. Tables 4.13 and 4.14 show the results of Toda-Yamamoto Granger-Causality Wald tests for SA and the USA respectively.

Table 4.13 SA Toda-Yamamoto Wald Causality Test

Test	Group 1 Variables	Group 2 Variables	DF	Chi-Square	Pr > ChiSq
1	HHD	HP CP INC IR GDP HC HS ER UR	9	19.76	0.019
2	HP	HHD CP INC IR GDP HC HS ER UR	9	16.92	0.050
3	CP	HHD HP INC IR GDP HC HS ER UR	9	65.56	<.0001
4	INC	HHD HP CP IR GDP HC HS ER UR	9	30.63	0.0003
5	IR	HHD HP CP INC GDP HC HS ER UR	9	20.95	0.013
6	GDP	HHD HP CP INC IR HC HS ER UR	9	7.65	0.569
7	HC	HHD HP CP INC IR GDP HS ER UR	9	50.88	<.0001
8	HS	HHD HP CP INC IR HC GDP ER UR	9	34.26	<.0001
9	ER	HHD HP CP INC IR HC GDP HS UR	9	12.80	0.172
10	UR	HHD HP CP INC IR HC GDP HS ER	9	10.86	0.285

Table 4.14 USA Toda-Yamamoto Wald Causality Test

Test	Group 1 Variables	Group 2 Variables	DF	Chi-Square	Pr > ChiSq
1	HHD	HP CP INC IR GDP HC HS ER UR	9	63.60	<.0001
2	HP	HHD CP INC IR GDP HC HS ER UR	9	30.35	0.0004
3	CP	HHD HP INC IR GDP HC HS ER UR	9	69.67	<.0001
4	INC	HHD HP CP IR GDP HC HS ER UR	9	38.57	<.0001
5	IR	HHD HP CP INC GDP HC HS ER UR	9	23.10	0.0060
6	GDP	HHD HP CP INC IR HC HS ER UR	9	16.52	0.0568
7	HC	HHD HP CP INC IR GDP HS ER UR	9	30.16	0.0004
8	HS	HHD HP CP INC IR HC GDP ER UR	9	28.39	0.0008
9	ER	HHD HP CP INC IR HC GDP HS UR	9	13.26	0.1513
10	UR	HHD HP CP INC IR HC GDP HS ER	9	60.02	<.0001

At the 5% significance level, the null hypothesis of no causality cannot be rejected for GDP, ER and UR for SA and ER for the US. This implies that there is unidirectional relationship running from GDP to HHD, from ER to HHD and from UR to HHD. The findings disagree with Nieto (2007) about UR who reported this variable to be the most significant determinant in the context of Spanish data. This analysis confirms these three variables to be weakly exogenous in the system of HHD in the context of SA. Only one variable seems not to be responding well in the case of the USA. Generally, there is a bidirectional relationship running from seven determinants using South African data and eight using USA data.

From the results, the study conclude that at the significance level of 5%, not all the variables are considered weakly exogenous to the whole system of equations. At this moment, the researcher confirms that the previously estimated VECM(1) is plausible. This process is proceeded with the diagnostic test of the model.

4.9 Model diagnostics results

The output below shows the results for checking the model for white noise residuals in terms of the univariate equation. In the third table, the second column contains the Durbin-Watson test statistics to test the null hypothesis that the residuals are uncorrelated. The third and fourth columns show the test statistics and their probability values to test the null hypothesis that the residuals have normality. The last two columns for ARCH (1) disturbances test the null hypothesis that the residuals have equal covariances (homoscedastic). The second table includes statistics for AR(1), AR(1,2), AR(1,2,3) and AR(1,2,3,4) models of residuals to confirm the null hypothesis that the residuals are uncorrelated.

Table 4.15 SA Model Diagnostics

Univariate Model ANOVA Diagnostics				
Variable	R-Square	Standard Deviation	F Value	Pr > F
HHD	0.508	0.015	9.42	<.0001

Portmanteau Test for Cross Correlations of Residuals			
Up To Lag	DF	Chi-Square	Pr > ChiSq
2	1	4.78	0.029
3	2	6.60	0.037
4	3	6.71	0.082

Univariate Model White Noise Diagnostics					
Variable	Durbin Watson	Normality		ARCH	
		Chi-Square	Pr > ChiSq	F Value	Pr > F
HHD	2.407	1.19	0.553	0.34	0.562

Univariate Model AR Diagnostics								
Variable	AR1		AR2		AR3		AR4	
	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F
HHD	4.74	0.032	2.68	0.074	2.13	0.102	1.73	0.151

Generally the model parameters from the South African data are stable as the probability value (0.0001) associated with the F-test on the table displaying univariate model ANOVA diagnostics shows. The model is also adequate as suggested by Ljung & Box (1978). This is emphasised by the probability associated with the Q-statistic (0.037<0.05 level of significance). This model was found not to be spurious as the value of R^2 (0.508) is less than the Durbin Watson test statistic (2.407) (Tanku, O'Neill & Rowen 2013). The square of portmanteau test for Box-Pierce statistic also concludes that the residuals of the model are not serially correlated up to lag 4, (0.037^2)<0.05 level of significance (Ljung & Box 1978). This is good for this study because the MINIC for cointegration analysis chose lag 3 associated with model 3 in the South African context. According to the output the covariances of the model are equal as proven by the probability value associated with the ARCH (1). This consequently does not reject the null hypothesis of homoscedasticity. Also the residuals are normally distributed as the normality associated probability suggests.

Table 4.16: USA Model Diagnostics

Univariate Model ANOVA Diagnostics				
Variable	R-Square	Standard Deviation	F Value	Pr > F
HHD	0.464	0.010	7.01	<.0001

Portmanteau Test for Cross Correlations of Residuals			
Up To Lag	DF	Chi-Square	Pr > ChiSq
2	1	1.36	0.243
3	2	2.91	0.233
4	3	2.93	0.403
5	4	9.54	0.049

Univariate Model White Noise Diagnostics					
Variable	Durbin	Normality		ARCH	
	Watson	Chi-Square	Pr > ChiSq	F Value	Pr > F
HHD	1.806	2.20	0.333	0.28	0.600

Univariate Model AR Diagnostics								
Variable	AR1		AR2		AR3		AR4	
	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F	F Value	Pr > F
HHD	0.52	0.473	0.57	0.567	0.97	0.410	0.72	0.579

The parameters of the model built using the American data set is also fit for further analysis as confirmed by the probability value of the F-test (Pesaran & Pesaran 1997). This model does not display conditions for spuriority as the R^2 (0.464) is less than the Durbin-Watson statistic (1.806) (Tunku, O'Neill & Rowen, 2013). The null hypothesis of no autocorrelation among the residuals is not rejected since the squared observed probabilities up to lag 3 is greater than 5 % level of significance. This is inconsistent with Ljung & Box's (1978) suggestions and concludes that the residuals are serially correlated up to lag 3. This is not advantageous as the MINIC selected lag one associated with model 1 using the US data. The hypotheses of homoscedasticity and non-normality of the residuals of the error terms cannot be rejected according to the output. This implies that the model is good and can be used for further analysis. In general, the two models proof to be good according to the diagnostic tests used. These models are used in the next section to assess the response of HHD to shocks due to innovations in the related determinants. Forecasts are also generated using these models.

4.10 Generalized Impulse Response Analysis Results

To examine the sign, speed and persistence of the responses of HHD to each of the determinants, quarterly impulse responses were estimated using the GIRF as discussed in chapter 3. The summary of the impulse responses is reported in Figure 15 to 23.

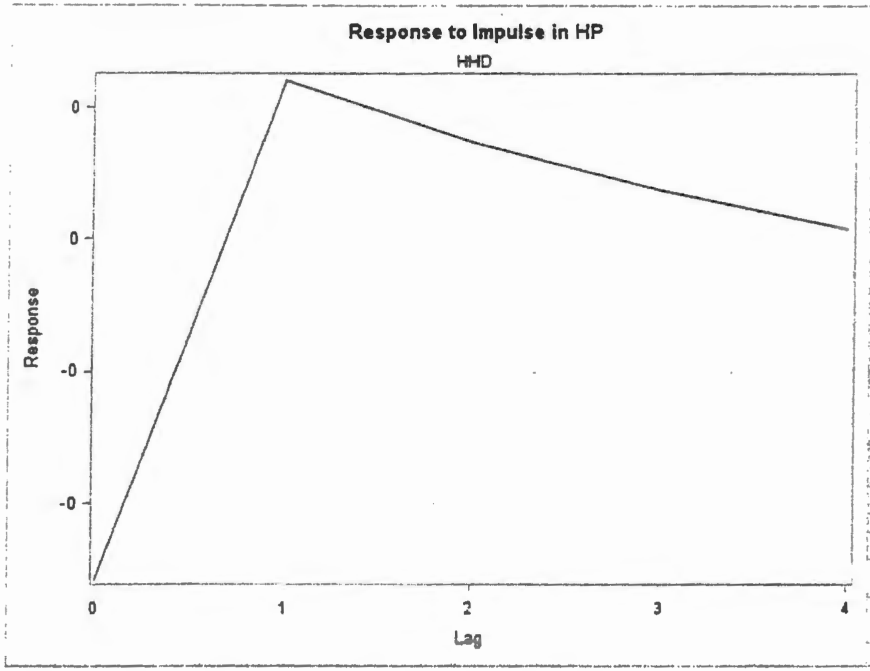


Figure 4.16a: Response to impulse in SA HP

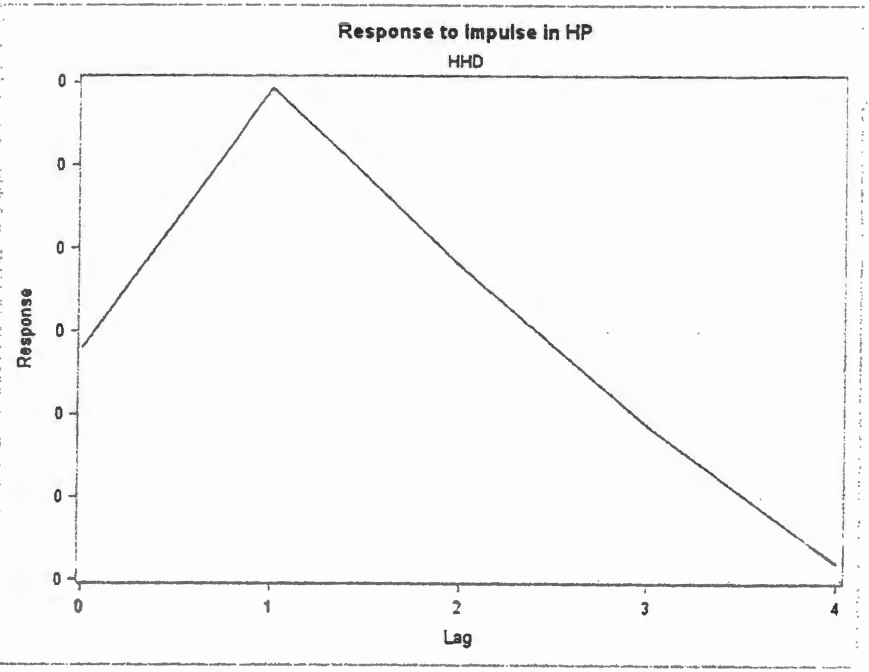


Figure 4.16b: Response to impulse in USA HP

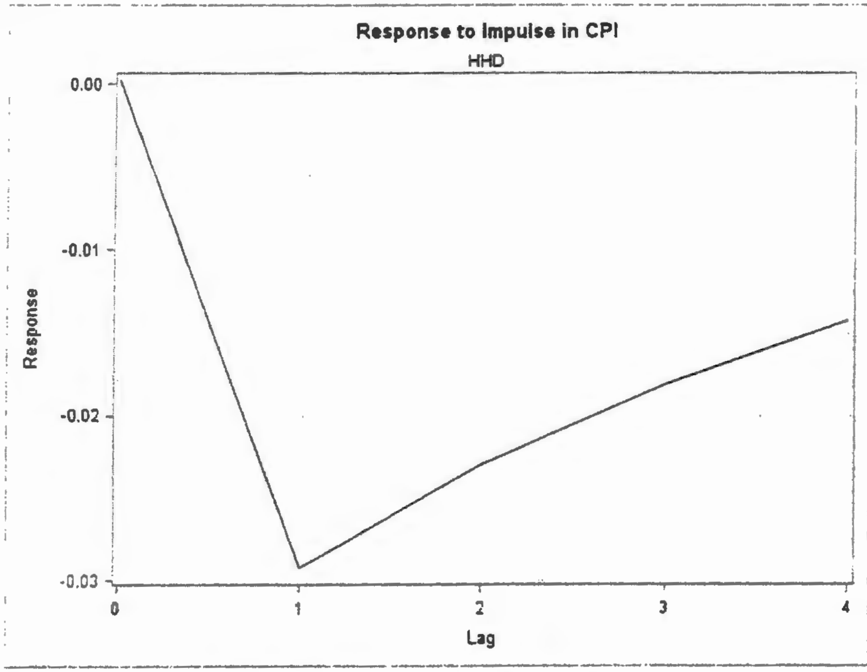


Figure 4.17a: Response to impulse in SA CP

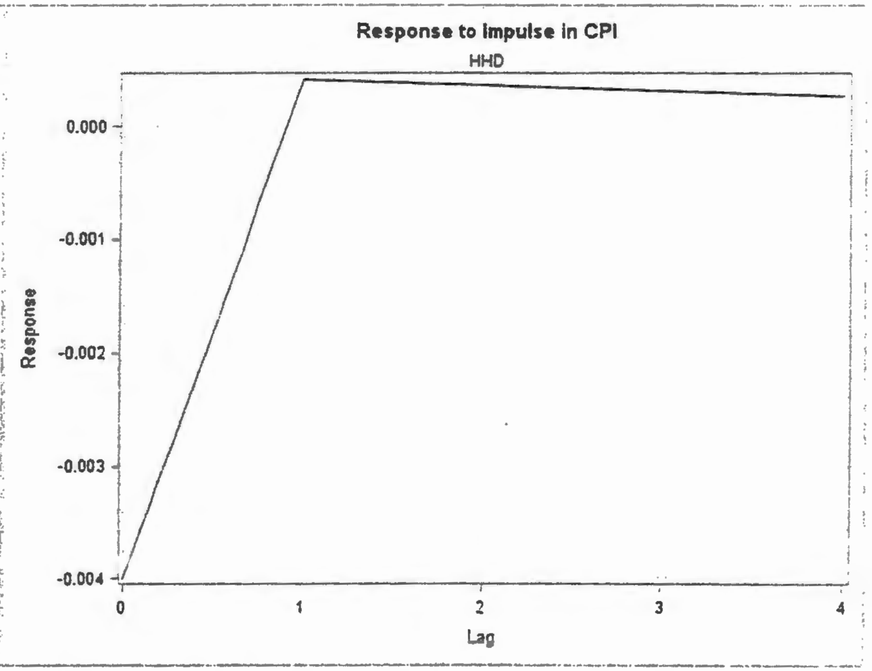


Figure 4.17b: Response to impulse in USA CP

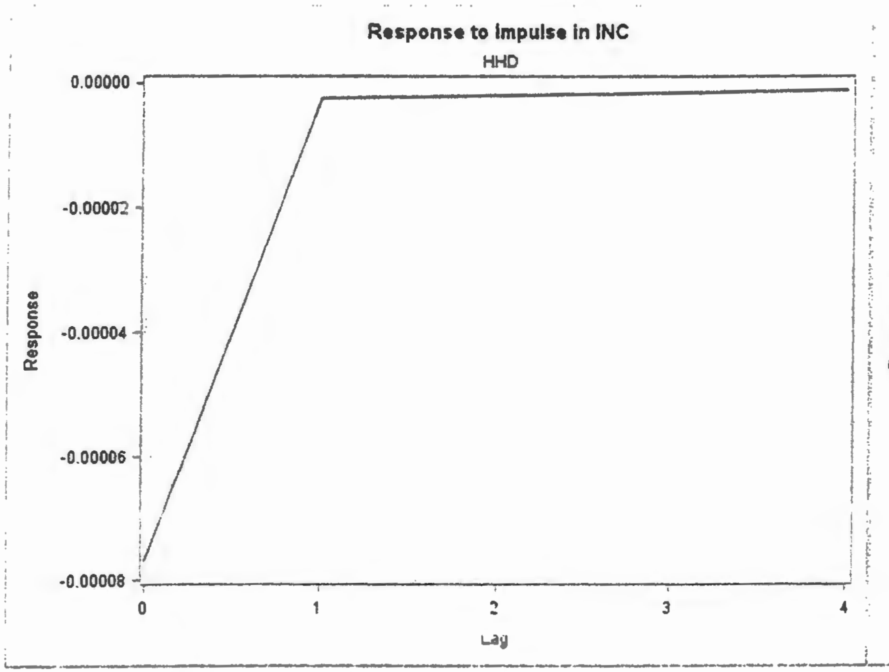


Figure 4.18a: Response to impulse in SA INC

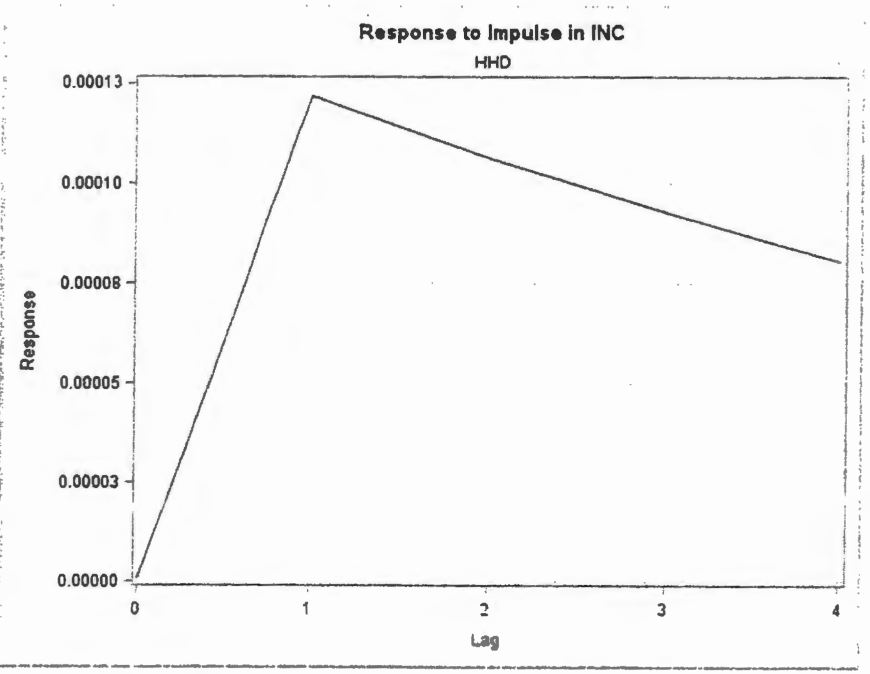


Figure 4.18b: Response to impulse in USA INC

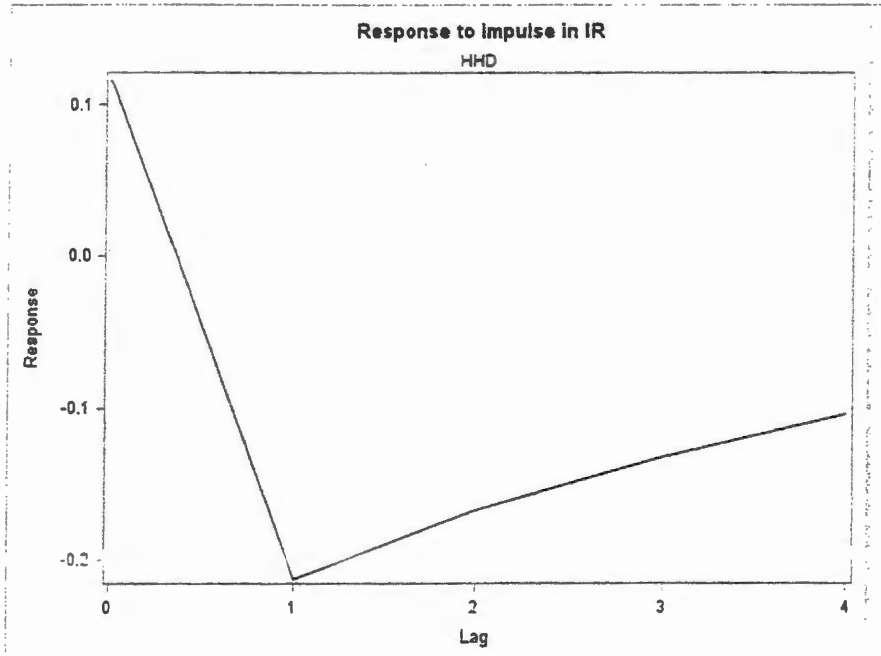


Figure 4.19a: Response to impulse in SA IR

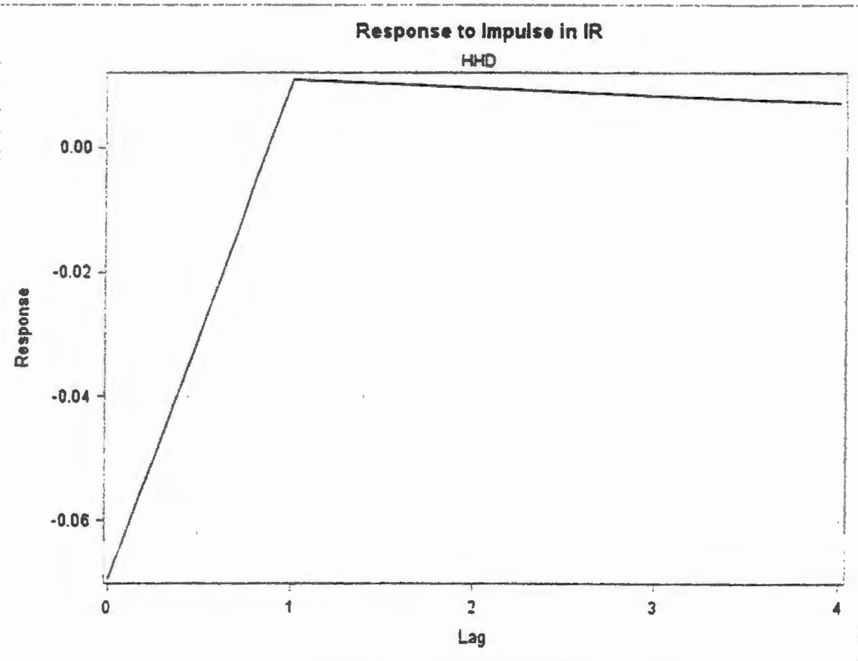


Figure 4.19b: Response to impulse in USA IR

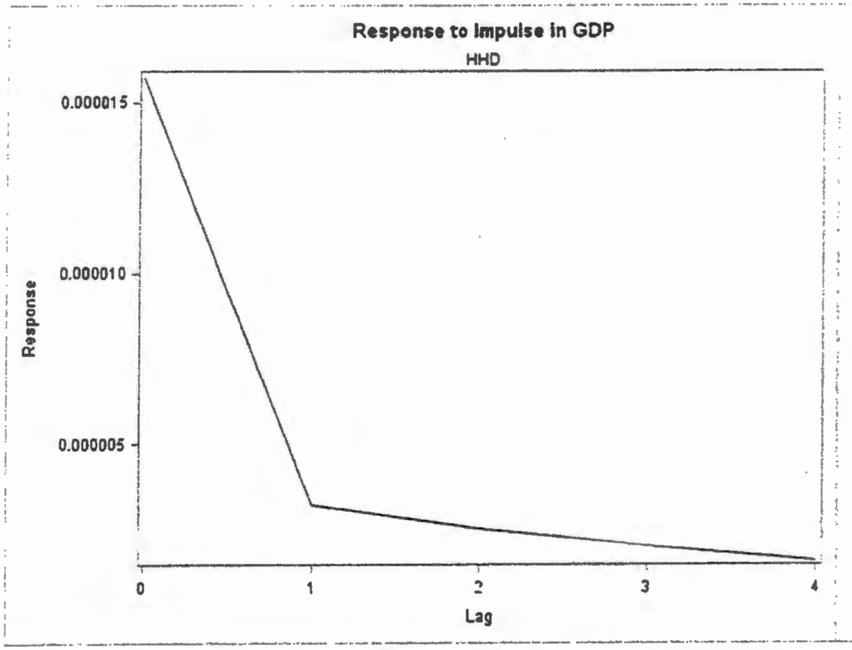


Figure 4.20a: Response to impulse in SA GDP

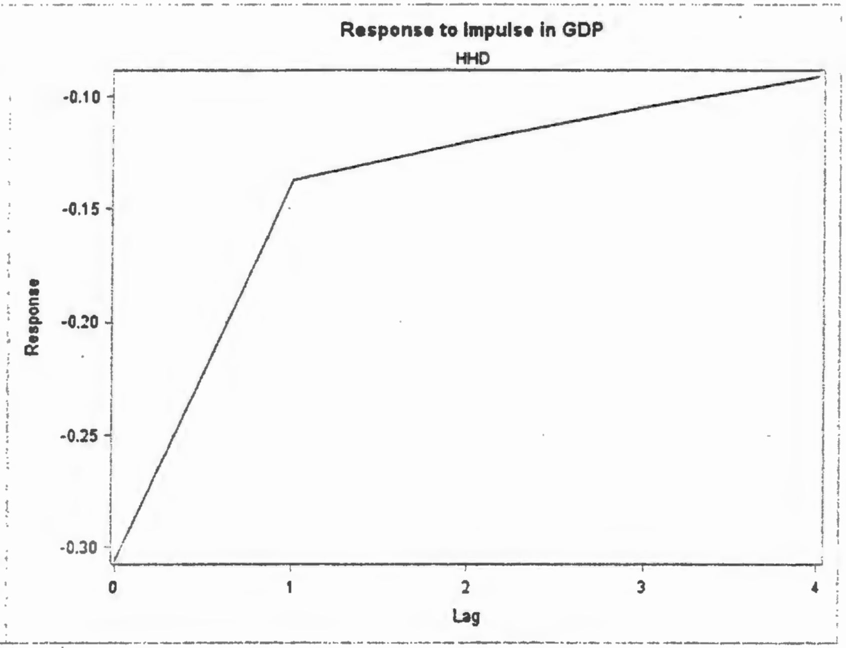


Figure 4.20b: Response to impulse in USA GDP

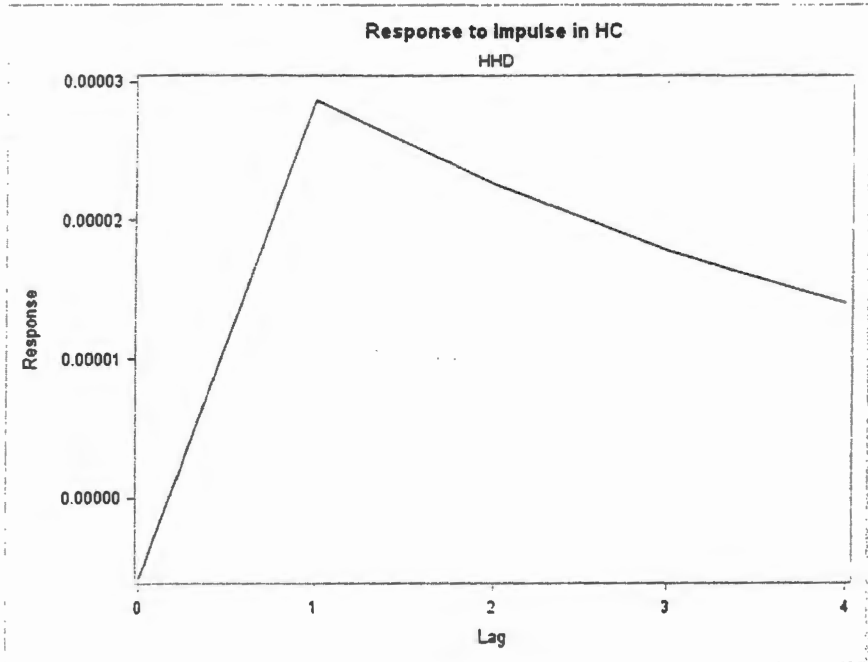


Figure 4.21a: Response to impulse in SA HC

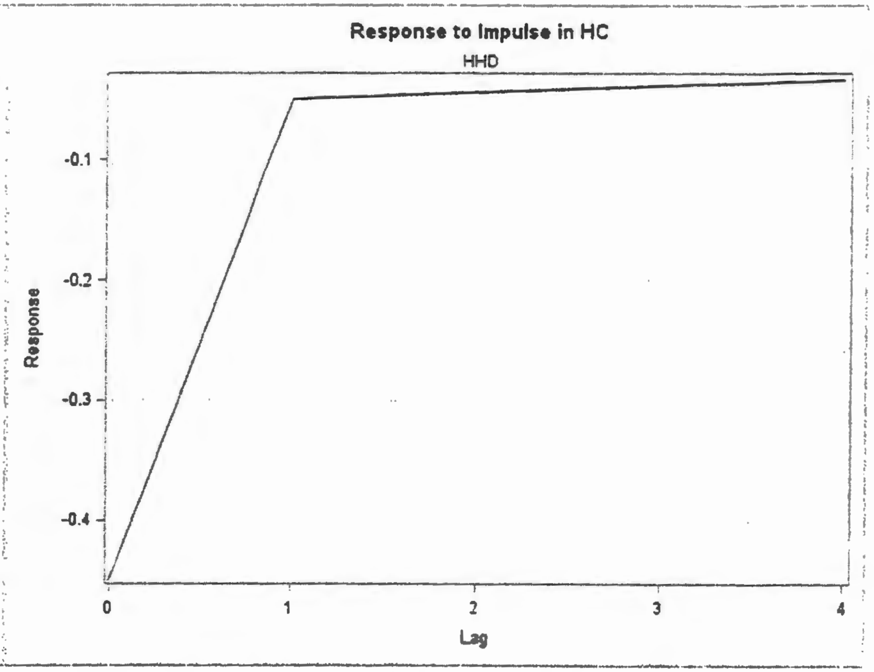


Figure 4.21b: Response to impulse in USA HC

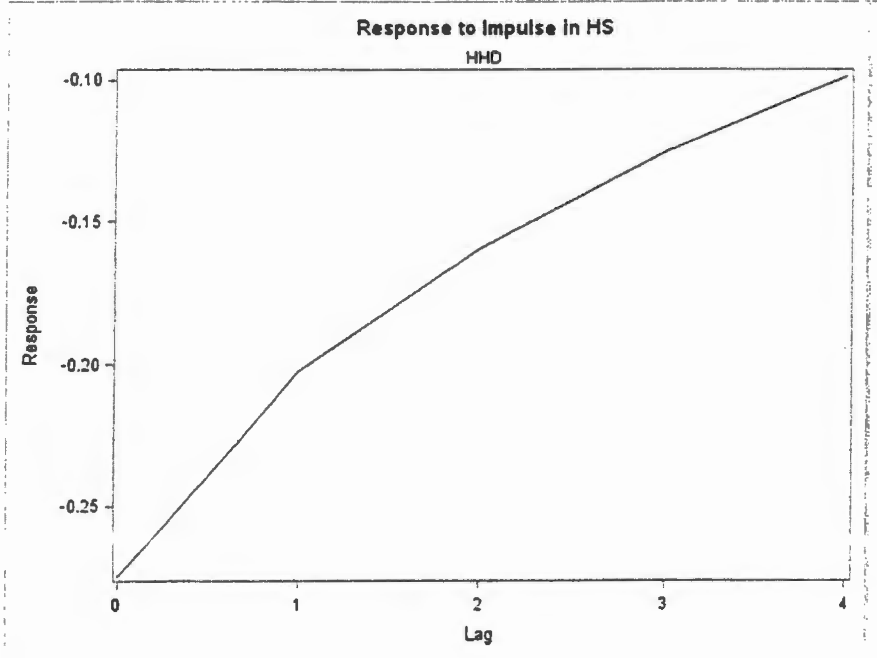


Figure 4.22a: Response to impulse in SA HS

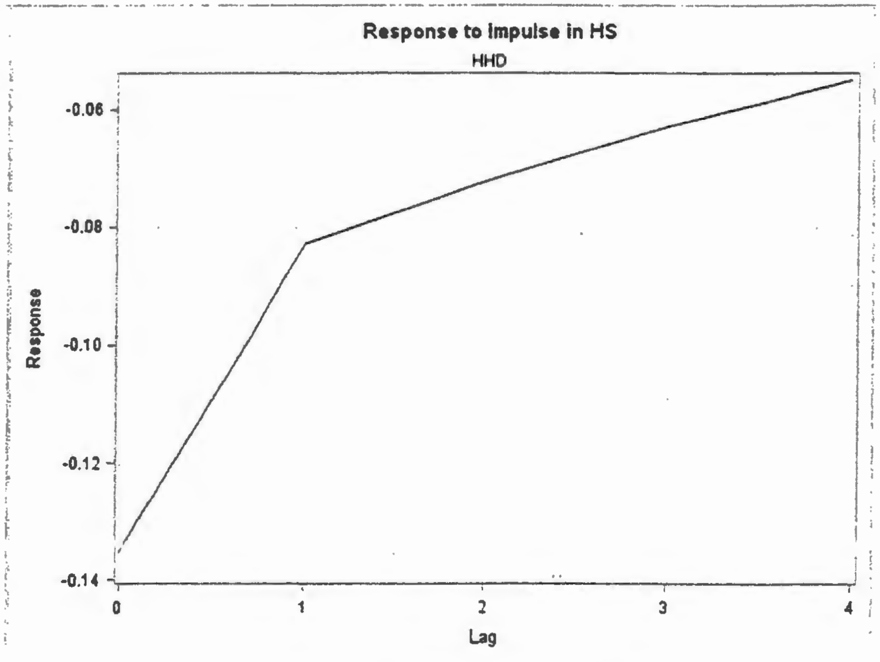


Figure 4.22b: Response to impulse in USA HS

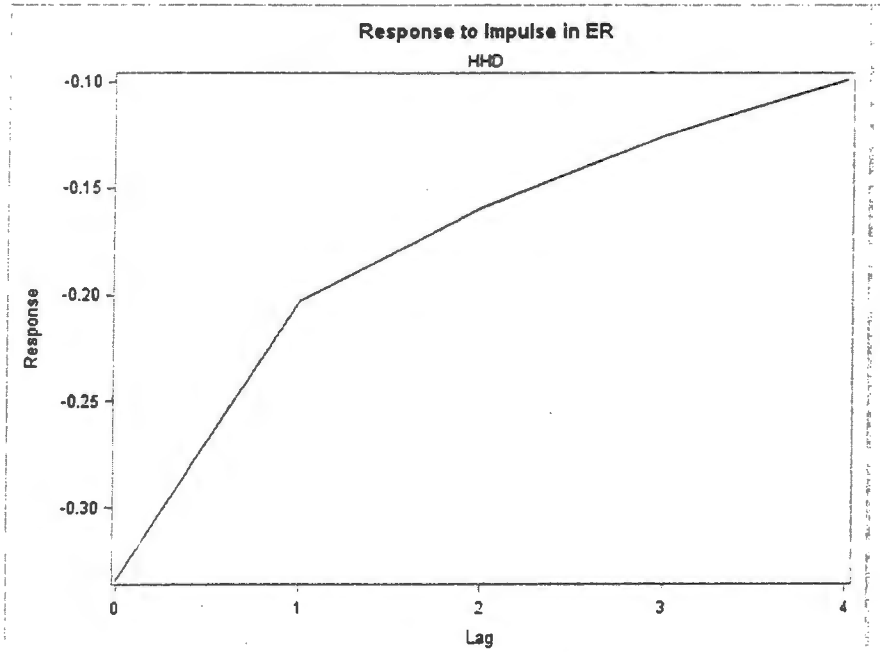


Figure 4.23a: Response to impulse in SA ER

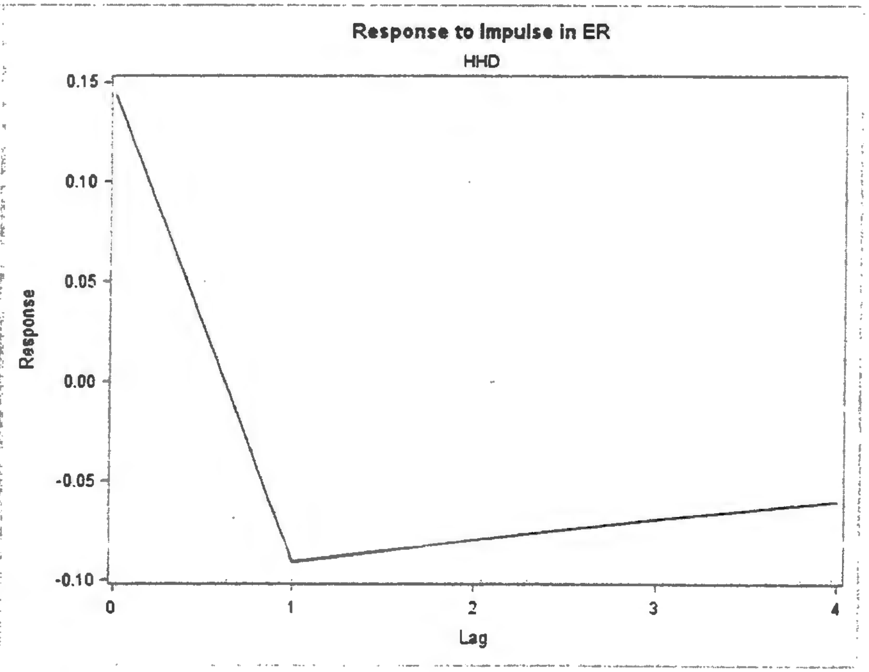


Figure 4.23b Response to impulse in USA ER

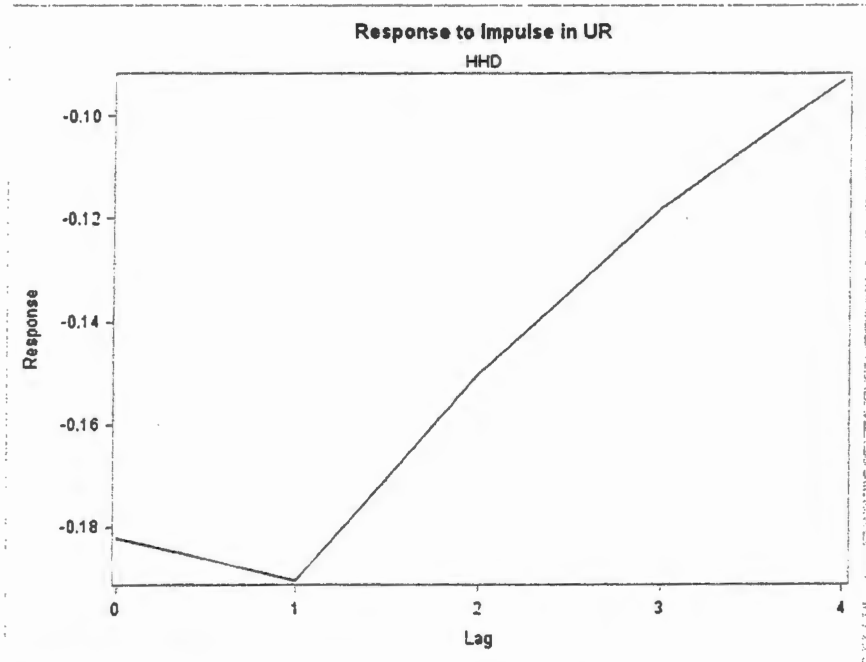


Figure 4.24a: Response to impulse in SA UR

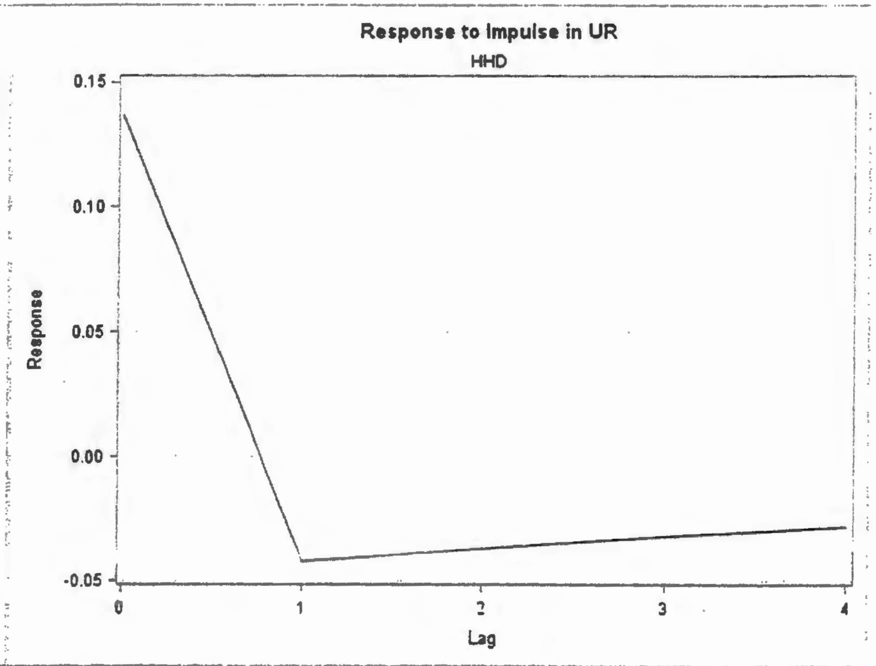


Figure 4.24b Response to impulse in USA UR

It is evident from Figures 4.16 to 4.24 that in the long-run, HHD will respond to different types of shocks. With regards to HP as Figure 4.16 reveals, it is evident that the response of HHD both in SA and the USA to a one-standard error shock is a sharp incline and a quick return to original state especially in a case of SA. Though HP has same effect to debts in these countries, the response seems to be quicker in the USA HP. HS will respond positively as a result of increased savings in the two countries. Debts in SA will be stabilised due to changes in INC, and the same can be concluded about CP and HC in the USA. However, HHD responds differently to shock with regard to IR, GDP and ER. Debts in the South African context started at a high peak and declined sharply with changes in CP and IR before going back to its original state in a very slow movement. The opposite is true for the American situation. The response to innovations in HHD to UR standard errors is more persistent remaining significantly positive in SA. There is no indication that this will return back to normal. In America this intervention will have a very slow effect.

It is clear from the analysis that the results obtained from the GIRF are in accordance with economic theory for most of the variables. It has also been confirmed from the impulses that HHD responds to innovations of all the determinants either negatively or positively, in some cases the response is almost stationary (For instance, the USA's CP and SA's INC, IR and HC.) After certain interventions, HHD either shoots up or down slowly or quickly. The impulses also confirm that the 2007-2009 financial crisis has contributed to the trends in household debts in the USA and SA. This also confirms Sims (1980) and Harris (1995) suggestions that if the variables are not be weakly exogenous, they must cause different responses to shock. Evidence can also be gathered from the analysis that some variables have more effect than others. More of the results on GIRF are given on Appendix B. These results reveal low correlations allowing us to associate shocks in HHD with other variables as suggested by Sims (1980).

4.11 Forecasting

This section provides the five-year forecasts for HHD for the period 13Q2 to 18Q1. These are generated using the VECM (1). The results are given in Figures 4.25 and 4.26.

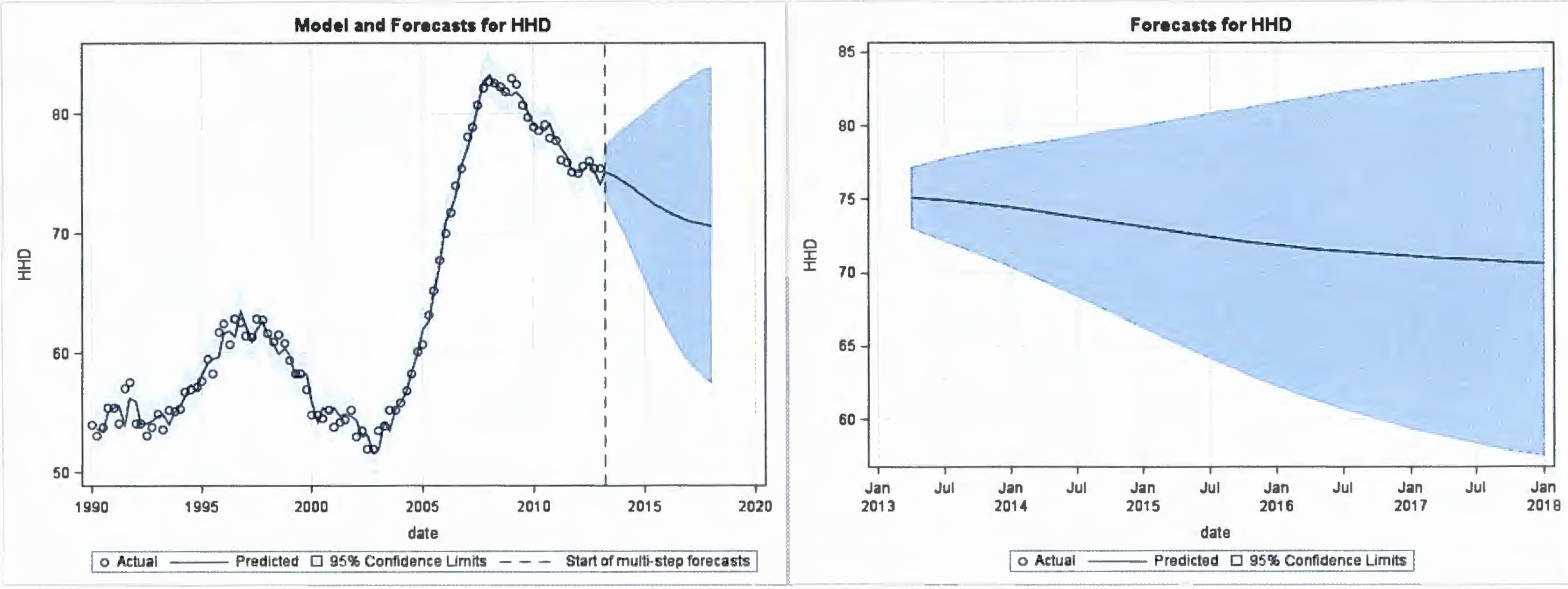


Figure 4.25: Forecasts of HHD SA

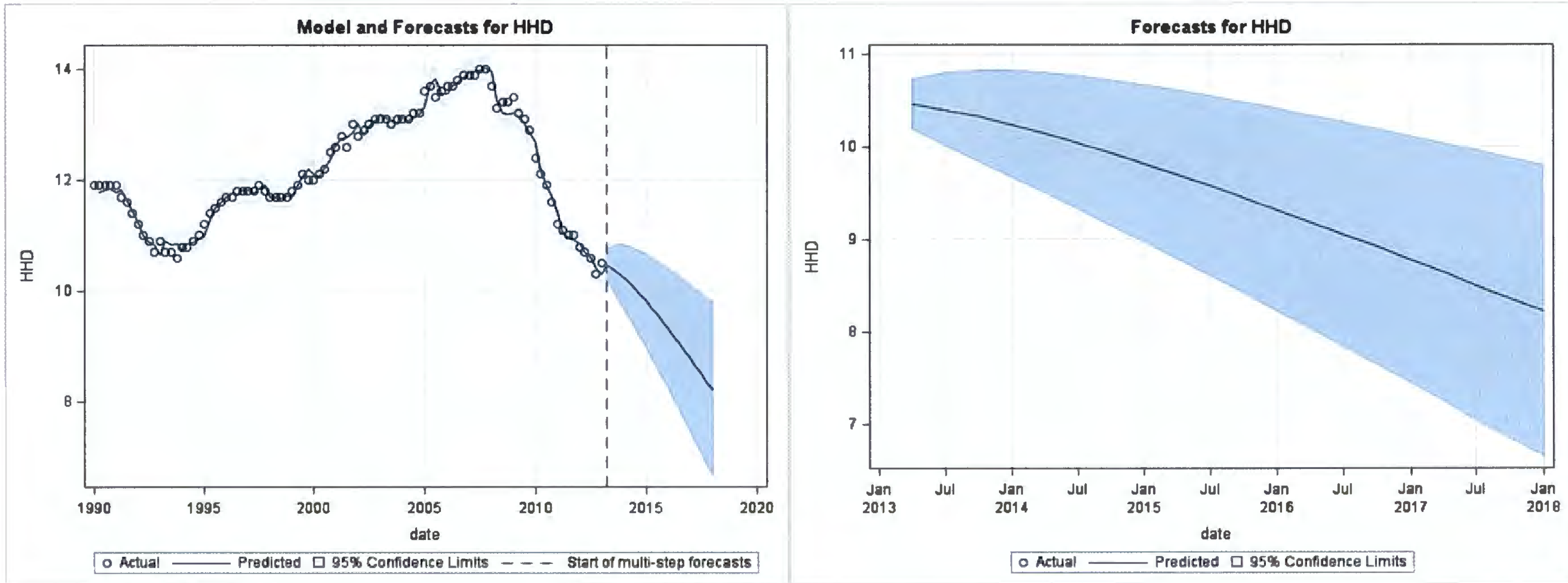


Figure 4.26: Forecasts of USA HHD

Reports show that in the first quarter of 2013 household debt in SA was about 75.4 % and estimated at 10.5 % in the USA. This is also illustrated by the dotted vertical lines in Figures 4.25 and 4.26. The figures also show the predicted rates of household debts from the second quarter of 2013 to the first quarter of 2018, a five-year period in counting. It is evident that HHD in both countries will in the long-run increase as the figures suggest. The prediction lines are shown moving exactly within the predicted confidence limits. During the financial crisis 2007-2009, households found themselves in deep debt and later on this was reduced gradually. The results clearly prove that by end of year 2018, debts in these countries will be estimated at a maximum of about 83.9 % in SA. The expected debt rate in the USA is 9.8 %. Appendix C gives the forecast values of the figures under discussion.

PART B: DISCRIMINANT ANALYSIS RESULTS

4.12 Discriminant analysis results

This section presents the results obtained through the DA framework. The analysis of data is in response to the problems posed in Chapter 1. This study intends to address two objectives through the DA procedure. Firstly, the results that help in determining the significant difference between SA and the USA using household debt determinants are interpreted. This again helps in identifying amongst the nine determinants (independent variables), which ones have a discriminatory power in these countries. The dependent variable used in this analysis is dichotomous, with subcategories, SA and the USA. The observations used in the analysis are mutually exclusive and exhaustive, *i.e.*, each belonging to only one group. Each determinant chosen has 93 observations. This sample size is chosen to guard against sampling error, and to cater for the discriminant and factor methods.

4.12.1 Confirmation of assumptions

Section 4.3 addressed some of the assumptions associated with DA such as multicollinearity, linearity, normality of residuals and sampling adequacy. All these assumptions were found to be acceptable according to the results in section 4.9. DA is vulnerable to unequal dispersion of variances. This section uses the Box's M test to evaluate this assumption as discussed in section 3.10.1. The null hypothesis is rejected if the probability associated with the test statistic is greater than 0.001 significance level. The PROC DISCRIM used to produce the results for this test is presented in Table 4.17 below.

Table 4.17:Test of homogeneity of within covariance matrices

Chi-Square	DF	Pr > ChiSq
1049.574	15	<.0001

Please note that the Chi-Square value 0.0001 is significant at the 0.001 level. This confirms the suitability of the within covariance matrices in the discriminant function. The variance-covariance matrix calculated from the data has equal dispersion. This is in agreement with Morison (1976) and Brace *et al.* (2009). Now that all the assumptions associated with the DA are satisfied, the study next uses the collected data to construct the model using Fisher's LDA. PROC STEPDISC is used in selecting the variables and this is preceded by PROC DISCRIM. The results obtained from these procedures are displayed in Table 4.18 and 4.19.

4.12.2 Variable selection results

Table 4.18a Forward Selection: Step 1

Statistics for Entry, DF = 1, 183				
Variable	R-Square	F Value	Pr > F	Tolerance
HP	0.0164	3.05	0.083	1.0000
CP	0.0224	4.20	0.042	1.0000
INC	0.0119	2.21	0.139	1.0000
IR	0.0001	0.01	0.916	1.0000
GDP	0.0071	1.31	0.254	1.0000
HC	0.0051	0.94	0.335	1.0000
HS	0.0014	0.25	0.618	1.0000
ER	0.0141	2.61	0.108	1.0000
UR	0.0034	0.62	0.432	1.0000

Multivariate Statistics					
Statistic	Value	F Value	DF	Den DF	Pr > F
Wilks' Lambda	0.977	4.20	1	183	0.041
Pillai's Trace	0.022	4.20	1	183	0.041
Average Squared Canonical Correlation	0.022				

In the first step of stepwise analysis the tolerance is 1.0 for each variable under consideration because no variables have yet entered the model. The variable CP is selected because its *F*-statistic, both for Wilks' lambda and Pillai's Trace is 4.20 is the largest among all variables. This implies that CP contributes the most to the improvement of the discriminating power of the model. This choice is confirmed by the R^2 (0.022) which is also greater than those of other determinants. Next the most significant determinant is selected and the results are summarised in Table 4.14b.

Table 4.18 b: Forward Selection: Step 2

Statistics for Entry, DF = 1, 182				
Variable	Partial R-Square	F Value	Pr > F	Tolerance
HP	0.0017	0.31	0.5798	0.1088
INC	0.0213	3.97	0.0479	0.0618
IR	0.0031	0.56	0.4549	0.8484
GDP	0.0582	11.25	0.0010	0.0644
HC	0.0814	16.13	<.0001	0.0676
HS	0.0001	0.02	0.8867	0.9021
ER	0.0010	0.19	0.6629	0.2093
UR	0.0263	4.91	0.0279	0.2161

Multivariate Statistics					
Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.898	10.34	2	182	<.0001
Pillai's Trace	0.102	10.34	2	182	<.0001
Average Squared Canonical Correlation	0.102				

Table 4.14b confirms HC to be the second most important variable according to the *F*-Statistic (16.13) and the R^2 (0.081) which are larger than those of other determinants. To show the next significant variable, the previous step is repeated until all the significant variables are entered in the discriminant model. To summarise, the study next presents the results explaining the contribution of the remaining variables. Table 4.18c gives a summary of all the variables in order of importance.

Table 4.18c: Forward Selection Summary

Forward Selection Summary						
Step	Entered	Partial R-Square	F Value	Pr > F	Wilks' Lambda	Pr< Lambda
1	CP	0.022	4.20	0.0418	0.977	0.0418
2	HC	0.081	16.13	<.0001	0.897	<.0001
3	GDP	0.474	163.59	<.0001	0.471	<.0001
4	INC	0.162	34.83	<.0001	0.395	<.0001
5	HS	0.083	16.36	<.0001	0.362	<.0001

Note that the results of the remaining variables not yet selected are given in Appendix D. Presented here is just a table summarising information about these variables. As Table 4.14c shows, with CP and HC already in the model, GDP is tested for removal before a new variable is selected for entry. Since GDP meets the criterion to stay, it is used as a determinant of household debts for variable selection. This variable is selected because its *F* statistic, 163.59, is the largest among all the seven remaining variables not in the model. This variable is also chosen because its associated tolerance, 0.0003[2], meets the criterion to enter. The process is repeated up to step 6. HS appears to be the variable with the least contribution hence it could not be included in a list. This variable is entered last in the model. Wilk's Lambda for each of the variables selected is significant. Hair *et al.* (2010) suggest smaller values of Wilk's' Lambda to conclude that the variable selected is a good predictor.

Only five variables have been identified as having a significant discriminatory power to differentiate between SA and the USA. However, all the variables are used to build a discriminant model following the Fisher's LDA approach. The next step uses PROC DISCRIM to help construct Fisher's LDA preceded by the communality estimates obtained from PROC FACTOR. Communality estimates are used in this study to confirm the credibility of the variables. Hatcher (1994) defines communality as the variance in observed variables accounted for by common factors or discriminants. The author recommends the use of this measure in the application of exploratory Factor Analyses. If for example, *m* factors were perfect predictors of household debts, y_t, then the error terms and the variances of these predictors are normally distributed with $\varepsilon_i = 0$ and $\sigma_i^2 = 1$. A large communality value signifies a strong influence by the underlying factors (Moroke & Mavetera 2013). A communality value that equals one is an indication that there are too few data analysed and a very low communality for the variable implies that this variable is an outlier. The results of communality estimates and Fisher's LDA are summarised in Table 4.20 and 4.21. The least significant variables are highlighted in Table 4.19. The results presented on the following table show the contribution of common determinants to discriminant function.

2 See appendix D

Table 4.19: Communalities

Determinant	Extraction
CP	0.952
INC	0.939
GDP	0.987
HC	0.223
HS	0.986
HP	0.985
IR	0.144
ER	0.845
UR	0.872

According to the results in Table 4.15, seven of the nine household debt determinants have high communality measures. This implies that of the entire variables gross domestic product, household savings and house prises have almost the same contribution in household debts in SA and the USA. Though household consumption and interest rates have been suggested by literature as determinants of household debts, their contributions appear to be less as opposed to others. Below is the output for Fisher's LDA constructed using the nine determinants of household debt.

4.12.3 Fisher's Linear Discriminant Model

This section discusses the results obtained from PROC DICRIM[3]. Unstandardized and standardised parameter estimates using the selected variables are shown in Tables 4.20 and 4.21 respectively.

Table 4.20: Linear Discriminant Function

Linear Discriminant Function for Group		
Variable	SA	US
Constant	-3.208	-0.305
HP	11.113	3.982
CP	33.589	30.605
INC	165.616	16.476
IR	-10.601	-4.906
GDP	75.853	-58.556
HC	-137.312	50.206
HS	-9.145	-0.321
ER	2.833	0.368
UR	1.297	-0.037

As Table 4.16 reveals, two discriminant functions have been constructed from the data. This model is presented according to equation [3.49] as follows:

[3] Refers to procedure for Discriminant analysis

$$Y_1 = -3.208 + \underbrace{11.113HP}_{+} + \underbrace{33.589CPI}_{+} + \underbrace{165.616INC}_{+} + ... + \underbrace{9.145HS}_{-} + \underbrace{2.833ER}_{+} + \underbrace{1.297UR}_{+}$$

$$Y_2 = -0.305 + \underbrace{3.982HP}_{+} + \underbrace{30.605CPI}_{+} + \underbrace{16.476INC}_{+} + ... + \underbrace{0.321HS}_{-} + \underbrace{0.368ER}_{+} + \underbrace{0.037UR}_{-}$$

The results reveal that house prices, consumer prices, household disposable income exchange rates and unemployment rates are associated with increased debts in SA and the USA. Also revealed is that interest rate and household savings reduce debts. Most of the signs of these models are in accordance with theory but the parameter estimates are misconstruing the contribution of variables. Since the first function has larger parameter estimates, it is standardised and the new parameters are expressed as correlation coefficients. These standardised coefficients are presented in Table 4.21.

Table 4.21: Standardised Discriminant Function

Determinant	Coefficient
HP	0.976
CP	0.969
INC	0.993
IR	-0.472
GDP	0.993
HC	0.992
HS	0.380
ER	0.919
UR	0.934

The standardised canonical discriminant model for HHD constructed using data set of period 90Q1 to 13Q1 is as follows;

$$Y = \underbrace{0.976HP}_{+} + \underbrace{0.969CPI}_{+} + \underbrace{0.993INC}_{+} + ... + \underbrace{0.380HS}_{-} + \underbrace{0.919ER}_{+} + \underbrace{0.934UR}_{+}$$

Literature is silent about the size of loading required for interpretation. However, Tabachnick & Fidell (2007:540) state: "By convention, correlations in excess of 0.33 may be considered eligible while lower ones are not." Based on this guideline, all the variables have coefficients greater than the thumb rule and are therefore regarded as having unique contribution to household debts. Interest rates and household saving have less contribution and all others contribute no less than 90 % to debts in SA and the USA. INC has highest coefficient implying that it has 99.5 % contribution to debts in these countries.

4.12.4 Comparison between SA and the USA

The parameters of the variables in the discriminant function have proven to have stable and valid parameters. Next the study use this function to determine if the difference between SA and the USA is significant or not. The Mahalanobis distance is used in this study together with the Hoteling's test according to equations [3.53] and [3.57] to help achieve this objective. Table 4.22 gives a summary of these statistics.

Table 4.22: Mahalanobis Distance Statistic

Squared Distance (D^2)			F Statistics, NDF=5, DDF=179 for Squared Distance to group		Prob>Mahalanobis Distance for Squared Distance to group	
Group	SA	USA	SA	USA	SA	USA
SA	0	7.138	0	35.078	1.0000	<.0001
USA	7.138	0	35.078	0	<.0001	1.0000

It is noted according to these results that there is a difference between SA and the USA in terms of some of household debt determinants. For instance, although remarkable similarities have been noted in terms of the six determinants, three marked differences exist between these countries. While disposable income and exchange rates are positively related to debts in the USA, the opposite applies to SA. Unemployment rates also affect these countries differently. This is verified by the Mahalanobis distance (7.138) between these countries. The critical value of *F* at 0.05 level of significance with 9 and 175 degrees of freedom is 1.88 and the associated observed probability is 0.0001. The observed *F*-statistic (35.078) is larger than the critical value at the significance of 0.05, suggesting the rejection of the hypothesis of no significant difference between SA and the US. This is a confirmation that though these countries share some of the similarities, they also have differences in relation to debt determinants. To affirm statistical difference between SA and the USA, the study uses the Hoteling's test. This test compares the countries on the basis of the vector of means calculated from the nine household debt determinants for each country. From equation [3.57], the Hoteling's test statistic is equal to:

$$T^2 = \frac{n_i n_j}{n_i + n_j} D^2 \sim T^2_{\alpha(p, n_1 + n_2 - 2)}$$

$$= \frac{(92)(93)}{92 + 93} 7.138$$

$$= 332.123$$

The critical value from the Hoteling's distribution at 0.05 significance level with 2 and 183 degrees of freedom is 6.113. The null hypothesis associated with this test is rejected, implying that there is difference between SA and the USA in terms of household debt determinants.

4.13 Summary

This chapter adopted the Johansen multivariate cointegration approach to construct the VAR(1) and VECM(1) models. Prior to building these models, the data were checked for inconsistencies and adequacy. The measures used proved that the data were adequate and did not contain outliers and missing values. The variables were tested to determine whether they satisfy the reliability condition with Cronbach's alpha. The movement of all nine variables was displayed in time series plots which showed that some variables are explained by trends and random walks, allowing an imposition of some restrictions in the models. These variables were subjected to a first difference, and they all were found to be stationary graphically and through the confirmation of the ADF, PP and KPSS tests. The convergence tests were also calculated for household debts in SA and USA.

The actual analysis of cointegration commenced by using the AIC and SBC to choose the optimal lag. The Fountis-Dickey maximum eigenvalue process was used in determining the number of cointegrating equations from the data set of each country. The information from the lag determination and the number of cointegrating ranks was used in building and choosing the VAR. This model was normalised setting the household debt to be equal to one to produce the VECM. To correct for the disequilibrium, the VECM was further adjusted and the product was stabilised coefficients explaining the short relationships between the variables. A modified Granger causality testing procedure by Toda & Yamamoto (1995) was used to check if the variables are weakly exogenous in the system or not. This test was also used to assess causal relationships between the ten variables. Having identified this causal relationship, the analysis was further continued on the final models by checking the parameters for stability. Furthermore, the models were used for forecasting household debts in SA and the USA. Forecasts of household debts were generated with the generalised impulse response functions.

Part B of this chapter used the discriminant and factor analysis methods. Firstly the stepwise method of discriminant analysis was used to help identify the variables in order of their importance and the communalities from factor method was used in determining the contribution of common variables to a discriminant model of household debts in SA and the USA. A model showing linear combinations of the nine household debt determinants for each country was constructed and was later standardised to give better and sensible results. Diagnostic testing was performed on the final model to confirm its validity and that it can be used for further analysis. Hoteling's test was computed from this model and was used in comparing SA and the USA.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

"By a small sample, we may judge the whole piece."

Miguel de Cervantes from Don QUIXOTE

5.1 INTRODUCTION

5.2 RESEARCH QUESTIONS AND CONCLUSIONS

5.3 RECOMMENDATIONS

5.4 SUMMARY

5.1 Introduction

In the main, this study investigated the extent of and compared household debt (HHD) in SA and the USA. Multivariate methods were applied to time series data collected for the period 90Q1 to 13Q1. Specifically, we identified the determinants that contributed the most to HDD and considered their relationships. In this context, a combination of econometric analysis and discriminant and factor analyses were employed. The estimation of the econometric and statistical models in this study has provided a better understanding of the relationship between the selected variables as well as made a critical comparison of HDD between the two countries possible. In particular, this study is novel in terms of incorporating multivariate techniques in analysing household debts. For instance, this study uses Fisher's LDA framework to compare SA and the US. Also, to the best of the researcher's knowledge, it is the first attempt to estimate a VECM for the HHD of the USA. Collectively, these novel results can guide policy makers in identifying which variables to target first in an attempt to address the high levels of HHD in both the SA and the US.

5.2 Research questions and conclusions

This section answers the research questions and gives a summary of findings obtained in Chapter 4. Also highlighted is how the study advances the knowledge in this area of research.

RESEARCH QUESTION 1:

What are the main determinants of HHD in SA and the USA in the period 90Q1 to 13Q1?

Conclusion 1:

This study considered the LCH by Ando & Modigliani (1963) and PIH by Friedman (1957) as a basis for selecting the determinants of household debt in SA and the USA. In particular, these theories give direction when identifying variables that influence household consumption and the level of borrowing. Numerous researchers, for instance Debelle (2004), Prinsloo (2002), van der Walt & Prinsloo (1995) and Aron & Muellbauer (2000) have adopted the same theories in their related studies. Dynan & Kohn (2007) also made suggestions about some of the variables affecting household debts.

Nine macroeconomic variables used in determining household debt levels in SA and the USA during the aforementioned period were included in this analysis. Specifically, these determinants of household debt comprise house prices (HP), consumer price (CP), disposable income (INC), interest rates (IR), gross domestic product (GDP), household consumption expenditure (HC), household savings (HS), exchange rates (ER), unemployment rates (UR). According to theories and literature, three of the nine determinants are positively related to HHD and the rest have a negative relationship. These findings played an integral role in estimating the models.

RESEARCH QUESTION 2:

What was the effect of the 2007-2009 financial crisis on HHD in SA and the USA?

Conclusion 2:

In this subsection, the time frame selected was the period before, during and after the 2007-2009 financial crisis. To better understand the levels of SA and the USA household debts, graphic measures were used. These plots are represented in Figures 4.2a and 4.2b for SA and the USA, respectively.

The conclusions made about the effect of the 2007-2009 financial crisis on HHD are based on the results displayed on time series plots. These plots reveal the three levels of household debts in SA and the USA. A few studies used the same approach to explore HHD levels in SA and the USA and others used the charts. These studies also analysed data on household debts collected over different periods and measured on different scales. For instance, household debts studies in SA by Meniago *et al.* (2013) used household data expressed in millions of rands collected from 85Q1–12Q1, while Aron *et al.* (1999) collected yearly data from 1975-2005. Prinsloo's (2002) data were also measured in millions of rands and it covered the period 97Q1 to 01Q4. In the USA context, Gartner (2013) analysed indexed household debts gathered between 1930 and 1939, Brown *et al.* (2013) used yearly data in trillions of dollars collected from 1999 – 2012 *(refer to Chapter 2 for other studies on the subject)*. Despite different measures and time spans used, these studies reported that the crisis had a positive impact on household debts. Not enough information could be gathered in case of the studies that used data set collected before the financial crisis such as the study by Gartner.

This study highlights that HHD-to-disposable income in SA had fluctuated between 52 % and 70 % prior to the crisis. From a USA perspective, the figures were estimated at about 11.9 % and 13.2 %. Despite the differences in the estimates, one can deduce from these statistics that the HHD levels in SA and the USA kept on increasing and these estimates significantly intensified during the financial crisis[4]. SA marked debt figure above the USA between 2007 and 2009. In 2006, SA National Credit Authorities (NCA) implemented credit regulations which were meant to discourage households from borrowing. This exercise was also enforced to help reduce debts in the country during the recession following the crisis. Seemingly, the introduction of the NCA in SA was not effective enough as people still found themselves in a pool of debt during the crisis and few years thereafter.

In the USA, the SMC triggered the 2007-2009 financial crisis. During that time, the determinants associated with HHD were seriously affected as also highlighted by Petersen *et al.* (2012) (*see Figures 4.3 to 4.11*). For instance, Allen & Giovannetti (2010) highlighted that interest rate cut policies adopted by the Federal Reserve Bank (FRB) and other central banks encouraged households to borrow more and buy houses. During this period, savings rates were also lower in the USA; this lead to increased household indebtedness. It was only few years after the crisis when the signs of deleveraging started showing. The findings of this study confirm the report Leaven & Valentia (2012) that indicates that the effects of the financial crisis are still lingering and in most cases they are on-going. This is accompanied by a slow decrease in household debts. As Figure 4.2 reveals, HHD remained high even a few

---

4 See figure 4.2

years after the crisis; confirming the extent to which these countries were affected during that period. This assertion is supported by the dynamics of some of the determinants in Figures 4.3 to 4.11. These plots provide a clue to the origin of the problem with HHD.

The researcher notes that although studies around this subject have been conducted, very few have used trends to explore HHD- to-percentage of disposable income especially in SA where the said data are available. Instead, descriptive measures such as charts and measures of central location were used.

RESEARCH QUESTION 3

Are the chosen determinants accurate and reliable measures of HHD?

Conclusion 3

In this study, the researcher needed to be sure that the selected variables were true measures of HHD and that the sample used is adequate enough to allow us to generalise the conclusions. In achieving this objective, the study used Cronbach's alpha and KMO as explained in Section 3.5.

The researcher notes that the selection of the variables in this study was undergirded by discussions in Chapter 2. Theories and literature dictate that these variables are true measures of HHD. To the researcher'sknowledge, no study has considered using any measure to verify if the variables classified as HDD determinants are true measures. Seemingly, thus far, Cronbach's alpha has only been used as a reliability measure for items in a questionnaire. KMO on the other hand has not been used before where time series data is analysed.

This measure was calculated for SA and the USA with the results being displayed in Table 4.3. The measure was found to be good and concurred with the opinions expressed in Kline (1999) and Cronbach & Shavelson (2004). This implies that the nine variables chosen for data analysis in this study are unidimensional[5], as suggested in Byrne *et al.* (1989), Chin (2000), Blaha (2001) and Diamantopoulos & Siguaw (2006). The sample selected for this study was also found to be adequate according to Kline (1974). This allows us to generalise the conclusions of made from the data and also to rely on the results obtained.

---

5 Means measure the same construct.

The researcher's quest is to obtain optimal results that may be useful in future and also add new knowledge to existing literature. The researcher therefore does not only rely on her intuition or just believe everything suggested by other authors without verifying. To be more definite, the study applies some of the underused measures such as Cronbach's alpha to measure the closeness of the determinants of household debt selected. With Cronbach's alpha, the researcher managed to prove that all the variables identified are true contributors to HHD. Knowledge has been on the information about the degree of reliability of the selected variables. The nine variables selected for this study are guaranteed to be real determinants of HHD in SA and the USA through the Cronbach's alpha. The strength of the uniqueness of these variables is about 85 %, well above the acceptable 70 % coefficient. This also proves that this measure is not only applicable to measuring items on questionnaire data. It can also be used as a measure for time series data. The KMO results proved that the sample selected represent about 83% and 77% for SA and USA respectively. These samples are more than the bench mark of 0.6 as suggested by Keiser (1974).

RESEARCH QUESTION 4

What is the relationship between HHD- to-disposable income and related determinants?

Conclusion 4

This study builds VEC models for the data collected over the period 90Q1 - 13Q1. The VEC model helps in determining the long-run relationships between time series variables. The Johansen multivariate cointegration approach, as discussed in Section 3.7, was adopted. The initial model incorporating all the nine variables was specified as in Equation 3.5.

The conclusion in this subsection is made in reference to the VEC Models for SA and the USA. For the USA, literature is silent about studies that construct VEC models of household debts using macroeconomic variables. Several studies did report on household indebtedness but the focus was more descriptive. Some studies used multiple regression analysis and others constructed vector auto-regression (VAR) models. One of the studies built a VEC Model on the effect of household debts on aggregate performance in the U.S. (*refer to Chapter 2 for description of these studies)*. In relation to South African household debts analyses, few variables were used as determinants of household debts and most of the data used did not cover the three periods. *Please refer to Subsection 1.4.2.5.* In most cases, household debts were estimated using household balance sheets rather than macroeconomic variables.

The Johansen cointegration approach was used to analyse relationships between variables in this study. The VAR models did not conform to economic theory. To obtain optimal results, the models were standardised on household debt as a dependent variable to obtain VECM (1). The signs of the coefficients of these models were found to be in accordance with economic theory but not in all the variables. Though all the variables were found to be moving together in the long-run, the magnitudes of their coefficients were also found to be ambiguous. Consequently, each model was subjected to the error correction mechanism to correct this overestimation and also obtain meaningful results. The study confirms that in the short run, about 13.78 % of disequilibrium is corrected every quarter before the entire system could be reinstated back to equilibrium. This is in accordance with Loganathan *et al.* (2010) study. A 1.54 % of modification may be effected in the USA system of household debts. Since the error correction terms of the ECM are both negative it means that household debt models in these countries is a true cointegrating vector according to Aziakpono (2006).

The ECM results confirmed that in the short-run, increases in household debts in SA can be attributed to household consumption expenditure, gross domestic product and house prices in that order. With reference to these findings, it is assumed that increased consumption expenditure by households is associated with increased household debts. Moreover, high inflationary rates will also lead to high debts. Economic growth may also be one of the leading factors of increased debts as it gives self-assurance to both households and the money lenders and thus debts. These are in accordance with the suggestions by LCH and findings by Philbrick & Gustafson (2010) respectively.

The remaining five determinants may not pose a threat; instead they may relieve households of their debts. Household disposable income was identified as having the most significant negative effect in the long-run, with savings identified as the second most significant negative determinant. These results are in accordance with theories and findings by Debelle (2004; 2004), Subhanij (2007), Ewell (2013), Kearns & Woods (2006), Prinsloo (2002), Kotzé & Smit (2008) except for the consumer prices. With reference to the US VECM, increased household debts in the long-run may be ascribed to increase in household consumption expenditure and personal disposable income. However, the remaining seven determinants report the opposite with gross domestic product marking the lowest effect. The results of this analysis from the USA perspective revealed some controversies in economic theory in terms of three determinants out of nine implying that these variables move together in the long-run but in an opposite direction. One other similarity highlighted in this analysis is that of the seven determinants negatively associated with household debts in the USA, five of them apply to SA. These countries household indebtedness is found to be positively and predominantly elevated by household consumption (*for the results on this subsection, please refer to Section 4.7*).

One imperative innovation about this analysis is the construction of the VEC model for household debts using macroeconomic variables specifically for the USA. This model highlighted two main determinants associated with increased debts and seven having adverse contribution in the long-run if interventions are not imposed. Researchers have learnt that household debts in the USA is generally attributed to similar determinants in the same direction as those in SA except for only three such as household income, gross domestic product and house prices. Another invention is the inclusion of quarterly unemployment data in the VEC model of household debts in SA with other determinants. No study has attempted to build a model of any kind using monthly or quarterly unemployment rates since only yearly data is accessible for SA databases.

RESEARCH QUESTION 5

Do SA and USA households differ significantly in relation to their debt?

Conclusion 5

In this section, the study used the multivariate technique, viz., the Toda-Yamamoto Granger causality, to investigate the causal relationships between the variables.

Granger (1988) suggests that if there is cointegration between the variables there should also be evidence of causality between them. Toda-Yamamoto causality has an advantage of handling multiple variables and is also effective in evaluating the hypothesis of weak exogeneity and parameter stability. This approach has also been found to be quick and effective in that one does not have to do repetition of analyses, instead all the variables are analysed concurrently. The researcher is interested in determining how many causal relationships are between the ten variables analysed.

In this section the researcher highlights the effectiveness of the VECM in capturing the causal effects of HHD variables in the long-run. As Laina (2011) reported, a unidirectional relationship was established running from household debts to GDP but not vice versa. The study analysed only two variables. Sinha & Sinha (2007) in their study analysed the three types of savings with GDP as a dependent variable. No evidence of causation could be established between GDP and household savings.

It is noted that of the few studies which used Toda-Yamamoto to analyse causal relationships, none used multiple variables let alone household debt determinants. This multivariate technique is being wrongly or unnecessarily used. Sinha and Sinha misused this technique; these authors should rather have applied techniques that only allow two variables such as the Engle-Granger causality test. It is evident that little attention has been paid to the causal analysis of the relationship between household debts using Toda-Yamamoto approach.

This subsection makes the conclusion on the basis of the results obtained using the said approach. With this test the researcher managed to ascertain the direction of relationships between the ten variables. In the context of the South African data, seven variables revealed bidirectional relationships with only gross domestic products, exchange rates and unemployment rates showing unidirectional causality *(refer to column 6 of Table 11 in the results chapter)*. Eight variables revealed a causal relationship in the American perspective except for gross domestic product and exchange rates. These findings are similar for SA with the sole exception of unemployment rates. One interesting finding concerns GDP and exchange rates which are found to be weakly exogenous and revealed no causal relationship with other variables in both the SA and the USA systems of household debts in SA and the USA. This may have been overlooked in the previous studies. These findings concur with Laina (2011) but contradict findings by Sinha & Sinha (2007) and this is one of the similarities gathered from the analysis. Another advanced knowledge is that the variables found not to be weakly exogenous satisfied stability conditions according to Pesaran & Pesaran (1997).

RESEARCH QUESTION 6

What are the forecasts of household debts in SA and the USA?

Conclusion 6

This subsection uses that GIRF[6] to help determine how the variable household debts will respond to shocks as a result of changes in related determinants. The results of this analysis are summarised as Figures 4.15 to 4.22.

The conclusions drawn in this paragraph are in relation to the results obtained through literature review and methods described in Chapter 3. Notably, very few studies made an attempt to do the estimation of household debts using related determinants, for example Meniago *et al.* (2013) also employed the GIRF. Some studies considered GIRF to study the response of consumer debt to changes in some of the variables used in this study (refer to Kim 2011).

The results obtained in Chapter 4 are briefly summarised here and conclusions drawn. The GIRF results revealed that household debts are subject to different types of shocks. Depending on the changes in each determinant, household debt is in one way or the other destined to be affected due to changes in related determinants. Figures 4.15 to 4.22 may be used as reference to these analyses. The results obtained were found to be in accordance with economic theory for most of the variables. However, household debts have proven not to be significantly responding to shocks in relation to changes in the USA consumer prices and SA's household income, interest rates and household consumption expenditure. Shocks will be negative in household debts in relation to house prices but positive as a result of changes in household savings and unemployment rates in both countries. This may be advantageous to SA and the USA in that, should house prices be lowered, debt rates may also be reduced. Forecasts obtained prove that in 2018Q1 there was a 95 % chance that SA households would be 86.34 % in debt if no interventions were imposed or if current strategies were not enhanced. About 10.52 % of debt rates are expected in the USA during that time. Please be reminded that a proxy has been used for HHD-to-disposable income for the USA, hence the stagy difference. Nevertheless, households for both countries will be more in debt in five years' time.

6 Refers to the generalized response function analysis

This have advanced the researcher's knowledge of the fact that household debts will still be high in the next five years in these countries according to the forecasts obtained. The researcher have also learnt that though literature have suggested that the variables used in the analysis are factors of household debts, changes in some of them may not be as malicious as changes in others. The results on the GIRF in some variables proved a warning by Leaven & Valentia (2012) about the effects of 2007-2009 financial crisis which are reported to be persistent and may continue doing even in the long-run.

RESEARCH QUESTION 7

Which of the determinants of household debt discriminate between the two countries?

Conclusion 7

The researcher in this subsection adopted Fisher's LDA approach to build a model that best describes linear combination of household debts in SA and the USA. The variables that best discriminate between these countries were chosen with the forward stepwise selection method and this choice was confirmed with the communalities from FA[7]. The results of this analysis are summarised on Section 4.12.

The researcher refers to literature as discussed in Chapter 2 to draw conclusions about the results obtained in Chapter 4. No prior studies have attempted to build a discriminant model of household debts with Fisher's LDA and FA. This is a completely new contribution.

This paragraph explains a better understanding of household debts with reference to the results obtained using LDA. Prior to the analysis, the assumption of homogeneity of variance-covariance matrices was checked and the Box's M test was not statistically significant at the 0.001 significance level (according to Brace *et al.* 2009). The scatterplot matrix revealed different kinds of relationships between variables and the determinant of the correlation matrices also proved multicollinearity as not posing a threat. A stepwise election method was used to select the variables to include in building a discriminant model. Of the nine variables, the procedure confirmed that only five[8] variables were significant predictors of household debts in SA and the USA. Nevertheless, others were included in the

---

7 Factor Analysis

8 The five variables selected are consumer prices, household income, gross domestic product, household consumption expenditure and household savings

discriminant model of household debts so as to capture their individual as well as their collective effects. This model was later used to come up with the tests used for comparing SA and the USA. The communality estimates confirmed that household consumption expenditure and interest rates may be the least preferred predictors to be included in the construction of LDA model for these countries. The eigenvalues proved that the variables associated with the first unstandardized discriminant function contribute almost 78% to household debts. This function was standardised to obtain the final Fisher's LDA model of household debts in SA and the USA. The initial model of household debts in SA and the USA revealed that six variables positively affect household debts identifying household disposable income and the gross domestic product as most significant determinants. However, these findings clash with suggestions by theory. The standardised function identified the entire variables especially the gross domestic product, savings and house prices as having the most significant contribution on household debts, with consumption expenditure and interest rates having the least but still significant contribution. All the parameters of standardised Fisher's LDA are & Pesaran's (1997) and Sims (1980) and Harris (1995) on weak exogeneity of the variables.

In this subsection, the researcher highlights how the use of Fisher's LDA in building the discriminant model has advanced the researcher's knowledge. The study has advanced knowledge about the most crucial determinants of household debts in SA and the USA collectively. It is now clear that household debts in SA and the USA are more associated with all the determinants save for consumption expenditure and interest rates. Knowledge has also been advanced in terms of the fact that the consumer prices, household income and household savings significantly increase household debts in SA and the USA. Gross domestic products also increase debts in SA but decrease it in the USA. Knowledge has also been enriched about the magnitude of contribution each determinant has in household debt. It is understood that household savings has less impact on household debts in SA than the USA and that household income contributes significantly to the problem.

RESEARCH QUESTION 8

Do SA and USA households differ significantly in relation to their debt?

Conclusion 8

One of the objectives of this study has been to use the Mahalanobis distance measure and Hoteling's test to assess the difference between SA and the USA. The results are summarised in Section 4.12.4.

This subsection makes conclusions based on the procedures discussed in section 3.11 and 3.12. This is the first study to attempt comparing SA and the USA on the basis of household debt determinants.

The Mahalanobis distance as calculated between SA and the USA was found to be different from zero confirming that these countries are somehow different but share some characteristics as highlighted in the previous sections.

Knowledge has been advanced in the sense that SA and the USA are different with reference to some of the household debt determinants. Advanced knowledge has also been gained that though all the household debts determinants have collective discriminatory power on these countries, four of them may not necessarily be as malicious.

5.3 Recommendations

This section formulates the recommendations for future studies on the basis of the findings of this study.

Recommendation 1:

Macroeconomic indicators were not the only variables identified as determinants of household debts. Other variables such as cross sectional variables especially the demographics of households were also acknowledged. Household balance sheets have also been used as determinants in several studies. This study therefore, based on these findings, recommend that these variables be incorporated in the analyses of household debts. This may help better determine which of these three groups of variables most inflate household indebtedness. The study also suggests the use of proxies for other macroeconomic variables that could not be included in the analysis due to their unavailability. Proxies of other factors not included in the analysis may be used in further studies of the issue under discussion. Based on the evidence gathered from using Cronbach's alpha and KMO as measures of reliability and sampling adequacy, this study recommends their applications not only where likert-scaled data is concerned. These measures may also be effective on time series analysis.

Recommendation 2:

Due to the effect of financial crisis on rising household debts in SA and the USA and the warnings by Leaven & Valentia (2012), this study recommends even harsher measures to be enacted that would benefit not only the consumers but also credit suppliers. Financial suppliers should be closely monitored and perhaps harsh penalties be decreed to those that do not adhere to the law. Reserve banks should in future think of other ways of assisting consumers other than reducing interest rates as this decision was reported to be one of the root causes of financial crises.

Recommendation 3:

For studies using several independent variables, it is important to be sure that these variables measure the same construct. Though the variables chosen for this study were suggested by theory and literature, the researcher needed a guarantee of their correlation. The results obtained provided assurance of these variables' reliability as measures of household debts. The researcher therefore suggest that researchers in future use this measure in their studies to check the extent of unity between the independent variables, especially those measured on metric scales and also when several variables are used in the analysis.

Recommendation 4:

This study accurately modelled household debts for SA and the USA using the VECM approach. Further analysis on the issue may be done using other techniques such as the CVAR. The results of these analyses may be compared to those of the current study. This may help policy-makers to advocate policies that may be adopted when dealing with household indebtedness. Furthermore, it is suggested that other variables affecting household debts such as those discussed in the previous section be included in the VEC models of SA and the USA. This may help in accurately and fairly marking the most important similarities and differences in these countries. The findings may add to existing literature on this subject and may also help data analysts to better make a distinction between the different multivariate methods used for modelling economic data. They will also understand that the analysis of household indebtedness is not only limited to univariate or bivariate measures. This may give guidance to authorities in the USA about the most hostile HHD determinants in the country.

Recommendation 5:

Estimations of future trends in household debts due to changes in related determinants were analysed. Based on the findings, more attention has to be paid to those variables with significant effects as discussed in Section 4.10. The GIRF results of house prices on household debts highlighted the fact that the effect of decreases in house prices in future on household net worth in SA and the USA will be good news. This therefore suggests that the governments of these countries should work together with housing agencies to try and come up with policy that will help in reducing house prices without perhaps affecting interest rates. The 95 % assurance that in the next five years household debts will still be high in these countries is a good warning. This calls for even more effective interventions to be put in place to avoid yet another crisis. Effective measures that may help in reducing personal debts should also be taken.

Recommendation 6:

The results obtained from the analysis in this subsection reveal a few discrepancies in relation to literature. Some of the signs of parameters are not in accordance with economic theories. Using the Econometric and Statistical Methods, all the variables were found to be good predictors of household debts for SA and the USA. The analysis also revealed that all the variables are stable and not weakly exogenous. Based on the discrepancies highlighted, this study recommends that future research be embarked upon where other multivariate methods such as Cluster Analysis, Logistic Regression and Artificial Neural networks will be used. The results obtained from these methods may be compared with those of the current study and may be used as evidence for verifying the results of the VECM and LDA.

5.4 Summary

This study applied multivariate econometric and statistical methods to analyse economic data collected from 1990Q1 to 2013Q1. The aims were mainly to assess relationships between household debts and related determinants, investigate the effect of 2007-2009 financial crisis on household debts in SA and the USA and also to compare these countries based on these household debt determinants. In general, households were deeply affected during the financial crisis with the five year forecasts giving extra warning. The VEC model of household debts in SA confirmed that in the long-run eight out of nine determinants will move together in the same direction except for the CP whilst the USA model revealed seven cointegrating relationships. House prices, income and GDP will move in the reverse direction. In general, the results confirmed that HDD in SA and the USA are similar with reference to most of the determinants and different in just a few.

CHAPTER 6

APPENDICES

Appendix A: VAR (1) MODEL

RSA

Model Parameter Estimates						
Equation	**Parameter**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variable**
HHD	**CONST1**	-1.30312	0.99847	-1.31	0.1955	1
	AR1_1_1	0.86272	0.03817	22.60	0.0001	HHD(t-1)
	AR1_1_2	0.02162	0.03059	0.71	0.4818	HP(t-1)
	AR1_1_3	0.05080	0.14393	0.35	0.7251	CP(t-1)
	AR1_1_4	0.49749	0.13520	3.68	0.0004	INC(t-1)
	AR1_1_5	-0.05997	0.02560	-2.34	0.0216	IR(t-1)
	AR1_1_6	-0.59461	0.15618	-3.81	0.0003	GDP(t-1)
	AR1_1_7	0.23653	0.16024	1.48	0.1438	HC(t-1)
	AR1_1_8	0.06741	0.03185	2.12	0.0374	HS(t-1)
	AR1_1_9	-0.00933	0.01238	-0.75	0.4533	ER(t-1)
	AR1_1_10	-0.09631	0.02207	-4.36	0.0001	UR(t-1)
HP	**CONST2**	2.19369	2.34598	0.94	0.3525	1
	AR1_2_1	-0.00256	0.08969	-0.03	0.9773	HHD(t-1)
	AR1_2_2	0.81055	0.07188	11.28	0.0001	HP(t-1)
	AR1_2_3	-0.55955	0.33816	-1.65	0.1019	CP(t-1)
	AR1_2_4	0.05662	0.31767	0.18	0.8590	INC(t-1)
	AR1_2_5	-0.09764	0.06015	-1.62	0.1084	IR(t-1)
	AR1_2_6	0.55464	0.36697	1.51	0.1346	GDP(t-1)
	AR1_2_7	-0.39272	0.37650	-1.04	0.3000	HC(t-1)
	AR1_2_8	-0.02873	0.07483	-0.38	0.7020	HS(t-1)
	AR1_2_9	0.01239	0.02910	0.43	0.6713	ER(t-1)
	AR1_2_10	-0.01838	0.05185	-0.35	0.7239	UR(t-1)
CP	**CONST3**	0.63695	0.47230	1.35	0.1812	1
	AR1_3_1	0.05534	0.01806	3.06	0.0030	HHD(t-1)
	AR1_3_2	-0.03553	0.01447	-2.46	0.0162	HP(t-1)
	AR1_3_3	0.66150	0.06808	9.72	0.0001	CP(t-1)
	AR1_3_4	0.05301	0.06395	0.83	0.4096	INC(t-1)
	AR1_3_5	0.03082	0.01211	2.54	0.0128	IR(t-1)
	AR1_3_6	0.22938	0.07388	3.10	0.0026	GDP(t-1)
	AR1_3_7	-0.21419	0.07580	-2.83	0.0059	HC(t-1)
	AR1_3_8	-0.01507	0.01506	-1.00	0.3201	HS(t-1)
	AR1_3_9	0.02507	0.00586	4.28	0.0001	ER(t-1)
	AR1_3_10	0.01785	0.01044	1.71	0.0910	UR(t-1)
INC	**CONST4**	0.35465	0.81510	0.44	0.6646	1
	AR1_4_1	-0.02846	0.03116	-0.91	0.3639	HHD(t-1)
	AR1_4_2	0.01990	0.02498	0.80	0.4279	HP(t-1)
	AR1_4_3	0.04226	0.11749	0.36	0.7200	CP(t-1)

	AR1_4_4	0.48319	0.11037	4.38	0.0001	INC(t-1)
	AR1_4_5	-0.00318	0.02090	-0.15	0.8794	IR(t-1)
	AR1_4_6	0.43765	0.12750	3.43	0.0009	GDP(t-1)
	AR1_4_7	0.02974	0.13082	0.23	0.8207	HC(t-1)
	AR1_4_8	-0.06853	0.02600	-2.64	0.0101	HS(t-1)
	AR1_4_9	-0.00209	0.01011	-0.21	0.8371	ER(t-1)
	AR1_4_10	0.00038	0.01801	0.02	0.9834	UR(t-1)

Type of Model	VAR(1) USA
Estimation Method	Least Squares Estimation

Model Parameter Estimates						
Equation	**Parameter**	**Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variable**
HHD	**CONST1**	-0.39623	1.35245	-0.29	0.7703	1
	AR1_1_1	0.90151	0.03110	28.99	0.0001	HHD(t-1)
	AR1_1_2	0.07142	0.03712	1.92	0.0579	HP(t-1)
	AR1_1_3	-0.01576	0.07506	-0.21	0.8343	CP(t-1)
	AR1_1_4	-0.02644	0.16742	-0.16	0.8749	INC(t-1)
	AR1_1_5	0.03365	0.01559	2.16	0.0339	IR(t-1)
	AR1_1_6	0.06450	0.14297	0.45	0.6531	GDP(t-1)
	AR1_1_7	-0.01809	0.08544	-0.21	0.8329	HC(t-1)
	AR1_1_8	0.02242	0.01084	2.07	0.0418	HS(t-1)
	AR1_1_9	-0.04979	0.01824	-2.73	0.0078	ER(t-1)
	AR1_1_10	-0.05314	0.01152	-4.61	0.0001	UR(t-1)
HP	**CONST2**	-2.00037	3.24488	-0.62	0.5393	1
	AR1_2_1	0.08161	0.07462	1.09	0.2774	HHD(t-1)
	AR1_2_2	0.68188	0.08905	7.66	0.0001	HP(t-1)
	AR1_2_3	-0.30213	0.18009	-1.68	0.0973	CP(t-1)
	AR1_2_4	0.76489	0.40169	1.90	0.0604	INC(t-1)
	AR1_2_5	-0.12647	0.03741	-3.38	0.0011	IR(t-1)
	AR1_2_6	-0.32413	0.34302	-0.94	0.3475	GDP(t-1)
	AR1_2_7	0.14667	0.20499	0.72	0.4763	HC(t-1)
	AR1_2_8	-0.07181	0.02601	-2.76	0.0071	HS(t-1)
	AR1_2_9	0.05984	0.04376	1.37	0.1753	ER(t-1)
	AR1_2_10	0.00379	0.02764	0.14	0.8912	UR(t-1)
CP	**CONST3**	4.12986	1.93377	2.14	0.0357	1
	AR1_3_1	-0.05356	0.04447	-1.20	0.2320	HHD(t-1)
	AR1_3_2	-0.00854	0.05307	-0.16	0.8726	HP(t-1)
	AR1_3_3	0.11027	0.10732	1.03	0.3072	CP(t-1)
	AR1_3_4	-0.01753	0.23938	-0.07	0.9418	INC(t-1)
	AR1_3_5	0.01159	0.02230	0.52	0.6046	IR(t-1)
	AR1_3_6	-0.45177	0.20442	-2.21	0.0299	GDP(t-1)
	AR1_3_7	0.70991	0.12216	5.81	0.0001	HC(t-1)
	AR1_3_8	0.00900	0.01550	0.58	0.5630	HS(t-1)
	AR1_3_9	0.03094	0.02608	1.19	0.2390	ER(t-1)
	AR1_3_10	0.00673	0.01647	0.41	0.6839	UR(t-1)

Appendix B: GIRF

RSA

Simple Impulse Response of Transfer Function									
Lag	**HP**	**CP**	**INC**	**IR**	**GDP**	**HC**	**HS**	**ER**	**UR**
0	-0.0000	0.00029	-0.00001	0.11639	0.00002	-0.00001	-0.27495	-0.33437	-0.18210
1	0.0000	-0.02920	-0.0000	-0.21305	0.00000	0.00003	-0.20253	-0.20270	-0.19059
2	0.0000	-0.02298	-0.0000	-0.16767	0.00000	0.00002	-0.15939	-0.15952	-0.15000
3	**0.0000**	**-0.01808**	**-0.0000**	**-0.13196**	**0.00000**	**0.00002**	**-0.12544**	**-0.12555**	**-0.11805**
4	0.0000	-0.01423	-0.0000	-0.10385	0.00000	0.00001	-0.09872	-0.09881	-0.09290

Accumulated Impulse Response of Transfer Function									
Lag	**HP**	**CP**	**INC**	**IR**	**GDP**	**HC**	**HS**	**ER**	**UR**
0	-0.000	0.001	-0.00008	0.11639	0.00002	-0.00001	-0.27495	-0.33437	-0.18210
1	0.000	-0.028	-0.00008	-0.09666	0.00002	0.00002	-0.47747	-0.53707	-0.37269
2	0.000	-0.051	-0.00008	-0.26433	0.00002	0.00005	-0.63687	-0.69659	-0.52269
3	**0.000**	**-0.069**	**-0.00008**	**-0.39629**	**0.00002**	**0.00006**	**-0.76231**	**-0.82214**	**-0.64074**
4	0.000	-0.084	-0.00008	-0.50015	0.00003	0.00008	-0.86103	-0.92094	-0.73364

USA

Simple Impulse Response of Transfer Function									
Lag	**HP**	**CP**	**INC**	**IR**	**GDP**	**HC**	**HS**	**ER**	**UR**
0	0.000	-0.0040	0.0000	-0.0695	-0.3057	-0.4502	-0.1352	0.14343	0.1364
1	0.000	0.0004	0.0001	0.0110	-0.1375	-0.0494	-0.0829	-0.0912	-0.0423
2	0.000	0.0003	0.0001	0.0096	-0.1198	-0.0431	-0.0722	-0.0795	-0.0368
3	0.000	0.0003	0.000	0.0084	-0.1044	-0.0375	-0.0629	-0.0692	-0.0321
4	0.000	0.0002	0.000	0.0073	-0.0910	-0.0327	-0.0548	-0.0603	-0.0280

Accumulated Impulse Response of Transfer Function									
Lag	**HP**	**CP**	**INC**	**IR**	**GDP**	**HC**	**HS**	**ER**	**UR**
0	0.0000	-0.0040	0.0000	-0.0695	-0.3057	-0.4502	-0.1352	0.1434	0.1364
1	0.0000	-0.0035	0.0001	-0.0585	-0.4432	-0.4996	-0.2182	0.0522	0.0941
2	0.0000	-0.0032	0.0002	-0.0488	-0.5631	-0.5428	-0.2904	-0.0273	0.0572
3	0.0000	-0.0029	0.0003	-0.0404	-0.6676	-0.5803	-0.3534	-0.0965	0.0251
4	0.0000	-0.0026	0.0004	-0.0331	-0.7586	-0.6131	-0.4083	-0.1569	-0.0029

Appendix C: Forecasts of HHD

RSA

Forecasts						
Variable	**Obs**	**Time**	**Forecast**	**Standard Error**	**95% Confidence Limits**	
HHD	**94**	**2013:2**	75.09557	1.07179	72.99490	77.19624
	95	**2013:3**	74.92598	1.47118	72.04252	77.80945
	96	**2013:4**	74.71373	1.79479	71.19600	78.23146
	97	**2014:1**	74.44441	2.11424	70.30058	78.58823
	98	**2014:2**	74.13156	2.45007	69.32951	78.93362
	99	**2014:3**	73.79219	2.80446	68.29554	79.28884
	100	**2014:4**	73.44182	3.17225	67.22432	79.65933
	101	**2015:1**	73.09340	3.54578	66.14380	80.04300
	102	**2015:2**	72.75706	3.91714	65.07961	80.43452
	103	**2015:3**	72.44028	4.27930	64.05301	80.82755
	104	**2015:4**	72.14808	4.62654	63.08023	81.21593
	105	**2016:1**	71.88335	4.95461	62.17249	81.59421
	106	**2016:2**	71.64719	5.26065	61.33652	81.95787
	107	**2016:3**	71.43925	5.54303	60.57512	82.30338
	108	**2016:4**	71.25803	5.80117	59.88795	82.62810
	109	**2017:1**	71.10120	6.03530	59.27224	82.93017
	110	**2017:2**	70.96586	6.24627	58.72340	83.20832
	111	**2017:3**	70.84877	6.43536	58.23570	83.46184
	112	**2017:4**	70.74654	6.60411	57.80271	83.69036
	113	**2018:1**	70.65579	6.75424	57.41773	83.89385

USA

Forecasts						
Variable	**Obs**	**Time**	**Forecast**	**Standard Error**	**95% Confidence Limits**	
HHD	**94**	**2013:2**	10.45897	0.14107	10.18247	10.73547
	95	**2013:3**	10.40021	0.20497	9.99848	10.80195
	96	**2013:4**	10.32668	0.25550	9.82591	10.82746
	97	**2014:1**	10.24087	0.29883	9.65518	10.82656
	98	**2014:2**	10.14513	0.33756	9.48353	10.80673
	99	**2014:3**	10.04138	0.37320	9.30992	10.77283
	100	**2014:4**	9.93113	0.40676	9.13389	10.72837
	101	**2015:1**	9.81556	0.43897	8.95519	10.67592
	102	**2015:2**	9.69556	0.47035	8.77370	10.61743
	103	**2015:3**	9.57186	0.50127	8.58939	10.55434
	104	**2015:4**	9.44502	0.53200	8.40231	10.48772
	105	**2016:1**	9.31548	0.56271	8.21260	10.41837
	106	**2016:2**	9.18363	0.59349	8.02042	10.34685
	107	**2016:3**	9.04978	0.62438	7.82601	10.27355
	108	**2016:4**	8.91420	0.65540	7.62964	10.19876
	109	**2017:1**	8.77713	0.68651	7.43160	10.12265
	110	**2017:2**	8.63877	0.71766	7.23219	10.04535
	111	**2017:3**	8.49934	0.74879	7.03173	9.96694
	112	**2017:4**	8.35899	0.77984	6.83053	9.88746
	113	**2018:1**	8.21790	0.81074	6.62888	9.80692

Appendix D Stepwise Discriminant Analysis Results

Forward Selection: Step 3

Statistics for Entry, DF = 1, 181				
Variable	**Partial R-Square**	**F Value**	**Pr > F**	**Tolerance**
HP	0.0561	10.76	0.0012	0.0338
INC	0.4518	149.19	<.0001	0.0026
IR	0.0001	0.01	0.9120	0.0642
GDP	0.4747	163.59	<.0001	0.0003
HS	0.0002	0.03	0.8603	0.0671
ER	0.0093	1.69	0.1948	0.0563
UR	0.0002	0.03	0.8685	0.0484

Multivariate Statistics					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
Wilks' Lambda	0.471681	67.58	3	181	<.0001
Pillai's Trace	0.528319	67.58	3	181	<.0001
Average Squared Canonical Correlation	0.528319				

Forward Selection: Step 4

Statistics for Entry, DF = 1, 180				
Variable	**Partial R-Square**	**F Value**	**Pr > F**	**Tolerance**
HP	0.0281	5.20	0.0238	0.0003
INC	0.1621	34.83	<.0001	0.0002
IR	0.0000	0.00	0.9720	0.0003
HS	0.0153	2.79	0.0966	0.0003
ER	0.0007	0.13	0.7155	0.0003
UR	0.0001	0.01	0.9111	0.0003

Multivariate Statistics					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
Wilks' Lambda	0.395205	68.87	4	180	<.0001
Pillai's Trace	0.604795	68.87	4	180	<.0001
Average Squared Canonical Correlation	0.604795				

Forward Selection: Step 5

Statistics for Entry, DF = 1, 179				
Variable	**Partial R-Square**	**F Value**	**Pr > F**	**Tolerance**
HP	0.0073	1.32	0.2528	0.0002
IR	0.0041	0.74	0.3901	0.0002
HS	0.0838	16.36	<.0001	0.0002
ER	0.0002	0.03	0.8624	0.0002
UR	0.0001	0.01	0.9170	0.0002

Multivariate Statistics					
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**	**Pr > F**
Wilks' Lambda	0.362103	63.07	5	179	<.0001
Pillai's Trace	0.637897	63.07	5	179	<.0001
Average Squared Canonical Correlation	0.637897				

Forward Selection: Step 6

Statistics for Entry, DF = 1, 178				
Variable	**Partial R-Square**	**F Value**	**Pr > F**	**Tolerance**
HP	0.0062	1.12	0.2920	0.0002
IR	0.0078	1.40	0.2378	0.0002
ER	0.0004	0.08	0.7845	0.0002
UR	0.0006	0.10	0.7481	0.0002

Appendix E Convergence analysis results

PP Unit Root Test of Household Debt Differential

Null Hypothesis: HHDIFF has a unit root

Exogenous: Constant, Linear Trend

Bandwidth: 6 (Newey-West automatic) using Bartlett kernel

			Adj. t-Stat	Prob.*
Phillips-Perron test statistic			-1.496796	0.8239
Test critical values:	1% level		-4.060874	
	5% level		-3.459397	
	10% level		-3.155786	

*MacKinnon (1996) one-sided p-values.

Phillips-Perron Test Equation	
Dependent Variable: D(HHDIFF)	
Method: Least Squares	
Sample (adjusted): 1990Q2 - 2013Q1	
Included observations: 92 after adjustments	

Variable	Coefficient	Std. Error	t-Statistic	Prob.
HHDIFF(-1)	-0.020271	0.020277	-0.999693	0.3202
C	0.895868	0.818608	1.094380	0.2767
@TREND("1990Q1")	0.008498	0.007622	1.114888	0.2679

R-squared	0.014378	Mean dependent var	0.247826
Adjusted R-squared	-0.007771	S.D. dependent var	1.258246
S.E. of regression	1.263125	Akaike info criterion	3.337120
Sum squared resid	141.9982	Schwarz criterion	3.419352
Log likelihood	-150.5075	Hannan-Quinn crite.	3.370310
F-statistic	0.649137	Durbin-Watson stat	1.618667
Prob(F-statistic)	0.524952		

KPSS Unit Root Test of Household Debt Differential

Null Hypothesis: HHDIFF is stationary
Exogenous: Constant, Linear Trend
Bandwidth: 7 (Newey-West automatic) using Bartlett kernel

				LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic				0.190636
Asymptotic critical values*:		1% level		0.216000
		5% level		0.146000
		10% level		0.119000

*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

KPSS Test Equation
Dependent Variable: HHDIFF
Method: Least Squares
Sample: 1990Q1 - 2013Q1
Included observations: 93

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	38.47072	1.343453	28.63571	0.0000
@TREND("1990Q1")	0.285606	0.025224	11.32269	0.0000

R-squared	0.584860	Mean dependent var	51.60860
Adjusted R-squared	0.580298	S.D. dependent var	10.07991
S.E. of regression	6.530207	Akaike info criterion	6.612025
Sum squared resid	3880.567	Schwarz criterion	6.666489
Log likelihood	-305.4592	Hannan-Quinn critr.	6.634016
F-statistic	128.2032	Durbin-Watson stat	0.037160
Prob(F-statistic)	0.000000		

REFERENCES

AFIFI, A.A., MAY, S., & CLARK, V.A., 2012, *Practical Multivariate analysis*, CRC Press, Chapman & Hall, USA, NY.

AFOLABI, J.A., 2008, 'Analysis of loan repayment among small scale farmers in South Western Nigeria', A Discriminant Analysis approach. *Journal of Social Science*, (17(1):83-88.

AKAIKE, H., 1974, 'A new look at the statistical model identification', *Automatic Control, IEEE Transactions*, 19(6):716- 723.

ALIMI, S.R. & OFONYELU, C.C., 2013, 'Toda-Yamamoto Causality test between money market interest rate and expected inflation', The Fisher hypothesis revisited. *European Scientific Journal,* 9(7).

ALLEN, F. & GIOVANNETTI, G., 2010, 'Fragile countries and the 2008-2009 crisis (EUI-RSCAS Working Papers 2010/13). Retrieved fromhttp://cadmus.eui.eu/dspace/bitstream/1814/13536/3/RSCAS_2010_13.pdf.

ALTMAN, E., 1968, 'Financial ratios, Discriminant Analysis and the prediction of corporate bankruptcy', *Journal of Finance,* 23(4):589-609.

ANDO, A. & MODIGLIANI, F., 1963, 'The life cycle hypothesis of savings', *Journal of American Economic Review,* 53(1).

ANTZOULATOS, A.A., 1996, 'Consumer credit and consumption forecast', *International Journal on Forecasting,* 12:439-453.

ANUNDSEN, K.A. & JANSEN, E.S., 2011, Self-reinforcing effects between housing prices and credit. Evidence from Norway, Discussion Papers 651, Research Department of Statistics Norway.

ARKAAH, I.P. & YAPRAK, A., 1988, 'Identifying target segments for foreign direct investment (FDI) attraction: An application of conjoint methodology', *International Marketing Review*, 5(3).

ARON, J., MUELLBAUER, J. & PRINSLOO J., 1999, '*Estimates of personal sector wealth for South Africa'*,(The centre for the study of African economies, working paper series 101, Retrieved from http://www.bepress.com/csae/paper101.

ARON. J. & MUELLBAUER, J., 2000, 'Personal and corporate saving in South Africa', *World Bank Economic Review*, 14(3):509–544.

ARON. J. & MUELLBAUER, J., 2002. 'Interest Rate Effects on Output: Evidence from a GDP Forecasting Model for South Africa', *IMF Staff Papers*, Palgrave Macmillan, 49 (Special i), pp.185-213.

AWH, R.Y. & WATERS, D., 2012, 'A Discriminant Analysis of economic, demographic and attitudinal characteristics of Bank charge-card holders, a case study', *Journal of Finance*, 29(3):973-998.

AZIAKPONO, M.J., 2006, 'Financial integration amongst the SACU countries: evidence from interest rate pass-through analysis', *Studies in Economics and Econometrics*, 30, No.2, 33-46.

BEN-KAABIA, M. GIL, J.M. & CHEBBI, H.E., 2002. 'The effect of long-run identification on impulse-response functions: An application to the relationship between macroeconomics and agriculture in Tunisia', *Agricultural Economics Review, Greek Association of Agricultural Economists,* 3(2).

BETZ, N.E., 1987, 'Use of Discriminant Analysis in counselling psychology research', *Journal of Counseling Psychology*, 34:393-403.

BLAHA, J., MERYDITH, S.J., WALLBROWN, F.H. & DOWD, E.T., 2001, 'Bringing another perspective to bear on the factor structure of the MMPI-2' *Measurement and Evaluation in counselling and development*, 33, 234-243.

BOONE, L. GIROUARD, N., & WANNER, I., 2001, 'Financial market liberalization, wealth and consumption (periodical style)', OECD Economics Department *Working Paper*, 308.

BOX, G.E.P. & PIERCE, D.A., 1970, 'Distribution of residual correlations in autoregressive-integrated moving average time series models', *Journal of the American Statistical Association*, 65:1509–1526.

BOX, G.E.P., 1954, 'Some Theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality of variance and of correlation between errors in the two-way classification', *Annals of Mathematical Statistics*, 25(3):484-498.

BOX, G.E.P. JENKINS, G.M. & REINSEL, G.C., 2008, *Time Series Analysis: Forecasting and Control,* Wiley Series in Probability and Statistics.

BRACE, N., KEMP, R. & SNELGAR, R., 2009, SPSS for Psychologists. Palgrave Macmillan; 4th Edition, Leeds, UK.

BROOKS, C., 2008, *Introductory econometrics for finance* (2.ed. reprinted. ed.). Cambridge University Press, Cambridge.

BROWN, M., HAUGHWOUT, A., LEE, D. & VAN DER KLAAUW, W., 2013, The financial crisis at the kitchen table: trends in household debt and credit, Current Issues in Economics and Finance, Federal Reserve Bank of New York, issue April.

BUSETTI, F., FORNI, L., HARVEY, A. & VENDITTI, F. 2007, 'Inflation Convergence and Divergence within the European Monetary Union', *International Journal of Central Banking*, 3(2): 95-121.

BYRNE, B. M., SHAVELSON, R. J., & MUTHÉN, B., 1989, 'Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance'. *Psychological Bulletin*, 105(3):456-466.

CALISTRI, A., in progress, 'Industrial Structures in the OECD: New evidence from the STAN database'.

CANADIAN ECONOMIC OUTLOOK, Royal Bank of Canada, June, 2012.

CHEN, G. RASMUSSEN, S. & REILLE, X., 2010, 'Growth and vulnerabilities in microfinance, Focus note, 61, CGAP, Washington, D.C.

CHIN, W.W., 2000, 'Partial Least Square for researchers: an overview and presentation of recent advances using the PLS approach', http://discnt.cba.uh.edu/chin/indx.html

CHOWDHURY, A.R., 1994. 'Stock market interdependencies: Evidence from the Asian NIEs', *Journal of Macroeconomics, Elsevier*, 16(4):629-651.

CHURCHILL, G.A. 1995, *Marketing Research: Methodological Foundations,* Dryden Press.

CLARK, A. E. & T. DANIEL., 2006, 'Forecasting South African house prices', *Investment Analyst Journal,* .64:27-33.

CLOYNE, J., 2013, 'Discretionary tax changes and the macroeconomy: new narrative evidence from the United Kingdom', *American Economic Review*, 103(4):1507-1528.

COSTA, P.T. JR. & McCRAE, R.R., 1992, 'NEO PI-R professional manual, Odessa, FL: Psychological Assessment Resources.

CRAMER, D., 2003., *Advanced quantitative data analysis,* Open University Press, Philadelphia, pp.254.

CRAWFORD, A. & FARUQUI, U., 2011, 'What Explains Trends in Household Debt in Canada?', *Bank of Canada Review*, 2011-2012, 3-15.

CRONBACH, L.J & SHAVELSON, R.J., 2004, 'My current thought on coefficient Alpha and Successor Procedures', *Educational and Psychological Measure,* 64(3):391-418.

CUSSEN, M., O'LEARY, B. & SMITH, D., 2012, 'The impact of financial turmoil on households. A Cross country comparison, *Quarterly Bulletin 2*.

CYNAMON, B.Z. & FAZZARI, S.M., 2008, 'Household Debt in the consumer age: source of growth risk of collapse', *Capitalism and Society*, 3(2).

D'AGOSTINO, R.B., 1971, 'An Omnibus Test of normality for moderate and large size samples', *Biometrika*, 58(2):341-348.

DANIELS, R., 2001, 'Consumer indebtedness among urban South African households: a descriptive overview', *Cape Town Development Policy Research Unit. (DPRU),* working Paper 01/55).

DAVIES, N., TRIGGS, C. & NEWBOLD, P., 1977, 'Significance levels of the Box-Pierce portmanteau statistic in finite samples', *Biometrika*, 64:517-522

DEBELLE, G., 2004, 'Household debt and macroeconomy", *BIS Quarterly Review*, March, 51-64, http://www.bis.org/publ/qtrpdf/r_qt0403e.pdf.

DEMYANYK, Y. & VAN HEMERT, O., 2008, 'Understanding the Subprime Mortgage Crisis', Working Paper, Federal Reserve Bank of St. Louis.

DIAMANTOPOULOS, A. & SIGUAW, J.A., 2006, 'Formative vs. Reflective Indicators in Measure Development: Does the Choice of Indicators Matter?' *British Journal of Management,* 17:263-282.

DICKEY, D.A. & FULLER, W.A. 1981, 'The likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica,* 49:1057-1072.

DICKEY, D.A. & FULLER, W.A., 1979, 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association,* 74:427-431.

DOYLE, P., SAUNDERS, J. & WONG, V., 1992, 'Competition in global markets: A case study of Japanese competition in the British market', *Journal of International Business Studies,* 23(3).

DRITSAKIS, N., 2007, 'Labour productivity, wages and unemployment: an empirical investigation for Greece using causality analysis', *Asian-African Journal of Economics and Econometrics,* 7(1):133-145.

DYNAN, K.E. & KOHN, D.L., 2007, 'Explore the effects of the ageing of the baby boom generation on the evolution of US household debt since 1983.

ELDER, J., 2003, 'An Impulse response function for a vector autoregression with Multivariate GARCH-in-Mean', *Economics Letters,* 79(1):21-26.

ELLIOTT, A.C., & WOODWARD, W.A., 2007, *Statistical analysis quick reference guidebook with SPSS examples,* Sage Publications, London.

ELWELL, C.K., 2013, 'Rebuilding Household Wealth: Implications for Economic Recovery', *Congressional Research Service*, www.crs.gov. pp.1-19.

ENDERS, W., 2005, *Applied Econometric Time Series*, Wiley, USA.

ENGLE, R.F., 1982, 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom nation', *Econometrica*, 50:987-1007.

ENGLE, R. & C. W. J. Granger. 1987. "Cointegration and Error Correction Estimates: Representation, Estimation and Testing," *Econometrica,* LV, 251-276.

FEIGNER, M.D. & SVERDLOV, L., 2002, 'The use of Discriminant Analysis to separate study population by treatment subgroups in a clinical trial with a Pentapeptide Antidepressant', *Journal of Applied Research*, 2(1):50-57.

FERGUSON, T. & GALBRAITH, J.K., 1998, 'The American wage structure, 1920-1947', *Levy Economics Institute*, Working Paper, 249.

FIELD, A., 2009, *Discovering statistics using SPSS*, SAGE publications Ltd. London.

FISHER. R.A., 1936, 'The utilisation of multiple measurements in taxonomic problems', *Annals of Eugenics*, 7(2):179-188.

FOUNTIS, N.G. & DICKEY, D.A., 1989, 'Testing for a unit root nonstationarity in Multivariate Autoregressive time series.' *Annals of Statistics*, 17:419-428.

FRIEDMAN, M., 1957, *A Theory of the Consumption Function.* Princeton University Press, New Jersey.

GALBRAITH, J.K., & CALMON, P., 1990, 'Relative wages and international competitiveness', Austin, Lyndon B. Johnson School of Public Affairs Working paper number 56, January.

GALBRAITH, J.K. & CALMON, P., 1994, 'Industries, trade and wages in M. Bernstein and D. Adler (*eds*)', *Understanding American economic decline*, Cambridge University Press, Cambridge.

GALBRAITH, J.K. & CALMON, P., 1996, 'Wage change and trade performance manufacturing industries, *Cambridge Journal of Economics*, 20:433-450.

GALBRAITH, J.K. & KIM, J., 1998, 'The legacy of the HCI: An empirical Analysis of Korean Industrial Policy', *Journal of Economic Development*,23(1):1-20.

GALBRAITH, J.K. & JIAQING, L., 1999, Cluster and Discriminant Analysis on Time-Series as a Research Tool. UTIP working paper. The University of Texas at Austin.

GARTNER, K., 2013, Household Debt and Economic Recovery Evidence from the U.S. Great Depression. *European Historical Economics Society.* EHES Working Papers in Economic History, No.36.

GEORGE C. & FERNANDEZ, J., 2000, Discriminant Analysis, A Powerful Classification Technique in Data Mining. Statistics and Data Analysis, Paper 247.

GHYMN, K., 1983, 'The Relative Importance of Import Decision Variables', *Journal of the Academy of Marketing Science*, 11(3).

GÓMEZ-ZALDÍVAR, M. & VENTOSA-SANTAULÁRIA, D. 2012, 'Regional Output Convergence in Mexico', *Latin American Journal of Economics,* 49(2): 217-236.

GRANGER, C.W.J., 1988, 'Some Recent developments in a Concepts of Causality', *Journal of Econometrics*, 39(1-2):199-211.

GUJARATI, D.N., 2003, *Basic Econometrics,* McGraw-Hill Inc. New York.

GUJARATI, D.N., 2006, *Essential of Econometrics*, McGraw-Hill Inc., New York.

GUJARATI, D.N., 2010, *Basic Econometrics,* McGraw-Hill Inc., New York.

GUJARATI, D.N., & PORTER, D.C., 2009, *Basic Econometrics*, McGraw-Hill, Australia.

HAIR J.F., BLACK W.C., BABIN B.J. & ANDERSON R.E., 2010, *Multivariate Data Analysis: A Global Perspective*, Prentice-Hall: Upper Saddle River, New Jersey.

HANSEN, B.E., 1992, 'Parameter Instability in Linear Models', *Journal of Policy Modeling*, 14(4):517-533.

HANSEN, J., 2003, 'Financial cycles and bankruptcies in the Nordic Countries', Sveriges Riksbank Working Paper Series, 149.

HARRIS, R.J.D., 1995, 'Using Cointegration analysis in econometric modelling. Harvester Wheatsheaf, Prentice Hall.

HARVEY, A.C. & CARVALHO, V., 2002, 'Models for Converging Economies', Cambridge Working Papers in Economics, No. 216: 1-33. Available at: http://www.econ.cam.ac.uk/dae/repec/cam/pdf/wp0216.pdf.

HARVEY, A.C. & BATES, D., 2003, 'Multivariate Unit Root Tests and Testing for Convergence', *Cambridge Working Papers in Economics,* No. 0301, University of Cambridge.

HASSAN, A.Y. & STRAZICICH, M.C., 2011, 'Global financial crisis in Africa. Is the Impact Permanent or Transitory? Time Series Evidence from North Africa.', *American Economic Review*, 101(3):577-581.

HASSAN, S.M., 2007, 'Parameters estimation of a Discriminant Function for some Higher Institutes Graduates in Egypt', Thebes Higher Institute for Management and Information Technology, Al Haram, Sakkra Rd., Giza, Egypt.

HASSAN. T., 1986, 'Discriminant Analysis of personality characteristics and occupation preferences of Nigerian Adolescents. Nigerian', *Journal of Guidance and Counselling.* 2:56-62.

HATCHER, L., 1994, *A step-by-step approach to using the SAS system for factor analysis and Structural Equation Modelling*, SAS Institute Inc., Cary, NC.

HATZINIKOLAOU, D. & POLASEK, M., 2005, 'The commodity-currency view of the Australian dollar: A multivariate cointegration approach', *Journal of Applied Economics*, Universidad del CEMA, 8(1):81-99.

HEILMANN, K., 2010, 'Stock market linkages-a Cointegration Approach', Dissertation, University of Nottingham.

HOLDEN, K., 1995, 'Vector autoregression modeling and forecasting', *Journal of Forecasting*, 14:159-166.

HOLMES, F., 2010, 'Identification of variables associated with group separation in descriptive Discriminant Analysis. Comparison of methods for interpreting structure coefficients', *Journal of Experimental Education*, 78(1).

HONG KONG MONETARY AUTHORITY, 2002, 'The Nexus of Consumer credit, household debt service and consumption', *Quarterly Bulletin*, November 2002, pp.35-48. http://www.info.gov.hk/hkma/eng/public/qb200211/fa2.pdf

HOTELING, H., 1951, 'A Generalised *T*-Test and measure of Multivariate dispersion', *Proceedings of the second Berkeley Symposium on mathematical statistics and probability*, 1: 23-41.

HURLEY, D.T., 2010, 'A generalized impulse response investigation of U.S. long and short-term interest yields and Asian holdings of U.S. treasuries', *Journal of International and Global Economic Studies*, 3(1):68-86.

HURST, E. & STAFFORD, F., 2004, 'Home Is Where the Equity Is: Mortgage Refinancing and Household Consumption', *Journal of Money, Credit and Banking*, Blackwell Publishing, 36(6):985-1014, December.

INTERNATIONAL MONETARY FUND (IMF) 2012, World *Economic Outlook*, Chapter 3: Dealing with Household Debt, pp.89-124, Washington D.C.

JABA, E., IEMNA, D.V., VIORICĂ, D. & BALAN, C.B., 2007, 'Discriminant Analysis in the study of Romanian Regional Economic Development', *Scientific Annals of the*

Alexander John Cuza University of Iasi: Economic Sciences Series. Alexander John Cuza, University of Iasi publications, Romania. ISSN: 0379-7864, 147-153.

JACOBSEN, D.H. & VATNE, B.H., 2011, 'The impact of house prices on household debt when controlling for home ownership', working paper, financial stability, Norges Bank.

JOHANSEN, S., 1991, 'Estimation and hypothesis testing of cointegrating vectors in Gaussian Vector Autoregressive Models', *Econometrica*, 59:1551-1580.

JOHANSEN, S. & JUSELIUS, K., 1990, 'Maximum Likelihood estimation and inference on cointegration with applications to the demand for money", *Oxford Bulletin of Economics and Statistics,* 52:169-210.

JOHANSEN, S., 1988, 'Statistical analysis of cointegration vectors', *Journal of Economic Dynamics and Control*, 12:231-54.

JOHNSON, W.B., 1979, 'The cross sectional stability of financial ratio analysis', *Journal of Financial and quantitative analysis*, 14(5).

KASISKA, M., 2007, 'A Discriminant Analysis of Factors Associated with the Adoption of Certified Organic Farming by Smallholders' Farmer in KwaZulu-Natal SA', *Conference proceedings*, pp.411-416.

KAYARANNI, D.A., 2003, 'Web-shoppers and non-shoppers: Comparability, Relative Advantage and Demographics', *Business Review,* 15(3):141-152.

KEARNS, A. & WOODS, M., 2006, 'The Concentration in property-related lending: a financial stability perspective', *Financial Stability Report CBFSAI*, pp. 133-144.

KEYNES, J.M., 1936, *The General Theory of Employment, Interest and Money*, Macmillan, London.

KIM, Y.K., 2011, 'The macroeconomic implications of households debt: An empirical analysis, Trinity College Department of Economics, Working Paper No. 11- 03.

KLECKA, W.R., 1980, *Discriminant Analysis. Quantitative Applications in the Social Sciences Series,* 19, Sage Publications, Thousand Oaks, CA.

KLINE, P., 1999, *The Handbook of Psychological testing* (2nd ed.). London: Routledge.

KOJUCHAROV, N., MARTIN, C.F., MARTIN, R.F. and XU, L., 2008, 'The Subprime Mortgage Crisis: Irrational Exuberance or Rational Error?', Working Paper, Federal Reserve Board of Governors, 3.

KOTZé, L. & SMIT, A.V.A., 2008, 'Personal finances: What is the possible impact on entrepreneurial activity in South Africa? Department of Business Management, Faculty of Economic and Management Sciences, University of the Free State. Available at Southern African Business Review, www.myunisa.ac.za.

KWIATKOWSKI, D., PHILLIPS, P.C.B., SCHMIDT, P. AND SHIN, Y., 1992, 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', *Journal of Econometrics* 54:159-78.

LAEVEN, L. & VALENCIA, F., 2012, 'Systemic banking crises database: An update', IMF working paper, 12/163.

LAINA, P., 2011, 'Dynamic Effects of total debt and GDP, A Time-Series Analysis of the United States', Masters' Dissertation, Department of Economics, Alto University.

LeSAGE, J.P., 1990, 'A comparison of the forecasting ability of ECM and VAR Models', *The Review of Economics and Statistics*, MIT Press, 72(4):664-71.

LINDHE, L., 2000, 'Macroeconomic indicators of credit risk in business lending', Economic Review 1, Sveriges Riskbank, pp.68-82.

LJUNG G.M. & BOX G.E.P., 1978, 'On a measure of lack of fit in time series models', *Biometrika*, 65:297-303.

LOGANATHAN, N., SUKEMI, M.N. & SANUSI, N.A., 2010, 'External Debt and Macroeconomics Performance in Malaysia: Sustainable or not?" *Global Economy and Finance Journal*, 3(2):122 -132.

LORGAT, M., 2003, 'Enabling worker finances: The effects of rising costs of living and debt on worker finances. Unpublished paper, South African Sociological Association, Durban.

LUDVIGSON, S., 1999, 'Consumption and Credit: A Model of Time-varying Liquidity Constraints', *The Review of Economics and Statistics*, 81(3):434-447.

MADDALA, G.S., 2001, *Introduction to Econometrics*, Wiley & Sons, Inc., New York.

MADUBEKO, V., 2010, 'The global financial crisis and its impact on the South African Economy, University of Fort Hare, SA.

MAHONEY, E.R., 1979, 'Sex education in the public schools. A Discriminant Analysis of Characteristics of Pro and Anti-individuals', *Journal of Sex Research,* 15(4).

MAKI, D.M., 2000, 'The growth of consumer credit and the household debt service burden', Board of Governors of the Federal Reserve System Finance and Economic Discussion Paper 2000-12, February.

MARK, Z. 2009, Financial Shock. FT Press. ISBN 978-0-13-701663-1.

MASHIGO, M.P., 2006, 'The debt spiral in the poor households in South Africa', *International Indigenous Journal of Entrepreneurship, Advancement, Strategy and Education,* 2(1).

MASON, J. W. & JAYADEV, A., 2012, 'Fisher dynamics in household debt: the case of the United States', 1929-2011.

MATERNA, R.D., 1988, 'An examination of internationalization in the U.S. aircraft industry in the 1980s: A case study of McDonnell Douglas Corporation', unpublished Ph.D. thesis, Georgia State University, McGraw-Hill, Atlanta.

MCCARTHY, J., 1997, 'Debt, delinquencies, and consumer spending', *Current Issues in Economics and Finance*, 3(3), Federal Reserve Bank of New York.

MENG, S., HOANG, T.N., & SIRIWARDANA, M., 2011, 'The determinants of Australia household debt: A Macro Level study', Business, Economics and public policy working papers, 2011-2014.

MENIAGO, C., MUKUDDEM-PETERSEN, J., PETERSEN, M. & MONGALE, I.P., 2013, 'What causes household debt to increase in South Africa?', *British Journal of Economic Modelling*, 1(33):482-492.

MODIGLIANI, F., 1966, 'The life-cycle hypothesis of saving, the demand for wealth and the supply of capital', *Social Research*, 33(2):160-217.

MODIGLIANI, F., 1986, 'Life cycle hypothesis, individual thrift and the wealth of nations', *The American Economic Review*, 76(3):297-313.

MOROKE, N.D. & MAVETERA, N., 2014, 'Exploratory Factor Analysis to investigate inter-relationships amongst serious crime in South Africa', *European Journal of Social Sciences,* 41 (8).

MORRISON, D.F., 1976, *Multivariate Statistical Methods,* McGraw-Hill, Inc. New York.

NACOS-BURDS, K.I., 2010, 'A comparative analysis of demographic and academic characteristics', Graduate theses and dissertations. IOWA State University.

NARAYAN, P.K. & SMYTH, R., 2009, Multivariate Granger causality between electricity consumption, exports and GDP: Evidence from a panel of Middle Eastern countries', *Energy Policy*, 37:229-236.

NAUDÉ, W., 2009, 'The financial crisis of 2008 and the developing countries', UNU-WIDER Discussion paper, No 2009/01, Retrieved from http://wider.unu.edu/dp2009-01.pdf.

NELSON, C.R., & PLOSSER, C.I., 1982, 'Trends and random walks in macroeconomic time series: some evidence and implications', *Journal of Monetary Economics,* 10:139-62.

NIETO, F., 2007, 'The determinants of household credit in Spain', *Working paper* series, No 0716, the Bank of Spain, Madrid. Available at http://ssrn.com/abstract

NIJS, V.R., DEKIMPE, M.G., STEENKAMP, J.E.M., & HANSSENS D.M., 2001, 'The category demand effects of price promotions', *Marketing Science*, 20(1):1-22.

NOBEL, A., 2003, *Time Series Econometric Cointegration and Autoregressive Conditional Heteroscedasticity*. The Royal Swedish Academy of Sciences, Stockholm, Sweden.

NORAT, M.A., 2005, 'Diagnostic checks in Time Series. *Journal of the Royal Statistical Society*, 168(1):256-256.

NWACHUKWU, T.E. & EGWAIKHIDE, F.O., 2007, 'An Error-Correction Model of the determinants of private saving in Nigeria', *A Paper presented at the African Economic Society (AES) Conference*, Cape Town, South Africa.

ÓLAFSSON, T.T. & VIGNISDÓTTIR, K.Á., 2012, 'Households position in the financial crisis in Iceland', Central Bank of Iceland. Working Paper, No. 59.

OSTERWALD-LENUM, M., 1992, 'A Note with Quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistic', *Oxford Bulletin of Economics and Statistics*, 54:461-472.

OZTUNA, D., ELHAN, AH., & TUCCAR, E., 2006, 'Investigation of four different normality tests in terms of type 1 error rate and power under different distributions', *Turkish Journal of Medical Sciences*, 36(3):171-1766.

PESARAN, M. & PESARAN, B., 1997, *Working with Microfit 4.0*. Oxford: Oxford University Press.

PESARAN, M. H. & SHIN, Y., 1998, 'Generalized impulse response analysis in Linear Multivariate Models", *Economics Letters*, 58:17–29.

PETERSEN, M.A., SENOSI, M.C. & MUKUDDEM-PETERSEN, J., 2012, *Subprime Mortgage Models,* Nova Science Publishers, New York, ISBN: 978-1-61761-132-2 (e-book) 2011, ISBN: 978-1-61728-0 (Hardcover).

PHILBRICK, P., & GUSTAFSSON, P., 2010, 'Australian household debt: An empirical investigation into the determinants of the rise in the debt-to-income ratio', http://biblioteket.ehl.lu.se/olle/papers/0003618.pdf

PHILLIPS, P.C.B. & PERRON, P., 1988, 'Testing for a unit root in Time Series Regression, *Biometrika*, 75(2):335-46.

PINCHES, G.E., & MINGO, K.A., 1973, 'A Multivariate Analysis of industrial bond ratings', *Journal of Finance*, 28(1).

PRINSLOO, J.W., 2002, 'Household debt, wealth and saving', South African Reserve Bank Quarterly Bulletin, No. 226, Pretoria.

RAMAKRISHNAM, K., 1986, 'Towards a theory of international banking: An analysis of international banking by U.S. banks', Ph.D. thesis, University of Texas, Dallas.

RAO, B.B., 2006, 'Estimating short and long-run relationships: A guide to applied economists Applied Economics', *Taylor and Francis Journals*, 39(13):1613-1625.

RENCHER, A.C., 2002, *Methods of Multivariate Analysis*. John Wiley & Sons. New York.

RINALDI, L. & SANCHIS, A., 2006, 'Household debt sustainability: What explains household non-performing loans? An empirical analysis', ECB Working Paper no. 570, January.

ROMER, C.D., & ROMER, D.H., 2010, 'The Macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks', *American Economic Review, American Economic Association,* 100(3):763-801.

ROYSTON, P., 1991, 'Estimating departure from normality', *Stat Med*, 10(8):1283-1293.

SCHWARZ, G., 1978, 'Estimating the dimension of a model', *The Annals of Statistics*, 6(2):461-464.

SHERRY, A., HENSON, R.K., & LEWIS, J.G., 2003, 'Evaluating the appropriateness of college aged norms for use with adolescents on the NEO personality inventory-revised assessment, 10:1-8.

SIMS, C.A., 1980, 'Macroeconomics and Reality', *Econometrica*, 48:1-48.

SINAI, T., & SOULELES, N.S., 2005, 'Owner occupied housing as a hedge against rent risk, *Quarterly Journal of Economics*, 120(2):763-789.

SINGH, A., SANDHU, H.S. & KUNDU, S.C., 2010, 'Investor's adoption of internet stock trading: A study', *Journal of Internet Banking and Commerce*, 15(1).

SINHA, D. & SINHA, T., 2007, 'Toda and Yamamoto Causality Tests between Per Capita Saving and Per Capita GDP for India', MPRA Paper 2564, University Library of Munich, Germany.

SIQUEIRA, E., OTUKI, T. & DA COSTA Jr. N., 2012, 'Stock return and fundamental variables: A discriminant analysis approach', *Applied Mathematical Sciences*, 6:113-116, 5719-5733.

SOUTH AFRICAN RESERVE BANK (SARB), 2009, Annual economic report, Pretoria.

SOUTH AFRICAN RESERVE BANK (SARB), 2010, Annual economic report, Pretoria.

SUBHANIJI, T., 2007, 'Some lessons from Securitisation Crisis', *Bangkok Post, Business New,* (13 July, 2007).

TABACHNICK, B.G. & FIDELL, L.S., 2007, *Using multivariate statistics,* Pearson Allyn & Bacon, Upper Saddle River, New Jersey.

TANG, C.F. & SHAHBAZ, M., 2009, 'Multivariate Granger Causality between electricity consumption, economic growth, financial development, population, and foreign trade in Portugal'.

The State of the Nation's Housing, 2008, Harvard University (http:/ / www. jchs. harvard. edu/ sites/ jchs. harvard. edu/files/ son2008. pdf)

THE WALL STREET JOURNAL, February 2013. US Edition.

TODA, H.Y. & YAMAMOTO, 1995, 'Statistical inference in Vector Autoregressions with possibly integrated processes', *Journal of Econometrics*, 66:225-250.

TODA, H.Y. & PHILLIPS, P.C.B., 1994, 'Vector Autoregressions and Causality: a theoretical overview and simulation study, *Econometric Reviews,* 13:259-285.

TREBBI, F. MIAN, A. & SUFI, A., 2012, 'Political constraints in the aftermath of financial crises', Research-based policy analysis and commentary from leading economists.

TROCHIM, W.M.K., 2006, Research Methods Knowledge Base. Webcenter for Social Research Methods.

TSAY, R.S., 2005, *Analysis of Financial Time Series,* Wiley, New Jersey.

TUNKU, S.T.A., O'NEILL C. & ROWAN N.J., 2013, 'The use of Cointegration and Error Correction Modelling to investigate the influence of diabetes and associated medical services expenditure on economic growth in Malaysia', *Journal of Asian Scientific Research;* 3(6):644-653.

TURKAN, S., POLAT, E. & GUNAY, S., 2012, 'Classification of domestics and foreign commercial banks in Turkey based on financial performances using Linear Discriminant Analysis, Logistic Regression and Artificial Neural Network Models', *Journal of European Economy*, 11, special issue. Ternopil National Economic University.

USLAY, C., TEACH, R. & SCHWARTZ, R.G., 2002, 'A cross-cultural analysis of attitudinal differences', *Journal of Research in Marketing and Entrepreneurship*, 4(2):101-118.

VAN der WALT, B. F. & PRINSLOO, J.W., 1995, 'The compilation and importance of household debt in South Africa", Occasional Paper No. 8, Pretoria: South African Reserve Bank.

VENABLE, W.M. & THOMPSON, B., 1998, 'Caretaker psychological factors predicting premature termination of children's counselling', *Journal of Counselling and Development,* 76(3):286-293.

VOGEL, R.H., 1976, 'Uses of managerial perceptions in clustering countries', *Journal of International Business Studies*, 7(1).

VOGELVANG, B., 2005, *Econometrics: Theory and Applications with EViews*. Prentice-Hall, New York.

WALSH, B.D., VACHA-HASSE, T., KAPES, J.T., DRESDEN, J.H., THOMPSON, W.A. & OCHOA-SHARGEY, B., 1996, 'The values scale: differences across grade levels for ethnic minority students', *Educational and psychological measurement*, 56:263-275.

WOODWARD, W.A., GRAY, H.C. & ELLIOT, A.C., 2012, *Applied Time Series Analysis,* CRC Press, Texas.

WUENSCH, K.L., 2001, Two Group Discriminant Function Analysis. http://core.ecu.edu

YILMAZER, T & DEVANEY, S.A., 2005, 'Consumer debt over the life-cycle', *Financial Services Review*, 14(4):285-304.

ZHUANG, H. & TURINETTI, E., 2011, 'Exploring determinants of U.S. Household Debt', *The Journal of Applied Business Research,* 27(6):85-92.