

**STOCHASTIC MODELS FOR
ASSET AND LIABILITY MODELLING
IN SOUTH AFRICA
OR ELSEWHERE**

Alexander James Maitland

A thesis submitted to the Faculty of Science, University of the Witwatersrand,
Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

February, 2010

DECLARATION

I declare that the published and unpublished articles and other elements that make up this thesis are the result of my own work. To the extent usually and reasonably expected, assistance and peer review were received from my supervisor and colleagues, and from editors and referees of the journals in which they were published. The thesis is being submitted for the Degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

February, 2010

ABSTRACT

Research in the area of stochastic models for actuarial use in South Africa is limited to relatively few publications. Until recently, there has been little focus on actuarial stochastic models that describe the empirical stochastic behaviour of South African financial and economic variables. A notable exception is Thomson's (1996) proposed methodology and model. This thesis presents a collection of five papers that were presented at conferences or submitted for peer review in the South African Actuarial Journal between 1996 and 2006. References to subsequent publications in the field are also provided. Such research has implications for medium and long-term financial simulations, capital adequacy, resilience reserving and asset allocation benchmarks as well as for the immunization of short-term interest rate risk, for investment policy determination and the general quantification and management of risk pertaining to those assets and liabilities.

This thesis reviews Thomson's model and methodology from both a statistical and economic perspective, and identifies various problems and limitations in that approach. New stochastic models for actuarial use in South Africa are proposed that improve the asset and liability modelling process and risk quantification. In particular, a new Multiple Markov-Switching (MMS) model framework is presented for modelling South African assets and liabilities, together with an optimal immunization framework for nominal liability cash flows. The MMS model is a descriptive model with structural features and parameter estimates based on historical data. However, it also incorporates theoretical aspects in its design, thereby providing a balance between purely theoretical models and those based only on empirical considerations.

DEDICATION

To my father and mother

For their inspiration and support,
and for the foundations they lovingly provided

ACKNOWLEDGEMENTS

My understanding of stochastic modelling issues and modelling has benefited greatly over the years from interactions with Paul Fatti, Glen Harris, Rob Howie, Jonathan Quadling, Mike Sherris, Rob Thomson, Jaco van der Walt, and David Wilkie, as well as anonymous referees whose suggestions helped make the papers more readable and complete. In particular, I would like to thank Paul Fatti for his encouragement and support in writing this thesis.

CONTENTS

	Page
DECLARATION	II
ABSTRACT	III
DEDICATION	IV
ACKNOWLEDGEMENTS	V
CONTENTS	VI
LIST OF TABLES	IX
LIST OF FIGURES	XI
TABLE OF MODEL VARIABLES	XIII
CHAPTER 1 – INTRODUCTION	
1. OBJECTIVE	1-1
2. THE DEVELOPMENT OF A STOCHASTIC MODEL FOR SOUTH AFRICA	1-3
3. THEORY AND MODELS	1-7
4. THE THESIS	1-12
CHAPTER 2 – A REVIEW OF THOMSON’S STOCHASTIC INVESTMENT MODEL	
FOREWORD TO THE PAPER	
1. INTRODUCTION	2-437
2. FORECAST STATISTICS FOR REAL AND NOMINAL RETURNS	2-438
3. FORECASTING WITH ARIMA PROCESSES	2-443
3.1. THE ARIMA(0,1,0) MODEL	2-446
3.2. NEUTRAL INITIAL CONDITIONS	2-448
4. TRANSFER FUNCTION MODELLING	2-449
4.1. PREWHITENING AND MODEL IDENTIFICATION	2-454
4.2. REAL AND NOMINAL RETURN DISCREPANCIES	2-455
5. RESIDUAL ANALYSIS	2-456
6. SUBSET PARAMETER STABILITY	2-457
7. CONCLUSION	2-458
REFERENCES	2-460
APPENDICES	2-461
ENDNOTE ON THE PAPER	

CHAPTER 3 – NON-STATIONARITY IN SOME SOUTH AFRICAN FINANCIAL AND ECONOMIC SERIES

FOREWORD TO THE PAPER

1. INTRODUCTION	3-632
2. ARMA MODELS	3-632
3. UNIT ROOT TESTS	3-635
4. INTERVENTIONS	3-639
5. UNIT ROOT TESTS REVISITED	3-641
6. GARCH EFFECTS	3-643
7. CONCLUSION	3-645
REFERENCES	3-645
ENDNOTE ON THE PAPER	

CHAPTER 4 – INTERPOLATING THE SOUTH AFRICAN YIELD CURVE USING PRINCIPAL COMPONENTS ANALYSIS: A DESCRIPTIVE APPROACH

FOREWORD TO THE PAPER

1. INTRODUCTION	4-129
2. BACKGROUND	4-131
3. PRINCIPAL COMPONENTS ANALYSIS	4-132
4. RMSE OPTIMALITY OF THE PRINCIPAL COMPONENTS APPROXIMATION	4-137
5. THE “SCREE” TEST	4-137
6. A DISCUSSION OF THE PCA RESULTS	4-138
7. RECONSTRUCTING THE YIELD CURVE	4-140
8. CONCLUSION	4-141
REFERENCES	4-142
APPENDICES	4-144
ENDNOTE ON THE PAPER	

CHAPTER 5 – A MULTIPLE MARKOV SWITCHING MODEL FOR ACTUARIAL USE IN SOUTH AFRICA

FOREWORD TO THE PAPER

1. INTRODUCTION	5-1
2. VARIABLES MODELLED AND TIME INTERVALS	5-3
2.1. VARIABLES MODELLED	5-3
2.2. TIME INTERVALS	5-5

2.3. TRANSFORMATIONS	5-7
3. TIME SERIES MODELLED	5-7
4. UNIT ROOT TESTS	5-12
5. UNIVARIATE MARKOV SWITCHING MODELS	5-14
5.1. INTRODUCTION	5-14
5.2. INFL	5-17
5.3. SINT	5-20
5.4. LINT	5-22
5.5. XSEQ	5-25
6. A MULTIVARIATE MARKOV SWITCHING FRAMEWORK	5-28
7. EMPIRICAL ESTIMATION OF THE MMS MODEL	5-31
8. CONCLUSION	5-35
REFERENCES	5-37
APPENDICES	5-40
ENDNOTE ON THE PAPER	

CHAPTER 6 – AN EMPIRICAL APPROACH TO IMMUNIZATION IN SOUTH AFRICA

FOREWORD TO THE PAPER

1. INTRODUCTION	6-119
2. PRINCIPAL COMPONENTS OF THE JSE-ACTUARIES YIELD CURVE	6-120
3. IMMUNIZATION USING PRINCIPAL COMPONENTS	6-124
4. OPTIMAL IMMUNIZATION	6-125
5. EMPIRICAL ANALYSIS	6-129
6. CONDITIONAL ARBITRAGE	6-135
7. CONCLUSION	6-137
REFERENCES	6-138
ENDNOTE ON THE PAPER	

CHAPTER 7 – CONCLUSION

1. RATIONAL EXPECTATIONS	7-1
2. INCORPORATING THE EFFICIENT MARKET HYPOTHESIS	7-4
2.1. AN OVERVIEW OF THE EFFICIENT MARKET HYPOTHESIS	7-5
2.2. EMH IN MODELS OF DIVIDEND YIELDS AND DIVIDEND GROWTH	7-6
3. CONCLUSION	7-12
COMPLETE LIST OF REFERENCES	7-16

LIST OF TABLES

	Page
Chapter 2 - A Review of Thomson's Stochastic Investment Model	
Table 1: The partial correlation matrix	2-453
Appendix tables:	
Table 1: Expected Nominal Returns	2-464
Table 2: Standard Deviation of Nominal Returns	2-464
Table 3: Skewness of Nominal Returns	2-465
Table 4: Correlation Matrix for Nominal Returns	2-465
Table 5: Expected Real Returns	2-466
Table 6: Standard Deviation of Real Returns	2-466
Table 7: Skewness of Real Returns	2-467
Table 8: Correlation Matrix for Real Returns	2-467
Table 9: Asymptotic Statistics	2-478
Chapter 3 - Non-stationarity in some South African financial and economic series	
Table 1: Parameter estimates and statistics for the equation for testing for unit roots	3-636
Table 2: Parameter estimates for testing for unit roots with a possible structural break	3-642
Chapter 4 - Interpolating the South African yield curve using principal components analysis: a descriptive approach	
Table 1: JSE-Actuaries Yield Curve (1986–1998)	4-133
Table 2: Yield correlations (1986–1998)	4-141
Chapter 5 - A Multiple Markov Switching Model for actuarial use in South Africa	
Table 1: Unit Root Tests	5-12
Table 2: MMS model parameters for the individual variables	5-31
Table 3: Residual correlation matrix for the MMS model	5-32
Table 4: Multiple switching transition matrix, P , for the MMS model	5-32
Table 5: States, ergodic probabilities and durations	5-33
Appendix tables:	
Table A.1.1. INFL MS-model estimates (1960:1-2006:2)	5-40
Table A.1.2. INFL MS-model estimates (1960:1-1998:4)	5-41
Table A.1.3. INFL MS-model estimates (1970:1-2006:2)	5-41
Table A.2.1. SINT MS-model estimates (1960:1-2006:2)	5-42

Table A.2.2. SINT MS-model estimates (1960:1-1998:4)	5-42
Table A.2.3. SINT MS-model estimates (1970:1-2006:2)	5-43
Table A.3.1. LINT MS-model estimates (1960:1-2006:2)	5-43
Table A.3.2. LINT MS-model estimates (1960:1-1998:4)	5-44
Table A.3.3. LINT MS-model estimates (1970:1-2006:2)	5-44
Table A.4.1. XSEQ MS-model estimates (1960:1-2006:2)	5-45
Table A.4.2. XSEQ MS-model estimates (1960:1-1998:4)	5-45
Table A.4.3. XSEQ MS-model estimates (1970:1-2006:2)	5-46

Chapter 6 - An Empirical Approach to Immunization in South Africa

Table 1: Variability explained by the first ten principal components	6-121
Table 2: Optimum immunized portfolios using parallel and principal component shifts	6-130
Table 3: The excess monthly coupon income per R100 nominal short-sale ... (Sep 98)	6-136

LIST OF FIGURES

	Page
Chapter 2 - A Review of Thomson's Stochastic Investment Model	
Figure 1: Mean historical nominal returns up to 1993 and forecast expected nominal returns	2-439
Figure 2: Mean historical real returns up to 1993 and forecast expected real returns	2-442
Figure 3: EQDY, forecast means & 95% confidence intervals - AR(1) model	2-444
Figure 4: PTDY, forecast means & 95% confidence intervals - transfer function model	2-444
Figure 5: EQDY, forecast means & 95% confidence intervals - ARIMA(0,1,0) model	2-446
Figure 6: Forecast distribution of Dividend Yields for lead times of 0-60 years	2-447
Appendix figures:	
Figure 7: EQDG, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	2-468
Figure 8: PTDG, 1977-1993, and forecast means & 95% confidence intervals, 1994-2003	2-468
Figure 9: Z_{PDRY} , 1967-1993, and forecast means & 95% confidence intervals, 1994-2003	2-469
Figure 10: PDRYZ, 1967-1993, and forecast means & 95% confidence intervals, 1994-2003	2-469
Figure 11: PDRY, 1967-1993, and forecast means & 95% confidence intervals, 1994-2003	2-470
Figure 12: PDRG, 1968-1993, and forecast means & 95% confidence intervals, 1994-2003	2-470
Figure 13: Z_{LINT} , 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	2-471
Figure 14: LINTZ, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	2-471
Figure 15: LINT, 1960-1993, and forecast means & 95% confidence intervals, 1994-2003	2-472
Figure 16: Z_{MINT} , 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	2-472
Figure 17: MINTZ, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	2-473
Figure 18: MINT, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	2-473
Figure 19: INFL, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	2-474
Figure 20: Cross-correlations for the ARIMA(0,1,0) Prewhitening Filter	2-479
Figure 21: Cross-correlations for the ARIMA(1,0,0) Prewhitening Filter	2-479
Chapter 3 - Non-stationarity in some South African financial and economic series	
Figure 1: EQDY, 1960-1993, forecast means & 95% confidence intervals, 1994-2003	3-633
Figure 2: EQDG, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	3-633
Figure 3: INFL, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003	3-634
Figure 4a: Recursive estimates of the intercept in equation (3)	3-640
Figure 4b: Recursive estimates of the coefficient of $INFL_{t-1}$ in equation (3)	3-640
Figure 5: EQDG, 1961-1993, one-step-ahead forecasts and residuals for the MA(1) process	3-644

Chapter 4 - Interpolating the South African yield curve using principal components

analysis: a descriptive approach

Figure 1: JSE-Actuaries Yield Curve (1986–1998)	4-133
Figure 2: Coefficients for the first three principal components of yield levels	4-134
Figure 3: Principal-component scores for yields (1986–1998)	4-135
Figure 4: Coefficients for the first three principal components of yield changes	4-136
Figure 5: Principal-component scores for yield changes (1986–1998)	4-136
Figure 6: Scree plot for yields (Y) and changes in yield (DY)	4-138

Chapter 5 - A Multiple Markov Switching Model for actuarial use in South Africa

Figure 1: <i>INFL</i> (“SAadj-X12 INFL”) and <i>INFL-YY</i> (“Y-on-Y INFL”)	5-8
Figure 2: <i>SINT</i>	5-10
Figure 3: <i>LINT</i>	5-11
Figure 4: <i>XSEQ</i>	5-11
Figure 5: Probability that <i>INFL</i> was in the high-inflation regime	5-18
Figure 6: Probability that <i>SINT</i> was in the high-interest regime	5-21
Figure 7: Probability that <i>LINT</i> was in the high-interest regime	5-24
Figure 8: Probability that <i>XSEQ</i> was in the volatile regime	5-27
Figure 9: Unconditional and scaled conditional distributions for <i>XSEQ</i>	5-27

Chapter 6 - An Empirical Approach to Immunization in South Africa

Figure 1: Coefficients for the first three principal components	6-122
Figure 2: Principal component 1 scores for yield changes (Feb 1986–May 2000)	6-122
Figure 3: Principal component 2 scores for yield changes (Feb 1986–May 2000)	6-123
Figure 4: Principal component 3 scores for yield changes (Feb 1986–May 2000)	6-123
Figure 5: JSE-Actuaries Yield Curve	6-129
Figure 6: Cash flows of assets and liabilities using Fisher-Weil and PC-immunization	6-131
Figure 7: Price movements following parallel shifts in the yield curve	6-132
Figure 8: Price movements following PC1 shifts in the yield curve	6-133
Figure 9: Price movements following PC2 shifts in the yield curve	6-134
Figure 10: Price movements following PC3 shifts in the yield curve	6-134

Chapter 7- Conclusion

Figure 1: Weighted 20-year expected price inflation, 1964-1993	7-2
Figure 2: One-year expected price inflation, 1964-1993	7-3

TABLE OF MODEL VARIABLES

Annual variables used in the Thomson (1996) model as well as in Chapters 3 and 7

$INFL_t$	the mean force of inflation in year t
$EQDY_t$	the natural logarithm of the All Share Dividend Yield per cent at time t
$EQDG_t$	the mean force of equity dividend growth in year t
$LINT_t$	the annual force of interest in year t on 20-year bonds as estimated by the JSE-Actuaries 20-year bond yield
$Z_{LINT,t}$	the “inflation-gain” estimate of $LINT_t$ (see Foreword to Chapter 2)
$LINTZ_t$	the error “inflation-gain” estimate at time t over the actual value of $LINT_t$
$MINT_t$	the annual force of interest in year t on money-market instruments as measured by the Ginsberg, Malan & Carsons money-market index
$Z_{MINT,t}$	the “inflation-gain” estimate of $MINT_t$ (see Foreword to Chapter 2)
$MINTZ_t$	the error “inflation-gain” estimate at time t over the actual value of $MINT_t$
$PDRY_t$	the natural logarithm of the yield per cent on direct property at time t
$Z_{PDRY,t}$	the “inflation-gain” estimate of $PDRY_t$ (see Foreword to Chapter 2)
$PDRYZ_t$	the error “inflation-gain” estimate at time t over the actual value of $PDRY_t$
$PDRG_t$	the mean force of direct property rental growth in year t as measured by the Dunlop Heywood investment property index
$PTDY_t$	the natural logarithm of the yield per cent on property unit trusts at time t
$PTDG_t$	the mean force of dividend growth on property unit trusts in year t

Variables used in Chapter 4 and the Multiple Markov Switching model of Chapter 5

$JAYC_{x_t}$	the yields for years $x \in \{00,01,\dots,25\}$ along the JSE-Actuaries Yield Curve
$INFL_t$	the seasonally adjusted, annualised force of inflation in quarter t
$LINT_t$	the annualised force of interest in quarter t on 20-year bonds as estimated by the JSE-Actuaries 20-year bond yield
$SINT_t$	the annualised the force of interest on 3-month Treasury bills at time t
$XSEQ_t$	the annualised, quarterly force of total return on equities in excess of $SINT_t$

CHAPTER 1 – INTRODUCTION

1 OBJECTIVE

As a profession, actuaries are called to make a wide variety of decisions regarding the future. For example, they are required to estimate the contribution rate required to meet accruing liabilities. Similarly, individuals must plan how much to save to meet future expenses, including those in retirement. Historically, such decisions have been based on deterministic assumptions about the future, such as dividend yields and dividend growth rates, a valuation rate of interest, expected returns on investments, as well as salary and price inflation. However, empirical evidence suggests considerable uncertainty in these values and, as Redington (1986) states:

“The briefest glance at the past tells us one fundamental actuarial lesson, that our strength lies in no way at all in the infallibility of our forecasts; it lies in our power to measure and deal with our own fallibility, to face and assess our own uncertainty.”

Since the early 1980s, a large number of papers have appeared in the actuarial literature dealing with the stochastic nature of these and other variables in order to both understand and better manage the associated risks. As Thomson (2004; p11) observes:

“In recent years, the major paradigm shift in actuarial science has arguably been in the quantification of risk: first in the treatment of investment returns as stochastic processes, and then in the adoption of the theories, models and exemplars of financial economics in the courses of study of the actuarial profession.”

This provides a richer understanding of the environment in which actuaries must solve problems and enables improved decision making by incorporating more formally this closer approximation to reality. Stochastic models have been developed to quantify certain aspects of risk, and the financial assumptions underlying these models have

become clearer as they have been compared with the theories and models of financial economics.

Nonetheless, research in the area of stochastic models for actuarial use in South Africa is limited to relatively few publications, and there has been very little focus on models that describe the stochastic behaviour of the South African financial and economic variables required for the financial decision making of institutions and individuals. Such research has implications for medium and long-term financial simulations, capital adequacy, resilience reserving and asset allocation benchmarks as well as for the immunization of short-term interest rate risk, for investment policy determination and the general quantification and management of risk pertaining to those assets and liabilities.

The objective of this thesis is to provide improved models and techniques for actuarial use in the quantification and management of financial risks relating to the assets and liabilities of financial institutions and individuals in South Africa. It is hoped that these models and techniques will be used to address some of the shortcomings evident in existing methods and practice.

Thomson's stochastic investment model for actuarial use in South Africa is reviewed from both a statistical and economic perspective. This review suggests that there are significant problems with that model that make its continued use difficult to justify. Alternative models are investigated and shown to provide a better description of the data. These address claims about the properties of the data, but are not put forward as candidates for projection purposes.

A new stochastic model for actuarial use in South Africa is then developed that improves the asset and liability modelling process and risk quantification. The interest rates modelled are justified with reference to the full set of yield curve risks. The model presented is a descriptive model, with structural features and parameter estimates based on historical data. However, it also incorporates theoretical aspects in its design, thereby

providing a balance between purely theoretical models and those based only on empirical considerations.

A new framework for the immunization of nominal liabilities is also proposed. A more detailed quantification of South African yield curve risks is provided, and this framework is tested against these risks. The theoretical bond portfolios constructed to immunize a set of liability cash flows are shown to improve the management of mismatch risk over those suggested by traditional methods. The optimization suggested is shown to be mathematically optimal given the hypothesized risks, and it is proposed here that this framework may also be suitable for use in other countries.

Finally, a tractable framework for the testing and implementation of the Efficient Market Hypothesis in models of dividend yields and dividend growth rates is presented. This can be used to address some of the theoretical concerns voiced about such models, and it is suggested that this be employed in the development of new models containing these variables.

2 THE DEVELOPMENT OF A STOCHASTIC MODEL FOR SOUTH AFRICA

Using data for the United Kingdom, Wilkie (1984, 1986) developed a stochastic model for inflation, equity dividends and dividend yields, and long-term interest rates. The relationship between these variables is described by a particular ‘cascade structure’ in which:

- inflation is a predictor of long-term interest rates, equity dividends and dividend yields; and,
- equity dividend yields are a predictor of long-term interest rates and equity dividends.

This cascade structure of the Wilkie Model is illustrated in Figure 1 below.

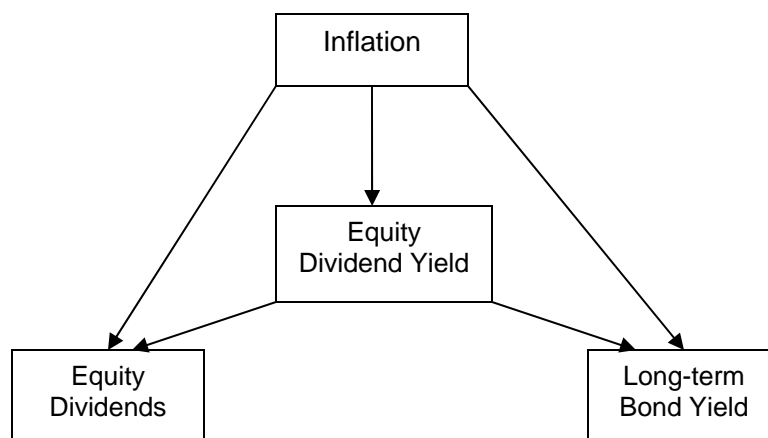


Figure 1. Structure of the Wilkie Model

Wilkie (1995) applied this cascade structure to certain other countries but its cascade structure has been found to be unsuitable as the basis for a general model. Claasen and Huber (1992) and Claasen (1993) found it to be unsuitable for South Africa, although no statistical justification was given in those papers. Carter (1991) found it to be unsuitable for Australia but then assumed a similar cascade structure despite statistical evidence to the contrary. Thomson (1996) found it to be unsuitable for South Africa and proposed a statistical methodology to identify the cascade structure. Thomson (*op. cit.*) then used his proposed methodology to identify and develop a stochastic model based on an analysis of South African data. This model will be referred to as the ‘Thomson Model’.

The Thomson Model is the first published and fully specified descriptive model for actuarial use in South Africa. The variables modelled can be grouped into two sets:

- a ‘core set’ comprising inflation, money market yields, long-term bond yields, equity dividend yields and equity dividend growth rates;
- A ‘property set’ comprising direct property rental yields and rental growth rates, and property trust dividends yields and dividend growth rates.

The core variables do not depend on the property variables, so it is possible to use that part of the model alone, if desired.

In addition, three artificial variables, called ‘modified variables’, were also introduced to assist in the modelling of money market yields, long-term bond yields and direct property rental yields (see Thomson (*ibid.*: 788; ¶4.2.2.7). Each of these represents a quantity equal to the difference between the observed variable and a factor representing a carried forward effect of inflation specific to that variable.

The use of modified variables introduces an unnecessary complexity into the model, and Maitland (1996) illustrates how to reformulate the model so that it is not necessary to construct such modified variables. After factoring out these modified variables, the relationship between the observed variables in the Thomson Model can be represented by the cascade structure shown in Figure 2. Arrows indicate the dependency of variables on one another, where a variable at the start of an arrow is used as a predictor of the variable at the end of that arrow.

It is clear from Figures 1 and 2 that the cascade structure of the variables common to each model is quite different. As Thomson (2004, Chapter 7: ¶2.1) points out:

“... it is not generally appropriate to adopt the structure of a model developed with reference to data from one country and merely reparameterise it for use in another.”

The methodology proposed for model structure identification by Thomson (1996: ¶4) represents the first published attempt to let the data determine the structure of an actuarial investment model. It is put forward as a purely data-driven approach that does not entertain preconceived economic relationships (see Thomson (*ibid.*: ¶4.1.1). Nonetheless, it does place certain restrictions on the structure that may otherwise not be motivated by the data. For example, dependent variables are assumed to be stationary even if the data suggests otherwise. This restriction is essentially motivated by economic considerations.

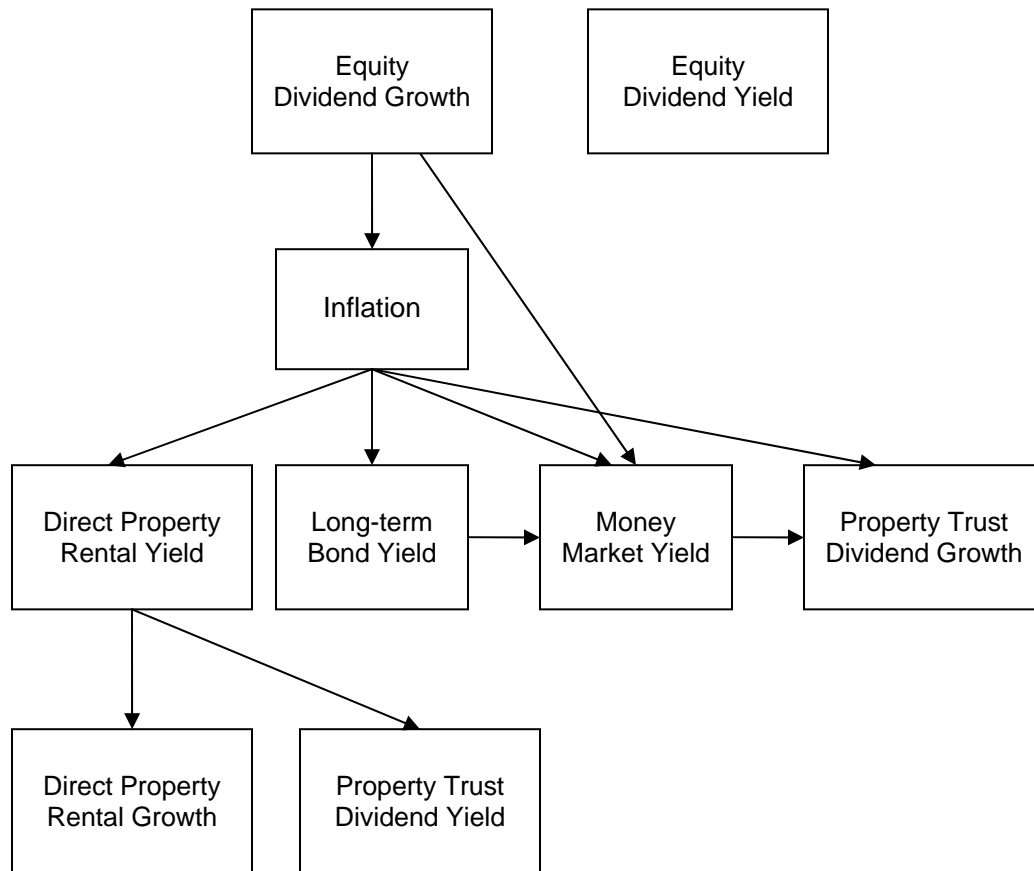


Figure 2. Structure of the Thomson Model

In reflecting on the estimated structure of the Thomson Model, Thomson (2004, Chapter 7: ¶2.1) suggests that while the structure of the model may need to be re-estimated from time to time, “the methodology for such restructuring does not have to be revisited”. He claims that the “strength of the approach lies not in the structure of the model itself, but in the methodology developed for the determination of that structure”.

The Thomson Model and Thomson’s (1996: ¶4) proposed methodology for model structure identification are reviewed by Maitland (1996). In that paper, it is shown that this methodology is essentially an unnecessary and misleading complication of the structure identification procedures proposed by Box & Jenkins (1970) for the

identification of ARIMAX transfer function models. It is also shown that Thomson's methodology can lead to the identification of incorrect structures and that the parameter estimates are inconsistent.

Thomson's (*ibid.*: ¶4) proposed methodology aims to identify a 'cascade structure' and so does not consider the possibility of feedback relationships between variables. This latter restriction reflects the limitations of transfer function models in general, and is a valid concern in the development of stochastic models with variables in which feedback relationships may exist.

One such example is the relationship between long-term interest rates and inflation: since markets are forward-looking, it seems reasonable to expect that long-term interest rates might at times predict future inflation; yet it is also conceivable that the market may not anticipate a change in the level of inflation, and that interest rates might adjust only after the change in the level of inflation has become apparent. Similar considerations apply to the relationships between: short-term interest rates and inflation; short- and long-term interest rates; as well as other variables. This highlights the importance of economic considerations in the identification of such models.

3 THEORY AND MODELS

Thomson (2004: 1-2) discusses the 'traditional approach to science' as a 'theory' comprising a set of logically consistent hypotheses in the form of laws constituting fundamental axioms, describing an unobservable underlying mechanism, from which empirical generalisations may be derived. Such derivations constitute an explanation and are strongly supported by empirical evidence. Hume's problem with this traditional approach to science is that empirical evidence cannot be used to establish the truth of claims about the future. Past observations cannot be the sole basis of generalizations, because they are specific to a historical period and location (Huber *et al.*, 1999: 380).

Thomson (*op. cit.*) then discusses the 'modelling approach to science' in which a 'theory' is defined as "merely a specification or definition of an abstraction or idealization of a

real system”. It is not a law describing the real system itself but rather a representation of the real system that satisfies or instantiates theory. (A system is defined as a set of processes and relationships that exhibit regularity.) The model is a homomorphic system and not isomorphic to the real system, so the truth about the real system and the degree of prediction error are unknown. Thomson (*ibid.*) argues that actuarial science fits more into the modelling approach to science than the traditional approach, since it is “more concerned with the practice of decision-making than with the pursuit of truth for its own sake”.

The modelling approach to science is clearly less ambitious than the traditional approach to science in that it does not attempt to justify models as fundamental laws. As Thomson observes (*ibid.*):

“[The modelling approach] does not claim to achieve more than is possible. The model provides a useful calculating device for the purposes of prediction and control ... and the approach can accommodate approximations.”

However, a fundamental concern common to both approaches to science is the concept of regularity. As Pemberton (1999: 154) argues:

“Actuarial science is concerned with the development of models which approximate the behaviour of reality and have a degree of predictive power ...”

Clearly, the processes and relationships should be ‘regular’ if the model representing them is to be used for projection purposes and decision-making. Hence, we find ourselves in the same circular argument as Hume’s problem elucidates: in both approaches, empirical evidence cannot be used to establish the validity of claims about the future.

While Hume’s arguments can lead to unproductive radical skepticism about everything, Hume was actually advocating a practical skepticism. As Howson (2000) states:

Hume's argument gives us no reason to suppose that relying on our scientific knowledge is in anyway misguided; it does not tell us that we are wrong to do so. It merely says that the attempt to show that there is any sound inductive reasoning to that knowledge from observation alone will fail. But it may well be that we are fully justified, *in terms of its truth or nearness to the truth*, in relying on it.

Howson (*ibid.*) interpreted Hume (1748) to say that:

“sound inductive inference must possess, in addition to whatever observational or experimental data is specified, at least one independent assumption (an inductive assumption) that in effect weights some of the possibilities consistent with that evidence more than others.”

That is, inductive inference must be backed not only by observations, but also by an independent inductive assumption. This suggests that, wherever possible, we should avoid relying on observation alone to justify the structure and our understanding of a system. Combining this idea with Ramsey's (1930) view on probabilistic reasoning, Howson (*ibid.*) concludes that:

“there is a genuine logic of induction which exhibits inductive reasoning as logically quite sound given suitable premisses, but does not justify those premisses.”

Financial economic theory provides a set of normative assumptions based on axioms derived from neo-classical economics (see Huang & Litzenberger, 1998) that can be relevant in formulating a stochastic economic model. These normative assumptions include the principle of no-arbitrage and the basic postulate that agents prefer more wealth to less, which underlies the efficient market hypothesis. As Huber *et al.* (*ibid.*: 387; ¶3.3.7) argue, this postulate:

“is generally not challenged, as agents are unlikely knowingly to allow others to benefit at their expense from arbitrage opportunities. Moreover, agents have a financial incentive to discover and exploit arbitrage opportunities. This suggests that inefficiencies are likely to be ephemeral.”

This led Huber *et al.* (*ibid.*) to suggest that:

“... long-term actuarial economic models should not assume that markets are inefficient. If the inefficiency could be shown to be true, then it would be exploited, and the model would eventually cease to be valid. Models incorporating inefficiencies are inherently unstable, and consequently unsuitable for long-term modelling.”

Nonetheless, it is clear that financial theories are not ‘law-like’ requirements or positive truths that state a model is true of some actual economic system but rather hypothetical statements that depend on the assumptions made. For example, there is nothing to prevent the existence of arbitrage opportunities in the market; indeed, it is the very action of rational agents maximizing utility that operate to exploit arbitrage opportunities and abnormal profits, thereby restoring prices via supply and demand to correctly reflect information available at that time.

What is critical, however, is the intended purpose of the model. If a model is intended to exploit opportunities for arbitrage and abnormal profits that may appear from time to time, it will need to approximate the reality of the system much more closely and will need to incorporate *all* relevant information at the time. This is likely to increase the model’s dimension and complexity enormously, or to make it so specific to the information related to the time period and the events surrounding it that it becomes useless as a more general model. Hence, it is unlikely that stochastic models intended for long-term actuarial modelling will be useful in exploiting abnormal profits unless these opportunities are a function of the basic variables used as part of that long-term stochastic model.

The stochastic models developed by Wilkie (1995) and Thomson (1996) model equity returns as a function of dividend yields and dividend growth rates. Both models are intended for long-term modelling and both assume the existence of abnormal profits on the basis of the level of dividend yields. Wilkie (1986b) and Smith (1996) note that, using expected returns calculated from the Wilkie (1986a) model and investing in the asset class with the highest expected return, it is possible to achieve excess returns of roughly 3% for virtually no extra risk by switching between bonds and equities. Maitland (1996) notes similar features of the Thomson (1996) model as well as the fact that the interpretation placed on the data by the model at that time was that equities were over-valued by 112%. Similar features exist in the models of Carter (1991) and the Finnish Insurance Modelling Group, as described by Ranne (1998). Whether such predictions are reliable or not is the source of much debate and controversy within the actuarial profession.

The market inefficiencies in these models are essentially a function of the mean reversion in dividend yields. While it may well be the case that equity markets are inefficient from time to time, the important fundamental question is: 'Are markets inefficient whenever the dividend yield deviates from its long-term mean as specified in the model?' This is the case whether the long-term mean is estimated from past data or whether it is based on an actuary's subjective judgement. Is it reasonable to assume that the dividend yield is a reliable indicator of market inefficiency, as suggested by these models? Huber *et al.* (*ibid.*) argue that it is not because if the inefficiency could be shown to be true, then it would be exploited.

An efficient market is one where prices respond rapidly to available information and where that information cannot be used to consistently achieve returns in excess of average market returns on a risk-adjusted basis. The Efficient Market Hypothesis (EMH) appears to be a reasonable assumption for long-term stochastic investment models yet none of the published models based on dividend yields and dividend growth rates appear to incorporate it. In each, the EMH is either dismissed without statistical testing and

proper consideration, or ignored completely. Even if the EMH is thought to be inappropriate for the model being used, it is surely useful as a point of departure, even if only to measure the extent and nature of the inefficiency assumed to exist. Yet none of these modellers do this.

It is suggested here that the exclusion of the EMH is often more by accident than design. In building statistical models based on historical data, linear time series modelling is used as a point of departure. In this framework, a relationship between variables is only considered to be significant if those coefficients in the time series equations are found to be significantly different from zero. The null hypothesis in this framework is simply a statistical concept based on the principle of parsimony rather than a hypothesis based on the considerations of financial economic theory.

Unfortunately, when modelling dividend yields and dividend growth rates within the standard time series framework, the EMH does not coincide with the default statistical null hypothesis. In fact, the EMH is a non-linear function of dividend yields and dividend growth rates, and a linear approximation is required before it can be considered within a linear time series analysis framework. Using the approach suggested in Chapter 7, it is recommended that all stochastic models containing dividend yields and dividend growth rates be reconsidered and tested against the EMH.

In Chapter 7, a further caveat is discussed: even if the null hypothesis of an efficient market is rejected within the proposed model structure, this does not imply a rejection of the EMH since what is actually being tested is the joint hypothesis that the market is efficient and that the model describing returns is appropriate.

4 THESIS

Financial economic theory provides a useful reference point for the development and application of stochastic models. However, as Huber *et al.* (1999: ¶3.1.2) discuss:

“...this theory is insufficient to define an actuarial economic model completely, and a number of auxiliary assumptions are required. No single theory is sufficient for the requirements of actuarial models.”

Some of the auxiliary assumptions and empirical features that must be modelled include: error distribution functions, which are critical to the quantification of risk; the levels of inflation and interest rates, as well as their temporal dynamics; monetary policy regimes and other regimes; the nature of yield curve risks; and other empirical features, including time-varying volatilities and the size and nature of the equity risk premium.

As discussed in the previous section, financial economic theories do not represent positive truths or fundamental laws; instead, they are ‘tendency statements’, which suggest how economic variables might behave under certain conditions. There is no guarantee that these conditions will prevail and economic theories are frequently found to have exceptions. (Huber *et al. ibid.*: ¶3.2.8).

Nonetheless, all the models presented in this thesis reference two key financial economic concepts:

- the principle of no-arbitrage; and,
- the Efficient Market Hypothesis.

In addition, other economic theories are referenced where these are thought to be relevant.

The principle of no-arbitrage states that price dynamics cannot allow for risk-free net profits in excess of the corresponding risk-free rate of interest. Arbitrage is the practice of taking advantage of a price discrepancy between two or more instruments in such a way as to make a risk-free profit at no cost. An arbitrage-free model is therefore one that does not permit the existence of arbitrage opportunities within the framework and variables of that model. The framework proposed in Chapter 6 also discusses the possibility of ‘conditional arbitrage’.

The chapters of this thesis are each self-contained, and have all been published in – or submitted to – journals, books or conferences.

Chapter 2 (Maitland, 1996) examines Thomson's (1996: ¶4) proposed methodology for model structure identification and reviews the Thomson Model from both a statistical and economic perspective. It is shown that the proposed methodology is essentially an unnecessary and misleading complication of the structure identification procedures proposed by Box & Jenkins (1970) for the identification of ARIMAX transfer function models. It is also shown that this methodology can lead to the identification of incorrect structures and that the parameter estimates are inconsistent.

In the Thomson Model, forecast statistics of real and nominal returns for each of the five asset classes and inflation are examined and are found to be inconsistent with the history. The model does not reference the Efficient Market Hypothesis and predicts return distributions that suggest the existence of abnormal profits. In addition, the interpretation placed on the data by the model at that time was that equities were over-valued by 112%. There is also considerable evidence of parameter instability, and key parameters defining the structure of the model also appear to be unstable. This review suggests that there are significant problems with the Thomson Model that make its continued use difficult to justify. These and other issues (see also Chapter 7) raise critical questions about the usefulness of the model as a tool for projection purposes.

Chapter 3 (Maitland, 1997) examines alternative descriptive models for inflation, equity dividend yields and dividend growth rates. Thomson (1996) assumed that these variables were covariance-stationary but this assumption was not tested statistically. This article formally tests these variables for unit roots against a number of alternative hypotheses. Specification and stability tests, recursive least squares tests and residual based tests, including tests for GARCH effects, are also carried out on each series. Intervention models are estimated for inflation and equity dividend yields, and a GARCH model is estimated for equity dividend growth rates.

In a recent article on mean reversion, Asher (2007) suggests that Thomson (1996) finds significant evidence of mean reversion of dividend yields in South Africa. The models estimated in Chapter 3 are shown to provide a better description of the data and suggest that there is little evidence for mean reversion in South African dividend yields. The models presented in this chapter are only intended as descriptive models and should not be used for projection purposes.

Chapter 4 (Maitland, 2002) uses Principal Components Analysis to determine the dimension of randomness in the yield curve, and suggests a methodology for estimating the full yield curve using a smaller number of yields from that yield curve. Dimension reduction facilitates the model building process and assists in the development of stochastic models in which other asset categories and economic variables are considered.

Chapter 5 (Maitland, *forthcoming*) introduces a new class of Markov switching models where switches in variables are not perfectly correlated. Maximum likelihood estimates of the parameters are derived and shown to require only the smoothed inferences obtained from a univariate analysis of the variables. The framework is used to estimate a Multiple Markov Switching (MMS) model of South African financial and economic variables, which can be used for various actuarial applications, especially those involving long-term projections. The model presented is a descriptive model, with structural features and parameter estimates based on historical data. However, it also incorporates theoretical aspects in its design, thereby providing a balance between purely theoretical models and those based only on empirical considerations.

The article in Chapter 6 (Maitland, 2001) presents an empirical approach to immunizing South African nominal liabilities in the presence of non-parallel yield-curve shifts. The results are compared with more common immunization strategies and illustrate the value in immunizing against non-parallel shifts in the yield curve. The methodology proposed is shown to be mathematically optimal given the hypothesized risks.

Each of the articles in Chapters 2 to 6 contains a survey of the literature relevant to the models and applications discussed and, where relevant, the sources of data used.

Chapter 7 returns to the financial economic concept of an efficient market referenced in the discussions of earlier chapters, and shows how the efficient market hypothesis can be incorporated into a model of dividend yields and dividend growth rates. Financial economic issues raised by some of the articles in earlier chapters are also discussed and conclusions are drawn regarding the use of such models.

CHAPTER 2

Foreword to the Paper

This chapter presents the paper by Maitland (1996). This paper reviews the Thomson Model from both a statistical and economic perspective, and examines Thomson's (1996: ¶4) proposed methodology for model structure identification. For ease of reference, a brief summary of the variables and equations of the Thomson Model is given below.

The choice of function used to model each variable is well motivated in Thomson (1996): the force of inflation, the force of interest and the force of dividend growth are modelled because forces are additive and the framework in which the variables are modelled is linear. The natural logarithm of the yield per cent is modelled in preference to the yield itself so that the average force of increase in the price index is a linear function of the force of dividend growth and logarithm of the yield. Further, the model cannot produce negative dividend yields. The variables initially modelled are as follows:

- 1) $INFL_t$, the mean force of inflation in year t ,
- 2) $EQDY_t$, the natural logarithm of the All Share Index Dividend Yield per cent at time t ,
- 3) $EQDG_t$, the mean force of equity dividend growth in year t ,
- 4) $LINT_t$, the annual force of interest in year t on 20-year bonds as estimated by the JSE-Actuaries 20-year bond yield,
- 5) $MINT_t$, the annual force of interest in year t on money-market instruments as measured by the Ginsberg, Malan & Carsons money-market index,
- 6) $PDRY_t$, the natural logarithm of the yield per cent on direct property at time t ,
- 7) $PDRG_t$, the mean force of direct property rental growth in year t as measured by the Dunlop Heywood investment property index,
- 8) $PTDY_t$, the natural logarithm of the yield per cent on property unit trusts at time t , and
- 9) $PTDG_t$, the mean force of dividend growth on property unit trusts in year t .

The Thomson Model in difference equation form is defined by the following set of equations. The variables $LINT$, $MINT$ and $PDRY$ are modelled as the sum of a unit gain function and an error process.

Equity model

$$EQDY_t = 0.310 + 0.810EQDY_{t-1} + 0.198\eta_t \quad (F2.1)$$

$$EQDG_t = 0.093 + 0.116\eta_t + 0.076\eta_{t-1} \quad (F2.2)$$

Inflation model

$$INFL_t = 0.008 + 0.899INFL_{t-1} + 0.088EQDG_t - 0.079EQDG_{t-1} + 0.077EQDG_{t-2} - 0.069EQDG_{t-3} + 0.020\eta_t \quad (F2.3)$$

Long-term interest rate model

$$Z_{LINT,t} = 0.006 + 0.126INFL_t + 0.85Z_{LINT,t-1} \quad (F2.4)$$

$$LINTZ_t = 0.010\eta_t + 0.006\eta_{t-1} \quad (F2.5)$$

$$LINT_t = Z_{LINT,t} + LINTZ_t \quad (F2.6)$$

2.4.4 Short-term interest rate model

$$Z_{MINT,t} = 0.004 + 0.141INFL_t + 0.85Z_{MINT,t-1} \quad (F2.7)$$

$$MINTZ_t = 0.008 - 0.091EQDG_t + 0.885LINTZ_t + 0.019\eta_t + 0.010\eta_{t-1} \quad (F2.8)$$

$$MINT_t = Z_{MINT,t} + MINTZ_t \quad (F2.9)$$

2.4.5 Direct property model

$$Z_{PDRY,t} = 0.486 + 0.559INFL_t + 0.74Z_{PDRY,t-1} \quad (F2.10)$$

$$PDRYZ_t = 0.680PDRYZ_t + 0.061\eta_t \quad (F2.11)$$

$$PDRY_t = Z_{PDRY,t} + PDRYZ_t \quad (F2.12)$$

$$PDRG_t = 0.096 + 0.545(PDRY_t - PDRY_{t-1}) + 0.068\eta_t + 0.041\eta_{t-1} \quad (F2.13)$$

Property trust model

$$PTDY_t = 2.547 + 0.598PDRY_t - 0.738PDRY_{t-4} + 0.104\eta_t \quad (F2.14)$$

$$PTDG_t = 0.077 + 1.721MINTZ_t - 0.967MINTZ_{t-2} + 0.053\eta_t \quad (F2.15)$$

A Review of Thomson's Stochastic Investment Model

James Maitland

Abstract

Thomson [1994] describes a stochastic investment model for actuarial use in South Africa. This paper reviews Thomson's model from both a statistical and economic perspective. It discusses the model building process and the limitations of Box-Jenkins models in modelling financial time series.

1. Introduction

Thomson [7] proposed a stochastic investment model for actuarial use in South Africa based on data from 1960 (or later) to 1993. The model was identified and estimated within the paradigm of Box-Jenkins ARIMA models in order to stochastically project the assets and liabilities of pension funds and life offices. A summary of the Thomson model is given in Appendix 1.

Although Lane [4] does review the Thomson model, his paper serves primarily as an introduction to the techniques involved in Monte-Carlo modelling and provides us with source code for the implementation of both the Thomson and Wilkie [8] models. He does not verify whether Thomson's model fits the data nor does he attempt to interpret what the parameters represent, and states "it is not the purpose of [his] paper to explore the statistical analysis of the time series in the model." This paper reviews Thomson's model from both a statistical and economic perspective and discusses the validity of the model as a forecasting tool: the purpose for which it was developed.

In section 2, forecast statistics of real and nominal returns for each of the five asset classes and inflation are examined using 1993 starting

conditions. Some forecast statistics differ considerably from historical figures. In section 3, the reasons for this are discussed in terms of the forecasts for the underlying variables and the nature of their ARIMA processes. This naturally leads to a discussion of the possible choices of ARIMA process and the consequences of such choices. It is here that the limitations of Box-Jenkins ARIMA models become apparent.

Section 4 extends this discussion to the identification and choice of transfer function. The unit gain function is shown to be a type of transfer function and maximum likelihood estimates of the parameters are calculated for consistency. A number of parameters are found to be insignificant.

In analysing the distribution of residuals, Thomson [7] uses residuals obtained from back-forecasts of each series. These are not the residuals used in the maximum likelihood estimation of model parameters and are not appropriate for determining the distribution of the error processes. In section 5, the one-step-ahead residuals and back-forecast residuals are discussed.

In section 6, the model parameters are estimated on subsets of the data. The AR(1) parameter for *INFL* changes considerably and other parameters are found to be unstable. The model structure is also found to be unstable : some of the parameters become insignificant and cross-correlations change. Even though subset testing leads to a much smaller sample size, the changes are considerable.

Section 7 summarises the results of the investigation.

2. Forecast Statistics for Real and Nominal Returns

In modelling equities, property trusts and direct property, two variables, namely the force of dividend growth and the log of dividend yields, have been modelled¹. This enables valuation of the assets at a future time-horizon on a basis consistent with that of the liabilities. Thomson [7] shows how these variables as well as the variables for long-term

¹ The model used corrects some minor typing errors and is slightly different to the model published in [7]. See Appendix 1.

interest bearing securities and money-market instruments can be used to calculate nominal investment returns for each asset class.

The distributions of nominal investment returns (with the exception of the distribution of money-market returns) are intractable, non-linear transformations of the variables modelled and it is necessary to evaluate them numerically. The lower order moments of these distributions for lead times of up to ten years have been estimated using Monte-Carlo simulation and the results are shown in Appendix 2 together with the sample moments of the historical nominal returns for each asset class.

To aid assimilation of these figures, a graph of some conspicuous forecast means from 1994 to 2003 is shown in figure 1. Expected nominal equity returns are particularly poor in the first few years and never rise above 17% while property trusts perform exceptionally well with an expected nominal return of 37% in the first year, decreasing to about 21% after ten years. (Direct property also performs well with an expected nominal return of about 21% throughout.)

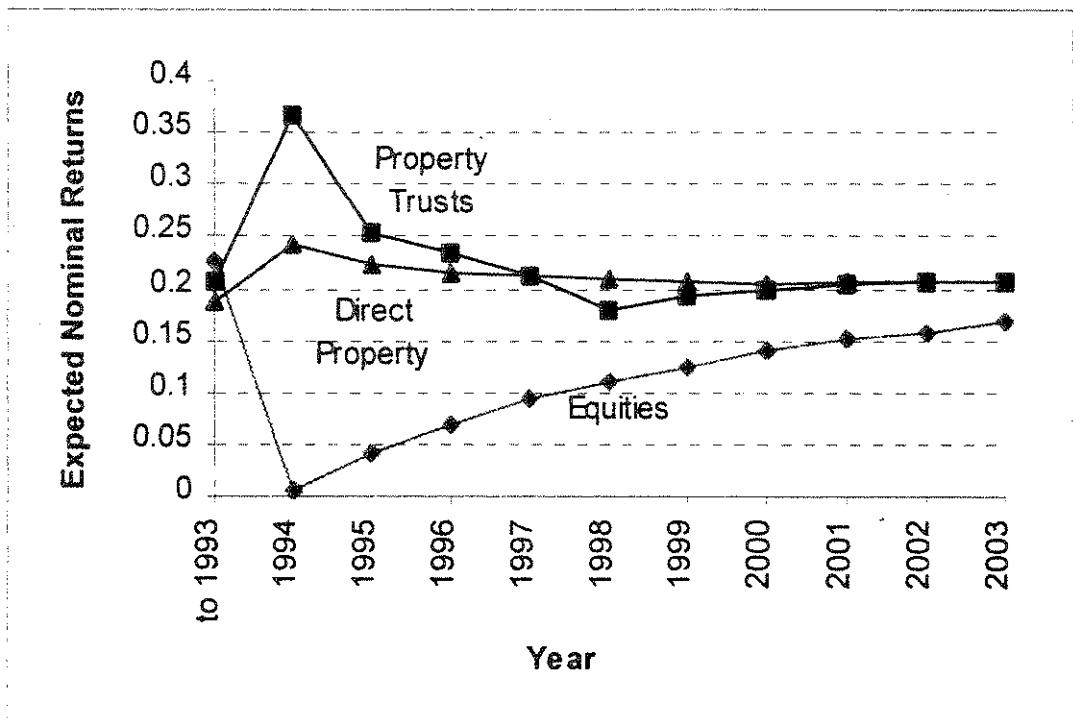


Figure 1. Mean historical nominal returns up to 1993 and forecast expected nominal returns, 1994-2003.

The historical nominal returns for equities and property trusts are 23% and 21% respectively. Why then do equities perform so poorly and property trusts so well? The answer to this will become clearer when the model for the underlying variables is examined in section 3.

For a model with weak stationarity, the historical and asymptotic first and second order moments are equal. However, the second order moments at short lead times are not comparable with the historical second order moments; the variances of the forecast distributions at short lead-times are generally less than their asymptotic variances and the correlations are distorted by the tendency of the forecast series to revert to their historical mean levels, a process known as mean reversion.

At a lead time of ten years, the forecast standard deviation of nominal equity returns is 30% compared with a historical nominal return of 26%. For property trusts, the corresponding figures are 22% historically with a forecast standard deviation of 22%.

the standard deviation of historical nominal equity returns is 26% while the forecast standard deviation at a lead time of ten years is 30%. For property trusts, the corresponding figures are 17% historically with a forecast standard deviation of 22%. The forecast distributions reach a near stable state at a lead time of ten years and one would expect these figures to bear a consistent resemblance to historical figures.

One should be careful comparing the expected returns in each forecast year with the historical mean returns because the historical figures are longitudinal while forecast means at a particular lead-time are cross-sectional. Ten-year forecast means are more useful for comparing with historical means as they represent the near stable-state distribution. Asymptotically, one would expect the forecast means to equal their historic mean level but, with equities, this not the case. The reason for this is the trend in the equity dividend yield series used in forecasting the returns. This is discussed in section 3.

After the oil shock in 1973, inflation rose to a new level of around 12% and the nominal returns of many assets increased in line with it. Historical statistics like the mean, the standard deviation and the skewness of nominal returns are therefore not as meaningful as the corresponding statistics for real returns. The correlations between

nominal returns are meaningful, however, as they measure the degree of “co-movement” between the variables and one would want a model to capture these correlations.

The historical correlation matrix and the forecast correlation matrix at a lead time of ten years are also shown in Appendix 2. The following forecast correlations differ considerably from their historical values:

	Historical Correlation	10-Year Forecast Correlation
$\rho_{EQ:PT}$	0.56	-0.08
$\rho_{EQ:LI}$	0.36	-0.06
$\rho_{EQ:INF}$	0.21	0.12
$\rho_{PT:MI}$	-0.01	0.23
$\rho_{PT:INF}$	0.25	0.00
$\rho_{PD:INF}$	0.20	0.02
$\rho_{LI:MI}$	0.31	-0.02
$\rho_{LI:INF}$	0.27	-0.10

using the following abbreviations:

EQ	equities
PT	property trusts
PD	direct property
LI	long-term interest-bearing securities
MI	money-market instruments
INF	inflation

It should be noted that all of these correlations err on the small side except for the correlation between property trusts and money-market instruments. The reasons for this will be discussed further in section 4.

The forecast expected real returns for each asset category from 1994 to 2003 are illustrated in figure 2 and are listed together with the historical

real returns in Appendix 3. It can be seen that equities again underperform in comparison to their historical real return of 12%. Property trusts and direct property investments on the other hand outperform their historical real returns of 6%.

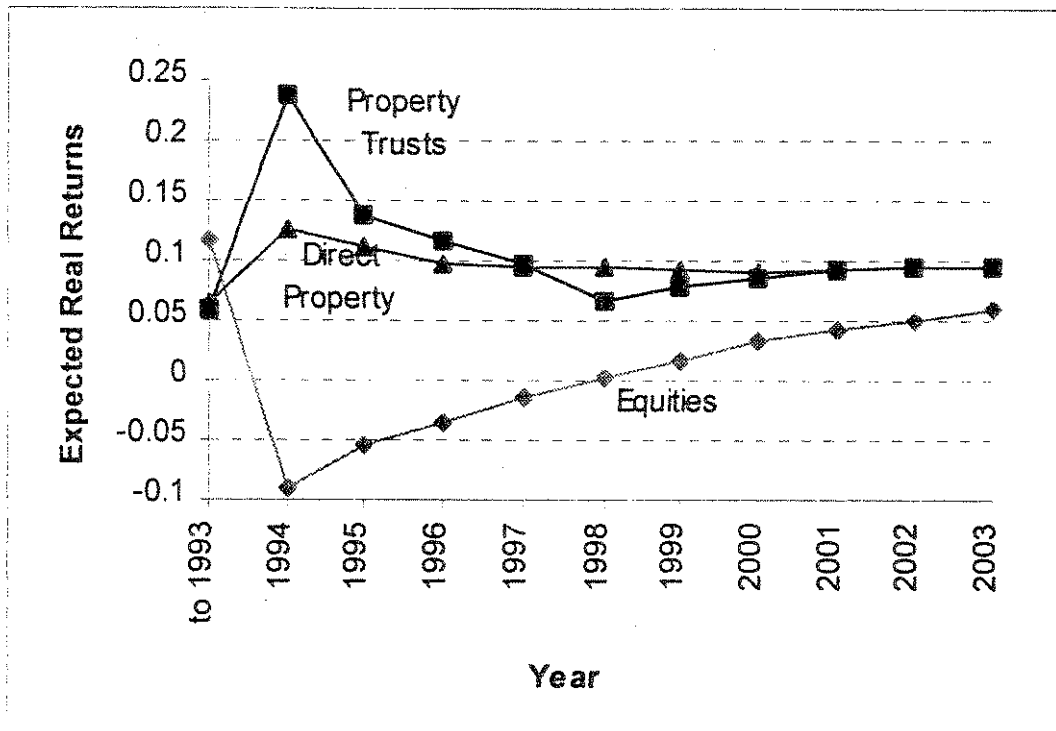


Figure 2. Mean historical real returns up to 1993 and forecast expected real returns, 1994-2003.

The accuracy of the stable-state forecast standard deviations of real returns in the model reflects the accuracy of the forecast first and second order moments of nominal returns and inflation. The correlation between nominal equity returns and inflation was under-estimated and we would expect the standard deviation of real equity returns to be over-estimated. This is indeed the case and the same is true of property trusts.

Some new discrepancies appear in the stable-state forecast correlation matrix for real returns (see Appendix 3) and all those present in the nominal return correlation matrix persist. Corresponding equity correlations continue to be under-estimated while the historical correlation between property trusts and money-market instruments of -0.1 is now over-estimated at 0.36. Direct property surfaces as the new

villain with the correlations between both short-term and long-term interest-bearing securities both being over-estimated. This is discussed further in section 4.

It should be noted that long-term interest-bearing securities have a negative real return in the first forecast year but this stabilises at around 5% after two years. The reason for this is discussed in section 3 and Appendix 5.

In this section, I have concentrated on the aspects of the model used in constructing mean-variance efficient frontiers, namely the means, the variances and the correlation matrix. In general, the higher the expected return, the lower the standard deviation of the return and the lower the correlation with other asset returns, the more likely it is for the asset to form part of some efficient portfolio. When considering real returns, if the nominal return and inflation are positively correlated, the variance of the real return can be expected to be lower than the variance of the nominal return.

3 Forecasting With ARIMA Processes

The results of section 2 highlight a number of discrepancies between historical and forecast return statistics but the most conspicuous must certainly be the excellent expected return on property trusts and the pitiful return on equities. These returns are modelled using the dividend yields $EQDY$ and $PTDY$, and the dividend growth rates $EQDG$ and $PTDG$ where

$$EQDY_t = \ln(\text{All Share Index Dividend Yield}_t \%)$$

$$EQDG_t = \text{Mean force of equity dividend growth in year } t$$

$$PTDY_t = \ln(\text{Property Unit Trust Index Div. Yield}_t \%)$$

$$PTDG_t = \text{Mean force of property trust div. growth in year } t.$$

and

$$EQDY_t = 0.310 + 0.810EQDY_{t-1} + 0.198\eta_t$$

$$PTDY_t = 2.547 + 0.598PTDY_{t-1} - 0.738PTDY_{t-2} + 0.104\eta_t$$

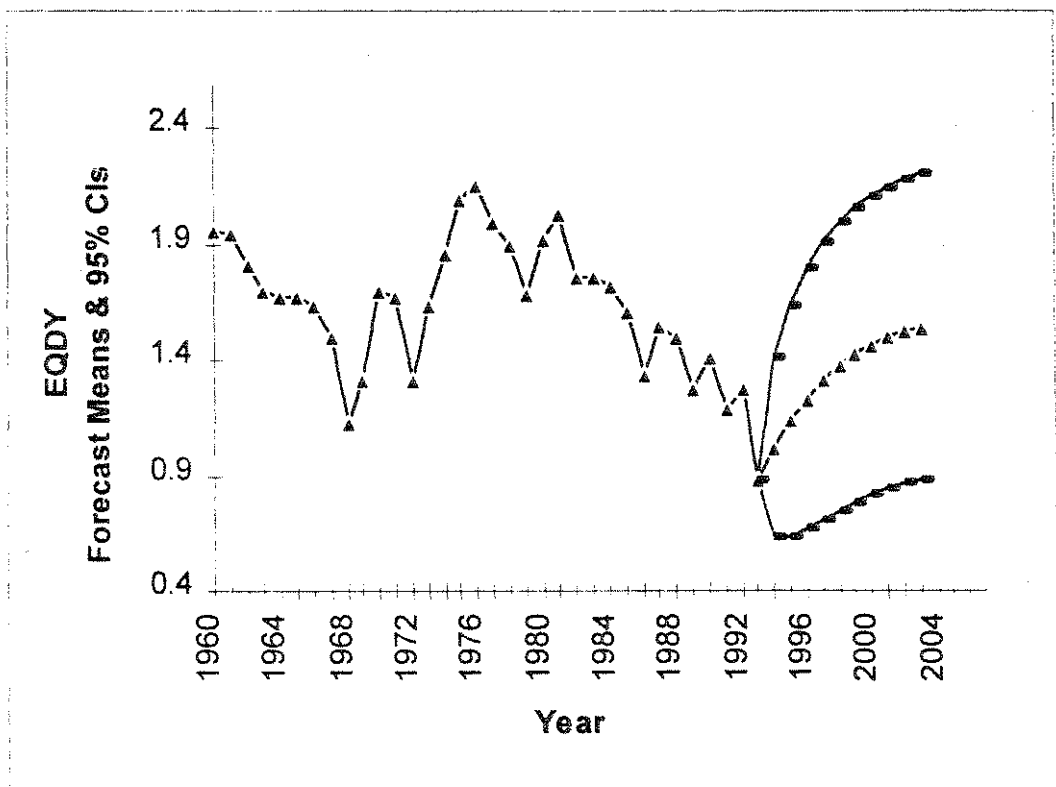


Figure 3. EQDY, 1960-1993, and forecast means & 95% confidence intervals, 1994-2003, using the above AR(1) model.

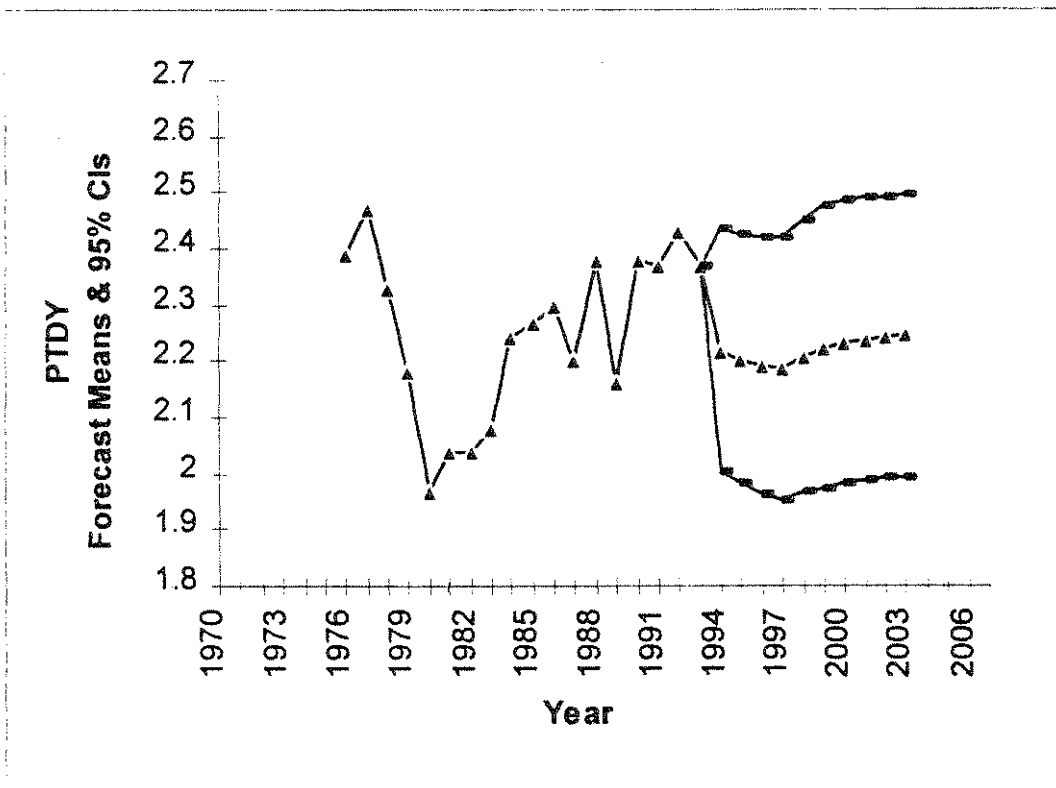


Figure 4. PTDY, 1976-1993, and forecast means & 95% confidence intervals, 1994-2003, using the above transfer function model.

A quick glance at the forecast means for *EQDY* in figure 3, and *PTDY* in figure 4 explains these discrepancies; rising yields imply falling prices and visa versa. So as not to obscure the discussion, the basic problems are considered in terms of the univariate series *EQDY* as it is independent of the rest of the model.

It is the nature of ARIMA models to revert to their asymptotic mean level and, if there has been no differencing, this will be the historic mean level. One needs to consider whether this level is appropriate. If it is not, the constant 0.310 can be adjusted to give a more suitable asymptotic mean level. (Note that the historic mean level is not 0.310, it is $0.310/(1-0.810)$. This is explained in Appendix 5 on asymptotic forecasts.) Forcing a change in the asymptotic mean level however has the effect of increasing the maximum likelihood estimate of $\phi_1=0.810$, further increasing the forecast variance.

A potentially more serious problem is that the process of mean reversion begins immediately with the first forecast. The effect of this if one is modelling a series of yields and the current yield is not at the asymptotic mean level is either a market crash (as with equities) or a market boom (as with property trusts). The interpretation placed on the data by the model is that equities are over-valued by 112%.

With an MA(q) process, the reversion is complete by lead-time q (as with *PTDY*). This explains the expected property trust return of 37% in the first forecast year. The speed of mean reversion in an AR process depends on the distance from the asymptotic mean level and the AR coefficients ϕ_1, ϕ_2, \dots but the expected forecast of the process only actually reaches this level after an infinite period of time. This partly explains the poor expected equity returns which are still increasing at a lead-time of ten years. Forecast means for the remaining model variables are illustrated in Appendix 4.

Another problem is that the series for *EQDY* is trended which accounts for the high value of $\phi_1=0.810$. It is this trend which increases the historic equity return above the asymptotic mean return and accounts for the lower asymptotic forecast return of 17%. The next subsection looks at the alternative ARIMA(0,1,0) model. (The parameter estimates for this model are not given in [7].)

3.1 The ARIMA(0,1,0) Model

Thomson [7] points out that although there is evidence that the series for EQDY should be differenced, the Maturity Guarantees Working Party of the Institute of Actuaries (MGWP) (1980; 146) [5] showed, in the context of dividend yields, that the resulting variance of long-term forecasts cast doubt on the validity of that assumption.

The differenced model for EQDY is

$$EQDY_t = EQDY_{t-1} + \mu_{EQDY} + \eta_t$$

where $\mu_{EQDY} = -0.033$ and $\eta_t \sim \text{i.i.d } N(0 ; 0.131^2)$. Hence,

$$EQDY_{1993+k} | EQDY_{1993} \sim N(0.88 - 0.033 * k; k * 0.131^2).$$

Figure 5 shows the forecast means and 95% confidence intervals for EQDY modelled as an ARIMA(0,1,0) process.

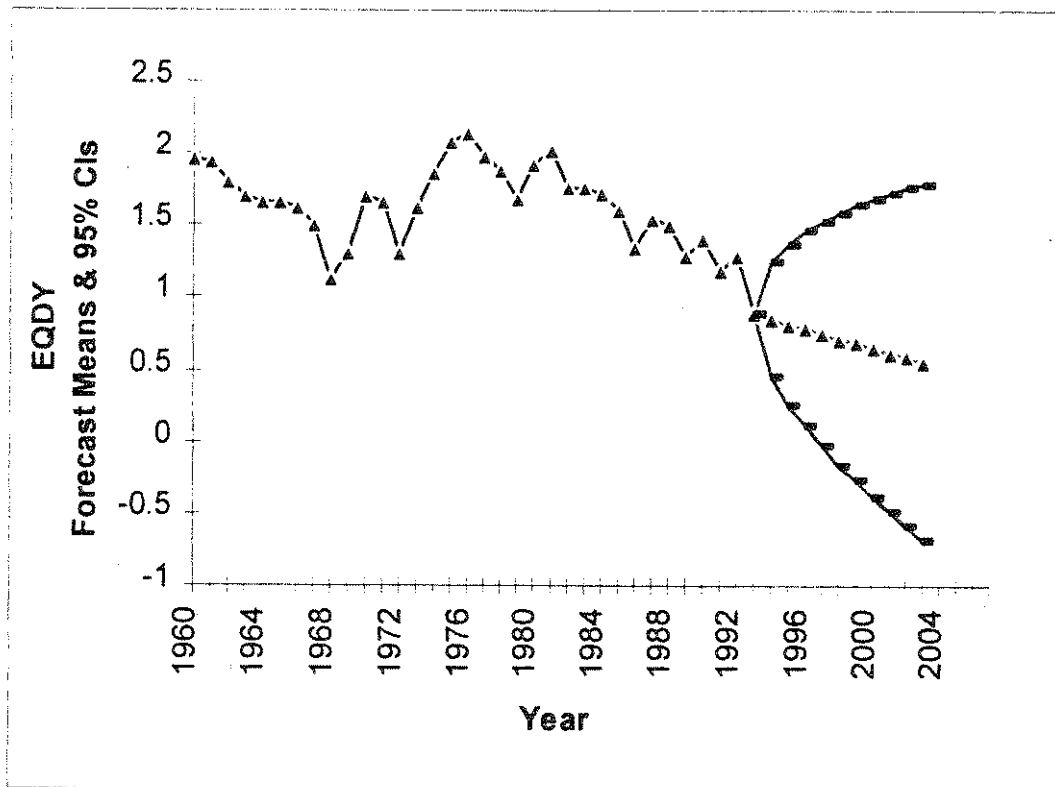


Figure 5. EQDY, 1960-1993, and forecast means & 95% confidence intervals, 1994-2003, using the ARIMA(0,1,0) model.

Since $EQDY_t = \ln(\text{Dividend Yield}_t * 100)$,

$\text{Dividend Yield}_{t+k} \sim \text{Lognormal}(0.88 - \ln 100 - 0.033 * k; 0.131^2 * k)$.

Hence

$$E[\text{Dividend Yield}_{t+k}] = (0.01) \cdot \exp(0.88 - 0.0244195 * k)$$

and

$$\text{Var}[\text{Dividend Yield}_{t+k}] = (0.01)^2 \cdot \exp(1.76 - 0.048839 * k) * \exp(0.017161 * k) - 1$$

Clearly, as $k \rightarrow \infty$, $\text{Var}[\text{Dividend Yield}_{t+k}] \rightarrow 0$. Thus when the series is decreasing and the variable has been log transformed, the resulting variance of long-term forecasts is bounded. Figure 6 shows the forecast distribution of Dividend Yields up to a lead time of 60 years. It should be noted that the variance increases before decreasing and that $E[\text{Dividend Yield}_{t+k}] \rightarrow 0$.

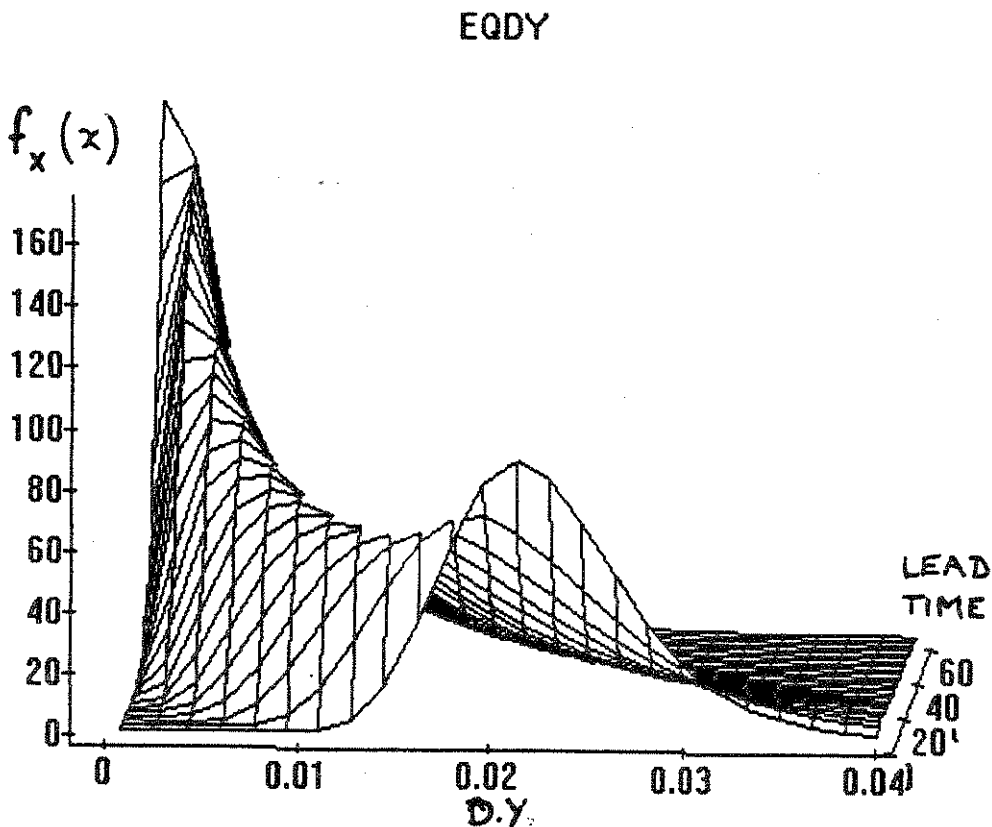


Figure 6. Forecast distribution of Dividend Yields for lead times of 0-60 years.

The size of the return ignoring dividends, however, does not depend on the absolute change in the dividend yield but rather on the change in the dividend yield relative to the current dividend yield. The coefficient of variation (C.V.) is therefore the appropriate measure of variability.

From the expected value and variance of the Dividend Yield given on page 11, the C.V. at lead-time k for the Dividend Yield when EQDY is modelled as an ARIMA(0,1,0) process is

$$[\exp(0.017161*k) - 1]^{0.5},$$

which tends to ∞ as k tends to ∞ . Thus, when a log-transformed series of dividend yields has a decreasing trend, it is the resulting coefficient of variations of long-term forecasts which cast doubt on the validity of a differenced model.

A more important issue must surely be the mean forecasts for dividend yields. Whether dividend yields in the past have been driven down in anticipation of high dividend growth or by local investors forced to invest within South Africa, or whether dividend yields of 2.4% are reasonable is a matter for debate. Nonetheless, standard ARIMA models allow for only two possibilities : the AR(1) model with rising expected forecast yields and, the ARIMA(0,1,0) model with falling expected forecast yields. They do not allow for expected forecast yields at the 1993 level of 2.4%.

3.2 Neutral Initial Conditions

Wilkie [8] defines the neutral initial condition for a variable as the long-run mean of that variable and the neutral initial condition for the error process as zero. The long-run means are the same as the asymptotic means and are discussed in Appendix 5 and evaluated for each variable in the Thomson model in table [9].

In his implementation of the Thomson model, Lane [4] makes use of neutral starting conditions for all variables (except the unit-gain variables) stating that "if current conditions are regarded as atypical then one does not wish to give too much weighting to them." This begs the question, "what is meant by atypical?" 1993 conditions can be compared with either their long-run means or with the forecasts for

1993 based on data from 1992 and before. If we have any faith in the forecasting ability of the model, the latter is preferable. These one-step-ahead residuals are discussed in section 5.

By using neutral initial conditions, we avoid the process of mean reversion which caused the low initial equity returns and high initial property trust returns. As desirable as this may appear, the use of neutral initial conditions undermines the validity of the model as a forecasting tool. If the one-step-ahead residual for a series is too large, it may be argued that the forecast of the current value rather than the current value itself is a more desirable starting value.

Using either neutral initial conditions or a forecast however completely ignores the current value. If we believe that the long-run mean is the "correct" level for the process, then the process of mean reversion from the current condition cannot be avoided and we must accept its consequences. ARIMA models assume that the time series has weak stationarity or that weak stationarity can be induced with a suitable degree of differencing. If we believe that the long-run mean is incorrect and that the series is not trended then a modelling tool other than the standard ARIMA model is required.

4 Transfer Function Modelling

The relationship between trended variables can be obscured by prewhitening. As Thomson [7] observes, "unless there are consistent lags between changes in the one variable and changes in the other, it is quite possible that the residuals of strongly correlated variables will not show any correlation."

In an attempt to overcome this obstacle, Thomson analysed the correlation matrix of the model variables and found significant correlations between the annual force of inflation, *INFL*, and each of *LINT* (the annual force of interest on long-term bonds), *MINT* (the annual forces of interest on money-market instruments) and *PDRY* (the natural logarithm of the yield percent on direct property). For the purposes of illustration, the model for *LINT* is shown below:

$$LINT_t = Z_{LINT,t} + LINTZ_t$$

where

$$Z_{LINT,t} = 0.006 + 0.126INFL_t + 0.85Z_{LINT,t-1},$$

and

$$LINTZ_t = 0.010\eta_t + 0.006\eta_{t-1}.$$

By rearranging the difference equation for $Z_{LINT,t}$,

$$\begin{aligned} Z_{LINT,t} &= 0.04 + 0.84\{(1-0.85)(INFL_t + 0.85INFL_{t-1} + 0.85^2 INFL_{t-2} + \dots)\} \\ &= \alpha[(1-k)/(1-kB)]INFL_t + \beta, \end{aligned}$$

where $\alpha=0.84$, $\beta=0.04$, $k=0.85$ and B is the backshift operator.

The factor $[(1-k)/(1-kB)]$ is the reason $Z_{LINT,t}$ is called a 'unit-gain' function. As Thomson explains, "each year's inflation is expected to ultimately give rise to a corresponding increase in $[Z_{LINT,t}]$, and] the time taken to do so depends on the value of k ." In fact, it ultimately gives rise to a proportionate increase, α , which in this case is 0.84.

The terms $Z_{LINT,t}$ and $LINTZ$ were estimated separately by first estimating the parameters α , β and k for $Z_{LINT,t}$. $LINTZ_t$ is an artificial variable constructed as the difference between $LINT_t$ and the unit-gain estimate of $LINT_t$, namely $Z_{LINT,t}$. It is what Thomson calls a modified variable. The parameter estimates for the unit-gain function were obtained by minimising the sum of $LINTZ_t^2$ while the process for $LINTZ_t$ was identified subsequently using standard prewhitening techniques.

In order to test the validity of the unit-gain functions, Thomson tested the significance of the ratio of the sample variance of the variables initially modelled to the sample variance of the modified variables. An F-distribution was used assuming that this ratio was the ratio of two independent chi-squared variables. The problem with this approach is that the error sum of squares of the modified variables is not a chi-squared random variable because the error process is auto-correlated. It is also questionable whether the numerator and denominator are independent.

Fortunately, we need not concern ourselves with these problems since the unit-gain parameters are easily estimated by maximum likelihood and their significance levels obtained in the usual way. The unit-gain functions and the modified variables can be estimated simultaneously by writing the models in transfer function form.

The equivalent transfer function model for *LINT* is

$$LINT_t = \mu + \omega_0/(1-\delta_1 B)INFL_t + (1-\theta_1 B)\eta_t$$

where $\eta_t \sim \text{i.i.d.}N(0, \sigma_\epsilon^2)$. The parameters α , β and k in the Thomson's unit-gain function correspond to $\omega_0/(1-k)$, μ and δ_1 respectively.

For consistency with the model estimation method used by Thomson, *LINT* has been centred and a zero intercept forced. This avoids maximum likelihood estimation of the parameter μ . The maximum likelihood estimates of the remaining parameters are

Parameter	Estimate	Approx. Std Error	T Ratio
ω_0	0.21692	0.06053	3.58
δ_1	0.70737	0.07654	9.24
θ_1	-0.72271	0.13713	-5.27

The data was centred by subtracting the value 0.10197.

The equivalent transfer function model for *MINT* is

$$MINT_t = \mu + \omega_{0,INFL}/(1-\delta_{1,INFL} B)INFL_t + \omega_{0,EQDG} EQDG_t + \omega_{0,LINTZ} LINTZ_t + (1-\theta_1 B)\epsilon_t$$

where $\epsilon_t \sim \text{i.i.d.}N(0, \sigma_\epsilon^2)$. The data was centred by subtracting the value 0.097818 and the maximum likelihood estimates are as follows :

Parameter	Estimate	Approx. Std Error	T Ratio
$\omega_{0,INFL}$	0.81887	0.15507	5.28
$\delta_{1,INFL}$	0.10040	0.15739	0.64
$\omega_{0,EQDG}$	-0.15957	0.03531	-4.52
$\omega_{0,LINTZ}$	0.69174	0.35915	1.93
θ_1	-0.67443	0.15427	-4.37

The parameter $\omega_{0,LINTZ}$ is not quite significant while the parameter $\delta_{1,INFL}$ is clearly not significant. $\delta_{1,INFL}$ is the heart of the unit-gain function and it is clear that, while two variables may be significantly correlated and the unit-gain account for a large degree of variability in the series, this does not guarantee its significance. When more variables are added, some of the variables act as surrogates for the unit-gain function. This shows that the unit-gain and error terms cannot be modelled separately.

The equivalent transfer function model for *PDRY* is

$$PDRY_t = \mu + \omega_0/(1-\delta_1 B)INFL_t + 1/(1-\phi_1 B)\varepsilon_t$$

where $\varepsilon_t \sim \text{i.i.d.}N(0, \sigma_\varepsilon^2)$. The data was centred by subtracting the value 2.078148 and the maximum likelihood estimates are as follows :

Parameter	Estimate	Approx. Std Error	T Ratio
ω_0	0.44369	0.27053	1.64
δ_1	0.82081	0.09979	8.23
ϕ_1	0.68058	0.14576	4.67

Since ω_0 is not significant, the term $\omega_0/(1-\delta_1 B)INFL_t$ is not significant which implies that the unit-gain function for *PDRY* is not significant.

For the purposes of developing the structure of the model, Thomson [7] used differenced input variables whenever the autocorrelation coefficients die out at high lags. In the context of the *LINT* unit-gain

function, he naturally found that “the time-series analysis effectively treats *INFL* and *LINT* as increasing functions of time and therefore disregards the correlation between them.” When attempting to model trended series within a stationary time-series model, it is necessary to avoid introducing time as a variable and differenced variables should not be used, even for the purposes of identifying the structure of the model. If a prewhitening filter includes a differencing factor and the input series is trended, the cross-correlation function of the prewhitened series gives the cross-correlations after removing a trend from the output series. If the output data are believed to have weak stationarity, this “detrending” is inappropriate and results in misleading cross-correlations. It is equivalent to over-differencing in univariate time series analysis.

The partial correlation coefficient between two variables Y_1 and Y_2 , eliminating Z_1, Z_2, \dots, Z_r measures the association between Y_1 and Y_2 after eliminating the effects of the variables Z_1, Z_2, \dots, Z_r [2]. For the variables initially modelled, the partial correlation matrix, eliminating the effect of time, is given in table 1 below. (The partial correlation is shown above and, the probability of obtaining a partial correlation greater in absolute value under the null hypothesis that the partial correlation is zero, is shown in brackets below.)

	<i>INFL</i>	<i>EQDG</i>	<i>EQDY</i>	<i>PTDG</i>	<i>PTDY</i>	<i>PDRG</i>	<i>PDRY</i>	<i>LINT</i>
<i>EQ</i> <i>DG</i>	0.34 (0.20)							
<i>EQ</i> <i>DY</i>	0.00 (1.0)	-0.20 (0.45)						
<i>PT</i> <i>DG</i>	0.20 (0.45)	-0.32 (0.22)	0.24 (0.38)					
<i>PT</i> <i>DY</i>	-0.27 (0.31)	-0.33 (0.21)	-0.26 (0.34)	-0.31 (0.24)				
<i>PD</i> <i>RG</i>	0.07 (0.80)	-0.16 (0.55)	0.56 (0.02)	0.37 (0.16)	-0.66 (0.01)			
<i>PD</i> <i>RY</i>	-0.42 (0.11)	-0.17 (0.53)	-0.38 (0.14)	-0.26 (0.33)	0.74 (0.00)	-0.41 (0.12)		
<i>LI</i> <i>NT</i>	0.59 (0.02)	0.00 (0.99)	0.54 (0.02)	0.32 (0.23)	-0.05 (0.85)	0.16 (0.56)	-0.46 (0.08)	
<i>MI</i> <i>NT</i>	0.26 (0.33)	-0.50 (0.04)	0.34 (0.19)	0.87 (0.00)	-0.04 (0.90)	0.18 (0.50)	-0.20 (0.45)	0.55 (0.03)

Table 1. The partial correlation matrix.

Only *LINT* is significantly correlated with *INFL* at the 5% (or even the 10%) level once the effects of time have been eliminated : the partial correlation matrix appears to be a better diagnostic tool than the correlation matrix when series are trended since it eliminates spurious correlations between variables which are induced by their common relationships with time.

4.1 Prewhitening and Model Identification

The cross-correlation function of two series which have been prewhitened by the univariate model of the input series is used to identify the relationship (or transfer function) between the two series [1]. (The cross-correlation function between the original series may show spurious cross-correlations due to the autocorrelation of the input series and so is not useful for identifying the transfer function.) A tentative model is then estimated by maximum likelihood and the significance of the parameters checked.

Using the EQDY (the natural logarithm of equity yields per cent) AR(1) filter (instead of differencing) to prewhiten the EQDG (the annual force of equity dividend growth) series, it becomes evident that equity dividend yields are significantly correlated with equity dividend growth rates. The cross-correlation at lag 1 is -0.434 which is significant at the 5% level. Contrary to Thomson's [7] findings, this gives credence to the conventional wisdom that a low dividend yield anticipates high dividend growth in the following year. The cross correlation functions using both an AR(1) and a differenced prewhitening filter are shown in Appendix 6.

The cross-correlation analysis relies on the fact that the input series is prewhitened. It is therefore important to check that the residuals of the input model are indeed i.i.d. $N(0, \sigma_e^2)$ random variables to avoid misleading cross-correlation coefficients. Since the maximum likelihood estimates are the parameter estimates which maximise the likelihood function, they are independent of the cross-correlations obtained in the prewhitening process. Therefore, the choice of input model for the purposes of determining the structure of the model need not be used in the final model : a model with white noise residuals can be used even if it is considered unsuitable for the final model. This

usually involves the use of longer-term lags but differenced input models should be avoided (as discussed above).

The decision of which series to use as the input series and which to use as the response series needs to be made. Thomson [7] fits univariate series to all the variables and examines the cross-correlation functions of each in turn. In this process, however, if the cross-correlation of X with input Y had already been examined, the cross-correlation of Y with input X was not analysed. This assumes that the cross-correlation at negative lags of X with input Y covers the possibility of X being an input for Y. However, if X is an input for Y, Y does not follow a univariate ARIMA process². It is this fact which requires that the two series be prewhitened both with X as input for Y and Y as input for X.

If a response series requires more than one input series, analysis of the cross correlation function may be misleading as an indication of the appropriate transfer function. "Any dependencies among two or more input series will confound the cross-correlations with the response series," as discussed in [6].

4.2 Real and Nominal Return Discrepancies

In section 2, a number of discrepancies between historical means and forecast means were presented for both real and nominal returns. These forecast mean returns are a direct consequence of the forecast means of the model variables since the model is stationary. The historical mean returns, however, depend both on the historical means of the model variables and whether the historical series are trended or not. For example, the high historical return on equities is not only a consequence of the high historical dividend growth rate but also the decreasing trend in historical dividend yields. Because the Thomson model is a stationary model, this trend is not continued and lower equity returns are the result. The process of mean reversion in fact reverses the trend until the series reaches its long-run mean resulting in particularly low initial returns, as discussed in section 3.

In the Thomson model, the real and nominal return forecast correlations (discussed in section 2) were found to be under-estimated

² See Appendix 7.

(except for the correlation between property trust returns and money-market instrument returns.) The correlations between the corresponding model variables are also under-estimated and, although the relationship between the model variables and their corresponding returns is non-linear, this is the cause of the low forecast return correlations.

The correlations for the model variables are under-estimated for two reasons. Firstly, the number of input variables a response variable can have is limited by the cascade structure of the transfer function model. Transfer function models cannot handle feedback and the prewhitening technique assumes that the input variables do not depend on past values of the response variable [6]. Secondly, the null hypothesis used in testing for significant cross-correlations between two prewhitened variables is that the cross-correlations are zero. With a short period of data, the confidence intervals are very wide and the power of the test is much reduced. This explains why the correlation structure of the property variables is so poorly modelled.

As discussed earlier in this section, the modified variable *MINTZ* (which is the error term after estimating the inflation unit-gain function) is not an appropriate model variable since the unit-gain function for *MINT* (the annual force of interest on money-market instruments) is not significant. The high correlation between property trust returns and money-market instrument returns is due to the variable *PTDG* (the annual force of property trust dividend growth). The model used for *PTDG* is

$$PTDG_t \approx 0.077 + 1.721MINTZ_t - 0.967MINTZ_{t-2} + 0.053\eta_t$$

which introduces the spurious correlation between the two returns.

5 Residual Analysis

Non-normal residuals may indicate that the error process follows a distribution other than the normal distribution which may indicate either the need for a transformation, that an explanatory variable has been omitted, or that the model is incorrectly specified [3].

Thomson [7] tests the skewness and kurtosis as well as the sample cumulative distribution function of the residuals for normality and finds

that the residuals of *EQDG*, *LINTZ* and *PDRYZ* are not normally distributed. In analysing the distribution of residuals, however, he uses residuals obtained from back-forecasts of each series. These are not the residuals used in the maximum likelihood estimation of model parameters and are not appropriate for determining the distribution of the error processes.

At most, back-forecast residuals give an indication of how well the model has performed historically since the time the back-forecast was made. It is misleading to analyse back-forecast residuals as a group since they represent forecasts at different lead times. For example, an AR(1) process has a forecast variance which increases to a stable level but the variance of the forecast distribution at small lead-times is much less than the variance at large lead-times.

6 Subset Parameter Stability

The parameters of the model have been estimated on two subsets of the data for the periods 1960 to 1975 and 1976 to 1993. The parameters of the modified variables *MINTZ* and *PDRYZ* change considerably from their estimates based on the full data set for the period 1960 to 1993. These estimates and corresponding t-ratios (shown in italics below) are as follows :

	1960 - 1993	1960 - 1975	1976 - 1993
θ_{Y1}	-0.531 <i>(-3.29)</i>	0.999	-0.999 <i>(-0.02)</i>
$\omega_{X1,Y0}$	0.885 <i>(2.62)</i>	1.561 <i>(11.64)</i>	0.400 <i>(0.72)</i>

Unstable parameter estimates for the variable *MINTZ*

	1960 - 1993	1960 - 1975	1976 - 1993
ϕ_{Y1}	0.680 <i>(4.79)</i>	0.118 <i>(0.31)</i>	0.878 <i>(8.94)</i>

Unstable parameter estimates for the variable *PDRYZ*

The notation used corresponds to the notation used by Thomson [7]. Both parameters for the variable *MINTZ* are insignificant for the period 1976 to 1993 and change considerably from their corresponding estimates for the period 1960 to 1975. The estimate for ϕ_{Y1} for the variable *PDRYZ* for the period 1960 to 1975 is also insignificant.

The parameter ϕ_1 for the variable *INFL* is unstable but remains significant over both sub-periods as shown below :

	1960 - 1993	1960 - 1975	1976 - 1993
ϕ_{Y1}	0.899 (12.47)	0.668 (2.55)	0.515 (2.10)
μ_Y	0.092	0.050	0.127

Unstable parameter estimates for the variable *INFL*

It is interesting to note that ϕ_{Y1} is lower in both sub-periods than over the whole period. The mean level of the series in the two sub-periods is quite different and the model cannot be said to be stationary. In fact, it is the change in the level of *INFL* following the oil shocks of '73 and '79 which causes the high value of $\phi_{Y1}=0.899$ when estimated over the period 1960 to 1993. The forecast mean for *INFL* in the Thomson model tends to the arbitrary level of 9.5% and clearly depends on the period of data used to estimate the parameters. Further, the forecast variance is too large because the value of ϕ_{Y1} has been over-estimated.

7 Conclusion

As a forecasting tool, the Thomson stochastic investment model suffers from a number of drawbacks due largely to the nature of ARIMA models. The assumption of weak stationarity imposed on the data by this class of models limits the forecast mean to either the historical mean or, if a differenced model is used, to a forecast mean which extrapolates the historical trend.

The unit-gain function is a type of transfer function and should be estimated simultaneously with the rest of the parameters in the model.

The correlation matrix used in the Thomson model to identify the unit-gain function is not appropriate and, although the partial correlation matrix performs somewhat better, the prewhitening technique should be preferred and a suitable prewhitening filter found. The unit-gain functions for *MINT* and *PDRY* are not significant and the modified variables *LINTZ*, *MINTZ*, and *PDRYZ* are redundant and result in a model that is over-parameterised and unstable.

The poor replication of the correlation structure is due partly to the short period for which property data is available and partly to the need for a cascade structure when modelling with Box-Jenkins ARIMA models. It should be noted that ARIMA models do not allow feedback; in order for a variable *X* to be used as an input for another variable *Y*, *Y* cannot be used to model *X*. Further, the effect of a surrogate variable *Z* for the input variable *X* is to eliminate *X* from the transfer function with output *Y*. Even though *X* may be significantly correlated with *Y*, the forecast correlation between *X* and *Y* will be zero.

The oil shocks of the 70's lead to a period of high, entrenched inflation in the 80's with inflation remaining above 10%. This change in the level of inflation invalidates the assumption of weak stationarity and results in an AR(1) parameter which is too high. The consequences are an arbitrary long-run mean of 9.5% and a forecast variance which is too large.

The interpretation placed on the data by ARIMA models is not necessarily incorrect and the models have found use in a large number of investment models, however, some of the discrepancies pointed out in this paper cannot be corrected by these models. A model using intervention analysis, breakpoint regression, GARCH effects, multivariate time series and the state-space model is currently being developed and will be the subject of a future paper.

References

1. Box, G.E.P. & Jenkins, G.M. (1976). *Time series analysis, forecasting and control. (Revised edition.)* Prentice Hall, New Jersey.
2. Johnson, N.L. & Kotz, S. (1972). *Distributions in statistics : continuous multivariate distributions.* John Wiley & Sons, Inc.
3. Kendall, M. & Ord, J.K. (1990). *Time Series. (Third edition.)* A Charles Griffin Title, Edward Arnold.
4. Lane, G.B. (1995). *A practical guide to stochastic asset modelling.* Presented to the 1995 Convention of the Actuarial Society of South Africa
5. Maturity Gaurantees Working Party of the Institute of Actuaries (MGWP) (1980). *Report of the MGWP.* Journal of the Institute of Actuaries 107, part 2, no. 435.
6. *SAS/ETS User's Guide. Version 6, second edition.* (1993). Copyright 1993 by SAS Institute Inc., Cary, NC, USA.
7. Thomson, R.J. (1994). *A stochastic investment model for actuarial use in South Africa.* Presented to the 1994 Convention of the Actuarial Society of South Africa.
8. Wilkie, A.D. (1994). *More on a stochastic asset model for actuarial use.* Presented to the Institute of Actuaries, 24 April 1995.

Appendix 1

The intended model for inflation is

$$INFL_t \approx 0.008 + 0.899INFL_{t-1} + 0.088EQDG_t - 0.079EQDG_{t-1} \\ + 0.077EQDG_{t-2} - 0.069EQDG_{t-3} + 0.020\eta_t$$

where $INFL_t$ is the mean force of inflation in year t. The intended model for property trust dividend growth is

$$PTDG_t \approx 0.077 + 1.721MINTZ_t - 0.967MINTZ_{t-2} + 0.053\eta_t$$

where $PTDG_t$ is the mean force of property unit trust dividend growth in year t.

The remainder of the model is identical to the model published in [7]. The remaining difference equations are given below with a brief summary of the acronyms used. The variables η_t are each independently and identically distributed $N(0,1)$ random variables.

The equity dividend yields and dividend growth rates are modelled as $EQDG_t$ and $EQDY_t$, respectively where $EQDG_t$ is the mean force of equity dividend growth in year t and $EQDY_t$ is the natural logarithm of the All Share Index Dividend Yield at time t per cent.

The difference equations for equities are as follows:

$$EQDG_t = 0.093 + 0.116\eta_t + 0.076\eta_{t-1}$$

$$EQDY_t = 0.310 + 0.810EQDY_{t-1} + 0.198\eta_t$$

Long-term interest rates are modelled as the sum of a unit gain function $Z_{LINT,t}$ and an error process $LINTZ_t$. The variable $LINT_t$ is the annual force of interest on 20-year bonds as estimate by the JSE-Actuaries 20-year bond yield.

The difference equations for long-term interest rates are as follows:

$$Z_{LINT,t} = 0.006 + 0.126INFL_t + 0.85Z_{LINT,t-1}$$

$$LINTZ_t = 0.010\eta_t + 0.006\eta_{t-1}$$

$$LINT_t = Z_{LINT,t} + LINTZ_t$$

Short-term interest rates are modelled as the sum of a unit gain function $Z_{MINT,t}$ and an error process $MINTZ_t$. The variable $MINT_t$ is the annual force of interest on money-market instruments as measured by the Ginsberg Malan & Carsons Money-market Index.

The difference equations for short-term interest rates are as follows:

$$Z_{MINT,t} = 0.004 + 0.141INFL_t + 0.85Z_{MINT,t-1}$$

$$MINTZ_t = 0.008 - 0.091EQDG_t + 0.885LINTZ_t + 0.019\eta_t + 0.010\eta_{t-1}$$

$$MINT_t = Z_{MINT,t} + MINTZ_t$$

Direct property dividend yields and dividend growth rates are modelled as $PDRY_t$ and $PDRG_t$ respectively. $PDRG_t$ is the mean force of direct property rental growth in year t as measured by the Dunlop Heywood investment property index and $PDRY_t$ is the natural logarithm of the yield per cent on direct property at time t . $PDRY_t$ is modelled as the sum of a unit gain function $Z_{PDRY,t}$ and an error process $PDRYZ_t$.

The difference equations for direct property are as follows:

$$Z_{PDRY,t} = 0.486 + 0.559INFL_t + 0.74Z_{PDRY,t-1}$$

$$PDRYZ_t = 0.680PDRYZ_t + 0.061\eta_t$$

$$PDRY_t = Z_{PDRY,t} + PDRYZ_t$$

$$PDRG_t = 0.096 + 0.545(PDRY_t - PDRY_{t-1}) + 0.068\eta_t + 0.041\eta_{t-1}$$

Property unit trust dividend yields and dividend growth rates are modelled as $PTDY_t$ and $PTDG_t$ respectively where $PTDY_t$ is the natural logarithm of the yield per cent on property unit trusts at time t . The corrected difference equation for $PTDG_t$ is given above and the difference equation for $PTDY_t$ is as follows:

$$PTDY_t = 2.547 + 0.598PDRY_t - 0.738PDRY_{t-1} + 0.104\eta_t$$

In order to start the simulation, the one-step-ahead residuals for both MA(q) and ARMA(p,q) processes are needed. They are -0.04587 for *EQDG*, -0.025108 for *LINTZ*, -0.00552 for *MINTZ* and 0.01254 for *PDRG*. These translate into the following N(0,1) errors which can be used directly in the difference equations:

$$\eta_{EQDG,1993} = -0.3954$$

$$\eta_{LINTZ,1993} = -2.5108$$

$$\eta_{MINTZ,1993} = -0.2905$$

$$\eta_{PDRG,1993} = 0.2265$$

Appendix 2: Historical & Forecast Nominal Return Statistics

Period	EQ	PT	PD	LI	MI	INF
to 1993	0.23	0.21	0.19	0.09	0.10	0.10
1994	0.01	0.37	0.24	0.07	0.13	0.10
1995	0.04	0.25	0.22	0.05	0.14	0.10
1996	0.07	0.24	0.22	0.16	0.14	0.11
1997	0.09	0.21	0.21	0.16	0.14	0.11
1998	0.11	0.18	0.21	0.16	0.14	0.11
1999	0.13	0.19	0.21	0.16	0.14	0.11
2000	0.14	0.20	0.21	0.15	0.14	0.11
2001	0.15	0.20	0.21	0.15	0.14	0.11
2002	0.16	0.21	0.21	0.15	0.13	0.10
2003	0.17	0.21	0.21	0.15	0.13	0.10

Table 1 : Expected Nominal Returns

Period	EQ	PT	PD	LI	MI	INF
to 1993	0.26	0.17	0.12	0.12	0.06	0.05
1994	0.25	0.18	0.09	0.08	0.03	0.03
1995	0.26	0.21	0.11	0.08	0.03	0.03
1996	0.27	0.21	0.11	0.10	0.03	0.04
1997	0.28	0.21	0.11	0.10	0.03	0.04
1998	0.29	0.21	0.11	0.10	0.03	0.04
1999	0.29	0.22	0.11	0.11	0.04	0.05
2000	0.30	0.22	0.11	0.11	0.04	0.05
2001	0.29	0.22	0.11	0.11	0.04	0.05
2002	0.30	0.22	0.11	0.11	0.04	0.05
2003	0.30	0.22	0.11	0.11	0.04	0.05

Table 2 : Standard Deviation of Nominal Returns

Period	EQ	PT	PD	LI	MI	INF
to 1993	0.70	0.77	0.04	0.70	0.61	-0.29
1994	0.71	0.37	0.23	0.34	0.04	0.04
1995	0.78	0.49	0.28	0.22	0.07	0.06
1996	0.78	0.56	0.23	0.29	0.07	0.06
1997	0.82	0.62	0.26	0.20	0.05	0.12
1998	0.80	0.60	0.26	0.22	0.10	0.16
1999	0.76	0.58	0.32	0.23	0.07	0.12
2000	0.80	0.55	0.25	0.19	0.10	0.11
2001	0.84	0.51	0.28	0.23	0.11	0.13
2002	0.81	0.65	0.26	0.16	0.10	0.10
2003	0.80	0.60	0.29	0.15	0.10	0.11

Table 3 : Skewness of Nominal Returns

	EQ	PT	PD	LI	MI	INF
EQ	1.00	0.56	0.20	0.36	-0.03	0.21
	1.00	-0.08	-0.01	-0.06	-0.18	0.12
PT	0.56	1.00	0.29	0.04	-0.01	0.25
	-0.08	1.00	0.09	-0.04	0.23	0.00
PD	0.20	0.29	1.00	0.01	0.16	0.20
	-0.01	0.09	1.00	0.02	0.05	0.02
LI	0.36	0.04	0.01	1.00	0.31	0.27
	-0.06	-0.04	0.02	1.00	-0.02	-0.10
MI	-0.03	-0.01	0.16	0.31	1.00	0.73
	-0.18	0.23	0.05	-0.02	1.00	0.50
INF	0.21	0.25	0.20	0.27	0.73	1.00
	0.12	0.00	0.02	-0.10	0.50	1.00

Table 4 : Correlation Matrix for Nominal Returns

Historical correlations are shown above with forecast correlations at a lead-time of ten years below.

Appendix 3: Historical & Forecast Real Return Statistics

Period	EQ	PT	PD	LI	MI
to 1993	0.12	0.06	0.06	0.00	0.01
1994	-0.09	0.24	0.13	-0.03	0.02
1995	-0.05	0.14	0.11	-0.05	0.04
1996	-0.04	0.12	0.10	0.05	0.03
1997	-0.01	0.10	0.10	0.05	0.03
1998	0.00	0.07	0.09	0.05	0.03
1999	0.02	0.08	0.09	0.05	0.03
2000	0.03	0.09	0.09	0.05	0.03
2001	0.04	0.09	0.09	0.05	0.03
2002	0.05	0.10	0.09	0.04	0.03
2003	0.06	0.10	0.09	0.04	0.03

Table 5 : Expected Real Returns

Period	EQ	PT	PD	LI	MI
to 1993	0.23	0.15	0.11	0.11	0.04
1994	0.22	0.17	0.09	0.08	0.04
1995	0.23	0.20	0.10	0.09	0.04
1996	0.24	0.20	0.11	0.11	0.04
1997	0.24	0.20	0.11	0.11	0.04
1998	0.25	0.20	0.11	0.11	0.04
1999	0.25	0.20	0.11	0.11	0.04
2000	0.26	0.21	0.11	0.11	0.04
2001	0.26	0.21	0.11	0.12	0.04
2002	0.26	0.21	0.11	0.12	0.04
2003	0.27	0.21	0.11	0.12	0.04

Table 6 : Standard Deviation of Real Returns

Period	EQ	PT	PD	LI	MI
to 1993	0.55	0.74	0.47	0.35	-0.68
1994	0.70	0.39	0.25	0.35	0.13
1995	0.77	0.51	0.29	0.31	0.16
1996	0.75	0.57	0.28	0.41	0.16
1997	0.79	0.64	0.29	0.35	0.07
1998	0.78	0.65	0.29	0.39	0.09
1999	0.76	0.57	0.30	0.41	0.17
2000	0.79	0.56	0.27	0.48	0.15
2001	0.85	0.55	0.32	0.48	0.15
2002	0.81	0.63	0.28	0.46	0.19
2003	0.81	0.62	0.33	0.48	0.17

Table 7 : Skewness of Real Returns

	EQ	PT	PD	LI	MI
EQ	1.00	0.54	0.15	0.33	-0.24
PT	1.00	-0.06	0.01	-0.01	-0.17
PD	0.54	1.00	0.24	0.03	-0.10
LI	-0.06	1.00	0.19	0.10	0.36
MI	0.15	0.24	1.00	0.01	0.12
	0.01	0.19	1.00	0.24	0.33
	0.33	0.03	0.01	1.00	0.21
	-0.01	0.10	0.24	1.00	0.35
	-0.24	-0.10	0.12	0.21	1.00
	-0.17	0.36	0.33	0.35	1.00

Table 8 : Correlation Matrix for Real Returns

Historical correlations are shown above with forecast correlations at a lead-time of ten years below.

Appendix 4: Forecast Means & 95% CIs for Model Variables

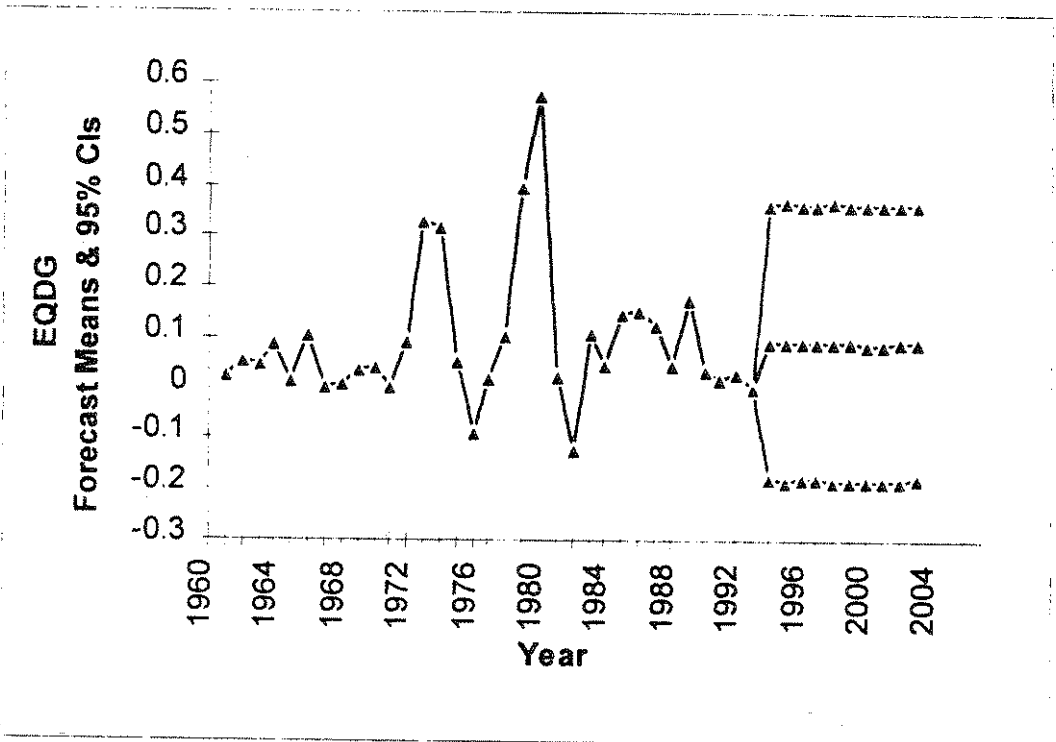


Figure 7. EQDG, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

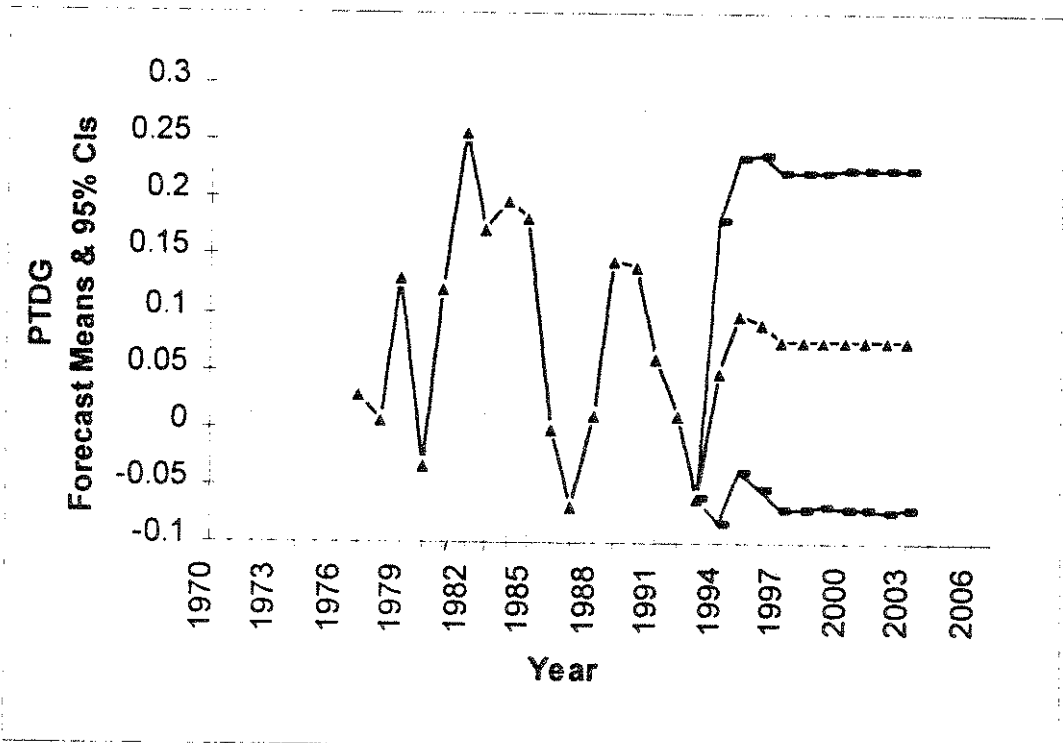


Figure 8. PTDG, 1977-1993, and forecast means & 95% confidence intervals, 1994-2003.

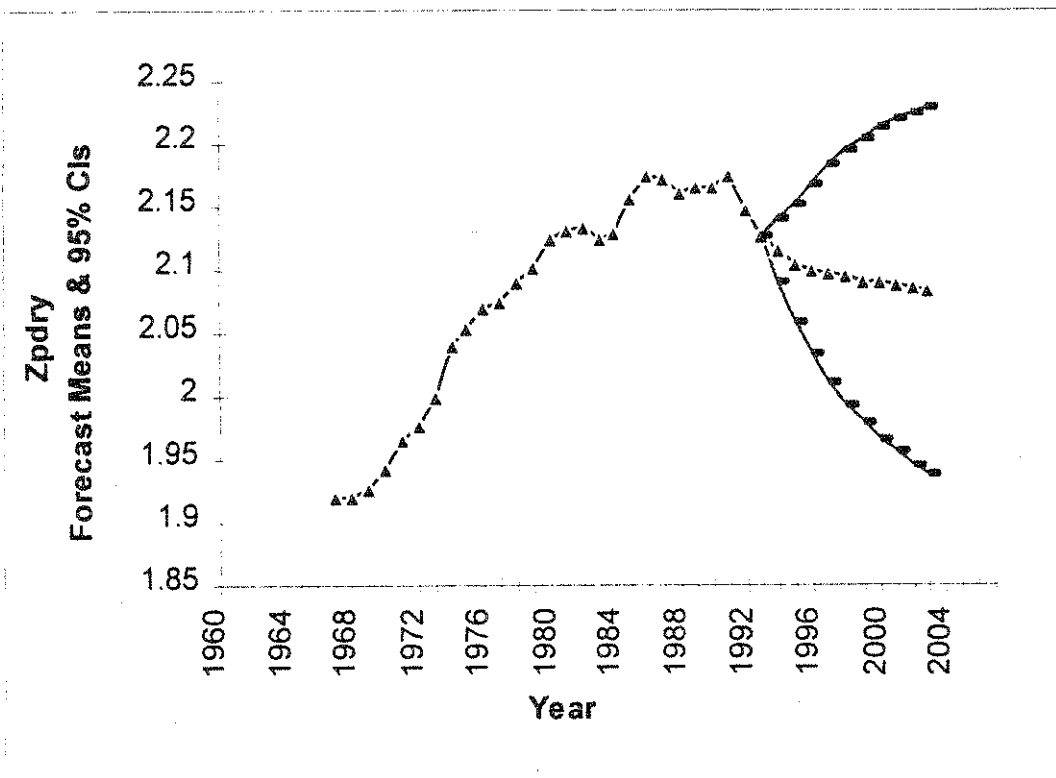


Figure 9. Z_{PDry} , 1967-1993, and forecast means & 95% confidence intervals, 1994-2003.

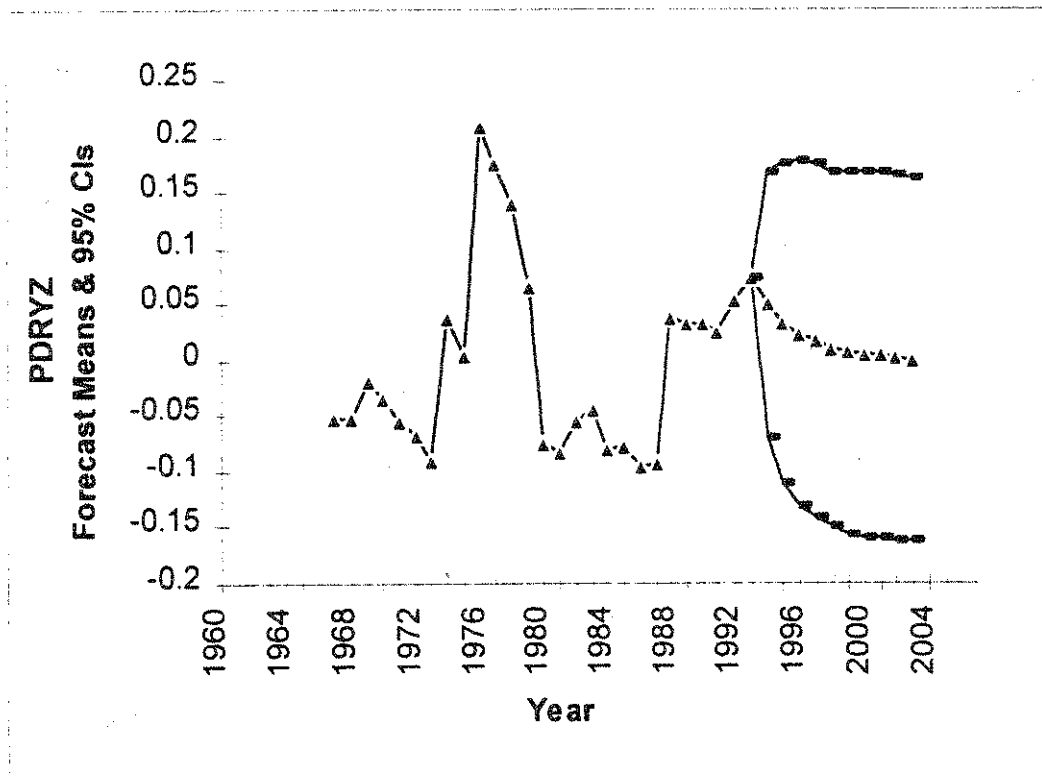


Figure 10. P_{DRYZ} , 1967-1993, and forecast means & 95% confidence intervals, 1994-2003.

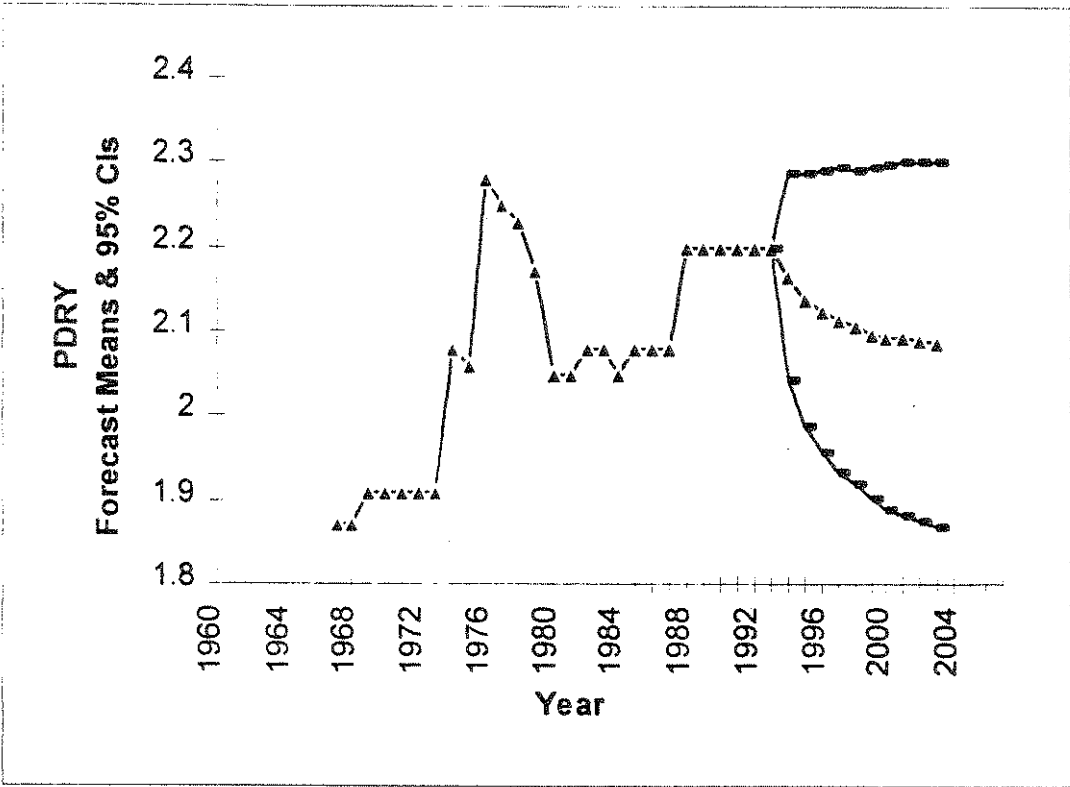


Figure 11. PDRY, 1967-1993, and forecast means & 95% confidence intervals, 1994-2003.

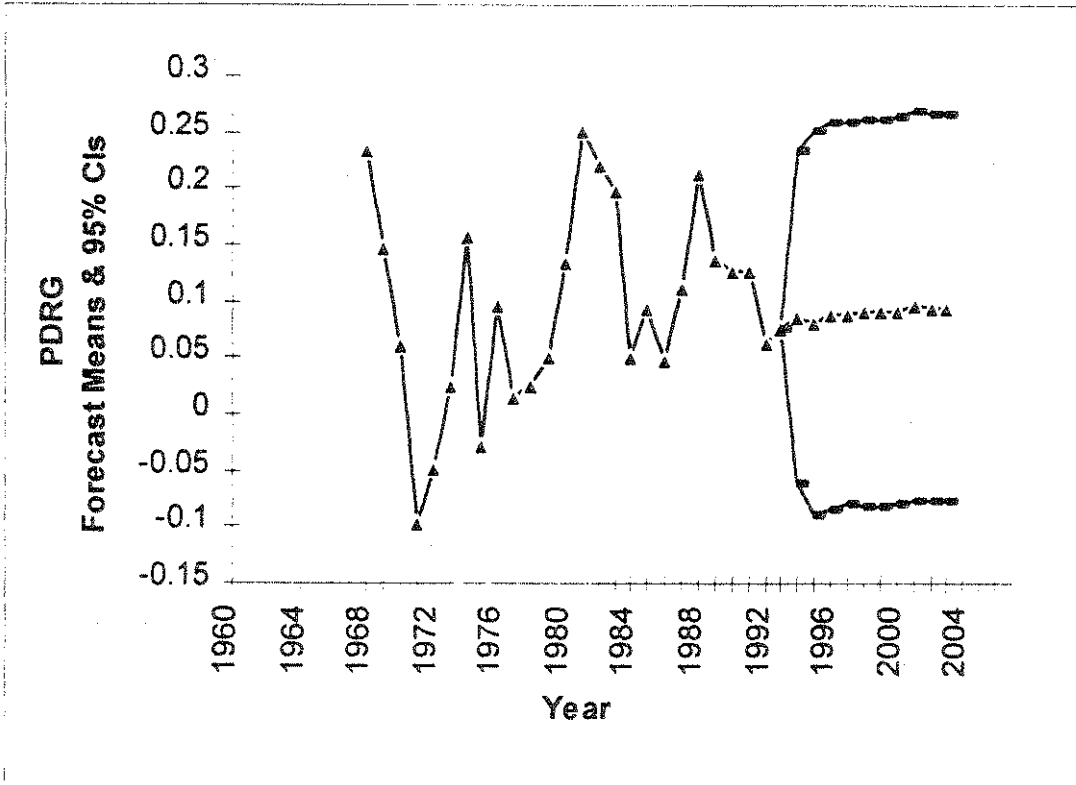


Figure 12. PDRG, 1968-1993, and forecast means & 95% confidence intervals, 1994-2003.

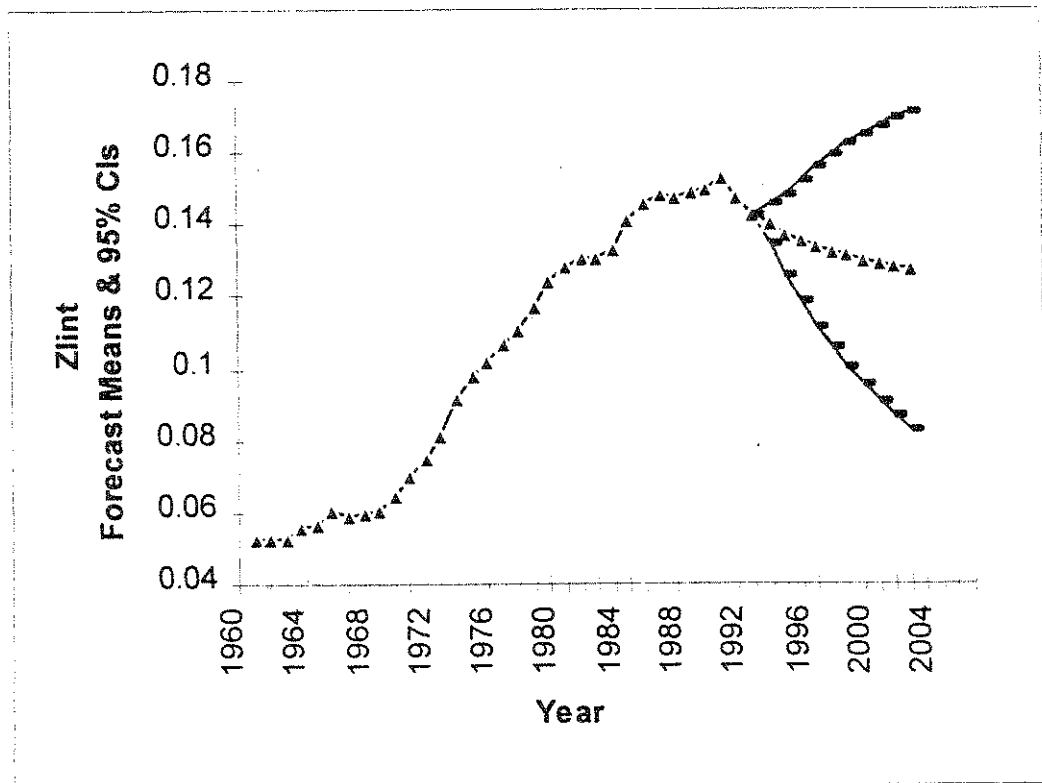


Figure 13. Z_{LINT} , 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

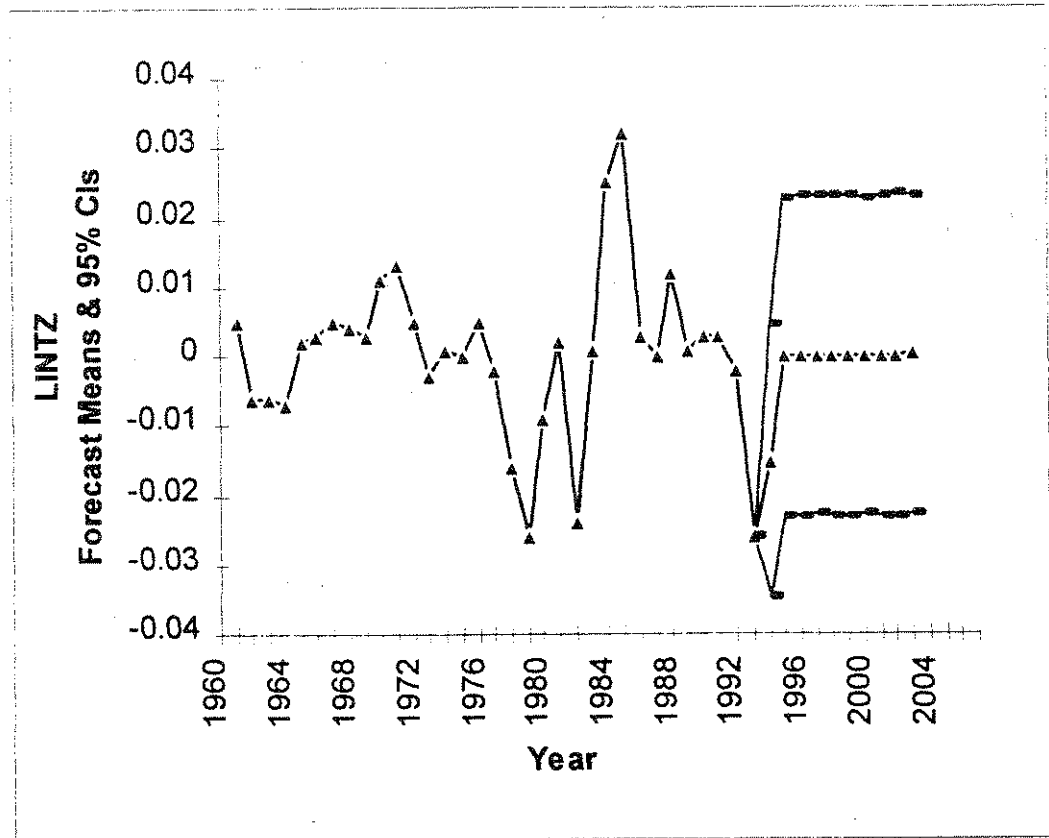


Figure 14. LINTZ, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

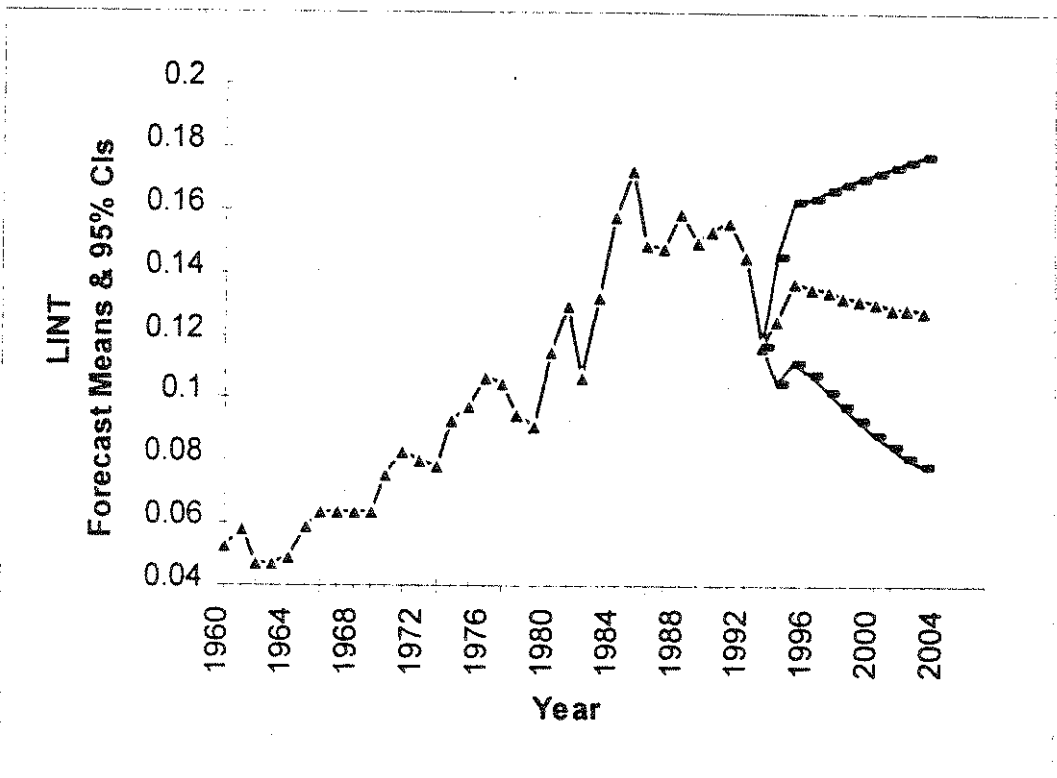


Figure 15. LINT, 1960-1993, and forecast means & 95% confidence intervals, 1994-2003.

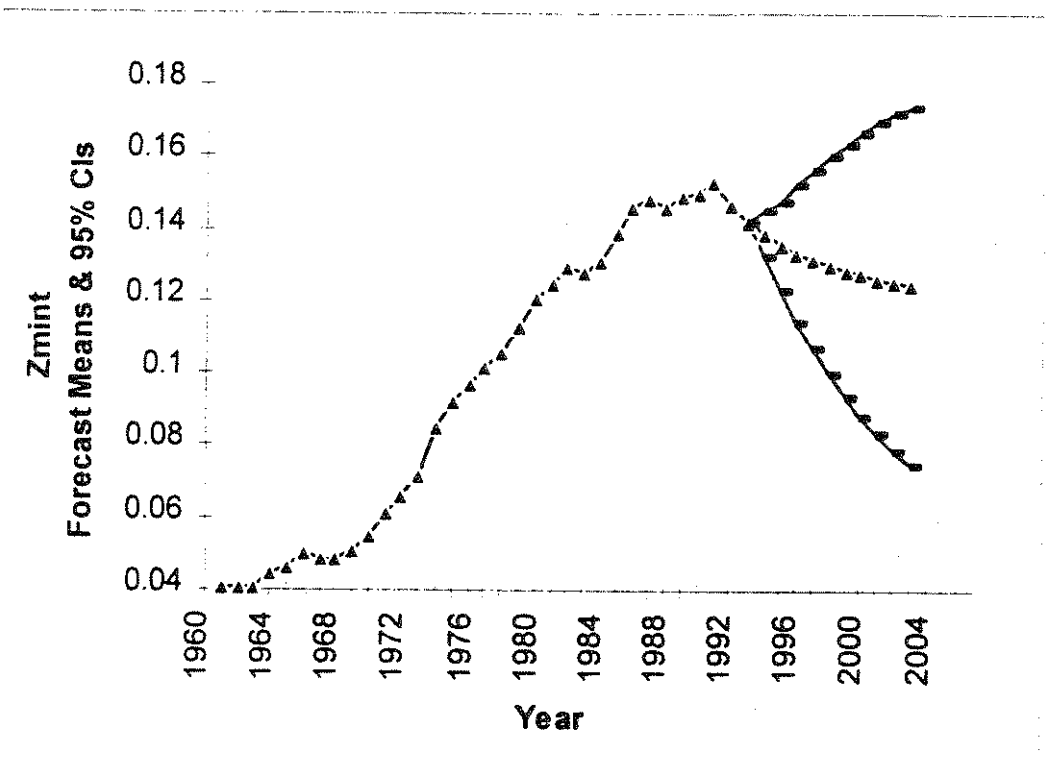


Figure 16. Z_{MINT} , 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

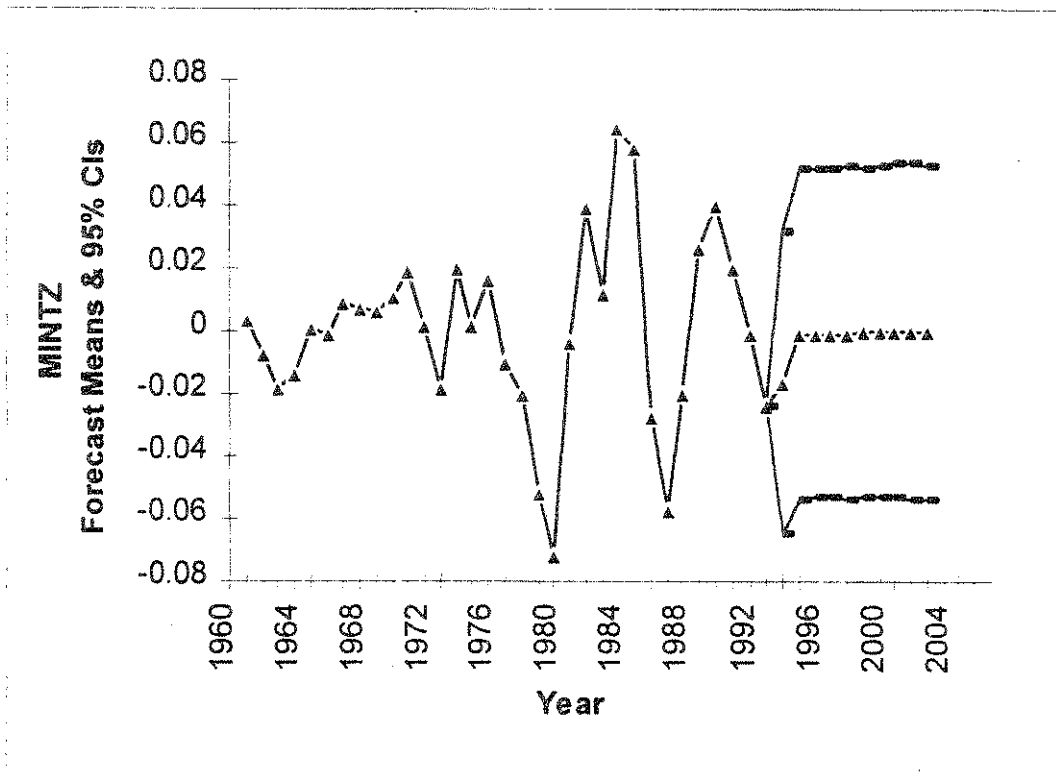


Figure 17. MINTZ, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

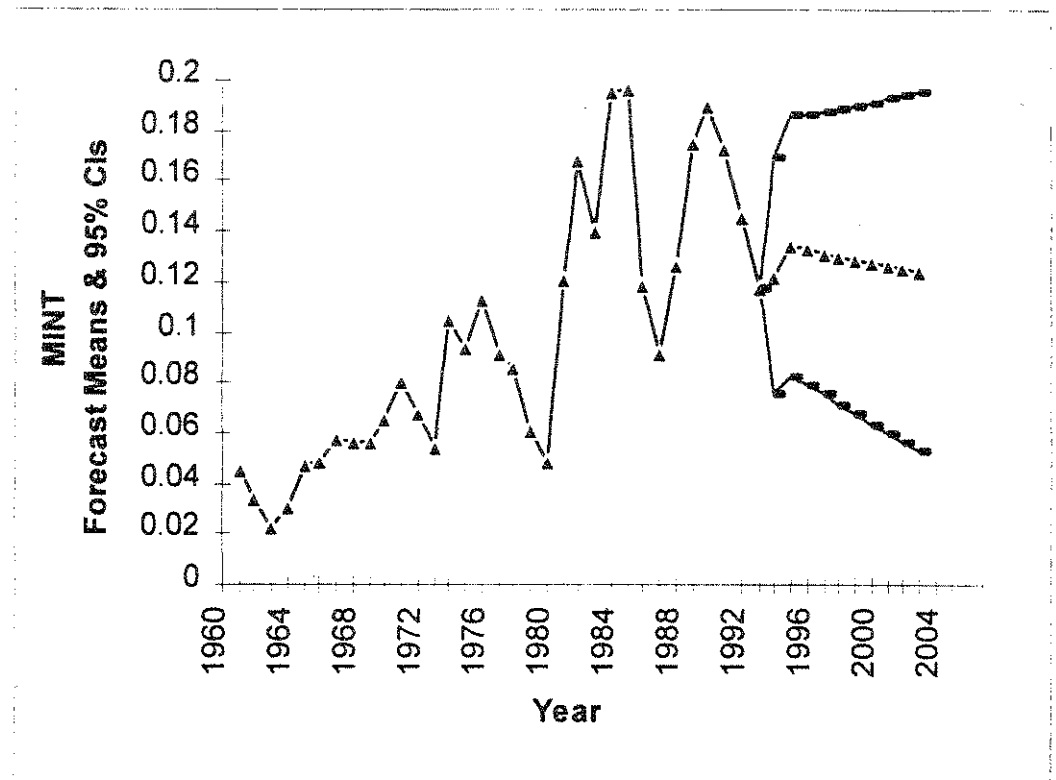


Figure 18. MINT, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

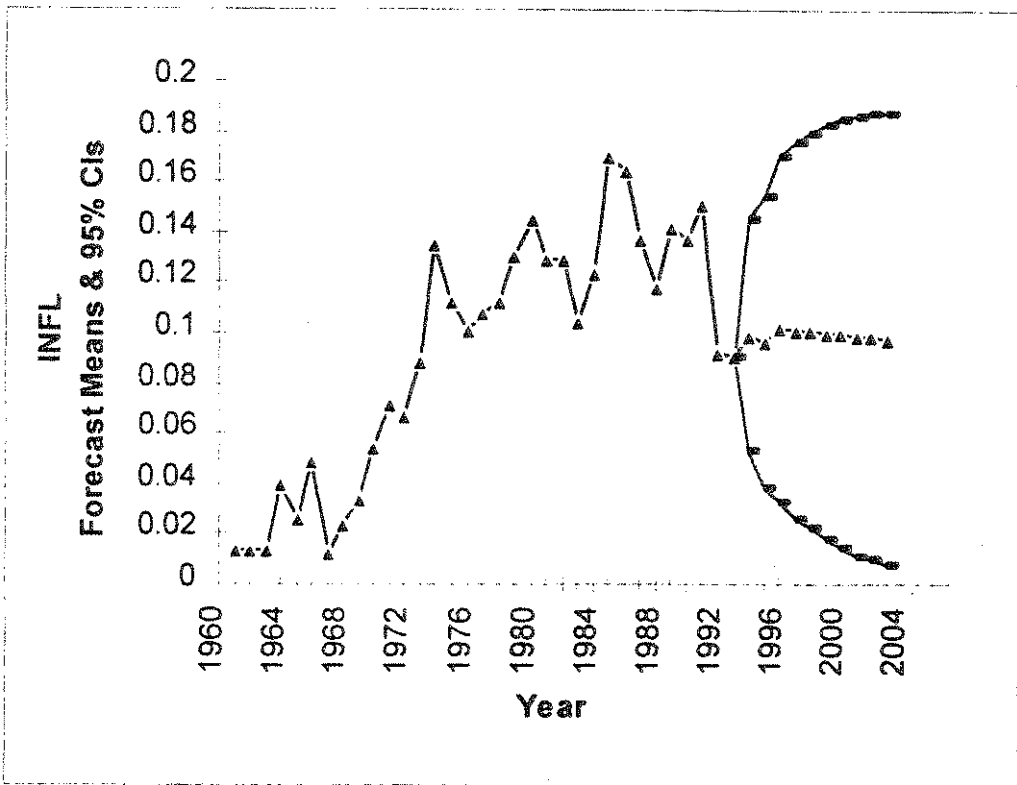


Figure 19. INFL, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

Appendix 5: Forecast statistics for model variables.

Given that Y_t follows as ARIMA(p,d,q) process and that $X_t = \nabla^d Y_t$, where $\nabla = (1-B)$ is the *difference operator* and B is the *backward shift operator*, X_t follows an ARMA(p,q) process. We can write the process for X_t in terms of the *autoregressive operator* of order p ,

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

and the *moving average operator* of order q ,

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q,$$

so that

$$X_t = \mu_x + \theta(B)/\phi(B) \varepsilon_t,$$

where $\varepsilon_t \sim \text{i.i.d } N(0, \sigma_\varepsilon^2)$ and μ_x is the mean of the series [1]. The process for X_t can also be written in terms of the difference equation

$$X_t = c + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$

where $c = (1 - \phi_1 - \phi_2 - \dots - \phi_p) \mu_x$. For $k \geq q$,

$$E[X_{t+k}] = c + \phi_1 E[X_{t+k-1}] + \phi_2 E[X_{t+k-2}] + \dots + \phi_p E[X_{t+k-p}]$$

since $E[\varepsilon_{t+j} | X_{t-1}, X_{t-2}, \dots] = 0 \quad \forall j \geq 0$. Because the process for X_t is stationary, $E[X_{t+k}] \rightarrow \gamma$, where γ is some finite number. Taking limits on both sides of the above equation gives

$$\gamma = c + \phi_1 \gamma + \phi_2 \gamma + \dots + \phi_p \gamma$$

Hence,

$$E[X_{t+k}] \rightarrow c/(1 - \phi_1 - \phi_2 - \dots - \phi_p) = \mu_x \text{ as } k \rightarrow \infty.$$

To calculate the variance of X_t , it is convenient to write X_t as a weighted sum of present and past values of the white noise process ε_t :

$$X_t = \mu_x + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots$$

Since ε_t is white noise,

$$\text{Var}(X_{t+k}|X_{t-1}, X_{t-2}, \dots) = \text{Var}(\varepsilon_{t+k}) + \psi_1^2 \text{Var}(\varepsilon_{t+k-1}) + \psi_2^2 \text{Var}(\varepsilon_{t+k-2}) + \dots$$

Now $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots$ are known and so have zero variance. Hence,

$$\text{Var}(X_{t+k}|X_{t-1}, X_{t-2}, \dots) = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_k^2) \sigma_\varepsilon^2$$

$$\text{As } k \rightarrow \infty, \text{Var}(X_{t+k}|X_{t-1}, X_{t-2}, \dots) \rightarrow (1 + \psi_1^2 + \psi_2^2 + \dots) \sigma_\varepsilon^2.$$

Calculation of the asymptotic mean and variance for an ARIMA(p,d,q) process is illustrated for p=0, q=0 and d=1 to avoid a large amount of messy calculation. The difference equation for Y_t following an ARIMA(0,1,0) process is

$$Y_t = \mu + Y_{t-1} + \varepsilon_t.$$

$$\Rightarrow Y_{t+k} = \mu + Y_{t+k-1} + \varepsilon_{t+k}.$$

$$\Rightarrow Y_{t+k} = (k+1)\mu + Y_{t-1} + \varepsilon_{t+k} + \varepsilon_{t+k-1} + \dots + \varepsilon_t.$$

Therefore,

$$\begin{aligned} E[Y_{t+k}|Y_{t-1}, Y_{t-2}, \dots] &= (k+1)\mu + Y_{t-1} \\ &\rightarrow -\infty \quad \text{for } \mu < 0, \\ &\quad Y_{t-1} \quad \text{for } \mu = 0, \\ &\rightarrow \infty \quad \text{for } \mu > 0, \text{ as } k \rightarrow \infty. \end{aligned}$$

Also,

$$\begin{aligned} \text{Var}[Y_{t+k}|Y_{t-1}, Y_{t-2}, \dots] &= \text{Var}[\varepsilon_{t+k}] + \text{Var}[\varepsilon_{t+k-1}] + \dots + \text{Var}[\varepsilon_t] \\ &= (k+1)\sigma_\varepsilon^2 \\ &\rightarrow \infty \quad \forall \mu, \text{ as } k \rightarrow \infty. \end{aligned}$$

These asymptotic results hold in general if Y_t follows an ARIMA(p,d,q) process with $d \geq 1$ since the series of ψ^2 weights does not converge. The asymptotic forecast means and standard deviations for each of the variables are shown in table 9.

In order to calculate the ψ weights for an ARMA(p,q) process, it is convenient to expand $\phi^{-1}(B)$ as an infinite series in powers of B. This is illustrated below with the bivariate process for *INFL*. If the difference equation is used, it is necessary to calculate the auto-correlation and cross-correlation functions for each variable used.

The process for *INFL* is

$$INFL_t = 0.008 + 0.899INFL_{t-1} + 0.088EQDG_t - 0.079EQDG_{t-1} + 0.077EQDG_{t-2} - 0.069EQDG_{t-3} + 0.020\varepsilon_t,$$

where $EQDG = 0.093 + 0.116\eta_t + 0.076\eta_{t-1}$, and ε_t and η_t are independent white noise processes.

Writing the process for *INFL* in terms of the white noise processes ε_t and η_t ,

$$INFL_t = [(0.08812 - 0.079217236B + 0.077494B^2 - 0.069664781B^3) * (0.11645631 + 0.075854982B) / (1 - 0.89897B)] \varepsilon_t + [0.01963711 / (1 - 0.89897B)] \eta_t.$$

Expanding this in powers of B gives

$$INFL_t = [0.01026213004 + 0.006684341063B + 0.009024665361B^2 + 0.00587830606B^3 + 0.92 * 10^{-10}B^4 + 0.827052 * 10^{-10}B^5 + \dots] \varepsilon_t + [0.1963711 + 0.01765317278B + 0.01586967273B^2 + 0.01426635969B^3 + 0.01282502937B^4 + 0.01152931666B^5 + 0.01036450979B^6 + \dots] \eta_t.$$

The variance at lead-time k is obtained by adding the variance of the ε process at lead-time k to the variance of the η process at lead-time k. The variance for each white noise process at lead-time k is obtained by summing the respective squares of the coefficients of all powers of B of order less than or equal to k.

Variable	Asymptotic Mean	Asymptotic Std Deviation
<i>EQDY</i>	1.6316	0.338
<i>EQDG</i>	0.0930	0.139
<i>PTDY</i>	2.2568	0.129
<i>PTDG</i>	0.0767	0.076
<i>Z_{PDRY}</i>	2.0732	0.074
<i>PDRYZ</i>	0.0000	0.083
<i>PDRY</i>	2.0732	0.111
<i>PDRG</i>	0.0960	0.088
<i>Z_{LINT}</i>	0.1197	0.023
<i>LINTZ</i>	0.0000	0.012
<i>LINT</i>	0.1197	0.026
<i>Z_{MINT}</i>	0.1158	0.028
<i>MINTZ</i>	-0.0005	0.029
<i>MINT</i>	0.1154	0.052
<i>INFL</i>	0.0949	0.049

Table 9. Asymptotic Statistics.

Appendix 6

The cross-correlation functions between EQDG and input EQDY using an ARIMA(0,1,0) and an ARIMA(1,0,0) prewhitening filter.

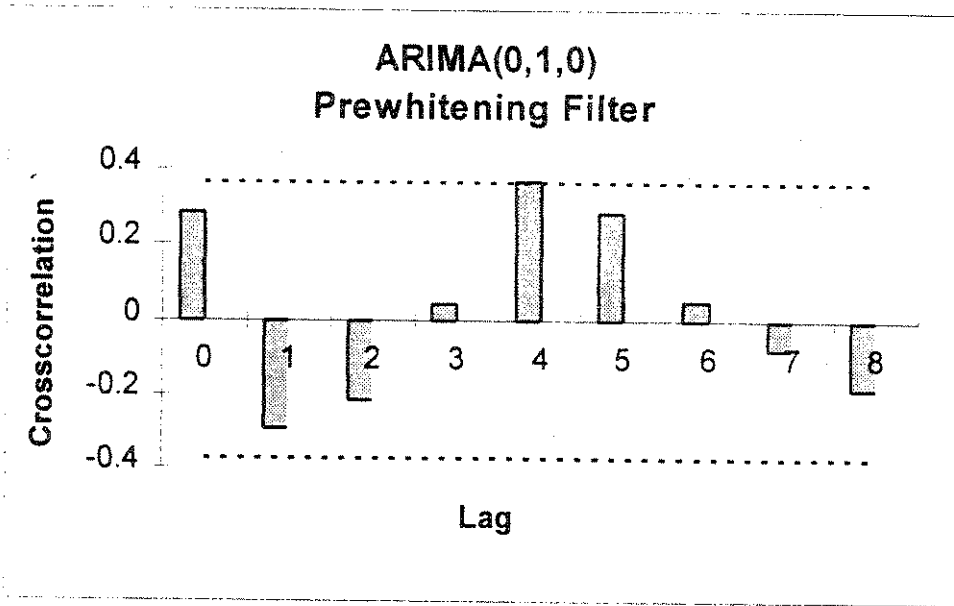


Figure 20

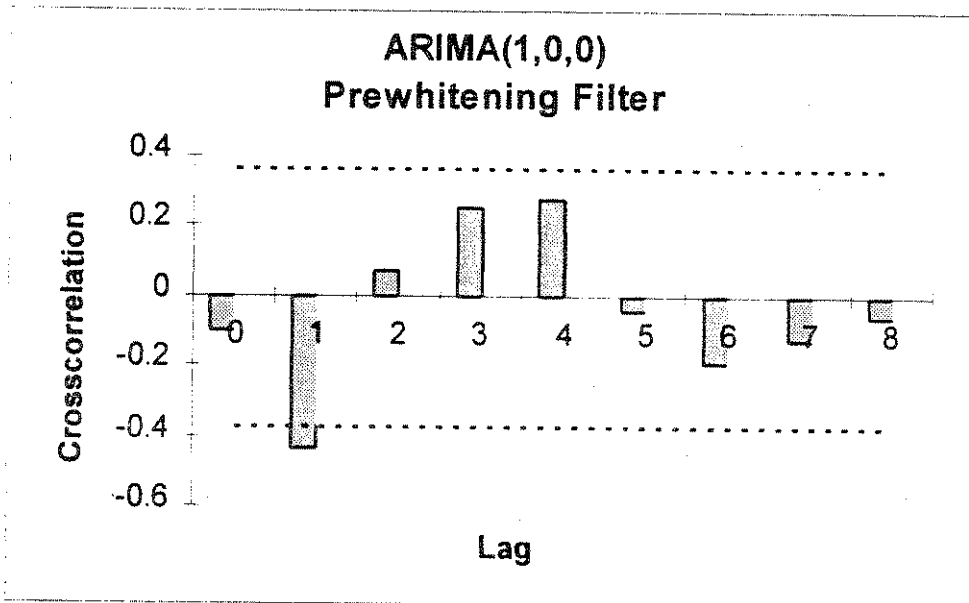


Figure 21

Appendix 7

Lemma If Y_t is modelled by an ARIMA transfer function with input X_t , where X_t is a univariate ARIMA process, then, in general, Y_t cannot be expressed as a univariate ARIMA process.

Proof

Let

$$Y_t = X_t + \varepsilon_t + \varepsilon_{t-1}, \quad \text{and} \quad X_t = \eta_t + 0.5\eta_{t-1},$$

where ε_t, η_t are i.i.d. $N(0,1)$ random variables. Then,

$$(i) \quad Y_t = \eta_t + 0.5\eta_{t-1} + \varepsilon_t + \varepsilon_{t-1}$$

Assume that Y_t can be expressed as a univariate ARIMA process. Then, by adding η_{t+k} to $\varepsilon_{t+j} \forall j=k$, we can write the process as

$$(ii) \quad Y_t = 2^{0.5} \alpha_t + (5^{0.5}/2) \alpha_{t-1}$$

where α_t i.i.d. $N(0,1), t=1,2,\dots$

Now, consider a *single* simulation of Y_{t+k} ($k=1,2,\dots$) with realisations $Y_{t+1}^*, Y_{t+2}^*, \dots$ given Y_t, Y_{t-1}, \dots obtained by simulating the values $\varepsilon_{t+1}^*, \varepsilon_{t+2}^*, \dots$ and $\eta_{t+1}^*, \eta_{t+2}^*, \dots$ given $\varepsilon_t, \varepsilon_{t-1}, \dots$ and $\eta_t, \eta_{t-1}, \dots$. From (i),

$$Y_{t+2}^* = \eta_{t+2}^* + 0.5\eta_{t+1}^* + \varepsilon_{t+2}^* + \varepsilon_{t+1}^*$$

From (ii),

$$Y_{t+2}^* = 2^{0.5} \alpha_{t+2}^* + (5^{0.5}/2) \alpha_{t+1}^*$$

where $\alpha_{t+1}^* = (\eta_{t+1}^* + \varepsilon_{t+1}^*)/(2^{0.5})$, so that

$$(5^{0.5}/2)\alpha_{t+1}^* = (10^{0.5}/4)(\eta_{t+1}^* + \varepsilon_{t+1}^*).$$

But,

$$(5^{0.5}/2)\alpha_{t+1}^* = 0.5\eta_{t+1}^* + \varepsilon_{t+1}^*.$$

Contradiction.

Endnote for Chapter 2

The paper presented in this chapter is the first (and only) paper (to date) to review of the Thomson Model. It has examined Thomson's (1996: ¶4) proposed methodology for model structure identification and found serious flaws that lead to incorrect conclusions and unnecessary structural complications, as well as inconsistent parameter estimates. However, even after the publication of Maitland (1996), it appears that these issues have not been given sufficient attention. Thomson (2004, Chapter 7: 6) states in reference to the methodology proposed by Thomson (1996: ¶4):

“... the structure of a model that has been developed in this manner needs to be redetermined from time to time ... However, the methodology for such restructuring does not have to be revisited.”

That is, Thomson (2004) recommends the model structure identification methodology proposed in Thomson (1996: ¶4), despite the fact that these flaws have been identified from a purely mathematical perspective.

It is recommended here that the original model identification procedures proposed by Box & Jenkins (1970) for the identification of ARIMAX transfer function models are more reliable, provided that the ARIMAX framework is appropriate for the variables being modelled. However, it is also suggested that this framework has limitations that make it inappropriate for use in the stochastic modelling of the variables considered. In particular, it does not allow for the modelling of feedback relationships between the variables.

The paper in this chapter has also reviewed the Thomson Model from a statistical and economic perspective. It finds that the forecast of the model are inconsistent with the history, and ascribes this in part to the assumption that an unstructured linear, mean-reverting time series model is an appropriate model to use. In Chapter 3, the assumption of a linear, mean-reverting model structure is considered with reference to alternative descriptive model structures.

There is also considerable evidence of parameter instability, and key parameters defining the structure of the Thomson Model also appear to be unstable. These issues are considered further in Chapter 3, and an alternative model that incorporates particular parameter instability is presented in Chapter 5.

Thomson (2004, Chapter 7:6) suggests that:

“As indicated in this author’s response to Maitland (*op. cit.* discussion: 481-5), it appeared that he had not adequately recognised the caveat that the decision-maker should reconsider the parameters of the model in the light of market information.”

It appears that the economic issues raised by Maitland (*ibid.*) have not been given sufficient attention. In particular, the Thomson Model does not reference the Efficient Market Hypothesis and predicts return distributions that suggest the existence of abnormal profits. There is simply no alternative set of parameters that could be chosen for the Thomson Model structure that would incorporate an efficient market, even if the decision-maker wished it so. As discussed in Section 3 of Chapter 1, the assumption of an efficient market in such models may not be an unreasonable assumption to make.

The paper presented in this chapter suggests that there are significant problems with the Thomson Model that make its continued use difficult to justify. The issues discussed raise critical questions about the usefulness of the model as a tool for projection purposes.

CHAPTER 3

Foreword to the Paper

Chapter 3 (Maitland, 1997) examines alternative descriptive models for inflation, equity dividend yields and dividend growth rates. These are the top three variables in the cascade structure of the Thomson Model, and are three of the five core variables modelled in that model.

Thomson (1996) assumed that these variables were covariance-stationary but this assumption was not tested statistically. This article formally tests these variables for unit roots against a number of alternative hypotheses. Specification and stability tests, recursive least squares tests and residual based tests, including tests for GARCH effects, are also carried out on each series. Intervention models are estimated for inflation and equity dividend yields, and a GARCH model is estimated for equity dividend growth rates.

Such a univariate analysis lays the foundation for subsequent multivariate modelling. In building a multivariate time series model, the purpose of developing univariate models for each of the variables is to guide subsequent multivariate modelling. Evidence of intervention effects, regime switching and stochastic volatility in univariate models suggests the need to consider such effects in subsequent multivariate modelling. In Chapter 5, a Multivariate Markov Switching framework is presented that considers these issues further.

NON-STATIONARITY IN SOME SOUTH AFRICAN FINANCIAL AND ECONOMIC SERIES

James Maitland

Department of Statistics and Actuarial Science
University of the Witwatersrand
Johannesburg, South Africa
Telephone: +27 11 486 1946
Facsimile: +27 088 011 486 1946
e-mail: jm007@iafrica.com

Abstract

This paper examines various forms of non-stationarity in some South African financial and economic time series useful for asset-liability modelling. It examines a number of univariate models for dealing with each type of non-stationarity and lays the foundation for a multivariate modelling of these series.

Keywords: stochastic models, non-stationary, GARCH, unit roots, markov-switching models.

1. Introduction

Thomson (1994, 1995, 1996) developed a stochastic investment model of inflation rates, short-term and long-term interest rates, dividend growth rates, dividend yields, rental growth rates and rental yields based on annual data from 1960 (or later) to 1993. The model was intended to be used for the long-term projection of assets and liabilities of a South African defined benefit pension fund.

Using the methods described by Box and Jenkins (1970), Thomson identified univariate models for each series. Analysis of the autocorrelation coefficients indicated that six out of the nine series should be differenced. These univariate models were then used to identify the transfer function cascade structure through the analysis of pre-whitened cross-correlation functions. In estimating the model's functional form and parameters, however, each variable was assumed to be stationary. In other words, no response variables were differenced in the transfer functions even though the univariate analysis indicated the need for differencing. Univariate models for variables with no inputs were revised, if necessary, by not differencing in an attempt to impose stationarity on the forecasts of the series.

This paper formally tests three of the nine series modelled by Thomson (1994, 1995, 1996) for unit roots against a number of alternative hypotheses. Specification and stability tests, recursive least squares tests and residual based tests, including tests for GARCH effects, are also carried out on each series. The analysis is based on annual data over the period 1960 to 1993. The series tested are

- *INFL*, defined as the mean force of inflation in year t ;
- *EQDY*, defined as the natural logarithm of the All Share Index Dividend Yields per cent; and
- *EQDG*, defined as the mean force of equity dividend growth in year t .

2. ARMA models

In the stochastic investment model developed by Thomson (1996), the variables *INFL*, *EQDY* and *EQDG* are at the top of the cascade structure and are modelled as follows:

$$(1) \quad EQDY_t = 0.310 + 0.810EQDY_{t-1} + 0.198\eta_{a,t},$$

$$(2) \quad EQDG_t = 0.093 + 0.116\eta_{b,t} + 0.076\eta_{b,t-1},$$

$$(3) \quad INFL_t = 0.008 + 0.899INFL_{t-1} + 0.088EQDG_t - 0.079EQDG_{t-1} + 0.077EQDG_{t-2} - 0.069EQDG_{t-3} + 0.020\eta_{c,t}$$

where $\eta_{a,t}$, $\eta_{b,t}$ and $\eta_{c,t}$ are independent and identically distributed $N(0,1)$ random variables for all t . The models were estimated by maximum likelihood after centering each series.

Each series is modelled as a stationary series. The asymptotic means for $EQDY$, $EQDG$ and $INFL$ are 1.632, 0.093 and 0.095 respectively while the corresponding asymptotic variances are 0.113998, 0.019232 and 0.002873. The input series $EQDG$ contributes 27.4% to the asymptotic variance of $INFL$. Using equations (1), (2) and (3), forecast means and confidence intervals for $EQDY$, $EQDG$ and $INFL$ are illustrated in Figures 1, 2 and 3.

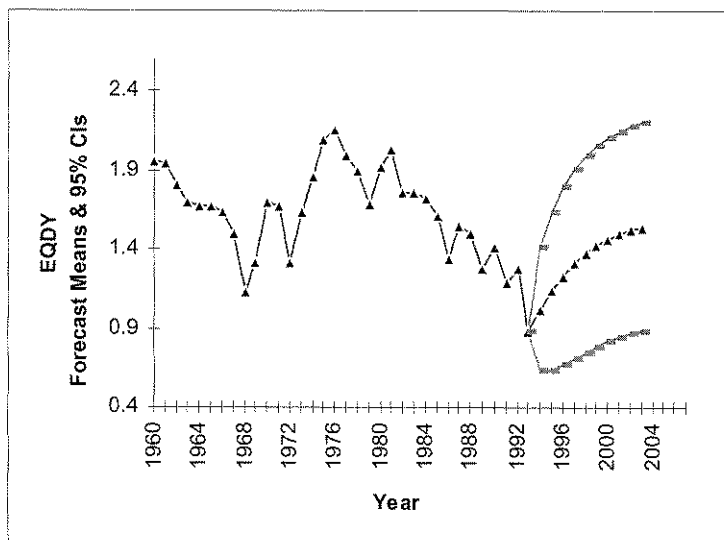


Figure 1. $EQDY$, 1960-1993, forecast means & 95% confidence intervals, 1994-2003.

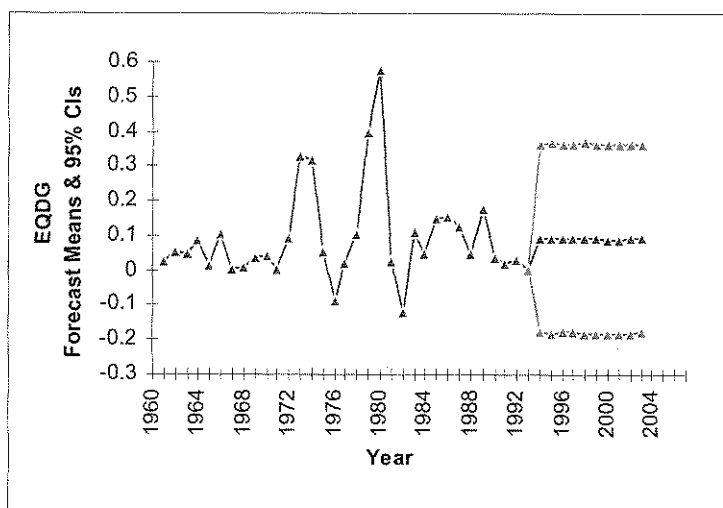


Figure 2. $EQDG$, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

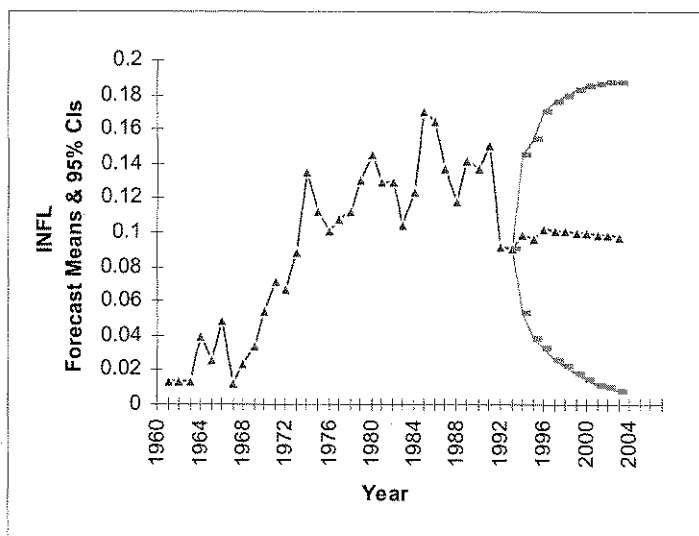


Figure 3. *INFL*, 1961-1993, and forecast means & 95% confidence intervals, 1994-2003.

Following the decision not to difference made by Thomson (1994, 1995, 1996) and using the methods described by Box and Jenkins (1970), the best univariate ARMA model for *INFL* is

$$(4) \quad INFL_t = 0.015899 + 0.853885 INFL_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim N(0; 0.022138^2)$. The model forecasts are stationary with an asymptotic mean and variance equal to 0.109 and 0.001809 respectively.

Both models for *INFL* (see equations (3) and (4)) have similar forecasts except that the forecast variance of the transfer function model (equation (3)) is larger than the forecast variance of the univariate AR(1) model. This is caused by the variance of the input variable *EQDG* and it is interesting to note that the standard deviation of the error process in equation (3) is not much less than the standard deviation of the error process in equation (4) despite the larger coefficient of $INFL_{t-1}$ in (3) compared with (4).

In building a multivariate time series model, the purpose of developing univariate models for each of the variables is to guide subsequent multivariate modelling. How best to proceed hinges on knowing whether the individual series are $I(0)$ or $I(1)$. The decision of not to difference has a profound effect on the model's long-term forecast means and confidence intervals. In fitting an ARMA model to a series, there is an implicit assumption of weak stationarity in the series: the asymptotic forecast mean and variance are constant and equal to the unconditional mean and variance. Further, each series is mean reverting so that if a variable is currently a long way from its asymptotic mean, mean reversion will affect short term forecasts as well, as illustrated by Figure 1.

On the other hand, the implication of a unit root in a time series is that shocks to the system are permanent, trends are stochastic and forecast variances increase linearly as the lead time of the forecast increases. In the section 3, each series is formally tested for unit roots.

3. Unit root tests

Under the null hypothesis that ρ equals one and for zero and non-zero values of α and β , Dickey and Fuller (1979, 1981) derive the limit distributions of the regression t test for ρ in the process $x_t = \alpha + \beta t + \rho x_{t-1} + \varepsilon_t$, when the equation being estimated is the above equation with various combinations of zero and non-zero values for α and β . $\{\varepsilon_t\}$ is assumed to be a sequence of independent normal random variables with mean zero and variance σ_ε^2 .

Said and Dickey (1984) generalise these results in the case of ARMA(p,q) error process by fitting an appropriate augmented Dickey Fuller regression. Perron (1988) shows that neither the Dickey-Fuller τ_μ statistic nor the Phillips-Perron $Z(\tau_\mu)$ statistic are capable of distinguishing a stationary process around a linear trend from a process with a unit root and drift since the rejection of a unit root is unlikely if the series is stationary around a linear trend and becomes impossible as the sample size increases. For this reason, it is advisable to begin one's tests for a unit root by fitting the equation

$$(5) \quad \nabla x_t = \alpha + \beta t + \phi x_{t-1} + \sum_{i=1}^l \theta_i \nabla x_{t-i} + \varepsilon_t,$$

where $\nabla x_t = x_t - x_{t-1}$, $\phi = (1-\rho)$, α and β are unrestricted, and, where l is less than or equal to some function of the sample size. The appropriate function of the sample size is discussed in Diebold and Nerlove (1990), Mills (1993), Said and Dickey (1984) and Schwert (1987).

Dickey and Fuller (1981) show that the statistic Φ_3 is the most powerful of the test statistics that permit the null model to contain a drift. Φ_3 is used to test the null hypothesis $(\alpha, \beta, \rho) = (\alpha, 0, 1)$ against the alternative hypothesis $(\alpha, \beta, \rho) \neq (\alpha, 0, 1)$. It should be tested prior to testing τ_t (which is used to test the null $\rho=1$ against the alternative $\rho < 1$) so that the alternative of a process which is stationary around a linear trend can be considered. For a full description of data-based, sequential testing procedures for unit roots, see Dolado, Jenkinson and Sosvilla-Rivero (1990), Mills (1993), Holden and Perman (1994) or Sherris, Tedesco and Zehnirith (1996).

The sequential testing procedure used in this paper is that of Holden and Perman (1994). This procedure is preferred to the sequential testing procedure described by Dolado, Jenkinson and Sosvilla-Rivero (1990) and Mills (1993) since it is more powerful, especially when the alternative hypothesis is a process stationary around a linear trend or an AR(1) process with a high value of ρ .¹

¹ To see this, consider the following cases:

1. τ_t lies between the non-standard τ_t critical value and the standardised normal critical value; and,
2. τ_μ lies between the non-standard τ_μ critical value and the standardised normal critical value.

Equation (5) has been fitted to each series with the appropriate value of l in each case. Parameters have been estimated by conditional maximum likelihood since these estimates are consistent even when the series is non-stationary. This is not true of the unconditional maximum likelihood estimates (Hamilton, 1994). The parameter estimates for equation (5) both with and without the trend term as well as the statistics Φ_1 , Φ_2 and Φ_3 (see Dickey Fuller (1981),) are reported in Table 1. Unless otherwise stated, all hypothesis tests are conducted at the 5% significance level.

$\nabla X_t =$	α	β	ϕ	θ_1	Φ_1	Φ_2	Φ_3
(1) <i>INFL</i>	0.014 (0.009)	0.001 (0.001)	-0.272 (0.160)			1.5237	2.0976
(2) <i>INFL</i>	0.016 (0.008)		-0.146 (0.079)		2.3533		
(3) <i>EQDY</i>	0.421 (0.257)	-0.004 (0.004)	-0.229 (0.137)			1.3768	1.6162
(4) <i>EQDY</i>	0.264 (0.220)		-0.179 (0.131)		1.3987		
(5) <i>EQDG</i>	0.075 (0.051)	0.001 (0.002)	-0.945 (0.193)	0.489 (0.170)		8.0661	12.0961
(6) <i>EQDG</i>	0.090 (0.028)		-0.936 (0.188)	0.481 (0.166)	13.248		

Table 1. Parameter estimates and statistics for equation (5) with standard errors in brackets below.

Line (1) in Table 1 shows the results of fitting equation (5) to *INFL* with a time trend; lagged first differences are not required. Φ_3 equals 2.0976 which is less than 5.91, the 10% critical value for a sample of size 25 given in Table VI by Dickey and Fuller (1981). Thus, we cannot reject $H_0: (\alpha, \beta, \rho) = (\alpha, 0, 1)$; there is no significant trend and the series possibly has a unit root with or without drift.

To support the conclusion that $\rho=1$ and since β can be assumed to be zero, we test τ_τ against the non-standard distribution described by Fuller (1976), Table 8.5.2. This table only gives critical values for a limited number of sample sizes however MacKinnon (1991) estimates response surface regressions in which critical values depend on the sample size. $\tau_\tau = -1.706450$ which is greater than the MacKinnon 10% critical value of -3.2109 so we cannot reject the null hypothesis of a unit root at this stage. This confirms the results of our previous test.

To establish whether the series has a non-zero drift, we carry out the F-test Φ_2 described in Dickey and Fuller (1981). This tests $H_0: (\alpha, \beta, \rho) = (0, 0, 1)$ against $H_1: (\alpha, \beta, \rho) \neq (0, 0, 1)$ but since we can assume that $\beta=0$ and $\rho=1$, it is effectively a test of $\alpha=0$. For *INFL*, Φ_2 equals 1.5237 which is less than the 10% critical value of 4.67 given by Dickey and Fuller (1981), Table V. We cannot reject the null hypothesis and so α is not significantly different from zero.

The power of the unit root test is improved by dropping the trend term in equation (5) (Holden and Perman, 1994); the results are shown in line (2) of Table 1. If $\alpha \neq 0$, τ_μ has a standardized normal distribution; otherwise, it has the non-standard distribution described in Fuller (1976), Table 8.5.2, middle panel, and MacKinnon (1991). For *INFL*, τ_μ equals -1.838761 which is greater than the MacKinnon 10% critical value of -2.6164 under the assumption that $\alpha=0$ so that we cannot reject the null hypothesis of a unit root.

To re-establish our conclusions regarding ρ and α , we carry out the F-test Φ_1 described in Dickey and Fuller (1981) which tests $H_0: (\alpha, \rho) = (0, 1)$ against $H_1: (\alpha, \rho) \neq (0, 1)$. For *INFL*, Φ_1 equals 2.3533 which is less than the 10% critical value of 4.12 given by Dickey and Fuller (1981), Table IV. Dropping both the intercept and time trend terms in equation (5) gives $\tau = -0.309757$ which is greater than the MacKinnon 10% critical value of -1.6213, confirming our previous results. Unit root tests on the differenced series are not reported here but confirm that no further differencing is required.

Based on samples of size 100, the empirical power of the tests Φ_3 , τ_τ , Φ_2 , τ_μ and Φ_1 for various values of α and ρ is given by Dickey and Fuller (1981), Table VII. With $\rho=0.8$ and $\alpha=0$, the empirical powers are 0.57, 0.46, 0.41, 0.71 and 0.78 respectively. It is clearly well worth re-establishing one's conclusions regarding $H_0: (\alpha, \rho) = (0, 1)$ by estimating equation (5) with β restricted to zero.

It should be noted that although the estimate of the intercept, α , in line (2) of Table 1 is almost significantly different from zero (with a p-value of 0.0655), the model being estimated is in fact the mean reverting model (since $\rho < 1$) so that α is actually a measure of the mean level of the series. When ρ is set equal to one, α measures the drift and is not significantly different from zero (with a p-value of 0.5578).

The preceding results suggest the model

$$(6) \quad INFL_t = INFL_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim N(0, 0.022972^2)$. That is, the mean force of inflation in year t is a random walk with zero drift. The time trend is stochastic and the forecast variance increases linearly with the lead time.

The results of fitting equation (5) to *EQDY* with no lagged differences are shown in lines (3) and (4) of Table 1. Φ_3 is not significant and τ_τ , which equals -1.668810, is also greater than the MacKinnon 10% critical value of -3.2081. There is no evidence against a unit root and we proceed by dropping the trend term. Φ_2 is not significant, nor is $\tau_\mu = -1.369578$ and Φ_1 is considerably less than the 10% critical value suggesting that the model $\nabla EQDY_t = \varepsilon_t$, $\varepsilon_t \sim N(0; 0.156381^2)$, is not inappropriate.

A quick glance at the series in Figure 1 however suggests that, at least after 1975, $EQDY$ is trended. The results of fitting equation (5) to $EQDY$ for the period 1975 to 1993 confirm this suspicion. Φ_3 equals 10.5742 which is significant at the 2.5% level. We are left with the following alternatives:

$$\begin{array}{ccc} \beta \neq 0 & \text{or} & \beta = 0 \\ \rho = 1 & & \rho \neq 1 \end{array} \quad \text{or} \quad \begin{array}{ccc} \beta \neq 0 & & \beta = 0 \\ \rho \neq 1 & & \rho = 1 \end{array}$$

We then test $\rho=1$ using $\tau_\tau \sim N(0,1)$. Since τ_τ equals -4.247466 which is significant, we are left with the two possibilities:

$$\begin{array}{ccc} \beta = 0 & \text{or} & \beta \neq 0 \\ \rho \neq 1 & & \rho = 1 \end{array}$$

In either case $\rho \neq 1$ so conventional test procedures can be used. We thus fit the model

$$(7) \quad EQDY_t = \alpha + \beta t + \rho EQDY_{t-1} + \varepsilon_t$$

and test the null hypothesis that $\beta=0$. β is significant but ρ is not. When the term $EQDY_{t-1}$ is dropped, the equation becomes

$$(8) \quad EQDY_t = 3.059 - 0.057t + \varepsilon_t$$

(0.136) (0.005)

where $\varepsilon_t \sim N(0;0.126935^2)$. Both terms are highly significant and this model seems to fit well over the sub-period 1975 to 1993. The year 1975 was chosen after inspection of the data and is thus highly correlated with the data. Christiano (1992) criticizes this method of choosing the break point since the finite sample and asymptotic distributions of the statistics depend on the extent of the correlation between the choice of breakpoints and the data, Perron (1994). The effects of structural breaks is considered in more detail in the next section.

Lines (5) and (6) in Table 1 shows the results of fitting equation (5) to $EQDG$ with and without the trend term respectively. Both Φ_3 and τ_τ , which equals -4.905872, are significant at the 1% level. However, since we can reject the null hypothesis that $\rho=1$, $EQDG$ does not have a unit root and we can use conventional test procedures to fit a model. The results of fitting the MA(1) model are discussed in section 6 together with tests for GARCH effects.

The tests described by Phillips and Perron (1988) are more appropriate than the augmented Dickey-Fuller tests if normality, autocorrelation or heterogeneity is a problem. Phillips (1987) shows that non-parametric test statistic $Z(\tau_\mu)$ has the same asymptotic power as τ_μ . However, if the underlying process has a unit root and a moving average structure, $\varepsilon_t - \theta\varepsilon_{t-1}$, $\theta > 0$, the finite sample properties are markedly worse for $Z(\tau_\mu)$ than for τ_μ . Agiakloglou (1996) shows that there are similar problems with the statistic proposed by Bierens (1993) in that the probability of a

type I error is much larger than suggested by the critical level of the test. If $\theta < 0$, the power of the $Z(\tau_\mu)$ statistic is greater than the power of the τ_μ statistic.

If one fits an ARIMA(0,1,1) model to the series, θ is negative suggesting that the Phillips Perron tests might be more powerful than the Dickey Fuller tests. Further, since the residuals of the MA(1) process and the series for *EQDG* are both found to be heterogeneous, the Phillips Perron tests are more appropriate. However, since the results of the Phillips Perron tests on *EQDG* are no different from the augmented Dickey Fuller test results, they are not reported here.

None of the Phillips-Perron tests were significant when applied to either *INFL* or *EQDY* supporting the results of the Dickey-Fuller tests reported above. There was, however, no evidence in either series of non-normality, autocorrelation or heterogeneity against the use of the Dickey-Fuller tests and these results are just mentioned in passing.

The purpose of the desired stochastic investment model is the long-term projection of assets and liabilities. Thus, both the short and long-run forecasts are important. The variance of the long-term forecasts resulting from the random walk model, however, casts doubt on the validity of a unit root in the models for *EQDY* and *INFL*. It is for this reason that Thomson (1994, 1995, 1996) does not difference either univariate series or response variables. Mean reversion in short-term forecasts, however, results in excessively high initial returns for property trusts and extremely low initial returns for equities in the Thomson model, as shown by Maitland (1996). This is equally undesirable and ARMA models do not provide the answer. One solution to the problem of a long term equilibrium when each series is I(1) is the use of cointegration but it is not the intention of this paper to identify cointegrating relationships and cointegration is not discussed further in this paper.

4. Interventions

The oil shocks of the 1973 and 1979 lead to a period of high, entrenched inflation in the 80's with inflation remaining above 10%. Chow (1960) suggests a forecast test and a breakpoint test to test for a structural change in parameters. The likelihood ratio statistic for the forecast test using *INFL* up to and including 1973 to forecast *INFL* from 1974 to 1993 is 51.0, with a p-value of 0.000264. The likelihood ratio statistic for the breakpoint test with a break point at the end of 1973 is 8.015, with a p-value of 0.018177. Both of these tests indicate the presence of a structural break in 1973.

Recursive estimates of the parameters are illustrated in Figures 4a and 4b and also indicate the presence of a structural break at the end of 1973. The structural break is caused by the change in the level of inflation following the 1973 oil shock and invalidates the assumption of weak stationarity. The results of fitting an AR(1) process are an AR(1) parameter which is too high, an arbitrary long-run mean of 9.5% and a forecast variance which is too large.

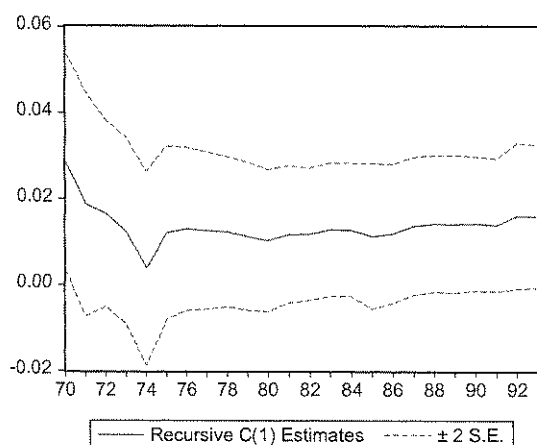


Figure 4a. Recursive estimates of the intercept in equation (3).

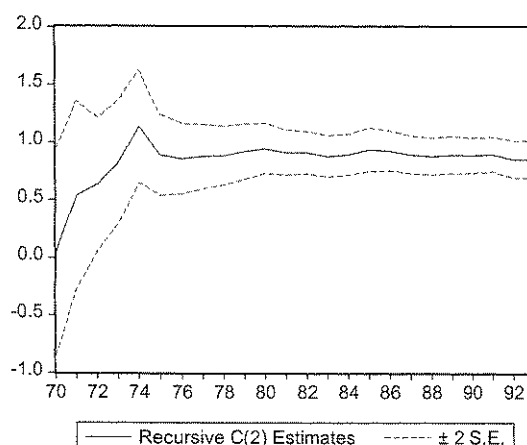


Figure 4b. Recursive estimates of the coefficient of $INFL_{t-1}$ in equation (3).

Box and Tiao (1975) introduce the concept of an intervention so that ‘aberrant’ events can be separated from the noise function and modelled as a separate component in the deterministic part of the general time series model. They model the intervention using either the “additive outlier” or “innovational outlier” model. In the additive outlier model, the transition caused by the intervention occurs instantaneously while in the innovational outlier model the transition between one regime and the next is gradual.

The general form of the intervention model can be expressed (see SAS/ETS (1993)) as the ARIMAX model

$$(9) \quad W_t = \mu + \sum_i \frac{\omega_i(B)}{\delta_i(B)} B^{k_i} X_{i,t} + \frac{\theta(B)}{\phi(B)} \varepsilon_t$$

where

- t is an element of $\{1, 2, \dots, T\}$ and indexes time;
- W_t is the response series;
- μ is the mean term;
- B is the backshift operator;
- $\phi(B)$ is the autoregressive operator;
- $\theta(B)$ is the moving average operator;
- $X_{i,t}$ is the i^{th} input variable at time t;
- k_i is the pure time delay for the effect of the i^{th} input series;
- $\omega_i(B)$ is the numerator polynomial for the transfer function of the i^{th} input series; and,
- $\delta_i(B)$ is the denominator polynomial for the transfer function of the i^{th} input series.

$X_{i,t}$ can either be another series or an intervention variable taking the form of a pulse function, $D(TB)_t$, a step function, DU_t , or broken trend function, DT_t , where

$$\begin{aligned} D(TB)_t &= 1 \text{ if } t = T_B, 0 \text{ otherwise;} \\ DU_t &= 1 \text{ if } t > T_B, 0 \text{ otherwise;} \\ DT_t &= t \text{ if } t > T_B, 0 \text{ otherwise;} \end{aligned}$$

and where T_B refers to the time of intervention.

In the case of a single intervention, if either $\phi(B)$, $\omega(B)$ or $\delta(B)$ are polynomials of degree greater than or equal to 1, the intervention is equivalent to the innovational outlier model; otherwise, the additive outlier model describes the intervention.

5. Unit root tests revisited

Perron (1989) shows that standard unit root tests which do not allow for the presence of a structural break have little power against the alternate of no unit root when the underlying series has a structural break but no unit root. The power of these tests decreases as the magnitude of the intervention variables increases. Perron extends the augmented Dickey Fuller regression to include a pulse term, a step term and a broken trend term as follows:

$$(10) \quad \nabla x_t = \alpha + \xi DU_t + \beta t + \gamma DT_t + \delta D(TB)_t + \phi x_{t-1} + \sum_{i=1}^l \theta_i \nabla x_{t-i} + \varepsilon_t$$

Under the null hypothesis, $\phi=0$ and $\gamma=\beta=0$; δ measures a one-time shock to the system which persists because of the unit root and ξ measures the change in drift following the intervention. Under the alternative hypothesis of a process stationary about a broken mean level or trend, ρ is less than one and at least one of β , ξ and γ is non-zero while δ should be close to zero.

The results of estimating equation (10) for *INFL* with an intervention in 1973 are shown in Table 2, line (1). Perron (1989), Table VI.B, provides critical values for the t-statistic from the least squares estimation of (10) under the null hypothesis of a unit root. With $\lambda \approx 0.4$ where $\lambda=T_B/T=13/33$ is the "break fraction" (see Perron (1989)), the t-statistic of -3.426504 is greater than the 10% critical value of -3.95. At this stage, we cannot reject the null hypothesis of a unit root process but an examination of the series for *INFL* suggests we should be estimating equation (10) with β and γ restricted to zero. This gives a more powerful test when the series in each regime are not trended or are without drift.

	α	ξ	β	γ	δ	φ	σ_ε
(1) <i>INFL</i>	-0.00116 (0.01633)	0.039785 (0.01909)	0.003465 (0.00208)	-0.00360 (0.002123)	0.019913 (0.02378)	-0.67559 (0.19717)	0.0199
(2) <i>INFL</i>	0.023018 (0.00828)	0.048736 (0.01811)	0	0	0.036873 (0.02200)	-0.56555 (0.17389)	0.0202
(3) <i>EQDY</i>	1.31497 (0.34375)	0.654388 (0.171344)	-0.02863 (0.014804)	-0.01165 (0.014286)	0.327018 (0.182049)	-0.70321 (0.166237)	0.1549
(4) <i>EQDY</i>	1.397165 (0.326676)	0.698910 (0.161441)	-0.03843 (0.008591)	0	0.388980 (0.164443)	-0.70849 (0.165115)	0.1539
(5) <i>EQDY</i>	1.457725 (0.350525)	0.643609 (0.171930)	-0.03678 (0.009216)	0	0	-0.73560 (0.177287)	0.1657

Table 2. Parameter estimates for equation (10) with standard errors in brackets below.

Under the null hypothesis of a unit root, Perron (1989), Table IV.B, provides critical values for the t-statistic of φ from the least squares estimation of (10) with γ restricted to zero. For *INFL*, the results of estimating equation (10) with $\beta=\gamma=0$ are shown in Table 2, line (2). The t-statistic of -3.252370 is greater than the 10% critical value of -3.44. Despite the fact that $\varphi=-0.56555$, the t-statistic is not large enough to reject the null hypothesis that $\varphi=0$ due to the small sample size. Nonetheless, it is worth fitting the intervention model assuming no unit roots and the results are as follows:

With $T_B = 13$ (1973-1960), the appropriate intervention model for *INFL* is the innovational outlier model

$$(11) \quad INFL_t = 0.023275 + 0.038309 DU_t + 0.514727 INFL_{t-1} + \varepsilon_t$$

(0.008533) (0.017528) (0.172309)

where $\varepsilon_t \sim N(0, 0.020864^2)$. Standard errors of the parameter estimates are shown in brackets below each estimate. All the parameters are significant at the 5% level.

After including the intervention term, the ϕ_1 parameter reduces to 0.514727 which is considerably lower than the estimate of 0.853885 in the univariate AR(1) model given in equation (4). The unconditional forecast variance within any one regime is 0.000592 compared with 0.001809 in (4), and the mean level of the series reflects the intervention rather than a weighted average of the mean levels before and after the intervention.

The results of estimating equation (10) for *EQDY* are shown in Table 2, line (3). The t-statistic for φ is -4.230129 which is less than the 5% critical value of -4.22 given by Perron (1989), Table VI. We are thus able to reject the null hypothesis of a unit root and conventional test procedures can be used. Since the parameter γ is not significant, there is no significant change in the trend of *EQDY* and to re-establish our conclusions regarding φ , we re-estimate equation (10) with γ restricted to zero. The results are shown in Table 2, line (4).

The t-statistic is -4.290899 which is less than the 2.5% critical value of -4.01 given in Perron (1989), Table IV, confirming our previous conclusion that the series does not have a unit root. Using standard testing procedures, all the coefficients are significant and the model appears to fit well. However, the term $D(TB)_t$ represents a single outlier and not a structural break. Removing this term results in the model whose coefficients appear in line (5) of Table 2. All the coefficients are highly significant and this model should be preferred above the model in line (4), Table 2, since it is more parsimonious and does not require the modelling of outliers when forecasting.

The intervention model (10) is a more appropriate description of the *INFL* and *EQDY* series in the period 1960 to 1993 than the AR(1) models but cannot be used directly for forecasting. The intervention date was chosen ex-ante and was not modified ex-post. It is unreasonable however to assume that changes to the structural components of *INFL* or *EQDY* are in anyway deterministic and, in order to forecast these series, α needs to be modelled as a random variable. The distinction between this random variable describing the intercept of the series and the random error sequence $\{\varepsilon_t\}$ is that changes in regime are relatively rare.

Hamilton (1993a, 1994) describes a Markov-switching model in which the parameters of an ARMA model in regime s_t are the outcome of an unobserved N-state Markov chain and are independent of the random error sequence $\{\varepsilon_t\}$. Diebold, Lee and Weinbach (1994) generalise this model to a model in which the probability of switching from one regime, s_t , to the next, s_{t+1} , can depend on a vector of observed variables as well as s_t . Garcia and Perron (1996) apply the Markov-switching model to the U.S. real interest rate from 1961 to 1986. Estimating a Markov-switching model for the mean level in *INFL*, however, would lead to a two-state Markov chain in which state two is an absorbing state. This is considered to be unrealistic and it is felt that, for the purposes of forecasting, an N-state Markov chain could be used with the number of states, N, the mean level of each state and the transition probabilities based on future expectations. This topic is extensive particularly in the framework of multivariate time series and is not discussed further in this paper.

6. GARCH effects

The residuals for the MA(1) model given in equation (2) are shown in Figure 5 together with the actual series and the one-step-ahead forecasts.

The Jarque-Bera statistic (see Jarque and Bera (1981)), is 3.241433 with a p-value of 0.197757 indicating that the residuals are not significantly non-normal. The Breusch-Godfrey serial correlation Lagrange multiplier test (see Breusch and Pagan (1978), Godfrey (1978)) shows no evidence of residual correlation. However, the correlogram of squared residuals shows significant spikes in the ACF and PACF at lag 1 and the Ljung-Box Q-statistic at lag 2 is 7.7103 with a p-value of 0.005 indicating significant serial correlation in the squared residuals. With one lag, the ARCH LM test developed by Engle (1982) has an F-statistic of 9.767517 (p-value = 0.0039) and a TR^2 of 7.8597 (p-value = 0.0051). This indicates the presence of autoregressive

conditional heteroscedasticity and is not surprising given the correlogram of squared residuals and the volatility clustering evident in the residuals in Figure 5.

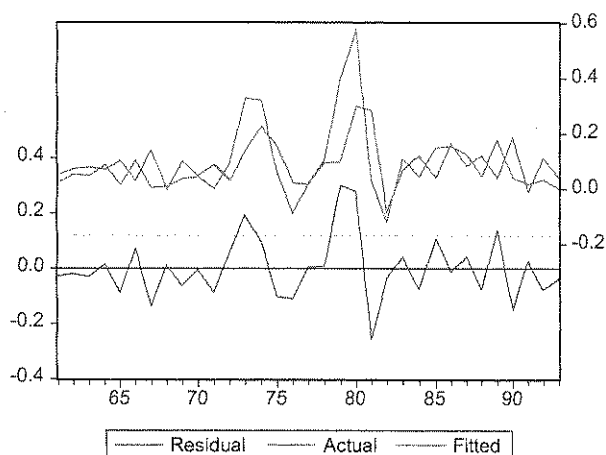


Figure 5. *EQDG*, 1961-1993, one-step-ahead forecasts and residuals for the MA(1) process.

Standard ARIMA models are designed to model only the mean of a series and assume that the variance is constant. In the GARCH model introduced by Engle (1982) and Bollerslev (1986), the variance is not constant but is conditional on previous squared residuals and previous estimates of the variance. Thus, the GARCH model allows one to model both the mean of a series and its variance. It is found that the MA(1)-GARCH(1,1) is a good description of the process underlying *EQDG*. The variance and mean equations for *EQDG* are given in equations (12a) and (12b) respectively.

$$(12a) \quad \sigma_t^2 = 0.000442 - 0.187215\varepsilon_t^2 + 1.198413\sigma_{t-1}^2$$

$$(0.00044) \quad (0.075028) \quad (0.072124)$$

$$(12b) \quad EQDG_t = 0.068119 + \varepsilon_t + 0.624301\varepsilon_{t-1}$$

$$(0.03127) \quad (0.134063)$$

where $\varepsilon_t \sim N(0; \sigma_t^2)$.

7. Conclusion

This paper has examined only three of the nine series used by Thomson (1994, 1995, 1996) in developing a South African stochastic investment model but it is felt that the results presented illustrate a number of the shortcomings of traditional Box-Jenkins time series models.

In order to adequately model long-term relationships within the framework of Box-Jenkins ARIMAX models, Thomson was forced to avoid differencing. One of the consequences of imposing weak stationarity on the model is that series are mean reverting and give rise to undesirable short term dynamics. If the series are in fact mean reverting, it would be possible to make large profits by judiciously buying and selling when the series are not equal to their unconditional means. This goes against the efficient market hypothesis. The purpose of a stochastic investment model is the long-term projection of assets and liabilities for which both short and long term forecasts are relevant. This suggests that a cointegration model might be a more appropriate tool for stochastic investment modelling.

The assumption of weak stationarity has a few more drawbacks. Firstly, because of the change in the level of inflation following the 70's oil shocks, both the transfer function model and the univariate model for *INFL* have an arbitrary long-run mean. The variance of the process is over-estimated, being much larger than the variance of *INFL* in either of the subperiods surrounding the intervention. Further, the importance of the previous year's inflation is over-estimated. For the *EQDY* series, the variance of the AR(1) model is unstable because the series is trended. The MA(1) model for *EQDG* assumes that the variance of dividend growth is constant. This is clearly not the case and results in forecast variances for equity returns which are larger than historical variances for most of the time. The important feature captured by the MA(1)-GARCH(1,1) model is the fluctuating variance of the growth rate. This has important consequences for short term risk management and the determination of excess volatility in the equity market.

This research lays the foundation for the multivariate modelling of these series and the nature of the relationships which should be built into a multivariate model. It also highlights the need to include time varying parameters into a multivariate model. These issues are critical in understanding the inter-relationships between the series and building a more realistic stochastic investment model for actuarial use in South Africa.

References

- Agiakloglou, C. (1996), 'A note on the unit root test based on the sample autocorrelations,' *The University of Piraeus Journal of Economics, Business, Statistics and Operations Research*, **46**, 97-106.
- Bierens, J.H. (1993), 'Higher-order sample autocorrelations and the unit root test,' *Journal of Econometrics*, **57**, 137-160.
- Box, G.E.P. and Jenkins, G.M. (1970), *Time series analysis, forecasting and control*, Holden Day, San Francisco.

Box, G.E.P. and Tiao, G.C. (1975), 'Intervention analysis with applications to economic and environmental problems,' *Journal of the American Statistical Association*, **70**, 70-79.

Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, **31**, 307-327.

Breusch, T.S. and Pagan, A.R. (1978), 'A simple test for heteroscedasticity and random coefficient variation,' *Econometrica*, **46**, 1287-1294.

Chow, G. (1960), 'Tests for equality between sets of coefficients in two linear regressions,' *Econometrica*, **28**, 591-605.

Christiano, L.J. (1992), 'Searching for break points in GNP,' *Journal of Business and Economic Statistics*, **10**, 237-250.

Dickey, D.A. and 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.

Dickey, D.A. and Fuller, W.A. (1981), 'Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root,' *Econometrica*, **49**, 1057-1072.

Diebold, F.X. and Nerlove, M. (1990), 'Unit roots in economic time series: a selective survey,' in Rhodes, G.F. and Fomby, T.B. (eds), *Advances in Econometrics*, Vol. 8. Greenwich, CT: JAI Press, 3-69.

Diebold, F.X.; Lee, J-H.; and Weinbach, G.C. (1994), 'Regime switching with time-varying transition probabilities,' in C.Hargreaves (ed.), *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press.

Dolado, J.J.; Jenkinson, T.; and Sosvilla-Rivero, S. (1990), 'Cointegration and unit roots,' *Journal of Economic Surveys*, **4**, 249-273.

Engle, R.F. (1982), "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," *Econometrica*, **50**, 987-1007.

EViews (1994), *User's Guide*, Quantitative Micro Software, Irvine, California, USA.

Garcia, R. and Perron, P. (1996), 'An analysis of the real interest rate under regime shifts,' *The Review of Economics and Statistics*, **78**, 111-125.

Godfrey, L.G. (1978), 'Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables,' *Econometrica*, **46**, 1293-1302.

Hamilton, J.D. (1994), *Time Series Analysis*, Princeton University Press.

- Hamilton, J.D. (1993a), 'Estimation, inference and forecasting of time series subject to changes in regime,' in G.S.Maddala; C.R.Rao; and, H.D.Vinod (eds), *Handbook of Statistics*, Vol. 11. New York: North-Holland.
- Holden, D. and Perman, R. (1994), 'Unit roots and cointegration for the economist,' in Rao, B.B. (ed.), *Cointegration for the Applied Economist*, MacMillan, London.
- Jarque, C.M. and Bera, A.K. (1981), 'An efficient large sample test for normality of observations and regression residuals,' *Working papers in Economics and Econometrics*, No. 47, Australian National University.
- Maitland, A.J. (1996), 'A review of Thomson's stochastic investment model,' *Transactions of the Actuarial Society of South Africa*, forthcoming.
- MacKinnon, J.G. (1991), 'Critical values for cointegration tests,' in Engle, R.F. and Granger, C.W.J. (eds), *Long-Run Economic Relationships: Readings in Cointegration*, Oxford University Press.
- Mills, T.C. (1993), *The econometric modelling of financial time series*, Cambridge University Press.
- Perron, P. (1988), 'Trends and random walks in macroeconomic time series: further evidence from a new approach,' *Journal of Economic Dynamics and Control*, **12**, 297-332.
- Perron, P. (1989), 'The Great Crash, the Oil Price Shock and the Unit Root Hypothesis,' *Econometrica*, **57**, 1361-1401.
- Perron, P. (1993), 'The Great Crash, the Oil Price Shock and the Unit Root Hypothesis: erratum,' *Econometrica*, **61**, 248-249.
- Perron, P. (1994), 'Trend, unit root and structural change in macroeconomic time series,' in Rao, B.B. (ed.), *Cointegration for the Applied Economist*, MacMillan, London.
- Phillips, P.C.B. (1987), 'Time series regression with a unit root,' *Econometrica*, **55**, 227-301.
- Phillips, P.C.B. and Perron, P. (1988), 'Testing for a unit root in time series regression,' *Biometrika*, **75**, 335-346.
- Said, S.E. and Dickey, D.A. (1984), 'Testing for unit roots in autoregressive-moving average models of unknown order,' *Biometrika*, **71**, 599-607.
- SAS/ETS (1993), *User's Guide. Version , second edition*. Copyright 1993 by SAS Institute Inc., Cary, NC, USA.

Sherris, M.; Tedesco, L.; and Zehnirith, B. (1996), 'Stochastic investment models: unit roots, cointegration, state space and GARCH models for Australian data,' *Research Paper Series*, Centre for Actuarial Studies, University of Melbourne, 33.

Schwert, G.W. (1987), 'Effects of model specification on tests for unit roots in macroeconomic data,' *Journal of Monetary Economics*, 20, 73-105.

Stock, J.H. (1993), 'Unit roots and trend breaks,' in Engle, R.F. and McFadden, D. (eds) *Handbook of Econometrics*, Amsterdam: North-Holland, Vol. 4.

Thomson, R.J. (1994), 'A stochastic investment model for actuarial use in South Africa,' Presented to the 1994 Convention of the Actuarial Society of South Africa.

Thomson, R.J. (1995), 'A methodology for the modelling of interest rates and other economic variables with reference to the money and capital markets of South Africa,' *Proceedings of the 5th AFIR Colloquium*.

Thomson, R.J. (1996), 'Stochastic investment modelling: the case of South Africa,' *British Actuarial Journal*, 2, 765-801.

Endnote for Chapter 3

In a recent article on mean reversion, Asher (2007) suggests that Thomson (1996) finds significant evidence of mean reversion of dividend yields in South Africa. However, the non-stationary models estimated in this chapter are shown to provide a better description of the data and suggest that there is little evidence for mean reversion in South African dividend yields.

The models identified in this chapter indicate the presence of non-linearity in the series modelled, and suggest that such effects be considered in the multivariate modelling of those series. However, the models make no reference to the Efficient Market Hypothesis and in no way consider appropriate multivariate structure for the joint modelling of these variables. The reader is referred to Chapter 7 for a discussion of the joint modelling of equity dividend yields and dividend growth rates in relation to the Efficient Market Hypothesis.

The intervention analysis discussed in Sections 4 and 5 assumes additional knowledge of a break point, and uses the oil shock of 1973 as the exogenous break point event. However, such a model requires knowledge of such exogenous events, and this together with the assumption of a single structural break is unsatisfactory as a probability law for projection purposes. Clearly, a model that allows for multiple structural breaks at unknown points in time is more useful for projection purposes.

The intervention model for equity dividend yields is particularly unsatisfactory as a model for projection purposes. There is no reason to assume that the downward linear trend should continue into the future or if the level of dividend yields might affect the probability of an intervention in future. In fact, by design the model says nothing about the probability of future interventions, which is clearly unsatisfactory for projection purposes.

Clearly, the models presented in this chapter are only intended as descriptive models and should not be used for projection purposes. However, they do suggest the need for

considering intervention effects and regime switching, as well as stochastic volatility. Such a univariate analysis lays the foundation for subsequent multivariate modelling and, in Chapter 5, a Multivariate Markov Switching framework is presented that considers these issues further.

The next chapter considers short and long-term interest rates, which are the remaining two of the five core variables in the Thomson Model.

CHAPTER 4

Foreword to the Paper

While the previous chapter considered three of the five core variables in the Thomson Model, this chapter considers short- and long-term interest rates, which are the remaining two of the five core variables in that model.

The paper in Chapter 4 (Maitland, 2002) uses Principal Components Analysis to determine the dimension of randomness in the yield curve. It proposes a methodology for estimating the full yield curve using a smaller number of yields from that yield curve, thereby reducing the number of yields required to estimate the full yield curve. Dimension reduction facilitates the model building process and assists in the development of stochastic models in which other asset categories and economic variables are considered.

The risk factors in the South African term structure of interest rates suggest that two key interest rate terms can be used to model most of the variability in yields. These two terms correspond to the two interest rate terms modelled by Thomson (1996), and provide justification for using those terms in a stochastic model for asset and liability modelling.

The two key terms identified in this chapter are used in the model presented in Chapter 5. In Chapter 6, a new framework for the immunization of nominal liabilities to Principal Component risk factors is presented.

INTERPOLATING THE SOUTH AFRICAN YIELD CURVE USING PRINCIPAL-COMPONENTS ANALYSIS: A DESCRIPTIVE APPROACH

By AJ Maitland

ABSTRACT

A principal-components analysis of the South African yield curve suggests that two factors explain most of the variability in both yields and changes in yields. This result is used to select which two interest rates to model and, given a model for these rates, how to use them to reproduce the entire curve. The objective of this paper is a methodology for interpolating the South African yield curve given a restricted number of yields on that curve, while at the same time minimising the number of yields from which to estimate the remainder of the curve. The interpolated curve can then be used for the purposes of discounting nominal future cash flows. Given values for the selected yields, this methodology provides the best fit to the remainder of the curve in the sense that it minimises the expected root-mean-squared error of the residuals. The paper does not provide a model for the evolution of the yield curve.

KEYWORDS

Principal components; par yield curve; descriptive yield-curve models; interpolation; South Africa

CONTACT DETAILS

James Maitland, Stochastic Solutions cc, P.O. Box 2484, Parklands, 2121, Johannesburg;
Tel: (0)11 486-1946; E-mail: stochsol@iafrica.com

1. INTRODUCTION

1.1 It is common practice in pension fund valuations for the actuary to use a single valuation rate of interest in calculating the present value of future asset and liability cash flows. The valuation rate is often assumed to be the long-term rate of interest that can be earned on future investments (Lee, 1986). Conveniently, this approach allows the use of standard actuarial commutation functions.

1.2 More recent application in asset and liability modelling (ALM) studies makes use of the simulated long-bond rate (20-year JSE-Actuaries Bond Yield) in calculating the valuation rate of interest. This provides a proxy for the market rate of interest and gives liability values closer to the market price than the traditional use of a constant discount rate. However, for nominal liabilities where the term of each cash flow is known, the long-bond rate can be a poor approximation to the term structure of interest rates and may give a present value quite different from the cost of the matching portfolio. According to the principle of no arbitrage, these two values should be identical (cf. Head *et al*, 2000: ¶5.2.1).

1.3 Ideally, the full yield curve should be available to place market values on nominal liabilities but, since it is not practical to model all maturities on the yield curve, Tilley (1992: 527) suggests the modelling of eight key yields and the use of linear interpolation to model intermediate maturities. In order to reduce the dimension of the problem further, Sherris (1995) suggests factor analysis to determine the dimension of randomness in the yield curve. In this paper, principal-components analysis (PCA) is used to determine the number of maturities, n , required to adequately describe the South African yield curve. The subset consisting of the first n principal components is then used to interpolate the entire yield curve, given a specific subset of n key points along the curve.

1.4 At this point, it is worth distinguishing between two sources of arbitrage. The first kind of arbitrage exists when two identical sets of cash flows have a different price. The situation discussed above in which liabilities are valued using a single discount rate instead of the term structure of interest rates is one example of this. The second kind exists where an immunised portfolio with a different cash-flow profile costs less than the dedicated portfolio, as discussed in Maitland (2001: ¶6.1).

1.5 In order to reduce the dimension of the model, at the expense of producing non-key yields that are not arbitrage-free, both the methodology suggested by Tilley (*op. cit.*) and that discussed in this paper give rise to arbitrage opportunities of the second kind¹. However, as Thorlacius (2000) points out:

For effective [ALM], accuracy and realism are important criteria and as such the simulation must provide an accurate representation of the probabilities of potential economic and market outcomes ... The problem comes in trying to create robust model characteristics that reflect those observed in the real world while at the same time confining the computational demands of the model. Structures that ensure arbitrage free interest rates tend to be too simple (and thus produce unrealistic scenarios) or require a large amount of computation power.

1.6 If the purpose of the ALM is to use it for static or dynamic decision-making purposes—for example, to optimise with respect to alternative portfolio selections at future simulation dates—this statement is well justified. It is also justified for static decision-making purposes where realistic probabilities of potential future outcomes are desired. In both cases, realistic market risk premiums and investor risk preferences are relevant to the decision-making process and a simplistic arbitrage-free model would not be useful for such purposes.

1.7 For any given number of key yields, n , the methodology proposed in this paper provides the best fitting (and hence most realistic) yield curve in the sense that it minimises the expected root-mean-squared error of the residuals. The computational burden is also minimal.

1.8 It should be noted that the JSE-Actuaries Yield Curve is itself not arbitrage-free in

the sense that successive curves can give rise to arbitrage opportunities of the second kind, since cubic spline interpolation is used to construct intermediate yields (cf. Section 6). Hence, one cannot hope that a parsimonious, best fitting model will be fully arbitrage-free. Nonetheless, the JSE-Actuaries Yield Curve is still used as the basis for pricing unsecuritised, nominal cash flows and as such provides an important tool for reducing arbitrage opportunities of the first kind. Likewise, the intention of the methodology proposed in this paper is to reduce the magnitude of arbitrage opportunities of the first kind by providing a realistic estimate of the yield curve while at the same time minimising the computational burden².

1.9 The question whether it is appropriate to sacrifice the condition of no arbitrage in order to create a functionally simpler model that produces more realistic scenarios depends on the application. The main argument against using a term-structure model with arbitrage is that it is possible to construct investment strategies that perform unreasonably well by exploiting arbitrage opportunities of the second kind. However, if such strategies³ are excluded and the model is simply used to reduce arbitrage opportunities of the first kind, such arbitrage opportunities should present no problem since they cannot be exploited. Since the objective is a reasonably accurate description of the par yield curve for the purposes of discounting nominal future cash flows while at the same time minimising the number of yields from which to estimate the remainder of the curve, the proposed methodology would appear acceptable⁴.

1.10 Having selected the key maturities, these can then be modelled as part of a larger set of variables including other asset categories and economic variables, for the purposes of modelling the assets and liabilities of a financial institution. The dynamic model for these n maturities is not discussed in this paper. Use of the interpolated curve should be limited to the discounting of future cash flows and should not be extended to infer the dynamics of interpolated yields, since, as shown by Maitland (2001), these may not be arbitrage-free. Hence, the model proposed is a descriptive model rather than an equilibrium or no-arbitrage model.

1.11 If the presence of arbitrage opportunities of the second kind still gives rise to concern, the technique described by Thorlacius (2000: ¶4) can be used to remove these. This technique essentially works by adding uncertainty in the form of an independent random variable to each interpolated rate, thereby breaking the arbitrage. The standard deviation of these processes can be made small enough that the original model is not significantly disturbed. Using this technique, the fit and statistical characteristics of the original model are broadly retained while providing a model that is arbitrage-free.

2. BACKGROUND

2.1 Before 1982 there was virtually no active secondary market in bonds. Prescribed-asset legislation forced pension and provident funds and insurance companies

to hold a certain percentage of their assets in respect of liabilities in government bonds, cash and other approved bonds. In the 1970s, insurance companies and pension funds held on average 41% of the long-term domestic marketable stock debt of the central government (compared with 47% by the Public Investment Commissioners); and 70% of local authorities' stock (Falkena *et al*, 1984: 129).

2.2 In the early 1980s an active secondary market in South African bonds began developing and has subsequently grown rapidly (McLeod, 1990). In 1986, the Johannesburg Stock Exchange (JSE) instituted a bond clearing-house and although the majority of bond trading was over the counter (OTC), a small number of trades were recorded on the JSE. Since some trades were recorded at each available maturity, and since these trades would have reflected yields traded OTC, the JSE-Actuaries Yield Curve can be considered to be a fair estimate of market yields prevailing at the time. In 1996, the bond exchange opened and the Financial Markets Control Act now requires all bond trades to be recorded by a recognised exchange.

3. PRINCIPAL-COMPONENTS ANALYSIS

3.1 A number of empirical studies by academic researchers and practitioners conclude that the short rate is non-stationary. A partial listing of these authors includes Stock & Watson (1988), Mills (1994: 68), Ang & Moore (1994), Johansen & Juselius (1992), Juselius (1995), and Pesaran & Shin (1996). In contrast, many theoretical models of the short-term interest rate assume stationarity and include a mean reversion term (cf. e.g. Vasicek, 1977; Brennan & Schwartz, 1982; Cox, Ingersoll & Ross, 1985), although non-stationary theoretical models also exist. Wilkie (1994) and Thomson (1996) both develop empirical models for interest rates assuming stationarity on the basis of economic rather than statistical arguments. This paper does not investigate issues of stationarity and, since there is no clear consensus, a PCA of both the levels and first differences of the South African yield curve is presented.

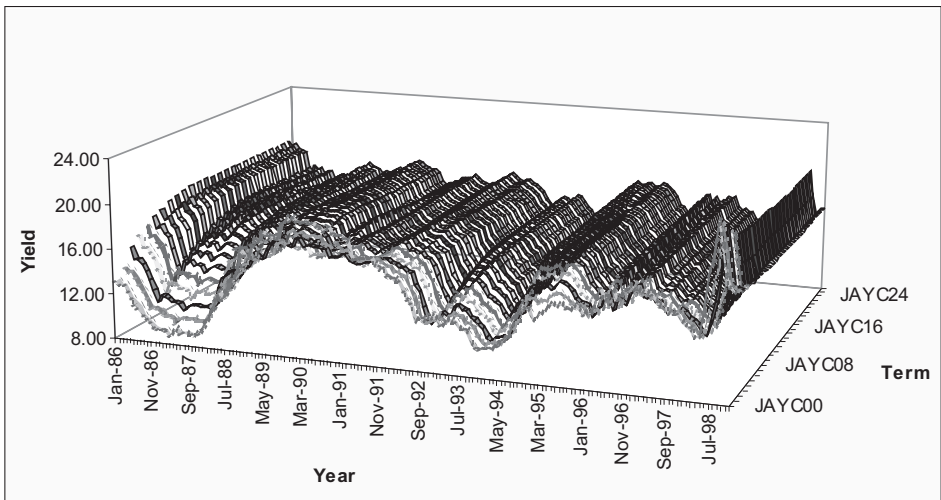
3.2 Let \mathbf{x} be a random d -vector with mean μ and covariance matrix Σ , and let $\mathbf{T} = (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_d)$ be an orthogonal matrix such that $\mathbf{T}'\Sigma\mathbf{T} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ are the eigenvalues of Σ . If $\mathbf{y} = \mathbf{T}'(\mathbf{x} - \mu)$, then $y_j = \mathbf{t}_j'(\mathbf{x} - \mu)$ is called the j^{th} principal-component score of \mathbf{x} and is the orthogonal projection of $\mathbf{x} - \mu$ in the direction \mathbf{t}_j (Seber, 1984: 176). Principal-components analysis explains the variance-covariance structure of the original variables through an orthogonal rotation of \mathbf{x} such that the first principal component gives the direction of maximum variation, the second gives the next largest direction of maximum variability orthogonal to the first principal component, and so on. If Σ is positive definite, d principal components are required to reproduce the total system variability completely, but much fewer principal components may explain a reasonable proportion of the total variability and hence reduce the dimension of the model with only a small loss of information.

3.3 We define the yields for annual terms from 0 and 25 years along the JSE-Actuaries Yield Curve (with the INET (1998) codes *JAYC00*, *JAYC01*... *JAYC25*) to be our 26-dimensional random vector. If yields are stationary, the moments of the level yields exist. Table 1 provides summary statistics for key yields at annual maturities from 0 to 25 years using monthly data from January 1986 to December 1998, while Figure 1 illustrates the yield curve over this period.

TABLE 1. JSE-Actuaries Yield Curve (1986–1998)

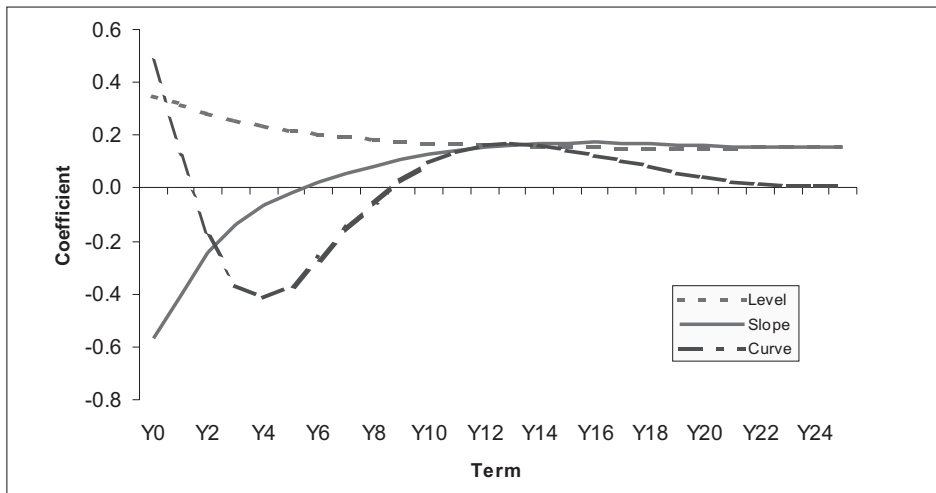
	mean	median	standard deviation	minimum	maximum
JAYC00	14,55	14,58	3,31	8,47	22,99
JAYC01	14,38	14,59	2,72	9,24	22,31
JAYC02	14,28	14,55	2,25	10,01	21,64
JAYC03	14,37	14,72	1,94	10,36	21,06
JAYC04	14,59	14,90	1,74	10,61	20,59
JAYC05	14,84	15,21	1,59	10,91	20,24
JAYC07	15,26	15,67	1,41	11,39	19,84
JAYC10	15,58	15,80	1,33	11,75	19,72
JAYC15	15,69	15,94	1,29	11,96	19,49
JAYC20	15,65	15,77	1,25	12,02	19,34

FIGURE 1. JSE-Actuaries Yield Curve (1986–1998)



3.4 A PCA on the covariance matrix of yields reveals that the first principal component explains 77,1% of the total variability in the yield curve, the first two together explain 98,4% and the first three together explain 99,4% of the total variability in the yield curve. Figure 2 illustrates the coefficients of each of the first three principal components by term to maturity.

FIGURE 2. Coefficients for the first three principal components of yield levels



3.5 The coefficients for the first principal component are all positive, so that an increase in the score of the first principal-component results in an increase in all yields. The first principal component can therefore be regarded as a level factor. Since the coefficients are not all equal, a change in the score of the first principal component does not result in a parallel shift; instead, the short end of the curve moves more than the long end.

3.6 The coefficients for the second factor are negative at the short end and monotonically increase to a positive value at the long end. Hence, a change in the score of the second principal-component results in an opposite effect on the two ends of the yield curve, and this factor can be viewed as causing a change or twist in the slope of the yield curve. The third principal component has a negative effect on medium yields and a positive effect on short and long-term yields and hence can be interpreted as a hump factor or butterfly. Figure 3 illustrates the principal-component scores for the first three principal components from January 1986 to December 1998.

3.7 The third principal component accounts for only 1% of the total variability and the remaining 23 principal components account for about 0,5% of the total variability. Hence, two principal components appear to capture most of the variability in the yield curve. This

FIGURE 3. Principal-component scores for yields (1986–1998)



is supported by the informal scree test illustrated in Figure 6 and discussed in Section 5. Section 7 discusses how these principal components can be used to reconstruct the entire yield curve.

3.8 So far, we have considered the covariance matrix of yields. However, if yields are non-stationary, then the population moments do not exist. We now define the monthly changes in yields for annual terms from 0 and 25 years along the JSE-Actuaries Yield Curve to be our 26-dimensional random vector and again use monthly yield data from January 1986 to December 1998.

3.9 A PCA on the covariance matrix of changes in yields reveals that the first principal component alone explains 92,8% of the total variability, the first two together explain 97,3% and the first three together explain 98,4% of the total variability. Hence, two principal components again appear to capture most of the variability in yield curve changes. Figure 4 illustrates the coefficients of the first three principal components by term to maturity, while the scree test illustrated in Figure 6 supports the choice of two principal components.

3.10 The first principal component affects all maturities by similar amounts and in the same direction. It can be interpreted as a level shift factor but not as a parallel shift factor since the coefficients are unequal. Unlike the levels PCA, the short end of the curve moves less than the long end in response to the score of the first principal component. The second factor has an opposite effect on short and long yields and can be viewed as a slope change or twist factor. The third principal component has a negative effect on medium

yields and a positive effect on short- and long-term yields, and hence can be interpreted as a curvature or butterfly factor. Figure 5 illustrates the principal-component scores for the first three principal components of yield curve changes for the period January 1986 to

FIGURE 4. Coefficients for the first three principal components of yield changes

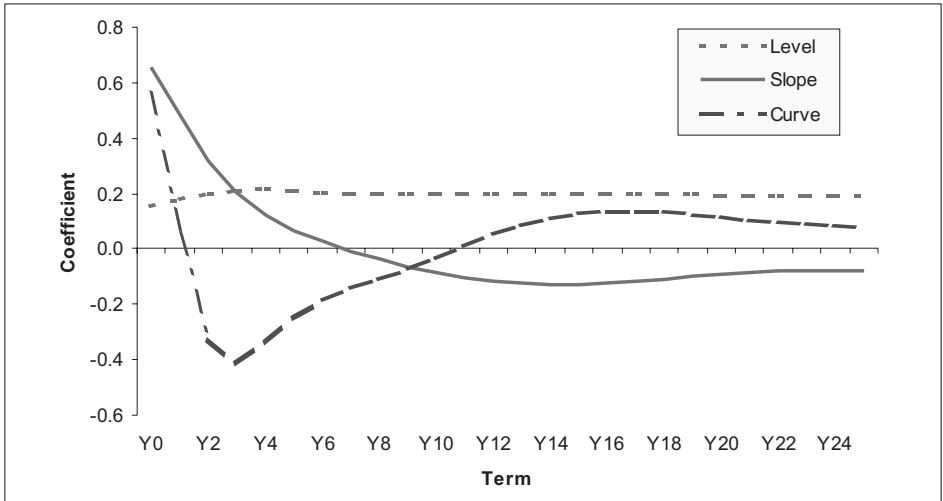
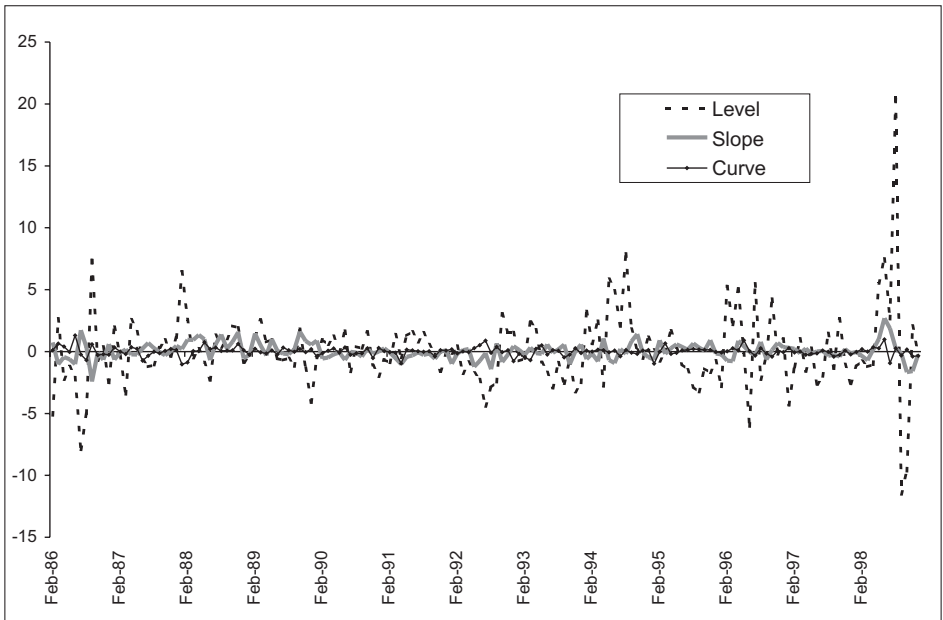


FIGURE 5. Principal-component scores for yield changes (1986–1998)



December 1998. International PCA results indicate the presence of similar components, although the proportion of variability explained by each component differs considerably from one market to another. Maitland (2001) gives a partial list of references to these results.

4. RMSE OPTIMALITY OF THE PRINCIPAL-COMPONENTS APPROXIMATION

4.1 The expected RMSE is commonly used as a simple descriptive measure of the fit of a particular model (Anderson *et al*, 1996: 59). For a maximum of n factors, it turns out that the first n principal components are optimal in the sense that they minimise the RMSE over all linear combinations of factors. A proof of this result, which, to the author's knowledge, has not been documented in the literature, is given in Appendix A.

4.2 In Section 7, a methodology is presented for interpolating the yield curve given a restricted number of yields and using the principal components. It should be noted that the principal components are optimal provided the key yields are known a priori. If the key yields are not known a priori but are projected using some stochastic model, then the RMSE may not be a valid measure of the fit of the forecast yield curve, since that may depend on the model for these key yields. It may be possible under certain conditions to separate the optimality of the projected yields from the interpolated curve whose optimality is conditional on the projected yields. However, an investigation into these conditions is beyond the scope of this paper and is left for future research.

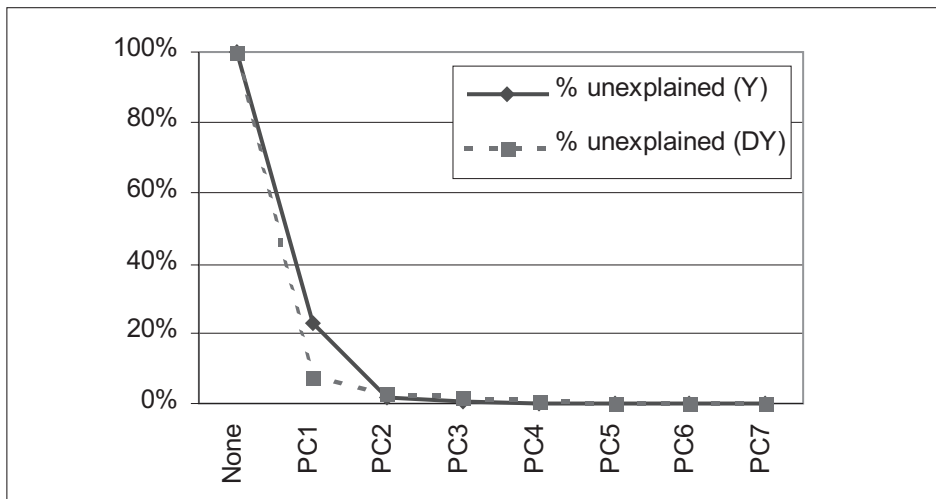
5. THE SCREE TEST

5.1 By plotting the root-mean-squared error (RMSE) of estimated yields against the number of parameters or factors included in the model, the marginal gain in explanatory power can be visually offset against the increase in the number of parameters. This informal statistical test is known as a 'scree' test (Cattell, 1965) and Monte Carlo studies (Tucker, Koopman & Linn, 1969) have shown that it is often superior in locating major common factors when minor factors are at play. In fitting parametric curves to yield-curve data, Chaplin (1998: 344–7) supports his choice of model using this test in preference to more formal statistical tests for assessing model fit.

5.2 PCA directly locates the factors and places them in order of importance, as discussed in Section 3. The total system variance as defined by the sum of the diagonal elements of Σ , $\sigma_{11} + \sigma_{22} + \dots + \sigma_{dd}$, is equal to $\lambda_1 + \lambda_2 + \dots + \lambda_d$. This follows since $\text{tr}(\mathbf{T}'\Sigma\mathbf{T}) = \text{tr}(\Sigma\mathbf{T}'\mathbf{T}) = \text{tr}(\Sigma\mathbf{I}_d) = \text{tr}(\Sigma)$, which follows from the properties of the trace operator and the fact that \mathbf{T}' is orthogonal. As shown in equation (3), the expected mean-squared error of yields (or changes in yields) from approximating the yield curve (or changes in the yield curve) with the first n principal components is given by $(\lambda_{n+1} + \dots + \lambda_d)/d$ (for the corresponding Σ). Hence, the ratio $(\lambda_{n+1} + \dots + \lambda_d)/(\lambda_1 + \dots + \lambda_d)$ gives an equivalent scree test.

5.3 Figure 6 illustrates the scree plot for the PCA of both the yields and changes in yields. In both cases, the improvement in fit from two to three parameters is insubstantial for the purpose at hand, bearing in mind the need for parsimony in dimension of the dynamic model. For certain applications a better description of yields may be preferred, in which case more principal components can be used to approximate the yield curve.

FIGURE 6. Scree plot for yields (Y) and changes in yield (DY)



6. A DISCUSSION OF THE PCA RESULTS

6.1 The traditional theory of immunisation as developed by Redington (1952) immunises a portfolio against parallel shifts in the yield curve. Parallel shifts imply the existence of arbitrage opportunities (cf. Boyle, 1978) and it is important to note that the first principal component does not represent an entirely parallel shift. However, for terms greater than five years, the first principal component does seem to represent a parallel shift, and for terms greater than 12 years, the second principal component also seems to represent virtually parallel shifts. Hence, the first two principal components, which represent 97,3% of the total variability, appear to indicate the regular occurrence of parallel shifts. However, this does not necessarily imply the existence of arbitrage opportunities at the long end of the curve, since, on average, 2,7% of the variability remains unexplained. Maitland (2001) shows how to identify arbitrage opportunities conditional on the absence of higher-order principal-component shifts.

6.2 Estimates of variances, covariances and correlations can be very sensitive to outliers and so we can expect principal components to have the same sensitivity. The extreme scores for the first principal component between August and October 1998

shown in Figures 3 and 5, and the corresponding large changes in the level of the yield curve evident from Figure 1, suggest the need for a PCA for sub-periods of the data. For the sub-period 1986 to 1997, the proportion of the variability explained by the first principal component of yield curve changes decreases from 92,8% (for the period 1986 to 1998) to 90,0%. It should be noted that although the extreme events occurred at points in time, the time-series properties of the scores are irrelevant for the purposes of this paper.

6.3 For level yields, the proportion of the variability explained by the first principal component reduces from 77,1% (for the period 1986 to 1998) to 76,1% for the sub-period 1986 to 1997. In both the yield and the differenced yield sub-period analyses, the principal components remain relatively unchanged, suggesting that the full-period analysis is relatively robust to the outliers from August to October 1998. A number of alternative sub-periods were considered and the results of the full period appeared to be relatively robust to the choice of sub-period.

6.4 In the above analyses, principal components are derived from the covariance matrix. If the variables in a PCA are measured on scales with widely differing ranges, it is preferable to use the correlation matrix (cf. Seber, 1984). Although the higher volatility of short rates compared with long rates results in an increased loading of the short rate on the first few factors, a PCA for both the levels yields and yield differences using the correlation matrix gives principal components and variability proportions that are similar to those obtained using the covariance matrices. Hence, the results of the PCA on the covariance matrix appear to be relatively robust to the lack of scaling. This is not too surprising given that the standard deviations of short and long yields are of the same order of magnitude.

6.5 One further point worth considering is the effect that the mathematical formulation of the JSE-Actuaries Yield Curve may have on the principal-components analysis. The curve is constructed in two steps (McLeod, 1990):

1. Using a form of cluster analysis, five cluster points are estimated and bonds are assigned to each cluster. The bonds in each cluster are then used to determine a weighted average term to maturity and a weighted average yield for their respective clusters. A sixth cluster with a maturity of 30 years and yield equal to the weighted average yield of the cluster with the highest weighted average yield is also determined.

2. Using these six cluster points, intermediate points along the curve are estimated using cubic spline interpolation.

6.6 Since the yield value of the sixth cluster is derived directly from one of the existing five cluster points, there are effectively five independent points along the curve. Hence, it is unlikely that more than five principal components would be required to reproduce most of the variability of the yield curve. The fact that two principal components capture most of the variability is a strong indication that the PCA is not constrained by the mathematical formulation of the yield curve.

7. RECONSTRUCTING THE YIELD CURVE

7.1 Using the principal components, \mathbf{T} , and the principal-component scores at time t , \mathbf{y}_t , of the level yields (or changes in yields), the level yields (or changes in yields) at time t , \mathbf{x}_t , can be reconstructed as $\mathbf{x}_t = \mathbf{T} \cdot \mathbf{y}_t + \boldsymbol{\mu}$. Since the first two principal components capture most of the variability in \mathbf{x} for a PCA of both the levels and first differences,

$$\mathbf{x}_t \approx y_{1,t} \cdot \mathbf{t}_1 + y_{2,t} \cdot \mathbf{t}_2 + \boldsymbol{\mu}.$$

7.2 For users deciding which variables to include in a stochastic model, it would be possible to model $y_{1,t}$ and $y_{2,t}$. However, since the relationship between these variables and the remaining variables in the stochastic model depends on the eigenvectors \mathbf{t}_1 and \mathbf{t}_2 , the resulting model may be difficult to interpret. Since a more direct and theoretically tractable relationship exists between actual yields and other stochastic variables, if any two yields (or changes in yields), $x_{a,t}$ and $x_{b,t}$, are modelled stochastically, these can be used to estimate $y_{1,t}$ and $y_{2,t}$, from which can be derived the full yield curve as explained above. More formally:

$$\text{Let } \mathbf{z}_t = \begin{bmatrix} z_{a,t} \\ z_{b,t} \end{bmatrix} = \begin{bmatrix} x_{a,t} \\ x_{b,t} \end{bmatrix} - \begin{bmatrix} \mu_{a,t} \\ \mu_{b,t} \end{bmatrix}. \quad (1)$$

$$\text{Then, } \mathbf{z}_t = \begin{bmatrix} t_{1,a} & t_{2,a} \\ t_{1,b} & t_{2,b} \end{bmatrix} \cdot \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \mathfrak{S} \cdot \tilde{\mathbf{y}}_t \quad \text{say,} \quad (2)$$

$$\text{so } \tilde{\mathbf{y}}_t = \mathfrak{S}^{-1} \cdot \mathbf{z}_t. \quad (3)$$

$$\text{Hence } \mathbf{x}_t = \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 \end{bmatrix} \cdot \mathfrak{S}^{-1} \mathbf{z}_t + \boldsymbol{\mu}_t. \quad (4)$$

7.3 Monthly data for the JSE-Actuaries Yield Curve for annual terms from 0 to 25 years are available from January 1986 onwards. Before this, only yields on 3- and 20-year bonds are available as well as the Alexander Forbes Money Market Index, from which can be derived a proxy for the short rate. These three series are available from 1960 onwards. Hence, if data for the full period from 1960 to 1998 are required for modelling purposes, it is possible to model only these three points on the yield curve.

7.4 If most of the variability in the yield curve could be explained by one principal component, the correlation between yields at different terms would be close to one and the yield at any term would be sufficient to reproduce the entire yield curve. Since two principal components are required to explain most of the variability in the yield curve, we require two terms, a and b , to reproduce the entire yield curve. These two terms should be chosen so that the absolute correlation between them is as small as possible in order to minimise the error in estimating $y_{1,t}$ and $y_{2,t}$. The correlation matrix for *JAYC00*, *JAYC03* and *JAYC20* is presented in Table 2. (Level yield correlations are shown below the diagonal and differenced yield correlations above.)

TABLE 2. Yield correlations (1986–1998)

	JAYC00	JAYC03	JAYC20
JAYC00	1	0,74	0,67
JAYC03	0,87	1	0,92
JAYC20	0,44	0,75	1

7.5 The correlations between *JAYC00* and *JAYC20* in Table 2 are less than the other correlations, suggesting that *JAYC03* can be dropped from the set of model variables. Figures 2 and 4 confirm this suggestion since the greatest differences between the coefficients of the first and second principal components are at the short and long maturities. Further, for most months between January 1986 and December 1998, *JAYC03* lies between *JAYC00* and *JAYC20*. Since the difference in term between *JAYC00* and *JAYC20* is the largest, errors in forecasting *JAYC00* and *JAYC20* have a smaller effect on the forecast error for *JAYC03* than any other pair of yields might have on the remaining yield.

8. CONCLUSION

8.1 The proposed methodology provides a way in which a yield curve can be interpolated from a restricted number of modelled yields, while at the same time minimising the number of yields from which to estimate the remainder of the curve. From a statistical perspective, the short rate and the long-bond yield should be used to reconstruct the South African yield curve, given the first and second principal components. Hence, for the purposes of reconstructing the yield curve, one need model only the short rate and the long-bond yield. If these variables are modelled as non-stationary variables, the yield curve can be reconstructed given forecast changes together with the yield curve at time zero. Otherwise, the yield curve can be reconstructed direct using equations (1) to (4). It should be noted that the optimality of the interpolated yield is conditional on the key yields being given a priori, as discussed in Section 4.

8.2 A number of other reasons exist for modelling the long-bond yield and the short rate as part of a larger set of variables, but a discussion of this is beyond the scope of this paper. However, the results in this paper give further credence to the choice of variables modelled by Thomson (1996). The methodology presented in this paper is not a framework for projecting financial and economic factors, but rather a methodology for interpolating the yield curve given these factors. It is suggested that this methodology be used to supplement the development of future stochastic investment models.

8.3 As discussed in Section 1, the interpolated curve should not be used to infer the dynamics of interpolated yields. Rather it should be used to value future cash flows in a more realistic manner. The purpose of interpolation is not to optimise with respect to

alternative terms, as described in Maitland (2001), but rather to revalue, at the end of a simulation interval, the bonds whose terms have shortened (by one interval) and to value the liabilities. For the purposes of interpolating arbitrage-free yields, readers are referred to the methodology of Heath, Jarrow & Morton (1992). Maitland (2001) provides a methodology for determining the number of principal components to include for the purposes of short-term risk analysis, based on the financial significance of additional principal components.

ACKNOWLEDGEMENTS

The author would like to thank the editor and two anonymous referees for helpful comments and suggestions on an earlier draft of this paper. Any errors or omissions remain the responsibility of the author.

NOTES

- 1 It should be noted that the use of a single, simulated discount rate for the valuation of nominal liabilities also implies the existence of arbitrage opportunities of the second kind, since this is equivalent to assuming parallel shifts in a flat yield curve (cf. Boyle, 1978).
- 2 For bonds that conform to the first and second principal components, arbitrage opportunities of the first kind are completely avoided by using the methodology proposed in Section 7.
- 3 i.e. strategies involving the short sale of any asset or assuming liabilities are marketable.
- 4 NB: The proposed model is not intended for use in derivatives pricing or for constructing dynamic strategies where the existence of arbitrage opportunities of the second kind is unacceptable. For arbitrage-free derivatives pricing, the interested reader is referred to the methodology of Heath, Jarrow & Morton (1992). For constructing dynamic strategies, the interested reader is referred to Maitland (2001).

REFERENCES

- Anderson N, Breedon F, Deacon M, Derry A & Murphy G (1996). *Estimating and interpreting the yield curve*. John Wiley & Sons
- Ang A & Moore D (1994). The Australian yield curve. Unpublished working paper, Macquarie University, Sydney
- Brennan MJ & Schwartz ES (1982). An equilibrium model of bond pricing and a test of market efficiency, *Journal of Financial and Quantitative Analysis* **17**, 75–100
- Boyle PP (1978). Immunization under stochastic models of the term structure. *JIA* **105**, 177–87
- Cattell RB (1965). Factor analysis: an introduction to fundamentals. *Biometrics* **21**, 190–215
- Chaplin GB (1998). A review of term structure models and their applications. *BAJ* **4**, 323–49

- Cox JC, Ingersoll JE & Ross SA (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407
- Darroch JN (1965). An optimal property of principal components. *Annals of Mathematical Statistics* **36**, 1579
- Falkena HB, Fourie LJ & Kok WJ (1984). *The Mechanics of the South African Financial System*. MacMillan, Johannesburg, 2nd edition
- Head SJ, Adkins DR, Cairns AJG, Corvesor DO, Cule DO, Exeley CJ, Johnson IS, Spain JG & Wise AJ (2000). Pension fund valuations and market values. *BAJ* **6**, 55–143
- Heath D, Jarrow R & Morton A (1992). Bond pricing and the term structure of interest rates; a new methodology. *Econometrica* **60**, 77–105
- INET (1998). INET Graphics and Database Service, Intelligent Network (Pty) Ltd
- Johansen S & Juselius K (1992). Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK. *Journal of Econometrics* **53**, 211–44
- Juselius K (1995). Do PPP and UIP hold in the long-run. *Journal of Econometrics* **69**, 211–40
- Lee EM (1986). *An introduction to pension schemes*. Institute of Actuaries, London
- McLeod HD (1990). The development of a market yield curve: the South African solution. *1st AFIR International Colloquium* (2), 197–212
- Maitland AJ (2001). An empirical approach to immunization in South Africa. *SAAJ* **1**, 119–38
- Mills TC (1994). *The econometric modelling of financial time series*. Cambridge University Press
- Pesaran MH & Shin Y (1996). Cointegration and speed of convergence to equilibrium. *Econometrica* **71**, 117–43
- Redington FM (1952). Review of the principles of life office valuations. *JIA* **78**, 286–315
- Seber GAF (1984). *Multivariate Observations*. Wiley
- Sherris M (1995). Interest rate risk factors in the Australian bond market. *Proceedings of the 5th AFIR Colloquium*
- Stock JH & Watson M (1988). Testing for common trends. *Journal of the American Statistical Association* **83**, 1097–107
- Thomson RJ (1996). Stochastic investment modelling: the case of South Africa. *BAJ* **2**, 765–801
- Thorlacius AE (2000). Arbitrage in asset modelling for integrated risk management. *Proceedings of the 10th AFIR Colloquium*
- Tilley JA (1992). An actuarial layman's guide to building stochastic interest rate generators. *Transactions of the Society of Actuaries* **44**, 509–64
- Tucker LR, Koopman RF & Linn RL (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrika* **34**, 421–59
- Vasicek O (1977). An equilibrium characterisation of the term structure. *Journal of Financial Economics* **5**, 177–88

APPENDIX A

OPTIMALITY OF THE FIRST n PRINCIPAL COMPONENTS

A.1 The expected RMSE is commonly used as a simple, descriptive statistical measure of the fit of a particular model (Anderson *et al*, 1996: 59). For a maximum of n factors, the first n principal components are optimal in the sense that they minimise the RMSE over all linear combinations of n factors.

A.2 PROOF

A.2.1 Since $\mathbf{y}_t = \mathbf{T}'(\mathbf{x}_t - \boldsymbol{\mu})$, the model for \mathbf{x}_t based on the first n principal components is

$$\hat{\mathbf{x}}_t = \boldsymbol{\mu} + \sum_{j=1}^n y_{j,t} \mathbf{t}_j; \quad (\text{A1})$$

where $\mathbf{x}_t - \hat{\mathbf{x}}_t = \boldsymbol{\varepsilon}_t = \sum_{j=n+1}^d y_{j,t} \mathbf{t}_j$.

A.2.2 In equation (A1), \mathbf{x}_t , $\boldsymbol{\mu}$, \mathbf{t}_j and $\boldsymbol{\varepsilon}_t$ are all $(d \times 1)$ vectors and $y_{j,t}$ is a scalar. The mean-squared error multiplied by d (MSE d) is given by:

$$\boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t = \left(\sum_{j=n+1}^d y_{j,t} \mathbf{t}_j \right)' \left(\sum_{j=n+1}^d y_{j,t} \mathbf{t}_j \right) = \sum_{j=n+1}^d y_{j,t}^2 \mathbf{t}_j' \mathbf{t}_j = \sum_{j=n+1}^d y_{j,t}^2 \quad (\text{A2})$$

since the eigenvectors, \mathbf{t}_j , are orthonormal. Hence, the expected MSE d is given by:

$$E[\boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t] = E \left[\sum_{j=n+1}^d y_{j,t}^2 \right] = \sum_{j=n+1}^d E[y_{j,t}^2] = \sum_{j=n+1}^d \lambda_j. \quad (\text{A3})$$

A.2.3 This follows since $V(\mathbf{y}_t) = E[(\mathbf{y}_t - E[\mathbf{y}_t])(\mathbf{y}_t - E[\mathbf{y}_t])'] = E[\mathbf{y}_t \mathbf{y}_t']$ as $E[\mathbf{y}_t] = \mathbf{T}'(E[\mathbf{x}_t] - \boldsymbol{\mu}) = \mathbf{0}$ so that $V(\mathbf{y}_t) = \mathbf{T}'E[(\mathbf{x}_t - \boldsymbol{\mu})(\mathbf{x}_t - \boldsymbol{\mu})']\mathbf{T} = \mathbf{T}'\boldsymbol{\Sigma}\mathbf{T} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$ and hence $E(y_{j,t}^2) = \lambda_j$.

A.2.4 Expressing equation (A3) in terms of the trace of the residual covariance matrix gives:

$$\begin{aligned} E[\boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t] &= \text{tr}\{E[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']\} = \text{tr}\left\{E\left[(\mathbf{x}_t^* - \hat{\mathbf{x}}_t^*)(\mathbf{x}_t^* - \hat{\mathbf{x}}_t^*)'\right]\right\}; \\ &= \text{tr}\left\{E\left[(\mathbf{x}_t^* - \Xi\boldsymbol{\psi}_t)(\mathbf{x}_t^* - \Xi\boldsymbol{\psi}_t)'\right]\right\} \end{aligned} \quad (\text{A4})$$

where * denotes that the variable has been centred, Ξ is a $(d \times n)$ matrix corresponding to the first n columns of \mathbf{T} and $\boldsymbol{\psi}_t$ is an n -component vector corresponding to the first n elements of \mathbf{y}_t .

A.2.5 Next, suppose we wish to select a set of n d -component vectors (factors) $\mathbf{S} = \{s_1, s_2, \dots, s_n\}$ and corresponding $(n \times m)$ matrix of weights $\xi = \{\xi_1, \xi_2, \dots, \xi_m\}$ for each time $t = \{1, 2, \dots, m\}$ to approximate the true curve at each time by a linear combination of these factors. Then the total expected MSE_d , τ_n , is given by:

$$\begin{aligned} \tau_n &= \sum_{t=1}^m \text{E}[\text{MSE}_d] = \text{tr} \left\{ \text{E} \left[\sum_{t=1}^m (\mathbf{x}_t^* - \mathbf{S}\xi_t)(\mathbf{x}_t^* - \mathbf{S}\xi_t)' \right] \right\} \\ &= \text{tr} \left\{ \text{E} \left[(\mathbf{x}^* - \mathbf{S}\xi)(\mathbf{x}^* - \mathbf{S}\xi)' \right] \right\}; \end{aligned} \tag{A5}$$

where \mathbf{x}^* is the $(d \times m)$ matrix corresponding to $\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_m^*$. Darroch (1965) has shown that τ_n is minimised with respect to \mathbf{S} and ξ if and only if

$$\begin{aligned} \mathbf{S}\xi &= \sum_{j=1}^n \mathbf{t}_j \mathbf{t}_j' \mathbf{x}^* \\ &\left(= \sum_{t=1}^m \sum_{j=1}^n y_{j,t} \mathbf{t}_j \right). \end{aligned} \tag{A6}$$

A.2.6 Hence, the model for \mathbf{x}_t based on the first n principal components gives the minimum value of τ_n , which is $m(\lambda_{n+1} + \dots + \lambda_d)$, as shown by equation (A3). Since the function $f: x \mapsto \sqrt{x/md}$ is monotonically increasing, minimising the RMSE is equivalent to minimising τ_n . This proves the result that the principal components are best in the sense that they minimise the RMSE over all linear combinations of variables.

Endnote for Chapter 4

The methodology proposed in this article provides a way in which a yield curve can be interpolated from a restricted number of modelled yields, while at the same time minimising the number of yields from which to estimate the remainder of the curve.

For any given number of key yields, it provides the best fitting (and hence most realistic) yield curve in the sense that it minimises the expected root-mean-squared error of the residuals. This optimality is proved to be a consequence of using the first n principal components to reconstruct the yield curve. The computational burden is also minimal.

The key interest rate terms identified are the zero- and twenty-year terms. These two terms have been identified from a statistical perspective but a number of other reasons exist for modelling the long-bond yield and the short rate as part of a larger set of variables. These two terms are those used in the model presented in Chapter 5, where additional justification for modelling these two rates is provided.

The results in this paper give further credence to the choice of variables modelled by Thomson (1996). However, the actual short-term interest rate modelled in Chapter 5 differs from that used by Thomson (*op.cit.*), and the justification for this can be found in that chapter.

CHAPTER 5

Foreword to the Paper

Chapter 5 (Maitland, forthcoming) presents a new stochastic model for South African equities, long and short-term interest rates, and inflation. This paper uses the analysis of Chapter 4 to justify modelling the zero-year and twenty-year nominal par yields, although the actual short-term interest rate modelled differs from that used by Thomson (1996). Also, instead of modelling equity dividend yields and dividend growth rates, the paper recommends modelling the total return on equities. The model is based on data using quarterly intervals instead of annual intervals, as this allows for a wider range of applications and improved risk analysis in the modeling of assets and liabilities.

This chapter also draws on the analysis of Chapter 3 in discussing the use of structural breaks to model certain non-linear effects. However, a regime switching approach is recommended in place of the structural break model discussed in Chapter 3. The regime switching model allows for the possibility of multiple structural breaks at unknown points in time. Hence, exogenous structural breaks are made endogenous and a probability law for such events is obtained.

Chapter 5 then extends the univariate regime switching analysis by considering the joint modelling of variables subject to regime switching. It introduces a new class of Markov switching models where switches in variables are not perfectly correlated. Maximum likelihood estimates of the parameters are derived and shown to require only the smoothed inferences obtained from a univariate analysis of the variables.

The model presented is a descriptive model, with structural features and parameter estimates based on historical data. However, it also incorporates theoretical aspects in its design, thereby providing a balance between purely theoretical models and those based only on empirical considerations.

A MULTIPLE MARKOV SWITCHING MODEL FOR ACTUARIAL USE IN SOUTH AFRICA

By AJ Maitland

ABSTRACT

This paper introduces a new class of Markov switching models where switches in variables are not perfectly correlated. Maximum likelihood estimates of the parameters are derived and shown to require only the smoothed inferences obtained from a univariate analysis of the variables. The framework is used to estimate a Multiple Markov Switching (MMS) model of South African financial and economic variables, which can be used for various actuarial applications, especially those involving long-term projections. Users may wish to set certain parameters in relation to future expectations rather than simply using estimates based on past data, but that process is not covered in this paper.

KEYWORDS

Multivariate, multiple Markov switching, long-term, financial projections, actuarial, stochastic model, time series models.

CONTACT DETAILS

Department of Statistics and Actuarial Science, University of the Witwatersrand.

P.O. Box 2484, Parklands, 2121, South Africa. Tel: +27 11 486 1946.

E-mail: james@stochasticsolutions.co.za

1. INTRODUCTION

Early actuarial stochastic models assume that the dynamic process for various economic and financial variables is linear. For example, Wilkie's (1986) model for UK inflation uses an AR(1) process to describe the data. Thomson (1996) uses a linear transfer function model to model inflation, with equity dividend growth as the

input. Other linear models include those of Carter (1991), Claassen (1993) and Sherris *et al* (1999).

One of the primary assumptions of such models is that certain key variables are stationary. For example, both Wilkie (1986, 1995) and Thomson (1996) assume that various yields and inflation are stationary, although standard unit root tests suggest that these variables may be integrated (see Maitland, 1997). The implication of a unit root in a time series is that shocks to the system are permanent, trends are stochastic and forecast variances increase linearly as the lead time of the forecast increases. Hence, stationarity is a necessary assumption for producing reasonable long-term projections.

Maitland (1996) shows that the Thomson (1996) transfer-function models suffer from a number of statistical problems and estimation errors. Mean reversion in certain variables of the model creates risk-adjusted returns that are unrealistic and gives rise to predictability that violates the Efficient Market Hypothesis (EMH). Maitland (1996, 1997) also shows that the Thomson (1996) model suffers from parameter instability and bias, which also makes it problematic for use in long term projections. Some authors argue that the EMH is unrealistic, but a more complete discussion of actuarial models and EMH is beyond the scope of this paper.

Later models include non-linear effects through the use of autoregressive conditional heteroscedastic (ARCH) models, which were introduced by Engle (1982). Such models include the inflation model of Wilkie (1995) and Hua (1994). Whitten and Thomas extend Wilkie's (1995) UK inflation model with further analysis using ARCH and threshold autoregressive (TAR) models. Harris (1994) defines an Exponential Regressive Conditional Heteroscedasticity (ERCH) model for Australian data.

Harris (1996) fits a Markov switching model to quarterly stock price returns and inflation. Markov switching models form another class of non-linear models and were first introduced by Goldfeld and Quandt (1973). They were popularized by the pioneering work of Hamilton (1989, 1990), who describes the likelihood function, regime inferences and an efficient estimation technique for fitting such models.

Krolzig (1997) develops a comprehensive framework for Markov switching vector autoregressions in which switches between the various components of the vector are perfectly correlated. In Harris (1999), an alternative switching vector autoregression framework is developed, and a vector switching model is estimated for Australian data. However, a vector switching model is not useful if switches between parameter values relating to the individual components of the vector of variables can occur at different points in time. In this case, neither framework provides a useful approach for jointly modelling the variables because parameter values of one variable may switch from one state to another without those of other variables also switching states.

This paper generalises the Markov switching framework by allowing the parameter values of individual series to switch at different points in time, while allowing for the joint modelling of the variables and state switching. It provides a new and parsimonious framework that allows individual variables to switch from state to another without all variables switching at the same time. The model is shown to provide a reasonable description of South African data. The framework presented also allows for easy application of normative assumptions (see Thomson, 2006) while retaining those descriptive aspects of the model that are still believed to be relevant to the future.

The model presented in this paper was initially presented to ASSA in 1999 and based on data from 1960Q1 to 1998Q4. Those parameter estimates are presented together with updated parameters based on data from the period 1960Q1 to 2006Q2.

2. VARIABLES MODELLED AND TIME INTERVALS

This section considers some of the data requirements for a stochastic asset-liability model. Many of the issues have been extensively covered by Thomson (1996) so this section focuses mainly on aspects where that approach has been modified or extended.

2.1 Variables modelled

As explained by Thomson (1996; 768) “sufficient variables should be modelled to enable the assets and liabilities of the financial institution to be simulated in such a way as to facilitate decision making.” Uncertainty in the liabilities may be due to a large number of random elements. For example, for a defined benefit pension scheme, wage inflation, price inflation and demographic effects are uncertain. However, since the demographic effects “have a lesser effect on the finances of the scheme and because they are not as strongly correlated with the variables used for simulating asset cash flows,” (Thomson, 1996; 770), their inclusion increases the dimension of the model unnecessarily. By marginalizing the distribution of the liabilities with respect to the demographic effects, we reduce the dimension of the model with only a small loss of information.

The variables specified in this paper facilitate a market-based approach to valuing liabilities. Section 9.1.1 of PGN 201 of ASSA (2003) states that: “the basis used to value the assets must be consistent with that used to value the liabilities...” If, instead of using discounted cash flow techniques, the market value of liabilities is used to measure the cost of the liabilities and assets are taken at market value, it is not necessary to model dividend yields and dividend growth rates.

In an attempt to minimize the dimension of the model and to simplify the analysis, this paper considers a model using only the following four variables:

- the inflation rate;
- the 0-year nominal yield;
- the 20-year nominal par yield; and
- the total return on equities.

Maitland (2002) shows how to construct a full yield curve given a model of the zero- and 20-year nominal par yields. Property, wage inflation and offshore asset classes have been excluded as well as CPI-linked yields, although the latter can be inferred from the inflation and nominal yield curve components of the model. Hence, the variables in this model represent only a subset of the variables required for a comprehensive asset and liability modelling exercise. However, it is believed that the

model presented can form the basis for a more comprehensive model of the assets and liabilities of a financial institution.

2.2 Time intervals

The intended purpose of a stochastic model largely dictates the minimum time interval between forecasts. The Thomson model (1996; 772) is developed for annual liability cash flows produced from demographic models based on annual age intervals and for comparison with revenue accounts prepared on an annual basis. Wilkie (1995) and the Finish Group (see Ranne, 1998) also use annual intervals although Wilkie (*op. cit.*) presents some results for quarterly and monthly intervals as well.

Sherris *et al* (1999; 238) consider annual cash flow projections to be a crude approximation to the timing of cash flows and hence prefer a quarterly model. For resilience reserving, capital adequacy and solvency testing, an annual model will tend to understate insolvency probabilities for two reasons. Firstly, solvency can only be assessed annually so that insolvency in the interim will not be detected if the fund has recovered by the following assessment. Secondly, since temporal aggregation tends to reduce excess kurtosis, an annual model may not capture large fluctuations such as stock market crashes and interest rate hikes that occur within the year (see Harris 1994; 36 & 38).

Thomson (1996; 772) states that “the use of quarterly data in the development of the model tends to accentuate the short-term relationships at the expense of longer-term relationships.” However, this is not inevitable if the model structure and span of the data allow for longer-term relationships. In the context of testing the stationarity of dividend yields, Wilkie (1995; 825-826) points out that, even with a large number of frequently sampled observations, a stationary process with high autocorrelation may appear to be non-stationary if the observation period (span) is too short. However, the problem in this context is that the span of the data is too short, not that the sampling frequency of the data is too high. The point Wilkie makes is that increasing the number of observations by sampling more frequently leads only to a marginal increase in power of unit root tests, whereas increasing the span of the data significantly increases the power of these tests (see Perron, 1991). Nonetheless, from

this perspective, a model developed using quarterly data from 1960 onwards should be no worse than one developed using annual data over the same period.

As Thomson (1996; 772) points out, a quarterly model can be used for comparison with investment performance results, which are often reviewed quarterly. Further, for some defined contribution funds, interim bonuses are declared for the following quarter based on investment returns to date and expected returns for the remainder of the year. The fund rules may not allow negative bonuses to be declared so that an investment reserve is required to cover shortfalls. In such cases, a quarterly model is required to estimate an appropriate investment reserve level and assess the impact of various bonus strategies. Another advantage of using quarterly intervals over annual intervals is that more data points lead to better parameter estimates. Also, annual figures can be derived from a quarterly model but quarterly figures cannot be derived from an annual model. Consequently, a quarterly time interval is preferred to an annual time interval and so quarterly data is used in this paper.

Possible complications from using quarterly intervals instead of annual intervals are that quarterly data may exhibit a relatively high kurtosis and may contain seasonal effects. As discussed above, for some applications it is important to capture high kurtosis in the data and so this should be modelled.

Financial series do not usually exhibit seasonal effects but such effects are likely in economic series such as the consumer price index. This seasonality may be caused by the use of interim price estimates for certain index constituents when actual prices are only available at the end of each year.

Since such seasonal effects are of little interest in the current context, modelling seasonality requires unnecessary additional parameters and model complexity. Hence, the quarter-on-quarter force of inflation series has been seasonally adjusted using the X-12-ARIMA method developed by the Statistical Research Division of the U.S. Census Bureau. The model used for the X-12-ARIMA seasonal adjustment is an ARIMA(1,0,1)x(1,0,0)₄ model (see the X-12-ARIMA Reference Manual by the Statistical Research Division (2000) for further details). The seasonally adjusted and annualised force of inflation is the inflation series modelled in this paper.

2.3 Transformations

The main purpose of transforming data is to enable the use of a simple model form rather than a more complicated one in the original data. The overriding consideration in the choice of transformation is that of linearity. If a non-linear model can be expressed, by suitable transformation of the variables, in linear form, it is said to be intrinsically linear (see Draper et al, 1981; 222).

Even if relations between variables turn out to be non-linear, linear modelling frameworks such as the transfer function model with autoregressive integrated moving average terms (ARIMAX) and the vector autoregressive moving average (VARMA) model classes provide a simple and parsimonious framework for model development and should be considered before moving to non-linear modelling frameworks. For this reason, functions that admit a linear relation between variables are highly desirable. For example, rates of growth are multiplicative whereas forces are additive and hence more linear. In this context, a logarithmic transformation of the rates is appropriate. In addition, a logarithmic transformation changes the range from $(-\infty, \infty)$ to $(0, \infty)$, allowing certain forces to be modelled with standard normal distributions while keeping their corresponding rates positive. This is desirable if the rate in question can assume only positive values. A broad literature review and discussion of the most appropriate functions to use for the variables modelled in the Thomson Model can be found in Thomson (1996; 773-777).

2.4 Time series modelled

The quarter-on-quarter force of inflation series is constructed by taking the natural logarithm of the ratio of the All Items Consumer Price Index at quarterly intervals. This series is seasonally adjusted (as discussed above) and then annualised to give the seasonally adjusted and annualised force of inflation series, *INFL*.

Figure 1 shows *INFL* together with the year-on-year force of inflation, *INFL-YY* at quarterly intervals from 1960Q1 to 2006Q2. It should be noted that modelling the year-on-year force of inflation at quarterly intervals is problematic in that it has a tendency to increase the autocorrelation between successive periods and to obscure temporal dependence in the series. This comparison is only shown for illustrative

purposes and for comparison with the more familiar year-on-year figures often published.

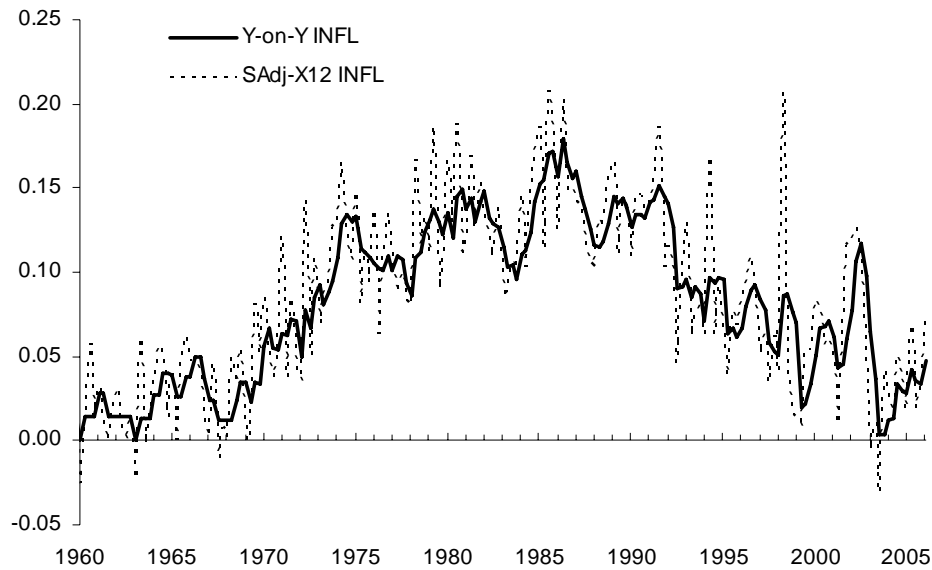


Figure 1. *INFL* (“SAdj-X12 INFL) and *INFL-YY* (“Y-on-Y INFL”)

For long-term interest bearing securities, the use of an average annual force of interest is well motivated by Thomson (1996; 776). The variable modelled is:

$$LINT_t = 2 * \ln(1 + JAYC20 / 200)$$

where *JAYC20* is the JSE-Actuaries 20-Year Bond Yield, as quoted under code *JAYC20* by INET (2006). NB. *JAYC20* is nominal yield convertible half-yearly.

For money-market instruments, Thomson (1996; 776) models the annual force of return on the Alexander Forbes Money-Market Index, as quoted under code *GMCI* by INET (2006). However, the Alexander Forbes Money-Market Index is constructed from the average monthly return on a portfolio consisting of 3-month NCDs with 1, 2 and 3 months to maturity. All information contained in *GMCI* at time *t* is available prior to time *t*, so *GMCI_t* does not belong in the information set at time *t*: unlike the yield curve, *GMCI* does not reflect rates available at the start of the period.

Ideally, the force of interest on 3-month Treasury bills at time t should be used to reflect the risk-free rate available at time t for the quarterly period from times t to $t+1$. However, since there is only a short history of Treasury bill rates, the 0-year nominal yield, as quoted under code *JAYC00* by INET (2006) is used as a proxy. The short-term interest rate modelled is called *SINT* and is defined as follows:

$$SINT_t = 2 * \ln(1 + JAYC00_t / 200) \quad \text{for } t \text{ in } \{1986Q1, \dots, 2006Q2\}$$

Since *JAYC00* is only available only from 1986 onwards, the annualized force of change in *GMCI* from time t to $t+1$ is used as a proxy for *JAYC00_t*, as follows:

$$SINT_t = 4 * \ln(GMCI_{t+1} / GMCI_t) \quad \text{for } t \text{ in } \{1960Q1, \dots, 1985Q4\}$$

Maitland (2002) shows that, in constructing a full arbitrage-free yield curve, the zero- and 20-year nominal par yields are the best yields to model to minimize the forecast error of the full yield curve.

For equities, it is preferable to model the excess equity return above the return on a risk-free asset (as modelled by the risk-free rate of interest) rather than the nominal equity return because risk-averse investors are typically interested in the additional returns they receive for taking on risk.

It could be argued that the real yield on a three-month CPI-linked bond is the appropriate risk-free hurdle rate for investors interested in accumulating real wealth. However, since such an instrument does not exist in South Africa, this approach is not particularly helpful. Arguably, for such investors, the three-month Treasury bill is the best proxy we have for a risk-free investment over the short term.

It could also be argued that, for an investor with longer-term liabilities, the return on an immunizing portfolio of longer-dated bonds is the relevant hurdle rate (see Maitland (2001) for the immunization framework that is mathematically optimal). This is indeed true. However, for the purposes of simplicity and without further knowledge of the segmentation of investor objectives, *SINT* is used as a proxy for the

three-month Treasury bill rate, and this is assumed to be the risk-free hurdle rate for each quarterly period.

The total return index for equities, $EQTRI$, is taken as the FTSE/JSE All Share Index, as quoted under code $J203TRI$ by JSE Information Products Sales Division (2005). This is available from 30 September 1995 onwards. Prior to this, a total return index for equities is constructed assuming dividends are uniformly distributed over the calendar year and using the variables ADY and $CI01$, where $CI01$ is the JSE-Actuaries All Share Index and ADY is the dividend yield per cent on that index, as quoted under codes $CI01$ and ADY by INET (1998).

The total, annualised, quarterly force of return on equities in excess of the risk-free rate, $XSEQ$, is then constructed as follows:

$$XSEQ_t = 4 * \ln(EQTRI_t / EQTRI_{t-1}) - SINT_{t-1}$$

The variables $SINT$, $LINT$ and $XSEQ$ are shown in Figures 2-4 below. The analysis in this paper uses data over the period from 1960Q1 to 2006Q2.

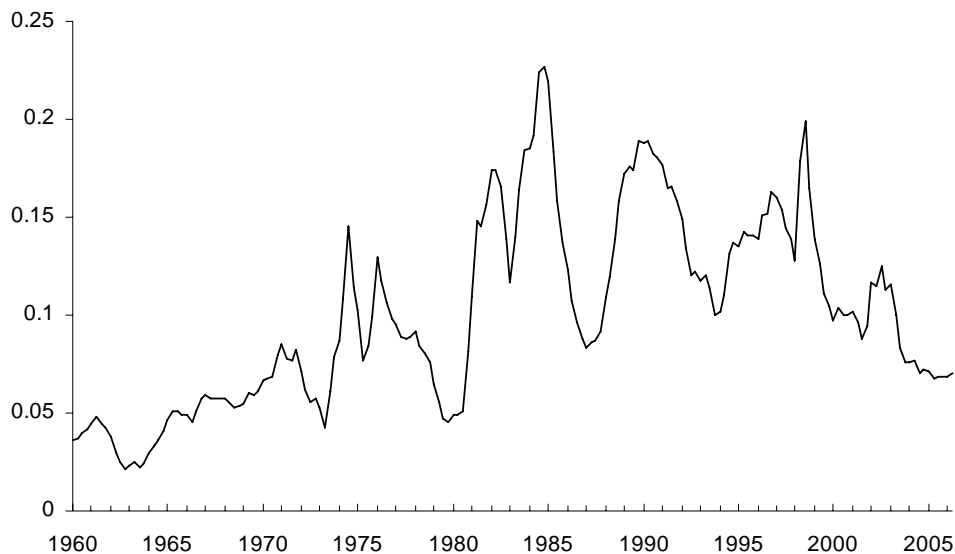


Figure 2: $SINT$

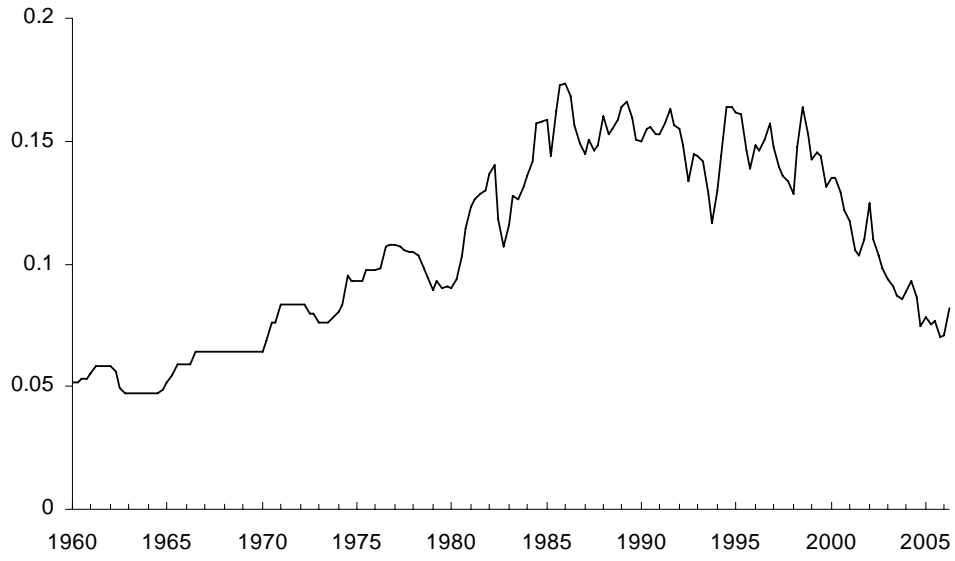


Figure 3: *LINT*

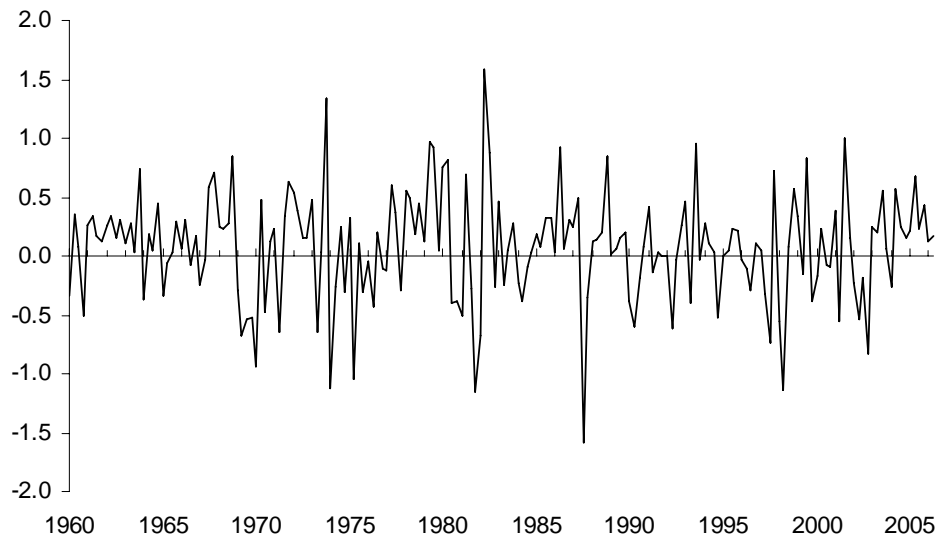


Figure 4: *XSEQ*

3. UNIT ROOT TESTS

In building a multivariate time series model, the purpose of developing univariate models for each of the variables is to guide subsequent multivariate modelling. How best to proceed hinges on knowing whether the individual series are stationary or non-stationary. Conventional time series estimation techniques based on classical assumptions about the distribution of the error terms can lead to incorrect inferences if the series are non-stationary. For example, if classical ordinary least squares is used to estimate the relationship between two non-stationary variables each containing a unit root, standard test statistics produce misleading inferences. This is known as the spurious regression problem (see Granger and Newbold, 1974).

Standard unit root tests generally test the null hypothesis of a unit root against the one-sided alternative of no unit root (see, for example, Hamilton, 1994, Chapter 17). The results of some standard unit root tests are shown in Table 1 below. These tests all include an intercept in the test equation and test for a unit root in the level series. In Table 1, ADF refers to the Augmented Dickey-Fuller test while PP refers to the Phillips-Perron test. For further details, see Maitland (1997), Dickey and Fuller (1979, 1981), and Phillips and Perron (1988).

	ADF	PP	KPSS
<i>INFL</i>	-2.982 (0.039)	-6.017 (0.000)	0.504 [0.463]
<i>SINT</i>	-2.857 (0.053)	-2.441 (0.132)	0.891 [0.463]
<i>LINT</i>	-1.534 (0.514)	-1.501 (0.531)	1.067 [0.463]
<i>XSEQ</i>	-13.202 (0.000)	-13.209 (0.000)	0.058 [0.463]

Table 1 - Unit Root Tests

For both the ADF and PP tests, the null hypothesis is that the series is non-stationary, and only if the series is sufficiently stationary is this assumption rejected. For each variable, the ADF and PP columns in Table 1 show the corresponding t -statistics and (p -values). These results suggest that the null hypothesis cannot be rejected at the 5%

level for *SINT* and *LINT*. The results also suggest that *INFL* and *XSEQ* do not contain a unit root since the null hypothesis is rejected at the 5% level for these two variables.

KPSS in the third column of Table 1 refers to the Kwiatkowski, Phillips, Schmidt & Shin (1992) test, which is a Lagrange Multiplier test that evaluates the null hypothesis that the series is stationary against the alternative that it is non-stationary. As a result, the KPSS test reverses the usual burden of proof. See Kwiatkowski, Phillips, Schmidt & Shin (*op. cit.*) for further details.

Cells in the third column of Table 1 show the LM-statistics and [5% critical values] for the KPSS test. An LM-statistic that is greater than the 5% critical value rejects the null hypothesis at the 5%-level. The KPSS result supports the findings of the ADF and PP tests that *XSEQ* does not contain a unit root. The results also suggest that the null hypothesis of stationary can be rejected at the 5% level for *INFL*, *SINT* and *LINT*. This supports the finding of the ADF and PP tests that *SINT* and *LINT* both contain a unit root but contradicts the earlier result for *INFL*.

The ADF and PP results for *INFL* in Table 1 above also contrast with the standard unit root test results shown in Maitland (1997), which used annual data for the force of inflation over the period 1960 to 1993. Rerunning the ADF and PP tests for *INFL* for the sub-period 1960Q1 to 1993Q4 shows that the null of a unit root cannot be rejected at the 5% level for the ADF test statistic but that it can be rejected for the PP statistic. Such mixed results are symptomatic of non-linear effects in the data.

The implication of a unit root in a time series is that shocks to the system are permanent, trends are stochastic and forecast variances increase linearly as the lead time of the forecast increases. Hence, whether or not a variable contains a unit root is critical for projection purposes.

The above results using data from 1960Q1 to 2006Q2 might suggest a multivariate model with *INFL* modelled as a stationary variable and *SINT* and *LINT* as non-stationary variables. However, such a mixed model would not make sense, particularly as *LINT* reflects the market's expectation of future inflation, the real rate of interest and an inflation risk premium. If inflation is stationary, a non-stationary

LINT would imply a non-stationary inflation risk premium and real interest rate. However, it is unreasonable to assume that real interest rates can wander off to any level, as implied by a forecast variance that increases linearly with time.

As discussed in Section 2.2, increasing the span of the data significantly increases the power of unit root tests (see Perron, 1991 and Wilkie, 1995, pages 825-826). Homer and Sylla (2005) show that for around four thousand years interest rates have remained around 5% or so (in non-inflationary times). This suggests that with an increased span of data, standard unit root tests on the above series will most likely indicate that they are stationary. However, even with the short span of data available, certain tests lead to more reasonable models than suggested by the above unit root test results.

Perron (1989) shows that standard unit root tests which do not allow for the presence of a structural break have little power against the alternate of no unit root when the underlying series has a structural break but no unit root. The power of these tests decreases as the magnitude of the intervention variables increases.

Using Perron's (1989) framework, Maitland (1997) shows that the null hypothesis of a unit root can be rejected once the possibility of a structural break is considered. However, Perron's framework does not entertain the possibility of multiple structural breaks at unknown points in time. For this reason, unit root tests that allow for the possibility of a deterministic structural break are not considered further in this paper. Instead, a modelling framework allowing for multiple structural breaks at unknown points in time is presented in the following section.

4. UNIVARIATE MARKOV SWITCHING MODELS

4.1 Introduction

Economic and financial time series can exhibit dramatic breaks in their behaviour, associated with events such as oil price shocks, changes in government policy, financial crises and shifts in investor expectations. If the behaviour for periods of time can be adequately described by autoregressive ($AR(p)$) models of the form:

$$y_t = c + \phi_1 y_{t-1} + K + \phi_p y_{t-p} + \varepsilon_t \quad (1)$$

with $\varepsilon_t \sim N(0, \sigma^2)$, then it would be reasonable to allow the parameters $c, \phi_1, \dots, \phi_p, \sigma$ of this model to change to accommodate such breaks. The encompassing model could then be described as:

$$y_t = c_{s_t} + \phi_{1,s_t} y_{t-1} + K + \phi_{p,s_t} y_{t-p} + \varepsilon_t \quad (2)$$

with $\varepsilon_t \sim N(0, \sigma_{s_t}^2)$, where s_t denotes the regime or state of the process at time t .

Since the determinants of these changes may be unobservable (as, for example, with a shift in investor expectations), or because one may simply not wish to include such determinants as factors in the model (the causes of financial crises are varied and inflation is not only influenced by oil price shocks), it is preferable to consider a probabilistic model to describe the occurrence of such breaks that give rise to changes in the parameters $c, \phi_1, \dots, \phi_p, \sigma$.

The simplest specification is that s_t is the realization of a Markov chain with the probability of a switch from state i to state j ($i, j = 1, 2, \dots, M$) being:

$$\Pr(s_t = j | s_{t-1} = i) = p_{ij}. \quad (3)$$

where $p_{i1} + p_{i2} + \dots + p_{iM} = 1$ for all $i \in \{1, \dots, M\}$. This assumes that the probability of a change in state or regime depends on the past only through the value of the most recent regime. The regimes are not observed directly but can be inferred through the observed behaviour of y_t .

The specification in equations (2) and (3) is non-linear and is referred to as Markov Switching (MS) model. Markov-switching regressions were first introduced by Goldfeld and Quandt (1973), and the likelihood function was first correctly calculated by Cosslett and Lee (1985). The pioneering work of Hamilton (1989, 1990) describes the likelihood function, regime inferences and an efficient estimation technique for

fitting such models. Krolzig (1997, Chapter 7) provides a useful discussion on model selection and model checking procedures for MS-models.

A model specification search is undertaken using the Schwartz information criterion (SC) and a likelihood ratio (LR) test. Given the number of regimes, standard asymptotic distribution theory holds for the SC concerning the number of autoregressive parameters and heteroscedasticity. (SC provides the most parsimonious model specification amongst the widely used Akaike Information Criterion (AIC), Hannan-Quinn (HQ) criterion and SC.) See Hamilton (1994) for details.

The LR test concerns the appropriate number of states in Equation (2), and follows a non-standard distribution. Unfortunately, equivalence in all regimes of the parameters $c, \phi_1, \dots, \phi_p, \sigma$ of Equation (2) implies that the Markov chain parameters p_{ij} are not identified under the null hypothesis of a single state ($M=1$).

As discussed in Garcia (1998), testing for the number of states in a regime switching framework is complex. Given some $M \geq 2$, the problem is that under any number of regimes smaller than M some transition probability (“nuisance”) parameters of the unrestricted model may take any value and are hence unidentified. The result is that the LR test fails to have a standard chi-square distribution with number of degrees of freedom equal to the number of restrictions imposed.

To overcome these complexities, the bounded likelihood ratio test proposed by Davies (1977) and recommended by Krolzig (1997) is used to test the null hypothesis of a single state $M=1$ (i.e. the “linear” model) against the alternative of two or more states ($M \geq 2$). This circumvents the problem of estimating nuisance parameters under the alternative hypothesis and derives instead an upper bound for the significance level of the LR test:

$$\Pr(LR > x) \leq \Pr(\chi_q^2 > x) + x^{q/2} \exp(-x/2) \cdot 2^{1-q/2} \cdot [\Gamma(q/2)]^{-1} \quad (4)$$

where $\Gamma(\cdot)$ is the standard gamma function and q is the number of nuisance parameters. (Note that for $M=2$, $q=2$.)

4.2 *INFL*

Maitland (1996) shows that the Thomson (1996) model suffers from parameter instability. In particular, he shows that the autoregressive parameter for inflation is much lower than that suggested by the Thomson model and the means quite different when estimated from the sub-periods 1960-1975 and 1976-1993.

Maitland (1997) also shows that it is this parameter instability that gives rise to the apparent unit root in the inflation series, and that the null hypothesis of a unit root can be rejected once the possibility of a single structural break is considered.

This assumption of a single structural break is unsatisfactory as a probability law that could have generated the inflation series. Furthermore, such features are not desirable for projection purposes and are not likely given the current framework of inflation targeting in South Africa (see Mboweni (2000) for details). Instead, a MS-model, which allows for multiple structural breaks at unknown points in time, is a more appropriate framework for modelling *INFL*.

A number of first order MS models with $M=2$ states have been estimated for *INFL*, allowing for switching in any combination of the intercept term (I), the autoregressive terms (A) and the variance of the residuals (H). The results are shown in Table A.1.1 of Appendix A. The null of a single state model ($M=1$) is rejected in favour of $M=2$, with the LR statistic being highly significant, even after applying Davies' (1977) correction. $M=3$ is rejected. In terms of the Schwartz Information Criterion (SC), the best model is the autoregressive model with switching only in the intercept term:

$$INFL_t = c_{s_t} + \phi_1 INFL_{t-1} + \varepsilon_t, \quad (5)$$

with $\varepsilon_t \sim N(0, \sigma^2=0.033^2)$, $c_1 = 0.0299(0.005)$, $c_2 = 0.0944(0.013)$ and $\phi_1=0.24(0.097)$. Standard errors are shown in parentheses. $s_t=1$ corresponds to a low-mean regime with a mean force of inflation, $c_1/(1-\phi_1)$, of about 4% while $s_t=2$ corresponds to a high-mean regime with a mean force of inflation of about 12%. The probability of remaining in state 1 given that the process is already in state 1 is $p_{11}=0.968$ while the probability of remaining in state 2 given that the process is already in state 2 is

$p_{22}=0.970$. The autoregressive parameter and the variance remain stable across regimes. The ergodic (unconditional) probabilities or stable-state probabilities (see, for example, Hamilton, 1994, Chapter 22) are both 0.5 for the low- and high-mean states, while the expected durations for the low- and high-mean states are both about 8 years.

Comparing this MS-model with the linear AR(1) model shown in Table A.1.1 of Appendix A, the problem of parameter bias becomes apparent. It can be seen that the autoregressive parameter of $\phi_1=0.67$ for the AR(1) model is much higher than $\phi_1=0.24$ for the MS-model. The bias is caused by the changing level of the series, as discussed by Maitland (1997).

The estimated probability (conditional on all the data) that the regime was in the high-inflation regime each quarter is shown in Figure 5. This should be compared with Figure 1 where it can be seen that a high inflation regime corresponds roughly to the periods 1973-1994, 1998:3 and the year 2002.

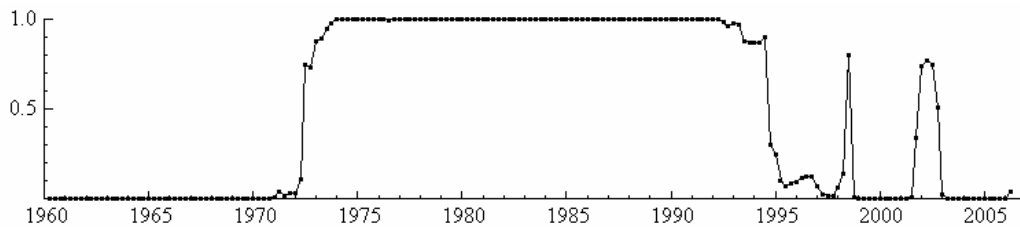


Figure 5. Probability that *INFL* was in the high-inflation regime

1973 corresponds to the first oil price shock, which led to a twenty year period of entrenched inflation, ending with the end of the apartheid era and the dismantling of many trade barriers, leading to increased international trade, decreased market power of domestic companies and downward pressure on real wages (see Aron & Muellbauer, 2000).

The third quarter of 1998 corresponds to the Russian debt crisis, and 2002 follows the dramatic fall in the Rand in 2001 following the Zimbabwe crisis and fears of contagion in the region.

With the introduction of inflation targeting in 2000 (see Mboweni (2000) for details), periods of persistent high inflation such as were experienced from 1973-1994 are arguably less likely to occur in future. In using the MS-model for projection purposes, the user is likely to lower the value of p_{22} , the probability of remaining in state 2.

However, the possibility of future inflation shocks (resulting from oil shocks, currency shocks, political crises, financial crises, wage pressures etc.) is not removed with the introduction of inflation targeting; nor can we rule out the possibility of a weak monetary policy regime at some point in the future. By simply adjusting the transition probabilities, the user is able to mimic stochastic projections under such scenarios.

Although certain aspects of the past are unlikely to repeat themselves in future, more stable aspects might still prove useful. For example, the user is able to retain the estimates for c_1 , c_2 , ϕ_1 and σ , unless more plausible values can be justified.

The values of ϕ_1 and σ are stable across both regimes and are therefore estimated from the full sample of data, so our confidence in these estimates should be higher. While inflation expectations from the real and nominal yield curves, and the inflation target band of between 3% and 6%, might suggest a slightly different values for c_1 and σ , the value of ϕ_1 can be retained, unless the user has reason to justify how inflation targeting might alter this dynamic.

In contrast to the MS-model for *INFL*, none of the parameters for *INFL* estimated in the Thomson (1996) model are useful for projection purposes. As discussed by Maitland (1996), Thomson's *INFL* autoregressive parameter is biased; the forecast mean tends to the arbitrary level of 9.5%, and clearly depends on the period of data used to estimate the parameters; and the forecast variance is inflated due to regime-switching in the underlying series.

The MS-model appears to be relatively stable when estimated over the sub-periods from 1960:1-1998:4 and 1970:1-2006:2. In both cases, the problem of parameter bias is again apparent in the corresponding linear models (Tables A.1.2-3 of Appendix A).

It is interesting to note that for the MSIH(2)-AR(1) model estimated over the sub-period 1960:1-1998:4, the parameter estimate for p_{22} is equal to one (see Table A.1.2 in Appendix A). This implies a permanent switch to the high-mean inflation regime, which is clearly unrealistic. The problem is that that span of data does not include periods where a switch from the high-mean regime to the low-mean regime occurred (or at least not with sufficient clarity to distinguish this given the higher volatility of the high-mean regime in this model). However, we know inflation has come down since then, and even if it had not, this is always a possibility. Hence, the parameter estimate $p_{22}=1$ from this subset model, while clearly a reasonable estimate given the data from that sub-period, is clearly not appropriate for forecasting. This illustrates the importance of applying judgement when setting parameters for projection purposes.

4.3 *SINT*

The results of fitting various first order MS-models to *SINT* are shown in Table A.2.1 of Appendix A. The null of a single state model ($M=1$) is rejected in favour of $M=2$, with the LR statistic being highly significant, even after applying Davies' (1977) correction. $M=3$ is rejected. In terms of the Schwartz Information Criterion (SC), the best model is the autoregressive model with switching in both the intercept term and the residual variance:

$$SINT_t = c_{s_t} + \phi_1 SINT_{t-1} + \varepsilon_t, \quad (6)$$

where $\varepsilon_t \sim N(0, \sigma_{s_t}^2)$, $c_1 = 0.0072(0.002)$, $c_2 = 0.0234(0.006)$ and $\phi_1 = 0.866(0.037)$. Standard errors are shown in parentheses. $s_t=1$ corresponds to a low-mean regime with a mean force of interest, $c_1/(1-\phi_1)$, of about 5.5% while $s_t=2$ corresponds to a high-mean regime with a mean force of interest of about 17.5%. $\sigma_1 = 0.0051$, $\sigma_2 = 0.0155$ so the high short-term interest regime is much more volatile, with volatility three times greater than that of the low short-term interest regime. The autoregressive parameter is stable across regimes.

The probability of remaining in state 1 given that the process is already in state 1 is $p_{11}=0.947$ while the probability of remaining in state 2 given that the process is

already in state 2 is $p_{22}=0.920$. The ergodic (unconditional) probabilities or stable-state probabilities are 0.601 for the low-mean state and 0.399 for the high mean state. The expected duration for the low-mean state is about 5 years while that for the high mean state is about 3 years. (Because of the probabilistic nature of the Markov-switching process, the series may remain in either state for as little as one quarter or much longer than the expected duration. Such asymmetry is not well captured by linear models.)

Comparing this MS-model with the linear AR(1) model shown in Table A.2.1 of Appendix A, the problem of parameter bias in the linear model again becomes apparent. The autoregressive parameter of $\phi_1=0.963$ for the AR(1) model is much higher than $\phi_1=0.866$ for the MS-model, with the bias again being induced by the changing level of the series. The estimated probability (conditional on all the data) that the regime was in the high-inflation regime each quarter is shown in Figure 6.

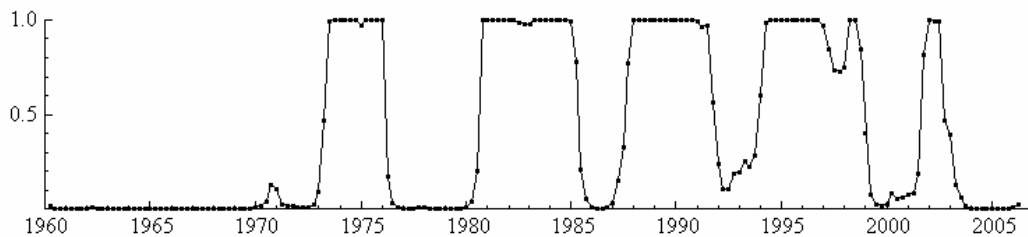


Figure 6. Probability that *SINT* was in the high-interest regime

The MSIH model for *SINT* appears to be very stable when estimated over the sub-periods from 1960:1-1998:4 and 1970:1-2006:2. In both cases, the problem of parameter bias in the autoregressive parameter is again apparent in the linear models.

The residuals from the AR(1) model for *SINT* exhibit a very high kurtosis of 5.4 and a Jarque-Bera statistic of 65.3, suggesting that the null hypothesis that the residuals are normally distributed can be rejected at the 99.999% level. Also, the Ljung-Box test statistics on the squared residuals indicate significant serial correlation structure in the volatility. This suggests fitting a GARCH model (see Engle (1982) and Bollerslev (1986) for details) to *SINT*.

The GARCH(1,1) model successfully removes all serial correlation in the squared residuals. However, the model suffers from extreme upward bias in the autoregressive parameter ϕ_1 , which is estimated to be 0.999. The kurtosis of the residuals increases to 7.1 and the Jarque-Bera statistic of 182.8 indicates an even more severe departure from normality than with the residuals from the AR(1) model. In contrast, the assumption of normality in the residuals from the MS-model in Equation (6) cannot be rejected.

The log-likelihood for the GARCH(1,1) model is 585.22, which is considerably less than the log-likelihood of 604.65 for the MS-model in Equation (6). Note, however, that the standard likelihood ratio test cannot be used to compare these two models because they are not nested.

4.4 *LINT*

The results of fitting various first order MS-models to *LINT* for the full period 1960:1-2006:2 as well as the sub-periods from 1960:1-1998:4 and 1970:1-2006:2 are shown in Tables A.3.1- A.3.3 of Appendix A. For both periods including the 1960s, the MSIH model is numerically unstable. Furthermore, the autoregressive parameter is very close to one when data from the 1960s is included in the estimation but drops when this period is excluded.

Examination of the data in Figure 3 reveals the cause of the instability. The earlier data is characterized by long stretches where the long-bond yield remains unchanged, interspersed with occasional jumps. For example, from December 1962 to September 1964, *LINT* is constant at 0.0469 while from September 1966 to March 1970, it remains constant at 0.064. The series was constructed by Dr James Greener “based on the coupon of the bonds issued in the primary market” using long-dated government bonds issued by the South African Reserve Bank for government. Since there were very few issues in the 1960s, the yield remained constant for long stretches at a time. These bonds were simply bought and held, largely by life offices, pension funds and particularly the Government Employees Pension Fund, which were all subject to Prudential Regulations forcing them to hold large volumes of government bonds. An active secondary market for these bonds did not develop until the early 1980s.

Such dynamics are not characteristic of ARIMAX, VARMA or MS stochastic processes and bias the autoregressive parameter ϕ_1 towards one. They are unlikely to be repeated in future if market forces continue to determine yields, and hence are not useful for projections purposes. Hence, data from the 1960s is excluded for estimation of the *LINT* MS-model parameters.

Using data for the period from 1970:1-2006:2, the null of a single state model ($M=1$) is rejected in favour of $M=2$, with the LR statistic being highly significant, even after applying Davies' (1977) correction. $M=3$ is rejected. In terms of the Schwartz Information Criterion (SC), the best model is the autoregressive model with switching in both the intercept term and the residual variance:

$$LINT_t = c_{s_t} + \phi_1 LINT_{t-1} + \varepsilon_t, \quad (7)$$

where $\varepsilon_t \sim N(0, \sigma_{s_t}^2)$, $c_1 = 0.0129(0.004)$, $c_2 = 0.0214(0.006)$ and $\phi_1=0.852(0.041)$. Standard errors are shown in parentheses. $s_t=1$ corresponds to a low-mean regime with a mean force of interest, $c_1/(1-\phi_1)$, of about 8.5% while $s_t=2$ corresponds to a high-mean regime with a mean force of interest of about 14.5%. $\sigma_1 = 0.0041$, $\sigma_2 = 0.0083$ so the high long-term interest regime is much more volatile, with volatility more than double that of the low long-term interest regime. The autoregressive parameter ($\phi_1=0.852$) is constant across regimes.

The probability of remaining in state 1 given that the process is already in state 1 is $p_{11}=0.973$ while the probability of remaining in state 2 given that the process is already in state 2 is $p_{22}=0.983$. The ergodic or stable-state probabilities are 0.391 for the low-mean state and 0.609 for the high mean state. The expected duration for the low-mean state is about 9 years while that for the high mean state is about 14 years.

Comparing this MS-model with the linear AR(1) model shown in Table A.2.3 of Appendix A, the problem of parameter bias in the linear model again becomes apparent. The autoregressive parameter of $\phi_1=0.965$ for the AR(1) model is much higher than $\phi_1=0.852$ for the MS-model, with the bias again being induced by the

changing level of the series. The estimated probability (conditional on all the data) that the regime was in the high-inflation regime each quarter is shown in the Figure 7. (These inferences are based on the parameters estimated from the period 1970:1-2006:2.)

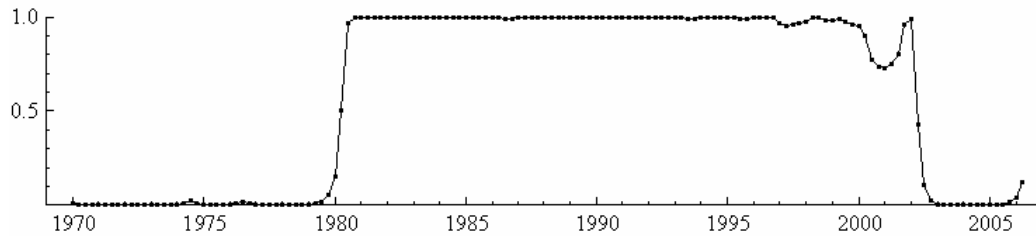


Figure 7. Probability that *LINT* was in the high-interest regime

The transition probability estimates for *LINT* are based on very few switches and so the precision of these estimates can be expected to be low. However, with the introduction of inflation targeting in 2000 (see Mboweni (2000) for details), periods of persistent high inflation such as were experienced from 1973-1994 would be expected to occur less often in future, and this would be expected to concentrate 20-year bond yields in the low mean regime. Hence, in using the MS-model for projection purposes, the user is likely to decrease the values of p_{12} and p_{22} , rather than rely on these parameter estimates.

The residuals from the AR(1) model for *LINT* for the period 1970:1-2006:2 exhibit a kurtosis of 3.5, a Jarque-Bera statistic of 3 and a p -value of 0.216, suggesting that the assumption of normality cannot be rejected. However, the Ljung-Box test statistics on the squared residuals indicate significant serial correlation structure in the volatility. This suggests fitting a GARCH model to *LINT*.

The GARCH(1,1) model successfully removes all serial correlation in the squared residuals. The Jarque-Bera statistic of 2.3 for the residuals suggests that the assumption of normality cannot be rejected. The log-likelihood for the GARCH(1,1) model is 521.36, which is slightly lower than the log-likelihood of 524.41 for the MS-model in Equation (7). However, the GARCH model suffers from upward bias in the autoregressive parameter ϕ_1 , and so cannot be recommended.

4.5 *XSEQ*

One of the earliest and most enduring models of the behaviour of stock price trajectories is the random walk model. Such a model implies that equity returns are independent and identically distributed (i.i.d.). The Black-Scholes option pricing theory extends this model by assuming that returns over any discrete time interval follow a lognormal distribution. These standard models are accommodated by modelling *XSEQ* as an i.i.d. normal random variable, i.e. the AR(0) model.

There is now a vast literature suggesting that the standard lognormal model is inadequate. Empirical studies of equity returns provide evidence of time varying volatility that the standard lognormal model is unable to capture. Engle (1995, p xii) states that “[t]he GARCH(1,1) model is [now] the leading generic model for almost all asset classes of returns,” and presents a collection of papers on variants of the ARCH model used in finance.

More recently, MS-models have been successfully applied to modelling equity returns. Harris (1996, 1999) introduced the regime switching lognormal model for equity returns, Bollen (1998) prices American and European options under this model, and Hardy (2001) successfully applies the approach of Harris (1999) and Bollen (*op. cit.*) to US and Canadian data.

The standard Jarque-Bera test is used to test the null hypothesis that *XSEQ* follows a normal distribution. *XSEQ* has a negative skewness of -0.24 and a kurtosis of 4.1 . The Jarque-Bera statistic is 10.41029 and has a p -value of 0.005 , which indicates a severe departure from normality. Superficially, this suggests fitting a model with a fat-tailed residual distribution.

The Ljung-Box test statistics on the squared residuals of the AR(0) model, however, indicate significant serial correlation structure in the volatility. This indicates the need for a model incorporating time-varying volatility rather than simply a model with a fat-tailed residual distribution. For this purpose, the ARCH(1) model turns out to be a better model for *XSEQ*, with the GARCH(1,1) model being over-parameterized. The log-likelihood for the ARCH(1) model is -118.0685 .

The results of fitting various MS-models to $XSEQ$ are shown in Table A.4.1 of Appendix A. The first order autoregressive term is not significant, as would be expected under the Efficient Market Hypothesis, and the results from this class of models are not shown. The null of a single state model ($M=1$) is rejected in favour of $M=2$, with a significant LR statistic of 17.16 and a p -value of 0.0034 after applying Davies' (1977) correction. $M=3$ is rejected. In terms of the Schwartz Information Criterion (SC), the best model is the autoregressive model with switching in both the intercept term and the residual variance:

$$XSEQ_t = c_{s_t} + \varepsilon_t, \quad (8)$$

where $\varepsilon_t \sim N(0, \sigma_{s_t}^2)$, $c_1 = -0.0333(0.093)$ and $c_2 = 0.1404(0.041)$. Standard errors are shown in parentheses. $\sigma_1 = 0.62057$ and $\sigma_2 = 0.30581$, so $s_t=1$ corresponds to a volatile, low-mean return regime with an effective annual volatility of 31%, while $s_t=2$ corresponds to a stable, high-mean return regime with an effective annual volatility of 15%. The model is stable when estimated over the sub-periods from 1960:1-1998:4 and 1970:1-2006:2. The log-likelihood for the MS-model in Equation (8) is -112.9, which is considerably higher than the log-likelihood of -118.0685 for the corresponding ARCH(1) model. Note, however, that the standard likelihood ratio test cannot be used to compare these two models because they are not nested.

The probability of remaining in regime 1 given that the process is already in regime 1 is $p_{11}=0.8259$ while the probability of remaining in regime 2 given that the process is already in regime 2 is $p_{22}=0.8853$. The ergodic or stable-state probabilities are 0.3972 for the volatile, low-mean state and 0.6028 for the stable, high-mean return state. The expected duration for the volatile, low-mean state is about 6 quarters while that for the high mean state is about 9 quarters. The estimated probability (conditional on all the data) that the regime was in the volatile, low-mean return regime each quarter is shown in Figure 8.

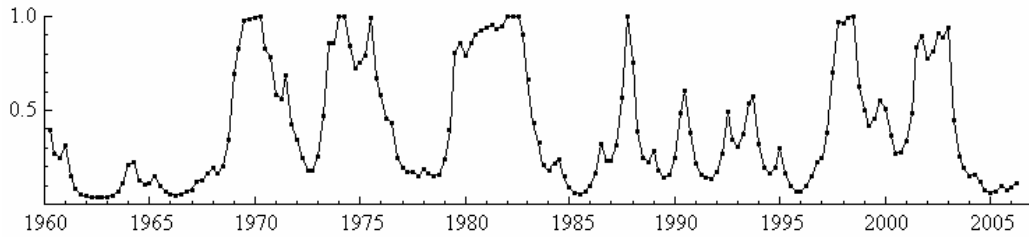


Figure 8. Probability that $XSEQ$ was in the volatile regime

Figure 9 shows the unconditional density function for $XSEQ$ and the joint density functions for each of the unobserved regimes, scaled by the probability of being in each of those regimes. Although the distribution of $XSEQ$ conditional on each regime is Gaussian, the mixing distribution exhibits higher kurtosis and negative skewness. However, unlike the unconditional distributions for $INFL$, $SINT$ and $LINT$, the unconditional distribution for $XSEQ$ is unimodal.

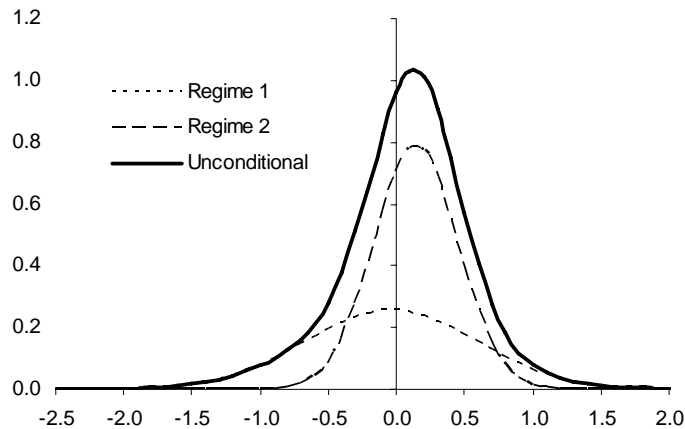


Figure 9. Unconditional and scaled conditional distributions for $XSEQ$

The MSIH(2)-AR(0) model of Equation (8) not only captures the time-varying volatility of quarterly equity returns but also captures a time-varying risk premium. Although the unconditional equity risk premium is about 7%, the risk premium in the volatile return regime is negative while that in the stable (less volatile) regime is strongly positive. This suggests that investors have been poorly rewarded for taking on risk when the equity market is in the volatile regime.

5. A MULTIVARIATE MARKOV SWITCHING FRAMEWORK

Harris (1996) fits univariate Markov switching models to quarterly equity returns and inflation data in Australia. In Harris (1999), he develops a regime switching vector autoregression model, in which a vector, comprising GDP, inflation, share price returns and changes in the long-bond yield, switches between two unobservable states. Although switching between states in Harris' (1996) univariate MS-models is not perfectly correlated, the correlation is high. Hence, this vector switching simplification provides a reasonable description of the joint process.

The Markov Switching Vector Autoregression (MSVAR) framework developed by Krolzig (1997) is restricted mainly to vector switching models in which switching between each univariate component of the vector is contemporaneously perfectly correlated. Krolzig also considers (see Krolzig, 1997, page 127) a bivariate model in which switching between each univariate series is not simultaneous, but where the lag in switching is fixed, which he refers to as intertemporally perfectly correlated regime shifts.

A problem arises when switching between states in the univariate MS-models is poorly correlated (see Figures 5-8 above). Suppose the number of variables is N and each variables is indexed $n \in \{1, 2, \dots, N\}$. Suppose also that the n^{th} variable has M_n possible states at time t , denoted s_t^n , with $s_t^n = j$, $j \in \{1, 2, \dots, M_n\}$. Then, if switching between states is less than perfectly correlated, the total number of states in the multivariate model is $M^{\text{Total}} = M_1 \cdot M_2 \cdot \dots \cdot M_N$.

Parsimony with respect to the number of regimes is extremely desirable in vector switching models since the number of observations feasible for the estimation of regime dependent parameters drops as the number of regimes increases.

For example, a four-variable system with each univariate series containing two regimes requires a total of sixteen regimes. Using the framework presented by Hamilton (1990) and Krolzig (1997), parameter estimates would be required for each

of sixteen autoregressive models (as described by Equation (2)), or worse, each of sixteen vector autoregressive models. Clearly, such an approach is not feasible.

A Tractable Multiple Markov Switching (MMS) Model Framework

Let S_t^* denote the joint state of the multivariate system at time t , and index this state as follows:

$$\begin{aligned}
S_t^* = 1 & \quad \text{if} \quad s_t^1 = 1, s_t^2 = 1, \dots, s_t^{N-1} = 1, s_t^N = 1; \\
S_t^* = 2 & \quad \text{if} \quad s_t^1 = 1, s_t^2 = 1, \dots, s_t^{N-1} = 1, s_t^N = 2; \\
& \dots \\
S_t^* = M_1 & \quad \text{if} \quad s_t^1 = 1, s_t^2 = 1, \dots, s_t^{N-1} = 1, s_t^N = M_N; \\
S_t^* = M_1 + 1 & \quad \text{if} \quad s_t^1 = 1, s_t^2 = 1, \dots, s_t^{N-1} = 2, s_t^N = 1; \\
& \dots \\
S_t^* = M^{Total} & \quad \text{if} \quad s_t^1 = M_1, s_t^2 = M_2, \dots, s_t^{N-1} = M_{N-1}, s_t^N = M_N. \tag{9}
\end{aligned}$$

Let $c_{S_t^*}^n$ denote the parameter $c_{S_t^*}$ for the n^{th} variable in state S_t^* , with similar notation for the autoregressive and residual standard deviation parameters $\phi_{1,S_t^*}^n, \dots, \phi_{p,S_t^*}^n, \sigma_{S_t^*}^n$ of Equation (2). Then place the following across-regime restrictions on the $c_{S_t^*}^n$ parameters and collect these into the reduced set $\theta^n = \left(c_{1,S_t^*}^n, \phi_{1,S_t^*}^n, \dots, \phi_{p,S_t^*}^n, \sigma_{S_t^*}^n \right)$:

$$c_{S_t^*=j_n}^n = c_{S_t^*=k}^n \quad \text{iff} \quad j \equiv k \left(\text{mod} \prod_{l=n}^N M_l \right) \quad \text{or} \quad \text{int} \left(\frac{j-1}{\prod_{l=n+1}^N M_l} \right) = \text{int} \left(\frac{k-1}{\prod_{l=n+1}^N M_l} \right), \tag{10}$$

with similar restrictions for the autoregressive and residual standard deviation parameters of Equation (2). This reduces the number of Equation (2) parameters to the same total number as would be estimated in estimating univariate MS-models for each of the variables.

Let S_t^* be the realization of a Markov chain with the probability of a switch from state i to state j ($i, j = 1, 2, \dots, M^{Total}$) being:

$$\Pr(S_t^* = j \mid S_{t-1}^* = i) = p_{ij}^*, \quad (11)$$

From the indexation in (9), we can find $i_n, j_n \in \{1, 2, \dots, M_n\}$ and $n \in \{1, \dots, N\}$, such that

$$\Pr(S_t^* = j \mid S_{t-1}^* = i) = \Pr(s_t^1 = j_1, \dots, s_t^N = j_N \mid s_{t-1}^1 = i_1, \dots, s_{t-1}^N = i_N), \quad (12)$$

Let Y_t^n represent the subset of our sample for the n^{th} variable (y_t^n, \dots, y_t^n) with $t \in \{1, 2, \dots, T\}$. Given the data to time t for all variables, $Y_t = \{Y_t^1, \dots, Y_t^N\}$, our joint inference about the unobserved states at times t and $t-1$ is:

$$\Pr(S_t^* = j, S_{t-1}^* = i \mid Y_t; \lambda) = \Pr(s_t^1 = j_1, \dots, s_t^N = j_N, s_{t-1}^1 = i_1, \dots, s_{t-1}^N = i_N \mid Y_t^1, \dots, Y_t^N; \lambda), \quad (13)$$

where $\lambda = \{P, \theta^1, \dots, \theta^N, \rho\}$, P^* is the multiple switching transition matrix of probabilities p_{ij}^* , and ρ is the vector of initial state probabilities across all variables (see Hamilton (1990) for further details).

If we assume that the inference in Equation (13) is independent for each of the variables, then it is shown in Appendix B that the maximum likelihood estimates for the transition probabilities satisfy:

$$\hat{p}_{ij}^* = \frac{\sum_{t=2}^T \prod_{n=1}^N \Pr\{s_t^n = j_n, s_{t-1}^n = i_n \mid Y_T^n; \lambda^n\}}{\sum_{t=2}^T \prod_{n=1}^N \Pr\{s_{t-1}^n = i_n \mid Y_T^n; \lambda^n\}} \quad (14)$$

where $\lambda^n = \{P^n, \theta^n, \rho^n\}$, and, for the n^{th} variable, ρ^n is the vector of initial state probabilities and P^n is the transition matrix. The maximum likelihood estimates for the parameters $c_{S_t^*}^n, \phi_{1, S_t^*}^n, \dots, \phi_{\rho, S_t^*}^n, \sigma_{S_t^*}^n$ and $\rho = \{\rho^1, \dots, \rho^N\}$ are the maximum likelihood estimates obtained from estimation of the univariate MS-model for the n^{th} variable.

The maximum likelihood estimate of the residual covariance matrix is the cross-product of the residuals from each of the variables in each state, with each factor weighted by the corresponding smoothed inferences (shown in the subject of the product in the denominator of Equation (14)). For details, see Appendix B.

The assumption of independence in the smoothed inferences for each of the variables in Equation (14) is not restrictive since it is exactly the assumption we made in fitting the univariate MS-models. However, this assumption does not imply that switching in any one variable is independent of switching in the other variables, and by reference to Equation (12) we find in general that:

$$\Pr(S_t^* = j | S_{t-1}^* = i) \neq \prod_{n=1}^N \Pr(s_t^n = j_n | s_{t-1}^n = i_n). \quad (15)$$

6. EMPIRICAL ESTIMATION OF THE MMS MODEL

Table 2 summarizes the univariate MS-model parameters obtained in Section 4 and the ordering, n , of the variables (from left to right) that are used in the MMS model.

	<i>INFL</i>	<i>SINT</i>	<i>LINT</i>	<i>XSEQ</i>
c₁	0.0299	0.0072	0.0129	-0.0333
c₂	0.0944	0.0234	0.0214	0.1404
φ₁	0.2415	0.8659	0.8523	-
σ₁	0.032995	0.005074	0.004082	0.62057
σ₂	-	0.015466	0.008319	0.30581
p₁₁	0.9682	0.9472	0.9730	0.8259
p₂₂	0.9688	0.9204	0.9827	0.8853
ℓ	350.4037	604.6501	524.4088	-112.935

Table 2. MMS model parameters for the individual variables

Calculation of the maximum likelihood estimate of the covariance matrix of the residuals is described in Appendix B. The corresponding correlation matrix of the residuals is shown in Table 3 below.

	<i>INFL</i>	<i>SINT</i>	<i>LINT</i>	<i>XSEQ</i>
<i>INFL</i>	1			
<i>SINT</i>	0.147	1		
<i>LINT</i>	0.211	0.179	1	
<i>XSEQ</i>	0.064	-0.157	-0.082	1

Table 3. Residual correlation matrix for the MMS model

Table 4 shows the maximum likelihood estimate of the transition matrix, P , using joint inferences for the period 1960:1 to 2006:2. The joint inferences for *LINT* are based on parameter estimates for the period 1970:1 to 2006:2 (see Section 4.4 for further details).

In keeping with Hamilton's (1990) approach, the format of the transition matrix, P , in Table 4 is that the transition probability p_{ij} occurs in row j and column i , where p_{ij} is defined as in Equation (11). Hence, the columns of P must sum to one.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.75	0.07	0.02	0	0.26	0.04	0	0	0.04	0	0	0	0.01	0	0	0
2	0.21	0.91	0.01	0.03	0.12	0.37	0	0	0	0	0	0	0	0	0	0
3	0	0	0.65	0.17	0	0	0.08	0.01	0	0	0.01	0	0	0	0	0
4	0	0	0.12	0.73	0	0	0.02	0.05	0	0	0	0.01	0	0	0	0
5	0.01	0	0	0	0.44	0.14	0.03	0	0	0	0	0	0.03	0.01	0	0
6	0	0	0	0	0.08	0.35	0	0.01	0	0	0	0	0	0.01	0	0
7	0	0	0.13	0.02	0.04	0.01	0.6	0.12	0	0	0	0	0	0	0.03	0
8	0	0	0.01	0.03	0	0.03	0.1	0.79	0	0	0	0	0	0	0.01	0.03
9	0.01	0	0	0	0	0	0	0	0.64	0.12	0	0	0.06	0.02	0	0
10	0.01	0.01	0	0	0	0	0	0	0.14	0.81	0	0	0.02	0.17	0	0
11	0	0	0.01	0	0	0	0	0	0.1	0.01	0.56	0.13	0	0	0.01	0
12	0	0	0	0.01	0	0	0	0	0	0.01	0.12	0.76	0	0	0	0.07
13	0	0	0	0	0.05	0.03	0	0	0.05	0.02	0	0	0.79	0.32	0.03	0
14	0	0	0	0	0	0.02	0	0	0	0.02	0	0	0.09	0.45	0	0
15	0	0	0.04	0	0.01	0	0.16	0	0.01	0	0.26	0.03	0.01	0	0.74	0.08
16	0	0	0	0.01	0	0	0.01	0.02	0	0	0.04	0.06	0	0	0.17	0.8

Table 4. Multiple switching transition matrix, P , for the MMS model

The MMS model is a discrete Markov process at discrete time points $t=1,2,3,\dots$ (in our case, quarter-ends) and these states are represented by S_t in Equation (9). Only one state change is possible from one discrete point in time to the next. For example, State 1 at time $t-1$ can switch to any one (but only one) of the 16 states at time t , i.e. given $S_{t-1}=1$, $S_t=I$, where I is any integer from 1 to 16.

Note that S_t also maps to one of the sixteen binary sets represented by the four univariate states ($s_t^1, s_t^2, s_t^3, s_t^4$). Hence it is possible for more than one of these univariate states to switch at the same time. An example of this can be seen in Table 4 where State 1, with binary representation (1,1,1,1), can go to States 2 (1,1,1,2), 3 (1,1,2,1), 5 (1,2,1,1) and 9 (2,1,1,1) as well as State 10 (2,1,1,2), where it can be seen that more than one of the univariate states has switched from 1 to 2.

The ergodic or stable-state probabilities and the expected duration (in quarters) for each state are shown in Table 5.

MMS State, S^*	s^1	s^2	s^3	s^4	Ergodic Pr.	Duration
1	1	1	1	1	0.084	3.9
2	1	1	1	2	0.235	11.0
3	1	1	2	1	0.030	2.9
4	1	1	2	2	0.032	3.7
5	1	2	1	1	0.008	1.8
6	1	2	1	2	0.003	1.5
7	1	2	2	1	0.037	2.5
8	1	2	2	2	0.053	4.7
9	2	1	1	1	0.037	2.8
10	2	1	1	2	0.068	5.3
11	2	1	2	1	0.033	2.3
12	2	1	2	2	0.064	4.2
13	2	2	1	1	0.057	4.7
14	2	2	1	2	0.014	1.8
15	2	2	2	1	0.116	3.8
16	2	2	2	2	0.130	4.9

Table 5. States, ergodic probabilities and durations

For example, Row 2 in Tables 4 and 5 corresponds to the joint state with low expected inflation, low expected short- and long-term interest rates and stable excess equity returns. Historically, this state had the highest duration and the system remained in this state the longest.

Row 16 in Tables 4 and 5 corresponds to the joint state with high expected inflation, high expected short- and long-term interest rates and stable excess equity returns. This state had the next highest duration and was the second most persistent state in the system.

Intuition might suggest that *SINT* and *LINT* should be in the same state at any one time. From the ergodic probabilities in Table 5, the probability that *SINT* and *LINT* are in the same state is calculated by summing rows 1,2,9 and 10, as well as rows 7,8,15 and 16. An alternative might be to model *LINT* as a ‘base’ variable, and then to model a variable ‘*SPREAD=SINT-LINT*’. However, using the tests outlined in Section 4, the null hypothesis of a single state for *SPREAD* is rejected in favour of a two-state MS-model. Similar results hold for the spreads of *SINT* over *INFL* and *LINT* over *INFL*. Hence, no reduction in states is possible from this approach.

For the *LINT* and *INFL*, another possible approach to modelling the dynamics of these variables is to model a series where *INFL* jumps first and then *LINT* jumps with a variable lag. Such a model has been proposed by Durland and McCurdy (1994). However, from a financial perspective, while past expectations in the bond market failed to adequately reflect future inflation, one would not want to force such a structure on a stochastic model as it would not permit instances where the bond market correctly anticipates future inflation and inflation shocks.

With inflation targeting and tight monetary policy, one might expect short-term and possibly long-term interest rates to rise with, or following, an increase in inflation. However, if the system is in states 9 or 10 (representing high expected inflation and low expected short- and long-term interest rates), it is more likely to remain in these states than move to states 13 or 14, which would equate to an increase in short term interest rates. (See columns 9 and 10 of Table 4 where it can be seen that the row 9 and 10 transition probabilities are much higher than those in rows 13 and 14.)

Clearly, the estimated transition matrix parameters are purely a description of past experience and are not appropriate for projection purposes, especially given the current framework of inflation targeting in South Africa.

However, while certain aspects of these dynamics may not be useful for projection purposes, it is possible to adjust these aspects while conditioning on other aspects that may be still be relevant. It is also possible to adjust the multiple switching transition matrix and regime specific parameters independently of one another. This makes the proposed framework very powerful as a tool for extracting those aspects of the past that one believes may be relevant to the future. It also makes market-consistent projections and stochastic scenario testing relatively simple.

Parameterization of the model for projection purposes is beyond the scope of this paper. Also, the effect of parameter uncertainty on projection distributions cannot be assessed in the proposed framework. Joint parameter uncertainty could be modelled using the MCMC approach, as in Harris (1999), but this left for future research.

7. CONCLUSION

The Multiple Markov Switching framework presented in this paper allows the modelling of Markov-switching variables where switches in variables are not perfectly correlated.

The approach to modelling Markov-switching variables follows Krolzig (1997) in recommending a bottom-up approach in which univariate Markov-switching models are used to identify states. The correlation, or otherwise, of switching between these states can then be used to guide the choice of vector-switching or multiple Markov switching for the multivariate model.

Maximum likelihood estimation of the parameters is shown to be relatively simple once the univariate Markov-switching models have been estimated.

The framework is used to estimate a Multiple Markov Switching (MMS) model of South African financial and economic variables. The variables estimated are by definition descriptive of a past that is unlikely to repeat itself. However, while certain dynamics may not be useful for projection purposes, it is possible to adjust these

aspects while conditioning on other features that may be still be relevant for projection purposes.

As part of such adjustments, it is also possible to include current market conditions so that projections are market consistent. Together with the framework suggested in Maitland (2002), it is possible to construct an arbitrage-free model of the local market. It is suggested that framework can be used for various actuarial applications, especially those involving long-term projections.

ACKNOWLEDGEMENTS

The author wishes to thank Professor David Wilkie and two anonymous referees for their extensive and helpful comments and suggestions on an earlier draft of this paper. The author would also like to thank Dr James Greener and Graham Smale at the JSE for their help in clarifying the nature of the bond yield data in the 1960s. The author retains sole responsibility for any errors or omissions that remain.

REFERENCES

- Aron, J. & Muellbauer, J. (2000), 'Inflation and Output Forecasting for South Africa: Monetary Transmission Implications,' *Working paper Series*, 2000-23, Nuffield College, Oxford University.
- Bollerslev, T. (1986), 'Generalized Autoregressive Conditional Heteroscedasticity,' *Journal of Econometrics*, **31**, 307-327.
- Bollen, N.P.B. (1998). 'Valuing options in regime switching models,' *Journal of Derivatives*, **6**, 38-49.
- Carter, J. (1991), 'The derivation and application of an Australian stochastic investment model,' *Transactions of the Institute of Actuaries of Australia*, **1**, 315-428.
- Cosslett, Stephen R., and Lung-Fei Lee (1985), 'Serial Correlation in Discrete Variable Models,' *Journal of Econometrics*, **27**, 79-97.
- Claasen, M.S. (1993), 'A systematic approach to the financial management and risk control of employee benefit funds,' *3rd AFIR International Colloquium*.
- Davies, R. (1977) 'Hypothesis testing when a nuisance parameter is present only under the alternative' *Biometrika*, **64**, 247-254.
- Dickey, D.A. and 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.
- Dickey, D.A. and Fuller, W.A. (1981), 'Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root,' *Econometrica*, **49**, 1057-1072.
- Draper, N.R. and Smith, H. (1981) *Applied Regression Analysis*. Second edition, Wiley.
- Durland, J.M. and McCurdy, T.H. (1994), "Duration Dependent Transitions in a Markov Model of U.S. GNP Growth," *Journal of Business and Economic Statistics*, **12**, 279-88.
- Engle, R.F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation,' *Econometrica*, **50**, 987-1007.
- Engle, R.F. (1995), *ARCH Selected Readings*, Oxford University Press.
- Garcia, Rene (1998), 'Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models,' *International Economic Review*, **39**, 763-788.
- Granger, C.W.J. and Newbold, P. (1974), 'Spurious regressions in Economics'

- Journal of Econometrics*, **2**, 111-120.
- Goldfeld, Stephen M., and Richard E. Quandt (1973), "A Markov Model for Switching Regressions," *Journal of Econometrics*, **1**, 3-16.
- Hamilton, J.D. (1989), 'A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,' *Econometrica*, **57**, 357-384.
- Hamilton, J.D. (1990), 'Analysis of time series subject to changes in regime,' *Journal of Econometrics*, **45**, 39-70.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Hardy, M.R. (2001). 'A regime-switching model of long-term stock returns,' *North American Actuarial Journal*, **52**, 41-53
- Harris, G. (1994) 'On Australian stochastic share return models for actuarial use,' *The Institute of Actuaries of Australia Quarterly Journal*, September, 34-54.
- Harris, G. (1996) "Market phases and cycles? A regime switching approach." *The Institute of Actuaries of Australia Quarterly Journal*, December, 28-44.
- Harris, G. (1999), 'Regime switching vector autoregressions: a Bayesian Markov Chain Monte Carlo approach,' *ASTIN*, **29**, 1, May 1999.
- Homer, S. and Sylla, R. (2005), *A History of Interest Rates*. Fourth Edition, Wiley.
- Hua, G. (1994), 'Modelling equity returns using a simple ARCH model,' *5th AFIR International Colloquium*, **2**, 537-565.
- INET Graphics and Database Service (1998). Intelligent Network (Pty) Ltd.
- INET Graphics and Database Service (2006). Intelligent Network (Pty) Ltd.
- JSE Information Products Sales Division (30 October 2005) "FTSE / JSE End of Day Indices Dissemination User Manual."
- Krolzig, Hans-Martin (1997), *Markov-Switching Vector Autoregressions : Modelling, Statistical Inference, and Application to Business Cycle Analysis*, Berlin: Springer.
- Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt & Yongcheol Shin (1992). 'Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root,' *Journal of Econometrics*, **54**, 159-178.
- Maitland, A.J. (1996). 'A Review of Thomson's Stochastic Investment Model.' *Transactions of the Actuarial Society of South Africa*, 1996.
- Maitland, A.J. (1997). 'Non-Stationarity in some South African Financial and Economic Series,' *7th International AFIR Colloquium*, **7**, 631-46.

- Maitland, A.J. (2001). An Empirical Approach to Immunization in South Africa. *South African Actuarial Journal*, **1**, 114-33.
- Maitland, A.J. (2002). Interpolating the South African Yield Curve using Principal Components Analysis: A Descriptive Approach. *South African Actuarial Journal*, **2**, 129-45.
- Mboweni, T.T. (2000), "A new monetary policy framework," Statement of the Monetary Policy Committee of South Africa, 6 April 2000.
- PGN 201 (2003) "RETIREMENT FUNDS - ACTUARIAL VALUATION REPORTS AND RELATED TOPICS," Actuarial Society of South Africa.
- Perron, P. (1989), 'The Great Crash, the Oil Price Shock and the Unit Root Hypothesis,' *Econometrica*, **57**, 1361-1401.
- Perron, P. (1991), 'Test consistency with varying sampling frequency,' *Econometric Theory*, **7**, 341-368.
- Phillips, P.C.B. and Perron, P. (1988), 'Testing for a unit root in time series regression,' *Biometrika*, **75**, 335-346.
- Ranne, A. (1998) "The Finish Stochastic investment model," Transactions of the 26th International Congress of Actuaries, **7**, 213-238.
- Sherris, M., Tedesco, L. and Zehnwirth, B. (1999) "Investment returns and inflation models: some Australian evidence," *British Actuarial Journal*, **5**, 1, 237-267.
- Statistical Research Division (2000). *X-12-ARIMA Reference Manual*. U.S Census Bureau, Washington DC, 20233-9100.
- Thomson, R.J. (1996) "Stochastic investment modelling - the case of South Africa," *British Actuarial Journal*, **2**, 765-802.
- Thomson, R.J. (2006) "A typology of models used in Actuarial Science," *South African Actuarial Journal*, **6**, 19-36.
- Wilkie, A.D. (1986). 'A stochastic investment model for actuarial use,' *Transactions of the Faculty of Actuaries*, **39**, 341-403.
- Wilkie, A.D. (1995). 'More on a stochastic asset model for actuarial use,' *British Actuarial Journal*, **1**, 777-964.

APPENDIX A
UNIVARIATE MS-MODEL PARAMETER ESTIMATES

In the tables that follow, MSIAH(M)-AR(p) refers to MS-models with M states and switching in any combination of the intercept term (I), the autoregressive terms (A) and the variance of the residuals (H). Standard errors are shown in parentheses. ℓ refers to the log-likelihood value, and “Davies” refers to the p -value as defined by Equation (4).

	MSI(2)-AR(1)	MSIH(2)-AR(1)	AR(1)
c₁	0.0299 (0.005)	0.0264 (0.006)	0.0274 (0.005)
c₂	0.0944 (0.013)	0.092 (0.012)	-
φ₁	0.2415 (0.097)	0.2376 (0.072)	0.6685 (0.055)
σ₁	0.032995	0.028097	0.0401
σ₂	-	0.036052	-
p₁₁	0.9682	0.9581	-
p₂₂	0.9688	0.9666	-
ℓ	350.4037	352.0626	333.4258
AIC	-3.7233	-3.7304	-
SC	-3.6188	-3.6086	-
Davies	0.000	0.000	-

Table A.1.1. *INFL* MS-model estimates (1960:1-2006:2)

	MSI(2)-AR(1)	MSIH(2)-AR(1)	AR(1)
c_1	0.032 (0.006)	0.0207 (0.005)	0.0319 (0.006)
c_2	0.0972 (0.012)	0.0845 (0.010)	-
ϕ_1	0.2134 (0.084)	0.2527 (0.080)	0.6431 (0.062)
σ_1	0.034518	0.02668	0.0413
σ_2	-	0.039215	-
ρ_{11}	0.9727	0.9775	-
ρ_{22}	0.9847	1.0000	-
ℓ	291.6738	295.1857	274.924
AIC	-3.6861	-3.7185	-
SC	-3.5683	-3.5811	-
Davies	0.000	0.000	-

Table A.1.2. *INFL* MS-model estimates (1960:1-1998:4)

	MSI(2)-AR(1)	MSIH(2)-AR(1)	AR(1)
c_1	0.0487 (0.007)	0.0504 (0.007)	0.0433 (0.008)
c_2	0.1001 (0.012)	0.1025 (0.011)	-
ϕ_1	0.2149 (0.082)	0.2022 (0.082)	0.5560 (0.069)
σ_1	0.035371	0.039314	0.0407
σ_2	-	0.031462	-
ρ_{11}	0.9823	0.984	-
ρ_{22}	0.9851	0.986	-
ℓ	273.72	275.3228	261.3556
AIC	-3.6674	-3.6757	-
SC	-3.5448	-3.5326	-
Davies	0.000	0.000	-

Table A.1.3. *INFL* MS-model estimates (1970:1-2006:2)

	MSI(2)-AR(1)	MSIH(2)-AR(1)	AR(1)
c₁	0.0078 (0.002)	0.0072 (0.002)	0.0040 (0.002)
c₂	0.0312 (0.003)	0.0234 (0.006)	-
φ₁	0.8609 (0.020)	0.8659 (0.037)	0.9626 (0.019)
σ₁	0.0082221	0.0050743	0.0124
σ₂	-	0.015466	-
p₁₁	0.9368	0.9472	-
p₂₂	0.8295	0.9204	-
ℓ	582.855	604.6501	550.2362
AIC	-6.2363	-6.4611	-
SC	-6.1318	-6.3392	-
Davies	0.000	0.000	-

Table A.2.1. *SINT* MS-model estimates (1960:1-2006:2)

	MSI(2)-AR(1)	MSIH(2)-AR(1)	AR(1)
c₁	0.0074 (0.002)	0.0061 (0.002)	0.0041 (0.002)
c₂	0.0313 (0.003)	0.0218 (0.004)	-
φ₁	0.8617 (0.023)	0.8794 (0.030)	0.9680 (0.020)
σ₁	0.0084873	0.0048662	0.0129
σ₂	-	0.015909	-
p₁₁	0.9299	0.9466	-
p₂₂	0.8429	0.9448	-
ℓ	481.1991	502.6978	455.0007
AIC	-6.1316	-6.3961	-
SC	-6.0138	-6.2587	-
Davies	0.000	0.000	-

Table A.2.2. *SINT* MS-model estimates (1960:1-1998:4)

	MSI(2)-AR(1)	MSIH(2)-AR(1)	AR(1)
c_1	0.0109 (0.003)	0.0118 (0.002)	0.0066 (0.003)
c_2	0.0346 (0.004)	0.0315 (0.004)	-
ϕ_1	0.8358 (0.029)	0.8169 (0.022)	0.9441 (0.027)
σ_1	0.0091264	0.0058739	0.0138
σ_2	-	0.014679	-
p_{11}	0.9107	0.9185	-
p_{22}	0.8292	0.9111	-
ℓ	440.8397	450.9521	418.6994
AIC	-5.9567	-6.0815	-
SC	-5.8341	-5.9385	-
Davies	0.000	0.000	-

Table A.2.3. *SINT* MS-model estimates (1970:1-2006:2)

	MSI(2)-AR(1)	MSIAH(2)-AR(1)	AR(1)
c_1	0.0031 (0.001)	0.0017 (0.004)	0.0023 (0.001)
c_2	0.0168 (0.002)	0.0066 (0.005)	-
$\phi_{1,1}$	0.9597 (0.012)	0.9803 (0.099)	0.9802 (0.013)
$\phi_{1,2}$	-	0.9503 (0.039)	-
σ_1	0.0051189	0.0024094	0.0066
σ_2	-	0.0082008	-
p_{11}	0.9343	0.9668	-
p_{22}	0.4288	0.9861	-
ℓ	678.0345	705.5335	667.1863
AIC	-7.2652	-7.5409	-
SC	-7.1608	-7.4016	-
Davies	0.000	0.000	-

Table A.3.1. *LINT* MS-model estimates (1960:1-2006:2)

	MSI(2)-AR(1)	MSIAH(2)-AR(1)	AR(1)
c_1	0.003 (0.001)	0.0016 (0.001)	0.0023 (0.001)
c_2	0.0171 (0.002)	0.0271 (0.008)	-
$\phi_{1,1}$	0.9649 (0.011)	0.9831 (0.015)	0.9846 (0.013)
$\phi_{1,2}$	-	0.8174 (0.057)	-
σ_1	0.0049612	0.00251	0.0065
σ_2	-	0.0081903	-
ρ_{11}	0.9387	0.9732	-
ρ_{22}	0.4207	0.9844	-
ℓ	572.2671	605.6214	561.7657
AIC	-7.3067	-7.7112	-
SC	-7.1889	-7.5542	-
Davies	0.000	0.000	-

Table A.3.2. *LINT* MS-model estimates (1960:1-1998:4)

	MSI(2)-AR(1)	MSIH(2)-AR(1)	AR(1)
c_1	0.0167 (0.005)	0.0129 (0.004)	0.0044 (0.002)
c_2	0.0284 (0.007)	0.0214 (0.006)	-
ϕ_1	0.8134 (0.054)	0.8523 (0.041)	0.9646 (0.020)
σ_1	0.0063606	0.0040818	0.0074
σ_2	-	0.0083189	-
ρ_{11}	0.9518	0.9730	-
ρ_{22}	0.9532	0.9827	-
ℓ	513.3791	524.4088	511.0785
AIC	-6.9504	-7.0878	-
SC	-6.8278	-6.9447	-
Davies	0.2034	0.000	-

Table A.3.3. *LINT* MS-model estimates (1970:1-2006:2)

	MSI(2)-AR(0)	MSIH(2)-AR(0)	AR(0)
c₁	-0.6606 (0.202)	-0.0333 (0.093)	0.0707 (0.034)
c₂	0.1473 (0.045)	0.1404 (0.041)	-
σ₁	0.40225	0.62057	0.4680
σ₂	-	0.30581	-
p₁₁	0.3361	0.8259	-
p₂₂	0.9311	0.8853	-
ℓ	-118.7513	-112.9345	-121.5166
AIC	1.3379	1.2858	-
SC	1.4249	1.3902	-
Davies	0.1368	0.0034	-

Table A.4.1. *XSEQ* MS-model estimates (1960:1-2006:2)

	MSI(2)-AR(0)	MSIH(2)-AR(0)	AR(0)
c₁	-0.6993 (0.209)	-0.0350 (0.107)	0.0601 (0.038)
c₂	0.1412 (0.048)	0.1172 (0.041)	-
σ₁	0.40582	0.65736	0.4773
σ₂	-	0.30653	-
p₁₁	0.3294	0.8384	-
p₂₂	0.9290	0.9004	-
ℓ	-102.2351	-95.8759	-104.7849
AIC	1.3837	1.3145	-
SC	1.4819	1.4323	-
Davies	0.1646	0.0025	-

Table A.4.2. *XSEQ* MS-model estimates (1960:1-1998:4)

	MSI(2)-AR(0)	MSIH(2)-AR(0)	AR(0)
c₁	-0.3085 (0.507)	-0.0396 (0.123)	0.0586 (0.041)
c₂	0.1615 (0.099)	0.1210 (0.054)	-
σ₁	0.4547	0.65933	0.4962
σ₂	-	0.33745	-
p₁₁	0.3398	0.7764	-
p₂₂	0.8158	0.8630	-
ℓ	-103.943	-99.5641	-104.348
AIC	1.4924	1.4461	-
SC	1.5945	1.5687	-
Davies	0.8471	0.0884	-

Table A.4.3. *XSEQ* MS-model estimates (1970:1-2006:2)

APPENDIX B
MAXIMUM LIKELIHOOD ESTIMATION

Suppose we have a sample of size T from a vector-valued autoregressive process $y_t \in \mathfrak{R}^n$, and suppose that the parameters, θ , of this process switch between a finite number of alternatives based on the unobserved states, s_t . If the probability of switching from one state to another can be expressed in the form of a Markov chain, then Hamilton (1990; p51) shows that the maximum likelihood estimates for the transition probabilities satisfy

$$\hat{P}_{ij} = \frac{\sum_{t=2}^T \Pr\{s_t = j, s_{t-1} = i \mid Y_T; \lambda\}}{\sum_{t=2}^T \Pr\{s_{t-1} = i \mid Y_T; \lambda\}} \quad (\text{B1})$$

Using the notation introduced in Section 5 of this paper, it follows that

$$\hat{P}_{ij}^* = \frac{\sum_{t=2}^T \Pr\{S_t^* = j, S_{t-1}^* = i \mid Y_T; \lambda\}}{\sum_{t=2}^T \Pr\{S_{t-1}^* = i \mid Y_T; \lambda\}}, \quad (\text{B2})$$

which, from Equation (13), can be written as

$$\hat{P}_{ij}^* = \frac{\sum_{t=2}^T \Pr(s_t^1 = j_1, \dots, s_t^N = j_N, s_{t-1}^1 = i_1, \dots, s_{t-1}^N = i_N \mid Y_T^1, \dots, Y_T^N; \lambda)}{\sum_{t=2}^T \Pr(s_{t-1}^1 = i_1, \dots, s_{t-1}^N = i_N \mid Y_T^1, \dots, Y_T^N; \lambda)}, \quad (\text{B3})$$

If we now assume that the joint smoothed-inference in Equation (B3) is independent for each of the variables, then

$$\Pr(s_t^1 = j_1, \dots, s_t^N = j_N, s_{t-1}^1 = i_1, \dots, s_{t-1}^N = i_N | Y_t^1, \dots, Y_t^N; \lambda) = \prod_{n=1}^N \Pr(s_t^n = j_{n1}, s_{t-1}^n = i_n | Y_t^n; \lambda^n)$$

and

$$\Pr(s_t^1 = j_1, \dots, s_t^N = j_N | Y_t^1, \dots, Y_t^N; \lambda) = \prod_{n=1}^N \Pr(s_t^n = j_{n1} | Y_t^n; \lambda^n), \quad (\text{B4})$$

This assumption is not restrictive since it is exactly the assumption we made in fitting the univariate MS-models in our bottom-up identification strategy. Substituting into Equation (B3) gives the desired result:

$$\hat{p}_{ij}^* = \frac{\sum_{t=2}^T \prod_{n=1}^N \Pr\{s_t^n = j_n, s_{t-1}^n = i_n | Y_t^n; \lambda^n\}}{\sum_{t=2}^T \prod_{n=1}^N \Pr\{s_{t-1}^n = i_n | Y_t^n; \lambda^n\}} \quad (\text{B5})$$

Generalising from Equation (2), let $\varepsilon_t^{n,j}$ denote the error term associated with regime $S_t^* = j$ for the n^{th} variable at time t , and collect these into a vector, $\varepsilon^{n,j}$, spanning the length of the time series. Let $\xi_{t|T}^j$ denote the smoothed regime probability, $\Pr(S_t^* = j | Y_T; \lambda)$, and collect these into a vector, ξ_T^j , spanning the length of the time series. Then, under the assumption of heteroscedasticity in either of the two variables under consideration, Krolzig (1997, p108) shows that the maximum likelihood estimate of the covariance in state j between variables $n=p$ and $n=q$ is:

$$\text{Cov}_j(y^p : y^q) = \left(\sum_t \varepsilon_t^{p,j} \cdot \varepsilon_t^{q,j} \cdot \xi_{t|T}^j \right) / T_j, \quad \text{where } T_j = \sum_t \xi_{t|T}^j. \quad (\text{B6})$$

This is the usual cross-product of error terms, weighted by the smoothed regime probability at each point in time and divided by the probability-weighted time that the series have been in regime j .

Krolzig (*op. cit.*) also shows that the maximum likelihood estimate of the covariance under homoscedasticity is the cross-product of error terms, weighted by the smoothed

regime probability at each point in time and then summed across all regimes before dividing by T.

$$Cov_j(y^p : y^q) = \sum_{j=1}^{M^{Total}} T_j \cdot Cov_j(y^p : y^q) / T. \quad (B7)$$

From the univariate analysis of the series, it may transpire that certain variables are heteroscedastic while others are homoscedastic. Let $S_i^* = u(k)$ denote the k^{th} union of states with constant variance for variables $n=p$ and $n=q$. Then the maximum likelihood estimate of the covariance for the k^{th} union of states is:

$$Cov_{u(k)}(y^p : y^q) = \sum_{u(k)} T_j \cdot Cov_j(y^p : y^q) / \sum_{u(k)} T_j. \quad (B8)$$

(Note: the summation is taken across all states that are homoscedastic.)

If the correlations are assumed to be constant across all states, an estimate of the correlation between variables $n=p$ and $n=q$ can be obtained as follows:

$$Corr(y^p : y^q) = \sum_{j=1}^{M^{Total}} T_j \cdot \left(Cov_j(y^p : y^q) / \sqrt{\sigma_j^p \cdot \sigma_j^q} \right) / T. \quad (B9)$$

Endnote for Chapter 5

The paper in this chapter has introduced a new class of Markov switching models for multivariate modelling. The model framework allows each individual variable from a vector of variables to switch at different times and for this reason is called a Multiple Markov Switching (MMS) model. In this respect, it differs from other multivariate Markov switching models discussed in the literature, where all elements of the vector of variables must switch at the same time, or with a fixed and known time-lag. In the MMS model, switching in individual variables is not perfectly correlated, nor is it completely independent; instead, the probability of a switch from one state to another is contingent on the states of other variables and the variable itself.

The MMS model presented in this chapter models the core South African financial and economic variables required for ALM. The regime switching model provides a better fit to each of the variables than a variety of other models, including linear time series models and generalized autoregressive conditional heteroscedastic models. Maximum likelihood estimates of the MMS parameters are derived and shown to require only the smoothed inferences obtained from a univariate analysis of the variables. Hence, the MMS framework is also tractable and computationally efficient. Regime switching models also avoid the problem of parameter bias found in linear models, which is a valuable feature, especially if the model is to be used for projection purposes.

The paper in this chapter was submitted to the South African Actuarial Journal in 2006. Since then, the following stochastic models for actuarial use in South Africa have appeared in the literature.

Howie (2007) presents a stochastic investment model for actuarial use in the United States, Britain and South Africa. Howie's model is a structural model of equities, bonds and cash, as well as inflation and economic growth. It is based on economic and financial theory and is split into an economic model and a financial market model. The economic variables of inflation, real interest rates and real economic growth are each first modelled as regressions on current and past values of each other, subject to non-circularity

constraints. Equities and bonds are then modelled as the discounted value of their projected dividend and coupon proceeds, discounted at a risk-adjusted rate. The discount rate is a linear function of inflation, the real interest rate and a risk premium. Structural parameters are estimated subject to theoretical constraints, but these structures and constraints are not tested empirically. For example, real earnings growth is assumed to be positively related to real economic growth and negatively to real interest rates. Time-varying volatilities are not considered although the financial market returns simulated by the model are non-normal and appear to exhibit significant leptokurtosis (fat-tails) with higher probabilities of severe down-market returns than are predicted by normal or log-normal distributions. The Howie model is a flexible generalised mixture model comprising a mean reverting component and a non-mean-reverting component for each of the economic variables. Given the proposed parameters however, the economic model for each country is non-stationary, which seems problematic for long-term projection purposes.

Thomson and Gott (2009) present a long-term equilibrium model of a local market of equity returns, yields on nominal and index-linked bonds, and inflation. This model is a theoretical model in that it relies on the abstract notion of an instrument earning an instantaneous risk-free real rate of interest, and various assumptions, including that market participants have homogenous expectations, that they are able to borrow or lend unlimited amounts at the same risk-free, and that they make decisions in mean-variance space. The model also assumes that means and variances are constant. Parameters are estimated using annual UK data, and the risk-free asset is taken to be a one-year index-linked bond, which does not generally exist in the South African market. In an earlier version of their paper, Thomson and Gott (2006) acknowledge that “due to the lack of data, it is difficult to parameterise the model descriptively for the South African market.”

Thomson (2008) presents descriptive univariate models of the real annual force of return on a proxy of the market portfolio for South Africa. The ‘market portfolio’ is taken to comprise listed South African equity and government bonds, aggregated in proportion to their dynamic market capitalisations. Various univariate time series models are used to

describe the market return as a multiple of a real risk-free return plus a constant plus an error term, with various possibilities in which the multiple, the constant and the error volatility may be time-varying. In conclusion, a basic model with no time-varying multiples, constants or volatilities is put forward. The univariate models presented in Thomson (2008) are not comparable with the MMS model of Chapter 5 since they are based on different data sets and model different variables.

Gott (2009) constructs a model for inflation, an equity index and the real and nominal yield curves for the purposes of pricing life-insurance embedded derivatives on a particular date. Parameters of the interest rate models are estimated with reference to historical correlations between changes in yields, and the model is calibrated to the South African swaptions data available as at 30 June 2007. The principal-component vectors are consistent with the findings of Maitland (2002), as well as the particular zero-coupon bonds selected as the drivers of the full yield curve. The equity component of the model uses the Heston (1993) model, with parameter estimates fitted to implied volatilities for the JSE TOP40 index futures options as at 30 June 2007. The equity model is not necessarily descriptive of real world equity returns and has not been tested against historical data. Its intended purpose is the market consistent pricing of equity-contingent claims on a particular date. The Gott model is not a descriptive model of real world returns or interest rates but rather a theoretical model calibrated to price various option contracts on a particular date.

Parameters for each of the models of Howie (2007), Thomson (2008) and Gott (2009) have been estimated using annual data from at most the last 20 years. In contrast, the MMS model presented in Chapter 5 is based on data 46 years of quarterly data.

Real world stochastic models are often used in the joint modelling of assets and liabilities of an individual or institution. Typically, the financial position is projected forward in time as a function of each scenario generated by the stochastic model via Monte Carlo simulation. Key financial indicators of interest to the decision maker (such as the surplus of assets over liabilities) are defined and estimated. For these financial indicators, certain

statistics such as the mean, standard deviation and partial moments can be calculated to indicate various measures of risk and reward, and the tradeoff between them. Finally, the tradeoff between risk and reward can be analyzed and optimized with respect to decision variables such as the asset allocation and other variables affecting the financial position. Alternatively, the decision-maker's utility function might be employed to measure and maximize expected utility.

The next chapter introduces an alternative approach to optimizing the asset allocation in respect of a particular class of liabilities. These liabilities take the form of a set of nominal liability cash flows where the size and timing of each cash flow is known with sufficient accuracy to implement an immunization process. Where immunization is possible, it provides a framework that is computationally more efficient than Monte Carlo simulation. It is also aimed at identifying a specific portfolio of bonds to immunize certain risks and optimize returns. Such an optimization within the Monte Carlo framework would require a massive increase in complexity and computational power.

CHAPTER 6

Foreword to the Paper

The paper in Chapter 6 (Maitland, 2001) considers asset and liability modelling in the specific case that the liabilities, or a subset of the liabilities, take the form of a set of nominal liability cash flows. It is assumed that the size and timing of each cash flow is known with certainty. The present value of these cash flows is calculated with reference to the nominal yield curve, and sufficient assets are put aside in a fund to ensure the liabilities are fully funded. Ideally, the cash flow proceeds from those assets would exactly match the liability cash flows; however, in practice, such assets are often not available and alternative strategies must be considered.

Ignoring credit risks, nominal bonds each provide a set of known cash flows whose present values can be calculated with reference to the nominal yield curve. However, even if the present value of the portfolio of bonds equals the present value of the liability cash flows to start with, a change in interest rates results in new present values, and generally these are no longer equal. In some cases, it is possible to remove this interest rate risk using dynamic investment strategies. “Immunization” is the name given to the construction of such strategies that protect or “immunize” the assets and liabilities of the fund against interest rate risk.

The paper in this chapter extends the application of Principal Components Analysis introduced in Chapter 4 to develop immunization strategies that protect against the major risk factors in the South African nominal yield curve. Where immunization is possible, it provides an optimization framework that is computationally more efficient than the Monte Carlo methods and optimization used with stochastic models such as that presented in Chapter 5.

AN EMPIRICAL APPROACH TO IMMUNIZATION IN SOUTH AFRICA

By AJ Maitland

ABSTRACT

This paper presents an empirical approach to immunizing South African nominal liabilities in the presence of non-parallel yield-curve shifts. The results are compared with immunization strategies based on Fisher-Weil duration and illustrate the value in immunizing against non-parallel shifts.

KEYWORDS

Immunization; South Africa; principal components; nominal liabilities; arbitrage

CONTACT DETAILS

James Maitland, Stochastic Solutions cc, PO Box 2484, Parklands, 2121, Johannesburg.

Tel: +27 11 486-1946. E-mail: stochsol@iafrica.com.

1. INTRODUCTION

1.1 In his seminal paper, Redington (1952) developed a theory of immunization under which the cash flow stream generated by an investor's nominal liabilities is protected against small changes in the valuation rate of interest. The immunization strategy is achieved by selecting a portfolio of assets with present value and Macaulay duration equal to those of the liabilities and convexity greater than that of the liabilities. Fisher & Weil (1971) adapted the measure of a bond's price elasticity developed by Macaulay (1938) to incorporate a term structure that is not flat and developed an immunization strategy based on this measure of duration. It is well known that these immunization strategies are valid only if term-structure shifts are parallel.

1.2 Consider representing the yields at d points along the yield curve as a single point in d dimensions, $\mathbf{c} = (c_1, c_2, \dots, c_d)$, each dimension corresponding to the maturity of one of the chosen points. A parallel shift in yields results in a new point in d dimensions equal to the original point plus a constant addition to each coordinate. Clearly, the points resulting from any parallel shift in yields will lie along the straight line passing through these two points, $\mathbf{c} + k\mathbf{1}$, where $\mathbf{1}$ is the unit vector and $k \in \mathfrak{R}$. Hence, parallel shifts are restricted to a shift in the direction of the unit vector.

1.3 A more general and empirically plausible model allows the term structure to shift in multiple directions. Using factor analysis, Litterman & Scheinkman (1991) provide empirical evidence that three factors are required to explain the term structure of US interest rates. Similar results are found by Sherris (1994) using Australian yield-curve

data, D'Ecclesia & Zenios (1994) using Italian bond-market data, Bühler & Zimmermann (1996) using Swiss and German interest rates and Feldman *et al* (1998) using real and nominal UK forward rates. A principal components analysis (PCA) on the covariance matrix of monthly changes in par yields for the JSE-Actuaries Yield Curve indicates that the first principal component explains 92,8% of the total variability and the first two principal components together explain 97,3%, while three principal components are required to explain 98,4% of the total variability, see Maitland (2000). More detail is given in Section 2 below.

1.4 Reitano (1991, 1992) developed a general framework for hedging interest-rate uncertainty that immunizes against term-structure shifts in multiple directions. Vectors whose elements correspond to rate changes at different maturity dates describe the shift in each direction. Each additional shift direction specified for immunization imposes at least one extra constraint to the portfolio selection problem. Barber & Copper (1996) use PCA to estimate the minimum number of fundamental directions in which to anticipate spot-rate changes. Unlike the decomposition of yield-curve shifts into parallel shifts, stylized slope changes and stylized curvature changes, or into key rate durations (see Ho, 1992), PCA provides the minimum number of components to explain any desired proportion of the total variability. Further, each subsequent principal component (PC) provides the direction of maximum variability orthogonal to the previous set (also referred to hereafter as a *fundamental* direction). Hence, the largest shifts are immunized as completely as possible and the effect of non-infinitesimal movements in any fundamental direction can be considered independently of non-infinitesimal movements in any other fundamental direction.

1.5 In this paper, the principal components analysis in Maitland (2000) is updated to include data to May 2000 and the work of Barber & Copper (1996) is extended by optimizing the immunized portfolio subject to principal component shifts. Hence, the suggested approach does not rule out the possibility of arbitrage. This is discussed further in Section 4. Hedging strategies immunized to an increasing number of principal component shifts are compared with hedging strategies based on Fisher-Weil duration and the risk of ignoring shape changes is illustrated using the yield curves of 30 September, 31 October and 30 November 1998. Finally, the optimization models introduced in Section 4 are used to identify “conditional” arbitrage opportunities in the September 1998 yield curve. The availability of conditional arbitrage opportunities then suggests a method for choosing the number of principal-component constraints required for immunization. This is discussed further in Section 6.

2. PRINCIPAL COMPONENTS OF THE JSE-ACTUARIES YIELD CURVE

2.1 Let \mathbf{z} be a random d -vector with mean μ and covariance matrix Σ , and let $\mathbf{T} = (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_d)$ be an orthogonal matrix (i.e. $\mathbf{T}\mathbf{T}' = \mathbf{T}'\mathbf{T} = \mathbf{I}$) such that $\mathbf{T}'\Sigma\mathbf{T} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$, where $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ are the eigenvalues of Σ . If $\mathbf{y} = \mathbf{T}'(\mathbf{z} - \mu)$, then $y_j = \mathbf{t}_j'(\mathbf{z} - \mu)$ is

called the j th principal component score of \mathbf{z} and is the orthogonal projection of $\mathbf{z}-\boldsymbol{\mu}$ in the direction \mathbf{t}_j , the j th principal component (Kendall, Stewart & Ord, 1983; 43.4). Hence, the scores at time t are that linear combination of the principal components required to reconstruct the yield curve at that time. Unpacking the matrix notation and letting \mathbf{z}_k represent the k th realization of the random vector \mathbf{z} , we can see that

$$\mathbf{z}_k = \boldsymbol{\mu} + y_{1,k}\mathbf{t}_1 + \dots + y_{d,k}\mathbf{t}_d,$$

from which it becomes clear that the variance of \mathbf{z}_k is

$$V(\mathbf{z}_k) = V(y_{1,k})\mathbf{t}_1\mathbf{t}_1^T + \dots + V(y_{d,k})\mathbf{t}_d\mathbf{t}_d^T,$$

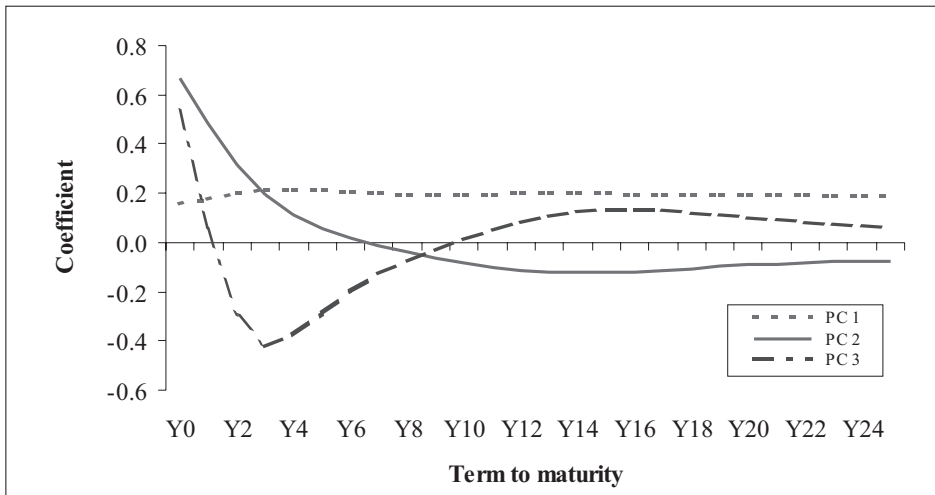
since the eigenvectors are orthogonal. Since $V(\mathbf{y}) = \mathbf{T}'\boldsymbol{\Sigma}\mathbf{T} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d)$, the scores are uncorrelated. Truncating this series to include only the first n terms gives an approximation to \mathbf{z}_k , the accuracy of which depends on the cumulative variability explained by those terms. The decision whether or not to include one more term is based on the incremental variability explained by the additional term.

2.2 Table 1 indicates the additional and cumulative proportions explained by the first ten principal components from the covariance matrix of monthly changes in par yields for the JSE-Actuaries Yield Curve for the period February 1986 to May 2000. It is clear that the first two principal components describe most of the variability of term-structure shifts but that immunization against higher-order shifts may be desired in order to further reduce risk. The last two columns of Table 1 give the months in which the minimum and maximum scores occurred for each of the first ten principal components and indicate months in which extreme exposure to the various risk factors could give cause for concern.

TABLE 1. Variability explained by the first ten principal components

PC No.	Eigenvalues	Incremental variability (%)	Cumulative variability (%)	Minimum score	Maximum Score
1	9,928712	92,415	92,415	Sep 98	Aug 98
2	0,519766	4,838	97,253	Sep 86	Jun 98
3	0,128539	1,196	98,450	Jan 88	Jun 86
4	0,088378	0,823	99,272	Jul 90	Oct 91
5	0,057733	0,537	99,810	Aug 86	Apr 86
6	0,016355	0,152	99,962	Jul 98	Jun 98
7	0,002045	0,019	99,981	Jan 90	Oct 91
8	0,001200	0,011	99,992	Feb 86	May 00
9	0,000346	0,003	99,995	Feb 86	Dec 90
10	0,000141	0,001	99,997	Feb 86	Jul 90

FIGURE 1. Coefficients for the first three principal components



2.3 Figure 1 illustrates the coefficients of the first three principal components by term to maturity. The first principal component affects all maturities by similar amounts and in the same direction. It can be interpreted as a level shift factor but not as a parallel shift factor since the coefficients are unequal. The second principal component has an opposite effect on short and long yields and can be viewed as a slope change factor or twist. The third principal component has a negative effect on medium yields and a positive effect on short- and long-term yields and hence can be interpreted as a curvature factor or butterfly. Figures 2 to 4 illustrate the principal component scores for the first three principal components of yield-curve changes for the period February 1986 to May 2000.

FIGURE 2. Principal component 1 scores for yield changes (Feb 1986–May 2000)

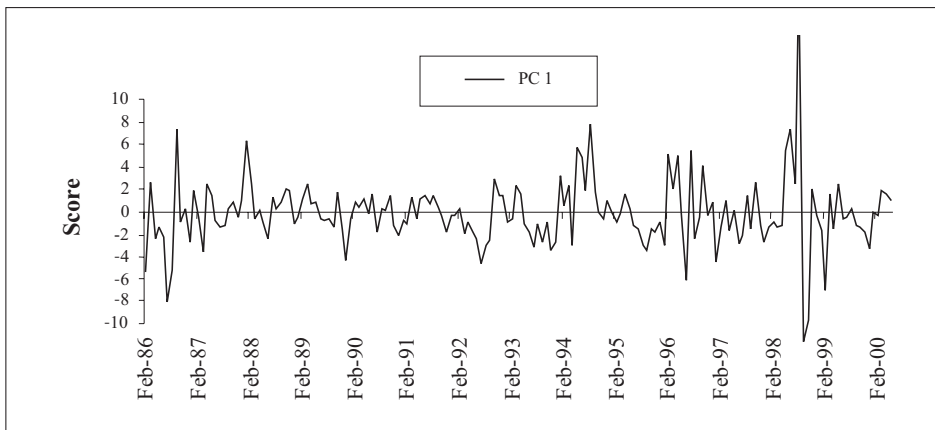


FIGURE 3. Principal component 2 scores for yield changes (Feb 1986–May 2000)

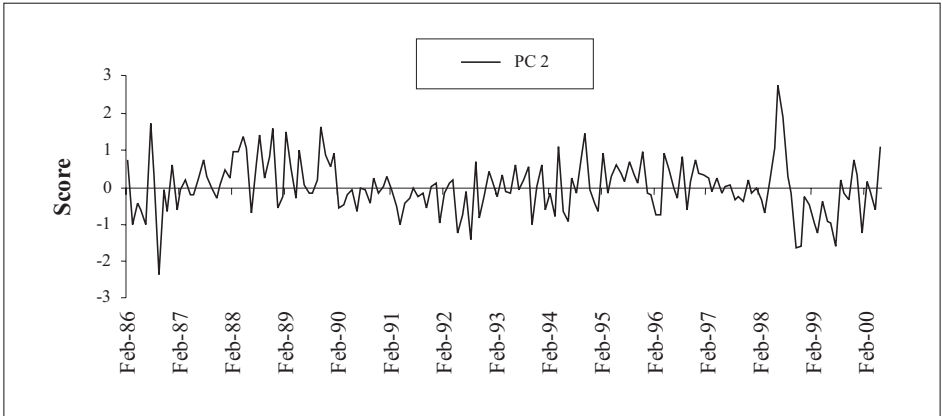
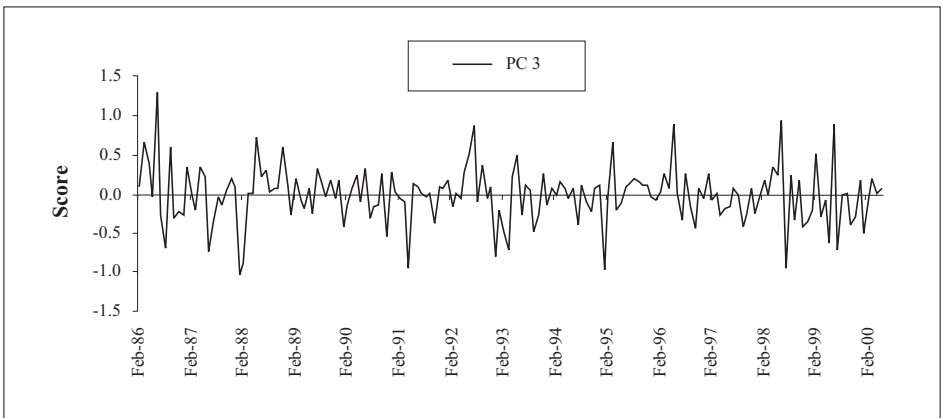


Figure 4. Principal component 3 scores for yield changes (Feb 1986–May 2000)



2.4 Estimates of variances, covariances and correlations can be very sensitive to outliers and so we can expect principal components to have the same sensitivity. The extreme scores for the first principal component between August and October 1998 shown in Figure 2 and the corresponding large changes in the level of the yield curve suggest the need for a PCA for sub-periods of the data. A number of alternative sub-periods have been considered and the results of the full period appear to be relatively robust to the choice of sub-period. In particular, the principal components are robust to outliers from August to November 1998, indicating that the shocks experienced over this period were of the same nature as previous shocks, despite their increased magnitude. The incremental proportions of the total variability explained by each of the first three principal components are also almost identical to those based on data to December 1998 and discussed in Maitland (2000).

3. IMMUNIZATION USING PRINCIPAL COMPONENTS

3.1 Given the par yield curve at time t , \mathbf{x}_t , and assuming that $\mathbf{z}_t = \mathbf{x}_t - \mathbf{x}_{t-1}$, the par yield curve one month forward is given by

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \sum_{j=1}^d y_{j,t} \mathbf{t}_j + \mu \quad (1)$$

3.2 In general, d principal components are required to reproduce all possible term-structure movements but the first $n < d$ principal components may explain a sufficient proportion of shifts. In this case,

$$\mathbf{x}_{t+1} \approx \mathbf{x}_t + \sum_{j=1}^n y_{j,t} \mathbf{t}_j + \mu \quad (2)$$

3.3 Barber & Copper (1996) immunize against spot-curve changes but it is equally possible to immunize against changes in the par-yield curve. It may be theoretically more transparent to analyse spot rates than par yields but the majority of domestic bonds currently in issue have coupons in excess of 10%, so par bonds are more representative of the market than zero-coupon bonds. Further, local market practitioners are often more familiar with par yields than spot rates and the results from a PCA of par yields are more intuitive than those from a PCA of spot rates. More importantly, a PCA analysis of the bootstrapped spot curves indicates that the first n components consistently explain a smaller proportion of the total variability than the corresponding number of components of the par-yield curve. This indicates that a more parsimonious linear model is possible using par yields than spot rates, even though both contain the same information. This is due to the non-linear relationship between par yields and spot rates.

3.4 In the subsequent analysis, it is assumed that par bonds are available at any time for annual maturities between 0 and 25 years and that both liability and coupon cash flows occur at annual intervals. In practice, a liquid market in par bonds at annual maturities along the curve does not exist. The spread of bonds in issue is lumpy (over 50% of the market capitalization being concentrated in just two bonds) so practitioners would probably wish to immunize using corporate debt as well as government bonds. However, for the purposes of this paper, credit and liquidity considerations are ignored. Hence, the empirical nature of the analysis refers to the use of empirical shifts in the yield curve and not to the use of actual bonds available in the market at the time. As a caveat to the subsequent analysis, it should be noted that the JSE-Actuaries Yield Curve is an artificial construct that may poorly reflect the yields of actual bonds traded, which may either disguise true arbitrage opportunities or create their illusion.

3.5 Since the PCA is based on monthly changes in yields, it is necessary to value bonds a month later following the change in yields. To simplify calculations, it is assumed that

the clean price of a bond of s years and eleven months is equal to the clean price of an $(s+1)$ -year bond with the same coupon. Hence, capital returns from the roll-down effect over the month are ignored and only those from fundamental shifts in the par yield curve are considered. The effect of this assumption is minor given that bonds are priced at par at the start of the month and that the yield on a bond of s years and eleven months is almost identical to the yield on an $(s+1)$ -year par bond. Apart from fundamental shifts, it is assumed that coupon income represents the only income generated over the month.

3.6 Let $B_{s,t}$ represent the present value (PV) at time t of a par bond maturing s years hence. If bonds are available at maturities from 0 to 25 years, then

$$V_{A,t} = \sum_{s=0}^{25} \alpha_{s,t} B_{s,t} \quad (3)$$

where $V_{A,t}$ represents the present value of the assets at time t and $\alpha_{s,t}$ represents the holding in bond $B_{s,t}$. Suppose $L_{s,t}$ represents the PV at time t of a liability cash flow payable s years hence. Then the PV of the liabilities at time t is

$$V_{L,t} = \sum_{s=0}^{25} L_{s,t} \quad (4)$$

3.7 At time t , the assets are immunized with respect to the liabilities relative to shifts described by the first n principal components if (see Barber & Copper, 1996)

$$\frac{\partial V_{A,t}}{\partial y_j} = \frac{\partial V_{L,t}}{\partial y_j} \quad \text{for } j = 1, 2, \dots, n. \quad (5)$$

3.8 In addition, Barber & Copper impose the wealth constraint:

$$V_{A,t} = V_{L,t}. \quad (6)$$

4. OPTIMAL IMMUNIZATION

4.1 Since the number of constraints is usually much less than the number of bonds, there exist a variety of portfolios from which to choose. The absolute-match portfolio will always satisfy constraints (5) and (6) but may require short positions in certain par bonds. It may be desirable to impose the non-negativity constraints $\alpha_{s,t} \geq 0 \quad \forall s \in S$, $S = \{0, 1, 2, \dots, 25\}$, although a solution under such constraints may not exist. For the moment, we will assume the existence of a feasible solution such that $\alpha_{s,t} \geq 0 \quad \forall s \in S$.

4.2 One alternative for optimizing the immunized portfolio is to drop the wealth constraint and minimize the capital required. It may be possible to hold a portfolio of assets that actually costs less than the absolute-match portfolio (i.e. $V_{A,t} < V_{L,t}$) and which

is also immunized with the respect to the liabilities. Ignoring capital returns from the roll-down effect, the only difference in monthly income generated from the immunized portfolio versus the absolute-match portfolio results from the difference in coupon income. To incorporate the coupon income earned on assets into the selection process, it is necessary to impose the additional constraint that this coupon income equals that from the absolute match. Since the monthly coupon income is proportionate to the par yield, this defines the following linear programming problem:

$$\begin{aligned}
 & \text{Minimize } V_{A,t} = \sum_{s=0}^{25} \alpha_{s,t} B_{s,t} \quad \text{w.r.t. } \{ \alpha_{s,t} \}_{s=0}^{25} \\
 & \text{subject to } \frac{\partial V_{A,t}}{\partial y_j} = \frac{\partial V_{L,t}}{\partial y_j} \quad \forall j = 1, 2, \dots, n \\
 & \alpha_t' \mathbf{x}_t = \omega_t' \mathbf{x}_t \\
 & \alpha_t \geq \mathbf{0}
 \end{aligned} \tag{7}$$

where α_t represents the column vector with elements $\alpha_{s,t}$ and ω_t represents the mix of bonds in the absolute-match portfolio, $\omega_t \geq \mathbf{0}$.

4.3 The optimum asset mix α_t^* , satisfies $1' \alpha_t \leq 1' \omega_t$ since, for $\omega_t \geq \mathbf{0}$, ω_t is a feasible solution. If $\alpha_t^* = \omega_t$, there is no arbitrage. If $\alpha_t^* \leq \omega_t$, $100 * 1' \alpha_t^* (= A^*$, say) represents the reduced capital required to immunize the liabilities. For a system subject only to infinitesimal shifts of the form $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \dots, \mathbf{t}_n$, the asset mix α_t^* is riskless in relation to the liabilities. In other words, for such shifts, the change in the nominal value of the absolute match will equal that in A^* and the coupon income from both portfolios will be the same. This follows from the first two constraints in equation (7). Hence, following such shifts, the *free capital*, defined as the difference between the present value of the liabilities and the assets (i.e. $100 * (1' \omega_t - 1' \alpha_t^*) = V_{L,t} - A^* = k$, say), will remain unchanged. (It is important to note that at the end of the month an amount of capital equal to k is still required for there to be sufficient funds to meet the liabilities. However, there are no constraints as to how this temporarily freed-up capital should be invested and the capital could simply be held in bank notes if desired.)

4.4 It is tempting to invest the free capital in the risk-free asset to generate additional funds with certainty (i.e. arbitrage profits conditional on the absence of shifts of the form $\mathbf{t}_{n+1}, \dots, \mathbf{t}_d$) and so improve the funding ratio, but this strategy is not optimal (see Model (8)). Also, since the portfolio is immunized only with respect to shifts of the form $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \dots, \mathbf{t}_n$, the apparent arbitrage profits represented by these additional funds are not generated with certainty, however unlikely such shifts may appear historically. "Conditional arbitrage" is discussed further in Section 6.

4.5 Another alternative for optimizing the immunized portfolio is to reintroduce the wealth constraint and maximize the coupon income. The immunization model then becomes:

$$\begin{aligned}
 & \text{Maximize Coupon Income} = \alpha_t' \mathbf{x}_t \quad \text{w.r.t.} \left\{ \alpha_{s,t} \right\}_{s=0}^{25} \\
 & \text{subject to} \quad \frac{\partial V_{A,t}}{\partial y_j} = \frac{\partial V_{L,t}}{\partial y_j} \quad \forall j = 1, 2, \dots, n \\
 & \quad \quad \quad V_{A,t} = V_{L,t} \\
 & \quad \quad \quad \alpha_t \geq \mathbf{0}
 \end{aligned} \tag{8}$$

4.6 If we assume that the value of the assets at time t is equal to the cost of the absolute-match portfolio and that $\alpha_{s,t} \geq 0 \quad \forall s \in S$, the optimum portfolio from Model (8) will generally give a higher return than a composite portfolio comprising the optimum portfolio from (7) together with an amount of capital equal to the free capital invested in the risk-free asset. The reason for this is that the composite portfolio satisfies the constraints of model (8) but forces certain funds into the risk-free asset in contrast to model (8). Again, no arbitrage exists if $\alpha_t^* = \omega_t$.

4.7 In the absence of arbitrage opportunities, it is always possible to solve a system of linear equations to determine an appropriate portfolio. For the models described above, this system of equations is simply the set of constraints specified in the respective linear programming problems. If the number of constraints is less than the number of bonds available, the immunized portfolio will not be unique. Since arbitrage opportunities may arise from time to time, optimization is required to determine the optimum portfolio. For a discussion on the number of constraints to consider, see Section 6.

4.8 The yield to maturity (YTM) might be considered to be an alternative objective function to the coupon income suggested in Model (8). The YTM on a single par bond is equal to its coupon, but the YTM of a portfolio of par bonds is a non-linear function of the YTM of each bond. Although it is possible to obtain a first-order approximation to the portfolio YTM as a linear function of α , this objective function is inappropriate. Since immunization is with respect to monthly changes in yields, bonds must be rebalanced monthly and are unlikely to be held to maturity. Further, unless the optimum portfolio is also the absolute-match portfolio, the immunized portfolio's constraints are violated with the passage of time, even in the absence of yield-curve shifts.

4.9 Model (8) is optimal only if $V_{A,t} = V_{L,t}$. For example, if a surplus is also to be invested in bonds without taking a position on interest-rate movements, the wealth constraint can be replaced with the more general wealth constraint $V_{A,t} = F \cdot V_{L,t}$, where F is the funding ratio. There is then no risk that the funding ratio will change, except to the extent that additional income is received.

4.10 The last constraint, $\alpha_t \geq \mathbf{0}$, in Model (7) and Model (8) is necessary to avoid short positions in certain bonds. So far, it has been assumed that no short positions are required in the absolute-match portfolio, i.e. $\omega_t \geq \mathbf{0}$. If the absolute-match portfolio contains short positions but these are not permitted in the immunized portfolio, a feasible solution to models (7) or (8) may not exist. If short positions are permitted in the immunized portfolio, they may still be limited to the extent required by the absolute-match portfolio. This suggests two possible modifications to the last constraint in models (7) and (8):

$$\begin{aligned} \text{(i)} \quad & \alpha_t \geq \min(\omega_t, \mathbf{0}) \\ \text{(ii)} \quad & [\min(\alpha_t, \mathbf{0})]' \mathbf{1} \geq [\min(\omega_t, \mathbf{0})]' \mathbf{1} \end{aligned} \quad (9)$$

4.11 Under (i), the short position in any bond may not exceed the corresponding short position in the absolute-match portfolio. Since ω_t is known *a priori*, models (7) and (8) remain linear programming problems. Under (ii), the total funds generated from short positions in the immunized portfolio may not exceed those generated from short positions in the absolute-match portfolio. Since optimization is with respect to α and constraint (ii) is non-linear in α , with (ii) replacing $\alpha \geq \mathbf{0}$, models (7) and (8) are no longer standard linear programming problems. However, they can still be solved with minor modifications to the simplex algorithm.

4.12 Another alternative when short sales are required in the absolute-match portfolio is to minimize the total funds generated from short positions (FGSP):

$$\begin{aligned} \text{Minimize } FGSP &= [\min(\alpha_t, \mathbf{0})]' \mathbf{1} \quad w.r.t. \left\{ \alpha_{s,t} \right\}_{s=0}^{25} \\ \text{subject to } & \frac{\partial V_{A,t}}{\partial y_j} = \frac{\partial V_{L,t}}{\partial y_j} \quad \forall j = 1, 2, \dots, n \\ & V_{A,t} = V_{L,t} \\ & \alpha_t' \mathbf{x}_t = \omega_t' \mathbf{x}_t \end{aligned} \quad (10)$$

4.13 In addition to the constraints suggested in models (7), (8) and (10), we may also wish to impose the second-order constraints:

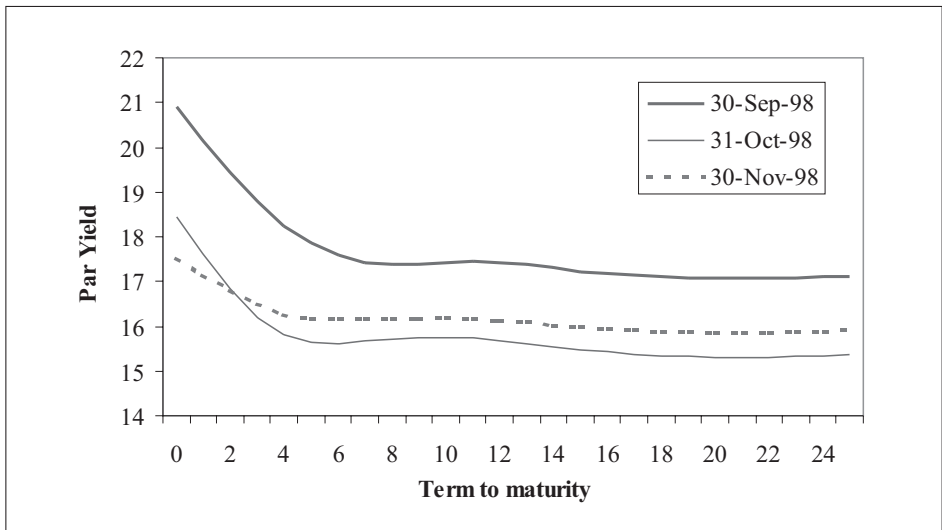
$$\frac{\partial^2 V_{A,t}}{\partial y_j^2} \geq \frac{\partial^2 V_{L,t}}{\partial y_j^2} \quad \forall j = 1, 2, \dots, n \quad (11)$$

4.14 These constraints ensure that the value of the assets is greater than or equal to that of the liabilities for non-infinitesimal shifts of the form $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \dots, \mathbf{t}_n$. This represents an additional source of arbitrage in the qualified sense that it is known with certainty that only the immunized fundamental shifts will occur. “Conditional arbitrage” is discussed further in Section 6.

5. EMPIRICAL ANALYSIS

5.1 In this section, the immunization strategy described in (8) is used to construct portfolios for a level stream of liability cash flows of R100 payable annually in arrear for five years. The term structures of 30 September, 31 October and 30 November 1998 shown in Figure 3 are used to illustrate these strategies.

Figure 5. JSE-Actuaries Yield Curve



5.2 The perfect-match portfolios as at 30 September and 31 October 1998 are shown in Table 2, together with the optimum immunized portfolios subject to one, two and three principal-component, partial-derivative constraints. The principal-component constraints use *ex-ante* estimates of the principal components so that the results provide an *ex-post* test of the data. The optimum immunized portfolios subject only to parallel shifts and built using standard duration-matching techniques, namely the Fisher-Weil duration, are also shown in Table 2. The holdings represent the rand investment in each par bond and maturities with zero holdings are omitted for brevity.

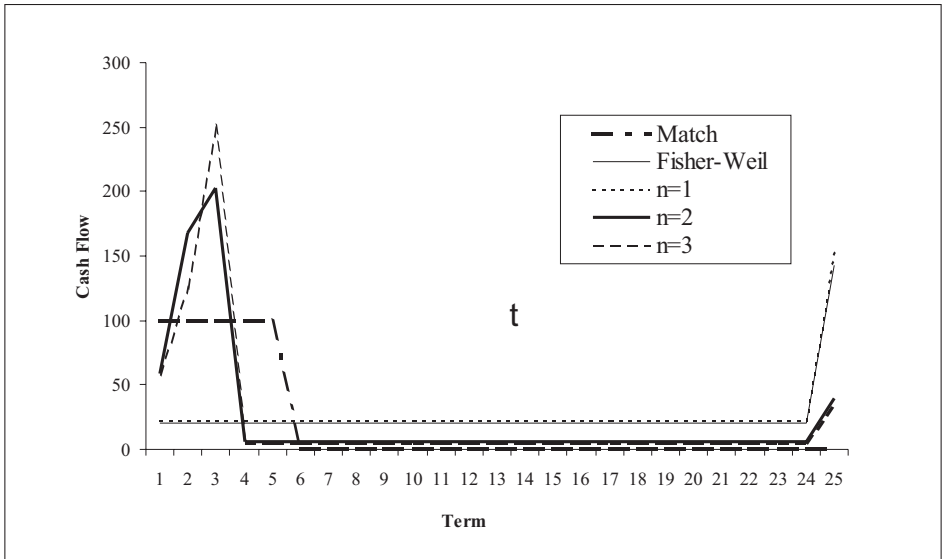
5.3 The no-arbitrage values of the liabilities at 30 September and 31 October 1998 are R309-71 and R327-12 respectively. In both cases it is assumed that the liability cash flows occur at annual intervals from the date of immunization. That is, for the portfolio identified on 30 September 1998, liability cash flows occur at the end of September in subsequent years, while for the portfolio identified on 31 October 1998, liability cash flows occur at the end of October in subsequent years. The future cash flows generated by each of the portfolios constructed for 30 September 1998 are shown in Figure 4. (Cash flows for portfolios constructed in October are not shown since they are almost identical.)

TABLE 2. Optimum immunized portfolios using parallel and principal component shifts

30 Sep 98					
Maturity	Match	Fisher-Weil	n=1	n=2	n=3
0	0,000	187,436	178,753	0,000	5,827
1	42,094	0,000	0,000	0,000	0,000
2	50,584	0,000	0,000	110,263	68,206
3	60,417	0,000	0,000	165,965	207,524
4	71,764	0,000	0,000	0,000	0,000
5	84,854	0,000	0,000	0,000	0,000
25	0,000	122,277	130,960	33,484	28,156
Monthly Coupon Income					
	4,826	5,009	4,982	4,861	4,856
31 Oct 98					
Maturity	Match	Fisher-Weil	n=1	n=2	n=3
0	0,000	205,764	195,080	0,000	0,000
1	46,755	0,000	0,000	0,000	0,000
2	54,993	0,000	0,000	133,811	108,799
3	64,249	0,000	0,000	158,578	180,885
4	74,657	0,000	0,000	0,000	0,000
5	86,468	0,000	0,000	0,000	0,000
25	0,000	121,358	132,042	34,733	17,457
Monthly Coupon income					
	4,437	4,717	4,689	4,462	4,453

5.4 The Fisher-Weil immunization strategy and the strategy immunized against changes in only the first principal component produce almost equivalent barbell strategies, since the first principal component is almost a parallel shift. In contrast, immunization against changes in the first two and the first three principal components produces strategies with cash flows that more closely match the liability cash flows. Clearly, as the number of constraints increases, the maximum coupon income decreases. The perfect-match portfolio is immunized against all changes to the par-yield curve but has the lowest yield.

FIGURE 6. Cash flows of assets and liabilities using Fisher-Weil and $n=1,2$ & 3 principal component immunization (30 Sep 98)



5.5 The advantage of the principal-component immunization approach is that it allows the portfolio manager to hedge against different types of risk while quantifying the sacrifice in yield. By dropping certain principal-component, partial-derivative constraints, the manager can take active risk in those yield-curve dynamics while hedging against others. For example, with the downward-sloping yield curves of September and October 1998, it may be reasonable to expect a downward shift in yields. Under this expectation, a manager might structure a barbell portfolio similar to the Fisher-Weil and single-principal-component portfolios described in Table 2.

5.6 The results are somewhat unexpected. For the portfolios of September 1998 when yields were relatively high, the drop in yields to 31 October appears to result from a roughly parallel shift (see Figure 3). This might suggest positive results from the barbell strategies. However, the post-shift present values of R327-51 and R327-57 for the hedges with two and three principal-component, partial-derivative constraints respectively exceed the present values of R323-30 for the Fisher-Weil and R324-27 for those with one. The PV of the liabilities a month later is R327-12, indicating substantial risks in the barbell strategies compared with the more fully immunized strategies. This paradox is resolved by noting that the score for the second principal component is relatively large and negative, indicating a flattening of the curve.

5.7 The benefits of the barbell portfolio one month hence for the October 1998 portfolios are again questionable. Under parallel shifts, the convexity effects of the

barbell strategy will produce excess funds; however, it is well known that parallel shifts are more the exception than the rule. Portfolio managers often use rules of thumb when assessing basis risk. One such rule is that when yields rise, shorter rates tend to rise faster than longer rates. Under such circumstances, a barbell portfolio immunized using the traditional Fisher-Weil duration will outperform the benchmark because of convexity effects, and because shorter rates may rise faster than longer rates. However, for the October 1998 portfolios, the post-shift present values of R323-11 for the Fisher-Weil and R322-75 for the single-principal-component hedges indicate poor immunity to the November shift since the present value of the liabilities at the end of November is R324-55. In contrast, the post-shift present values of R325-07 and R325-09 for the hedges with two and three principal-component, partial-derivative constraints respectively are roughly equivalent to the present value of the liabilities a month later. The reason for the poor results from the Fisher-Weil and the single-principal-component hedges is that the score for the second principal component is again relatively large and negative, indicating a further flattening of the curve.

5.8 Figures 5 to 8 show the price movements in response to parallel and fundamental yield-curve shifts for the absolute-match portfolio and each of the four immunized portfolios constructed for 30 September 1998. Considering each type of shift in isolation, price movements for the four portfolios relative to price movements in the absolute-match portfolio illustrate how well each portfolio is immunized against that type of shift.

FIGURE 7. Price movements following parallel shifts in the yield curve

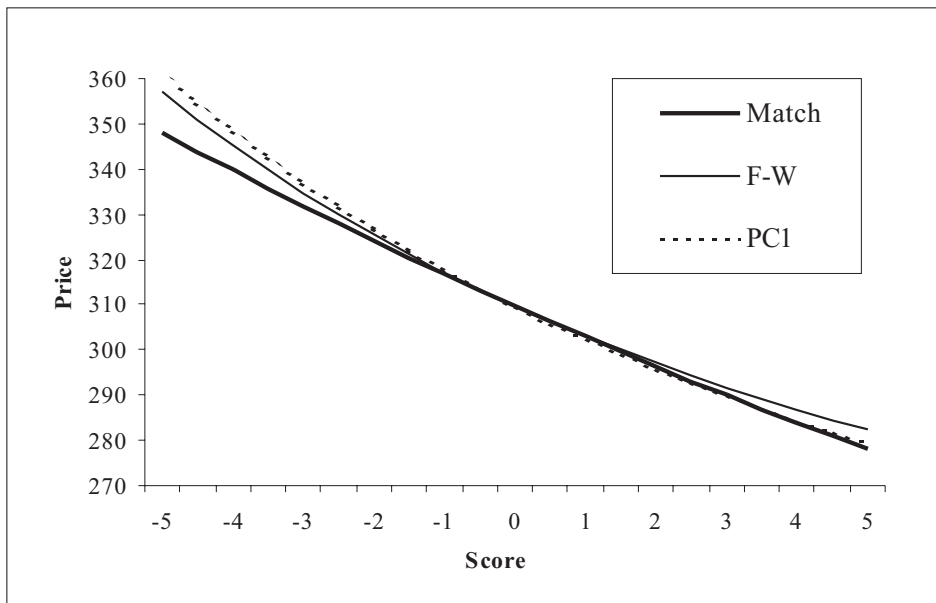
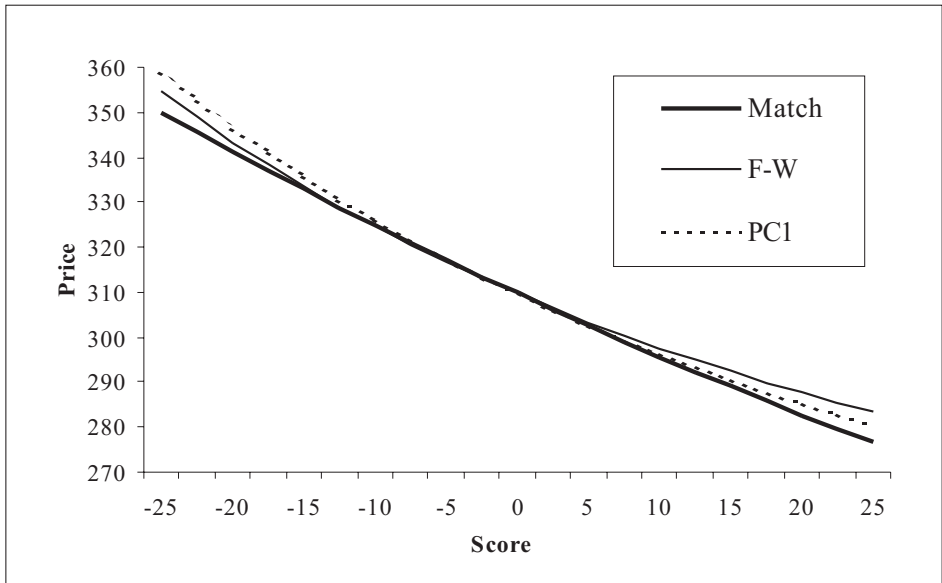


FIGURE 8. Price movements following PC1 shifts in the yield curve



5.9 The graphs of price movements for the two portfolios immunized against shifts in the first two and the first three principal components have been omitted from figures 5 and 6 for clarity. Following parallel or PC1-type shifts, prices for these two portfolios lie between the prices for the absolute-match portfolio and those for the portfolio immunized against a shift in only the first principal component.

5.10 While all portfolios are fairly well immunized against parallel and PC1-type shifts, the Fisher-Weil and single-principal-component hedges are poorly immunized against PC2 and PC3-type shifts. This is obvious to some extent since shifting the curve in a way that is not anticipated by the Fisher-Weil hedge, for example, will result in poor performance from that hedge. The extent to which the unanticipated shift affects the hedge portfolio depends on how exposed the hedge is to that type of shift. What is clear from Figures 9 and 10 is just how exposed the Fisher-Weil and single-principal-component hedges are to PC2 and PC3-type shifts. For September 1998, the tracking errors for the Fisher-Weil and single-principal-component hedges are 12% and 9% respectively, while for the hedges immunized against PC2 and PC3-type shifts the corresponding error is 1% in both cases.

5.11 Although the price graphs in Figures 5 to 8 are specific to the portfolios constructed for 30 September 1998, they give a good indication of price movements for 31 October 1998 portfolios since the corresponding portfolios are similar. The 30 September 1998 and 31 October 1998 portfolios immunized against PC2-type shifts are marginally

FIGURE 9. Price movements following PC2 shifts in the yield curve

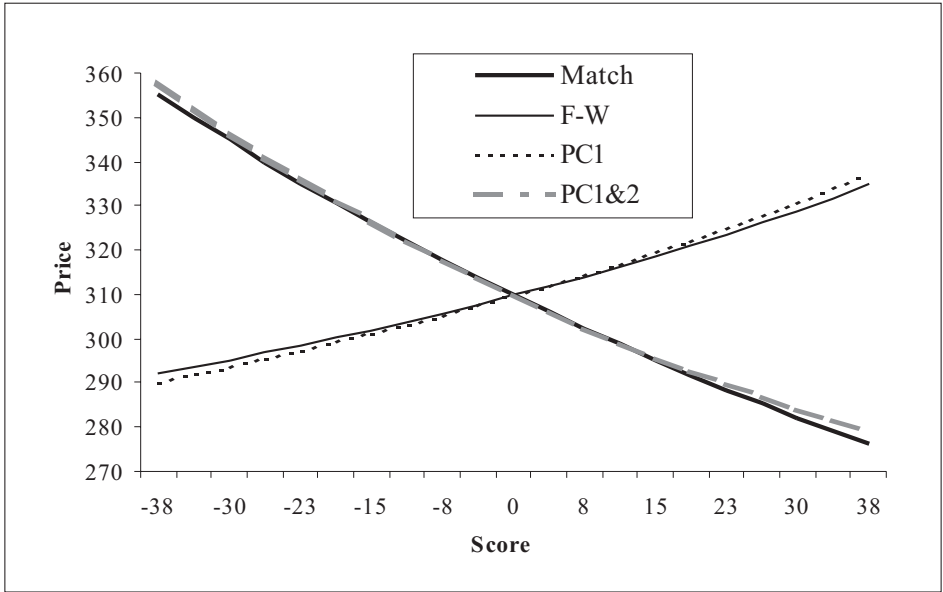
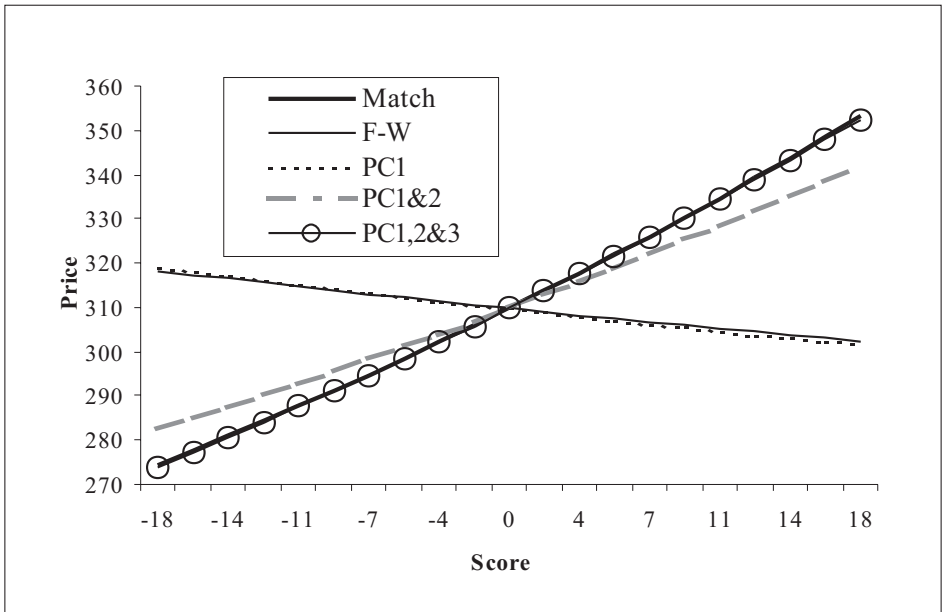


FIGURE 10. Price movements following PC3 shifts in the yield curve



exposed to PC3-type shifts. This accounts for most of the difference in PVs for the two- and three-principal-component hedges following the shocks of October and November 1998. However, since these errors are relatively small and since the error for the PC1&2 hedge shown in Figure 10 is relatively small, a fund manager hedging this liability stream might not be too concerned about immunizing against PC3-type shifts.

5.12 The first two principal components illustrated in Figure 1 are both roughly level beyond a maturity of ten years. This might suggest that for a liability cash flow with maturity greater than ten years, an immunized portfolio subject to both the first two principal-component constraints would be equivalent to an immunized portfolio subject to only the first. However, since an immunized portfolio may contain bonds of any maturity, including bonds with a maturity of less than ten years, these two immunized portfolios will not be equivalent in general. Hence, the five-year annuity considered in the above example is also illustrative of the risks faced by alternative nominal liabilities.

6. CONDITIONAL ARBITRAGE

6.1 For any nominal liability cash flow stream, the coupon income from the optimum portfolio given by Model (8) will be at least as great as that from the absolute-match portfolio since the latter also satisfies the constraints of Model (8). For a system subject only to infinitesimal shifts of the form $t_1, t_2, t_3, \dots, t_n$, an immunized portfolio subject to the first n principal-component, partial-derivative constraints is risk-free in relation to the liabilities.

6.2 If short selling is permitted, the short sale of an s -year par bond creates a nominal liability, which can be immunized in the same manner. If the optimum portfolio consists of bonds other than the s -year par bond and if the coupon income from the optimum portfolio is greater than that from the s -year par bond, an arbitrage opportunity exists. By short selling the s -year par bond and using the proceeds to purchase the optimum portfolio, a monthly risk-free profit equal to the difference between the monthly coupon income from the optimum portfolio and that from the s -year par bond is created. It should now be clear why short sales must be limited: unlimited short sales give rise to infinite arbitrage profits and result in an unbounded objective function.

6.3 Table 3 illustrates the maximum arbitrage profits possible from the short sale of R100 nominal of each s -year par bond for optimum portfolios immunized against shifts in one, two and three principal components, while the last row gives the maximum arbitrage profits possible from the short sale of any par bond. For the short sale of any s -year par bond, as the number of principal-component partial-derivative constraints increases, the maximum possible arbitrage profit from the optimum immunized portfolio decreases.

6.4 Since the optimum portfolios generating the arbitrage profits illustrated in Table 3 are only immunized against at most the first three fundamental shifts, the apparent

arbitrage profits represented by these additional funds are not generated with certainty. For fundamental shifts of the form t_4, t_5, \dots, t_d , such profits are not guaranteed. However unlikely such shifts may appear historically, there is always the possibility that they may occur in future. Hence, such arbitrage opportunities might be referred to as “conditional arbitrage”.

TABLE 3. The excess monthly coupon income per R100 nominal short-sale of each s -year par bond produced by the optimum, immunized portfolio subject to 1, 2 and 3 principal-component, partial-derivative constraints (Sep 98).

s	n=1	n=2	n=3
1	0,019	0,000	0,000
2	0,035	0,000	0,000
3	0,049	0,000	0,000
4	0,060	0,000	0,000
5	0,067	0,000	0,000
6	0,071	0,025	0,000
7	0,068	0,043	0,000
8	0,059	0,048	0,000
9	0,045	0,042	0,000
10	0,030	0,029	0,000
11	0,016	0,014	0,000
12	0,008	0,004	0,000
13	0,005	0,000	0,000
14	0,006	0,000	0,000
15	0,007	0,000	0,000
16	0,007	0,000	0,000
17	0,009	0,000	0,000
18	0,010	0,003	0,000
19	0,010	0,006	0,000
20	0,010	0,008	0,000
21	0,009	0,008	0,000
22	0,008	0,007	0,000
23	0,006	0,006	0,000
24	0,003	0,003	0,000
25	0,000	0,000	0,000
Max	0,071	0,048	0,000

6.5 For 30 September 1998, this section shows that negligible arbitrage opportunities exist if one conditions on three principal-component constraints. Hence, for this period, it may be imprudent to condition on less than three PC constraints when selecting an immunized portfolio, since if the market did not think such shifts possible the arbitrage opportunities illustrated for $n=1&2$ in Table 3 should not exist. Although this logic is applicable to a stream of liability cash flows in general, certain portfolios may be only marginally exposed to PC3-type shifts so that the PC3-constraint can effectively be ignored, as discussed in Section 5.

6.6 It should be noted that any portfolio other than the absolute-match portfolio always contains risk since it is actually possible for the yield curve to move in any number of ways in reality. Since the scientific method is based on the presumption that the past will in some sense be like the future, the above analysis conditions on what was historically likely.

7. CONCLUSION

7.1 Principal components analysis provides a parsimonious description of historical South African yield-curve dynamics. In this paper, a variety of models have been introduced to immunize against such dynamics. These models are optimal in two senses: firstly, they minimize the number of constraints required to immunize against any desired proportion of the total variability and, secondly, they maximize the income in excess of that produced by the absolute-match portfolio. The optimization models can be used to identify optimal portfolios for immunizing any nominal cash flow stream, the short sale of any existing bond or a portfolio of bonds. Hence, these models can also be used for enhanced index tracking.

7.2 The immunization strategies presented in Section 5 accentuate the substantial risk of using the traditional Fisher-Weil duration as the only measure of risk. Clearly, the more fully immunized strategies bear less risk and illustrate the importance of immunizing against shifts other than parallel shifts. Further, it is likely that optimization will maximize exposure to non-immunized shifts. Hence, the increased income from optimum portfolios with fewer principal-component, partial-derivative constraints should always be weighed against the investor's risk tolerance to non-immunized shocks and the market's risk premium for these shocks as implied by the yield curve at that time.

ACKNOWLEDGEMENTS

The author would like to thank Jaco van der Walt for stimulating discussions leading to the development of the minimum capital immunization strategy and for introducing him to the paper by Barber & Copper (1996). The author would also like to thank Rob Thomson and Keith Guthrie for helpful comments on an earlier draft of this paper and Jonathan Quadling for programming assistance. Helpful comments from two anonymous referees and a research grant from the Actuarial Society of South Africa are gratefully acknowledged.

REFERENCES

- Barber JR & Copper ML (1996). Immunization using principal components analysis, *The Journal of Portfolio Management*, 22 (3) , 99–105
- Bühler A & Zimmermann H (1996). A statistical analysis of the term structure of interest rates in Switzerland and Germany, *The Journal of Fixed Income*, 6, 55–67
- D'Ecclesia RL & Zenios SA (1994). Risk factor analysis and portfolio immunization in the Italian bond market, *The Journal of Fixed Income*, 4, 51–9
- Feldman KS, Bergman B, Cairns AJG, Chaplin GB, Gwilt GD, Lockyer PR & Turley FB (1998). Report of the fixed interest working group, *BAJ*, 4, 213–383
- Fisher L & Weil RL (1971). Coping with the risk of interest rate fluctuations: returns to bond holders from a naïve and optimal strategy, *Journal of Business*, 44 (3), 408–31
- Ho TSY (1992). Key rate durations: measures of interest rate risk, *The Journal of Fixed Income*, 2, 29–44
- Kendall M, Stewart A & Ord JK (1983). *The Advanced Theory of Statistics*, Vol. 3, Charles Griffen
- Litterman R & Scheinkman J (1991). Common Factors Affecting Bond Returns, *Journal of Fixed Income*, 1 (1), 54–61
- Macaulay FR (1938). Some theoretical problems suggested by the movement of interest rates, bond yields, and stock prices in U.S. since 1856, National Bureau of Economic Research, New York
- Maitland AJ (2000). Interpolating the South African Yield Curve, *10th AFIR Colloquium*, 469–82
- Redington FM (1952). Review of the principles of life office valuations, *JIA*, 78, 286–315
- Reitano RR (1991). Multivariate immunization theory, *TSA XLIII*, 392–438
- Reitano RR (1992). Non-parallel yield-curve shifts and immunization, *Journal of Portfolio Management*, 18, 36–43
- Sherris M (1995). Interest rate risk factors in the Australian bond market, *5th AFIR Colloquium*, 859–70

Endnote for Chapter 6

The methodology proposed in this paper provides an empirical framework for the immunization of South African nominal liabilities in the presence of non-parallel yield-curve shifts. It extends the traditional approach to immunization, which is valid only if yield-curve shifts are parallel.

Principal components analysis provides a parsimonious description of historical South African yield-curve dynamics. In this paper, a variety of new models have been put forward to immunize against such dynamics. These models are optimal in two senses: firstly, they minimize the number of constraints required to immunize against any desired proportion of the total variability of interest rate risk and, secondly, they maximize the income in excess of that produced by the absolute-match portfolio. The optimization models can be used to identify optimal portfolios for immunizing any nominal cash flow stream.

The immunization strategies presented are shown to substantially reduce the risk when compared with using the traditional Fisher-Weil duration as the only measure of risk. The more fully immunized strategies bear less risk and illustrate the importance of immunizing against shifts other than parallel shifts.

It should be noted that any portfolio other than the absolute-match portfolio always contains risk since it is possible for the yield curve to move in any number of ways in reality. Since the scientific method is based on the presumption that the past will in some sense be like the future, the above analysis conditions on what was historically likely, and the optimization framework suggested in this chapter is mathematically optimal given the estimated risks.

CHAPTER 7 – CONCLUSION

The first two sections of this chapter consider some additional theoretical challenges raised by the structure of the Thomson Model, which the design of the MMS Model proposed in Chapter 5 avoids. In particular, the Thomson Model is not consistent with rational expectations, and the model does not consider the Efficient Market Hypothesis. A tractable framework for the testing and implementation of the Efficient Market Hypothesis in models of dividend yields and dividend growth rates is presented. The final section concludes.

1 RATIONAL EXPECTATIONS

In the Thomson model, the variables $LINT$ and $MINT$ are modelled as the sum of a unit gain function and an error process (see equations F2.4-6 and F2.7-9 in the Foreword to Chapter 2). Thomson (1996: 794) refers to $LINTZ$ as “the real yield on long-term fixed interest securities” which implies that Z_{LINT} must reflect the expected inflation over the future lifetime of the bond. Since it is also possible to calculate expected future inflation directly from the model for $INFL$, Z_{LINT} can be compared with the optimal estimate of Z_{LINT} obtained from this model.

The Rational Expectations Hypothesis (REH) is defined by Begg (1982: 30):

The Rational Expectations Hypothesis states that the unobservable subjective expectations of individuals are exactly the true mathematical conditional expectations implied by the model itself.

If the combined model for $INFL$ and $LINT$ is to be internally consistent with the REH, then the two estimates for Z_{LINT} should be the same since, under the REH, investors do not knowingly make systematic forecasting errors.

Although Z_{LINT} reflects the expected inflation over the lifetime of the bond under Thomson’s interpretation of $LINTZ$, it is not simply the average force of inflation over the

next 20 years. Since $LINT$ is the force of interest on a 20-year par bond, Z_{LINT} should be compared with expected future inflation using the following weighted-average:

$$\sum_{k=1}^{20} \left(\frac{C_{t+k}}{C_t^*} \right) \left(\frac{1}{k} \sum_{j=1}^k E_t(INFL_{t+k}) \right) \quad (\text{Eq 7.1.1})$$

where C_{t+k} is the coupon at time $(t+k)$ for a 20-year par-bond issued at time t , and,

$$C_t^* = \sum_{k=1}^{20} C_{t+k} \cdot \quad (\text{Eq 7.1.2})$$

Figure 7-1 illustrates the comparison between Z_{LINT} from Equation (F2.4) and the optimal estimate of Z_{LINT} , Z_{LINT}^* , obtained from the $INFL$ model through Equation (7.1.1). The Thomson Model contradicts the REH since $Z_{LINT} \neq Z_{LINT}^*$ and implies that long-term bond yields have consistently overestimated inflation since 1976.

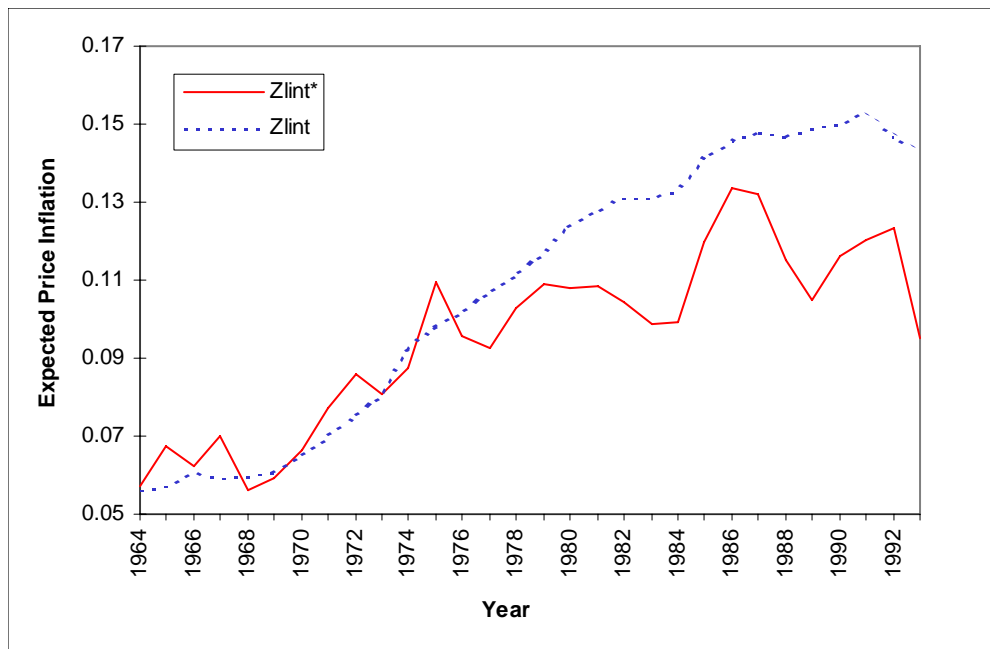


Figure 7.1 Weighted, 20-year expected price inflation for the period 1964-1993, using the estimates Z_{LINT} and Z_{LINT}^*

The comparison illustrated in Figure 7.1 is based on parameter estimates obtained from the full sample (1960-1993) so that the forecasts of future expected inflation are *ex-post* forecasts. Huang & Litzenger (1988) refine the REH to state that investors do not knowingly make systematic *ex-ante* forecasting errors. Under this form of the REH, forecasts of expected future inflation at time t should be based on parameter estimates obtained from data up to time t only. However, Thomson assumes *a priori* that the model parameters and structure are constant over the period of data (1960-1993) on which the model is based. Under this assumption, *ex-post* and *ex-ante* forecasts are expected to be equal and parameter estimates based on the full sample should be preferred.

Thomson (*ibid.*: 794) also refers to $MINTZ$ as the “future real interest rate on money-market instruments” which implies that Z_{MINT} is the money-market’s expectation of inflation for the following year. A comparison with the optimal estimate of inflation for the following year from Thomson’s inflation model is illustrated in Figure 7.2. Clearly, Thomson’s short-term interest rate model also contradicts the REH.

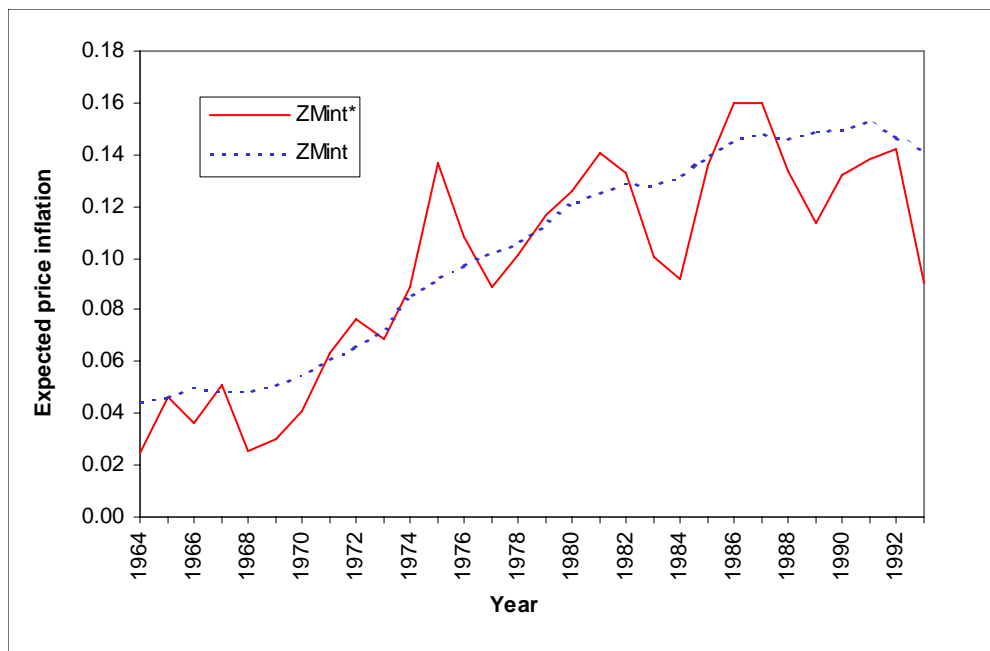


Figure 7.2 One-year expected price inflation for the period 1964-1993, using the estimates Z_{MINT} and Z_{MINT}^*

2 INCORPORATING THE EFFICIENT MARKET HYPOTHESIS

A vast literature exists with claims that the market is not efficient. For a general survey of empirical work on “event studies” that seek to determine if stock prices respond efficiently to information, see Fama (1998). One paper of particular interest to actuaries is the paper by Fama and French (1988), who found that ten-year returns were a function of the level of dividend yields, with relatively low dividend yield being followed by lower returns and visa versa. This feature is captured by both the Thomson and Wilkie models. However, Wilkie (1986b) and Smith (1996) also note that, using expected returns calculated from the Wilkie (1986a) model and investing in the asset class with the highest expected return, it was possible to achieve excess returns of roughly 3% for virtually no extra risk by switching between bonds and equities. Maitland (1996) notes similar features of the Thomson (1996) model as well as the fact that the interpretation placed on the data by the model is that equities are over-valued by 112%.

Despite the historic evidence for certain market inefficiencies, recent studies find evidence that confirms the overpowering logic for the self-destruction of predictable patterns in returns in that previously documented predictability disappeared at a time when a consensus was emerging that predictable patterns were present. For example, Malkiel (2003) notes that the use of dividend yields to predict future returns has been ineffective since the mid-1980's: “Dividend yields have been at the three percent level or below continuously since the mid-1980s, indicating very low forecasted returns. In fact, for all 10 year periods from 1985 through 1992 that ended June 30, 2002, realized annual equity returns from the market index have averaged [a relatively high return of] approximately 15 percent.” Similarly, Huber (1997) notes that if Wilkie's (1986b) dynamic strategy had been followed over the interval 1983-95, then an excess return of 2% less than the return on equities would have been achieved. The same factors that alerted authors to the success of their trading rules may have brought traders to exploit these factors and hence lead to the demise of these prediction rules.

2.1 AN OVERVIEW OF THE EFFICIENT MARKET HYPOTHESIS

The efficient market hypothesis (EMH) is concerned with the behaviour of asset prices over time. When the term ‘efficient market’ was first introduced into the economics literature, it was defined as a market which ‘adjusts rapidly to new information’ (Fama et al 1969). To the extent that investors can respond profitably to new information, prices should adjust rapidly to reflect this new information to the point where asset prices reflect their new perceived value.

In an efficient market, when information arises, the news spreads very quickly and is incorporated into asset prices without delay. Thus, neither technical analysis, which is the study of past asset prices in an attempt to predict future prices, nor even fundamental analysis, which is the analysis of financial information such as dividend yields, company earnings, etc., help investors pick “undervalued” assets to achieve expected returns greater than those that could be obtained by holding a randomly selected portfolio of assets with comparable risk.

If investors as a whole did not respond rapidly to new information about changes in the value of assets, it would be possible for one of these investors to implement profitable trading strategies by using this information before other investors did. However, the possibility of earning expected returns in excess of those required to compensate for the risk of such investments motivates investors to respond quickly to new information. Hence, the profit motive works to increase market efficiency and it should not be possible to earn abnormal profits by predicting future price changes on the basis of past information.

A more comprehensive definition of market efficiency is provided by Malkiel (1992):

A capital market is said to be efficient if it fully and correctly reflects all information in determining security prices. Formally, the market is said to be efficient with respect to some information set, Ω_t , if security prices would be unaffected by revealing that information to all participants. Moreover, efficiency

with respect to an information set, Ω_t , implies that it is impossible to make economic profits by trading on the basis of Ω_t .

Three forms of the EMH are entertained in the literature on the basis of the variables contained in the information set, Ω_t . The “weak” form of the EMH includes only past and current asset prices (and possibly dividends) in Ω_t , the “semi-strong” form includes all publicly available information while the “strong” form includes all public and private information. For the purposes of this section, we will consider a weak form of the EMH with only the $EQDY$ and $EQDG$ (see Chapter 2) included in the information set, Ω_t .

2.2 EMH IN MODELS OF DIVIDEND YIELDS AND DIVIDEND GROWTH

If (in the set of variables being modelled) there is no compelling evidence against the EMH, there should be no reason to assume it does not hold. However, in the models of Wilkie (1986a, 1995) and Thomson (1996) amongst others, the EMH is not tested and market inefficiency is incorporated almost by default. This section explores the reason why the EMH is not fundamental to the methodology proposed by Thomson (1996) and why a large number of actuarial stochastic models based on dividend yields and dividends or dividend growth rates do not incorporate the EMH.

An alternative methodology is presented, which incorporates the EMH into the fundamental structure of these models and allows the EMH to be tested. As a result, the reliability of such models can be tested statistically and appropriate parameters chosen for developing robust benchmarks and risk management strategies for the assets and liabilities of a financial institution.

Suppose R_t is the one-period rate of return defined by

$$R_t = (P_t + D_t - P_{t-1})/P_{t-1}, \quad (7.2.1)$$

where P_t is the price at time t and D_t is the dividend paid out over the period $t-1$ to t . Define DY_t as dividend yield at time t , i.e. $DY_t = D_t/P_t$, and DG_t as the dividend growth factor for the period $t-1$ to t , i.e. $DG_t = D_t/D_{t-1}$. Then, using Equation 7.2.1, the one-period force of return over the period $t-1$ to t is:

$$\begin{aligned}
 F_t &= \ln(1 + R_t) = \ln\left(\frac{P_t + D_t}{P_{t-1}}\right) = \ln\left(\frac{D_t/P_{t-1}}{D_t/P_t} + \frac{D_t}{P_{t-1}}\right) = \ln\left(\frac{DG_t \cdot D_{t-1}/P_{t-1}}{D_t/P_t} + \frac{DG_t \cdot D_{t-1}}{P_{t-1}}\right) \\
 &= \ln\left(\frac{DG_t \cdot DY_{t-1}}{DY_t} + DG_t \cdot DY_{t-1}\right) = \ln(DG_t) + \ln(\cdot DY_{t-1}) + \ln\left(\frac{1}{DY_t} + 1\right) \\
 &= \ln(DG_t) + \ln(\cdot DY_{t-1}) - \ln(DY_t) + \ln(1 + DY_t)
 \end{aligned} \tag{7.2.2}$$

Using Equation 7.2.2 and the equity variables $EQDY_t$ and $EQDG_t$ defined in Chapter 2 and in Thomson (1996), we can write F_t as

$$\begin{aligned}
 F_t &= EQDG_t + EQDY_{t-1} - EQDY_t + \ln(1 + DY_t) \\
 &= EQDG_t + EQDY_{t-1} - EQDY_t + \ln(1 + \exp(EQDY_t)/100)
 \end{aligned} \tag{7.2.3}$$

Equation 7.2.3 is non-linear in the variable $EQDY_t$ and so does not lend itself to linear time series analysis in its current form. However, using the approximation

$$1 + a \cdot e^x \approx e^{\alpha x + \beta} \tag{7.2.4}$$

we can construct the following linear approximation to Equation 7.2.3 for suitable values of α and β in the range of $EQDY_t$:

$$\begin{aligned}
 F_t &= EQDG_t + EQDY_{t-1} - EQDY_t + \alpha \cdot EQDY_t + \beta \\
 &= EQDG_t + EQDY_{t-1} - (1 - \alpha) \cdot EQDY_t + \beta
 \end{aligned} \tag{7.2.5}$$

Values for α and β for the empirical range of $EQDY_t$ are discussed later in this section.

Suppose that r_t is the return for the period $t-1$ to t on a zero-coupon government bond with one period to maturity, where the period to maturity of the bond is exactly the same as the holding period of the equity index. (Note that the risk-free rate, r_t , is known at the start of the period, i.e. at time $t-1$.) Under the EMH, if investors are risk averse, they should price equities so that the expected return on equity index over the following period equals the risk-free rate in addition to a (possibly) time-varying risk premium, λ_t . That is, $E[F_t | \Omega_{t-1}] = r_t + \lambda_t$, where both r_t and λ_t are assumed known at time $t-1$. Taking expectations in Equation 7.2.5 conditional on Ω_{t-1} and setting $h_t = r_t + \lambda_t - \beta$ gives:

$$h_t = E[EQDG_t | \Omega_{t-1}] - (1 - \alpha) \cdot E[EQDY_t | \Omega_{t-1}] + EQDY_{t-1} \quad (7.2.6)$$

Equation 7.2.6 provides a necessary condition for the EMH using the linear approximation to model equity returns as a function of $EQDY$ and $EQDG$.

Clearly, it is possible for h_t to be a function of some or all of the values contained in Ω_{t-1} and it is not true in general that returns are non-predictable under the EMH. Under no arbitrage, if Q_t is the ‘stochastic discount factor’ or ‘pricing kernel’ that accounts for a time-varying risk premium, $E[Q_t R_t | \Omega_{t-1}] = 0$, which implies that

$$E[R_t | \Omega_{t-1}] = \frac{-Cov(Q_t : R_t | \Omega_{t-1})}{E[Q_t | \Omega_{t-1}]} \quad (7.2.7)$$

Hence, as Granger and Timmermann (2002) point out, “predictability of returns ... need not violate the EMH. Forecasting models that work because they predict the conditional covariance of returns with the pricing kernel ... scaled by its conditional mean are not ruled out.” The process generating the risk premium is model dependent and cannot be observed. Hence, tests of the EMH can only be conducted under the auxiliary hypothesis that the model for the risk premium is correct. This is important since empirical tests of market efficiency are necessarily joint tests of market efficiency and a particular asset-price model. When the joint hypothesis is rejected it is logically possible that this is a

consequence of deficiencies in the particular asset-price model rather than in the EMH. This is the ‘bad model’ problem (Fama 1991).

While the risk-free rate may vary from period to period with some degree of predictability, it may be less reasonable to assume that changes in the risk premium can be predicted with any degree of certainty. Hence, unless there is significant evidence for a time-varying risk premium and significant information in Ω_{t-1} to predict it, it is reasonable to start by assuming that $E[\lambda_t | \Omega_{t-1}] = \lambda$.

In the analysis that follows, we assume both a constant risk-free rate and a constant risk premium for simplicity, i.e. $h_t = h \quad \forall t$. Hence, Equation 7.2.6 with $h_t = h$ provides a necessary condition for the EMH under the simplest form of model for equity returns and the risk-free rate. (The unrealistic assumption of a constant risk-free rate is easily relaxed by including the realised values of r_t in equations 7.2.6 and 7.2.8 so that one is effectively working with excess returns above the risk-free rate.) We also assume that the market cannot forecast $EQDY_t$ using past values of variables other than $EQDY$ itself, and that the AR(1) process is a reasonable model for $EQDY$, i.e. $EQDY_t = c + \phi_1 \cdot EQDY_{t-1} + \varepsilon_t$.

Therefore, since $E[EQDY_t | \Omega_{t-1}] = E[c + \phi_1 \cdot EQDY_{t-1} + \varepsilon_t | \Omega_{t-1}] = c + \phi_1 \cdot EQDY_{t-1}$, solving for $E[EQDG_t | \Omega_{t-1}]$ in Equation 7.2.6 gives:

$$E[EQDG_t | \Omega_{t-1}] = \varphi \cdot EQDY_{t-1} + \kappa \quad (7.2.8)$$

where $\kappa = (h + (1 - \alpha)c)$ and $\varphi = -(1 - (1 - \alpha)\phi_1)$. This suggests a natural transfer function for $EQDG_t$ in which the value φ is based on the estimates α and ϕ_1 . Since α and ϕ_1 are known prior to estimating Equation 7.2.8, the value of φ provides a testable null hypothesis for the joint test of market efficiency and this stochastic model of $EQDY$ and $EQDG$. In general φ does not equal zero, so it is important to realize that the standard time series null hypothesis for transfer function models is not appropriate for this transfer function model.

Suitable values of α and β must be chosen so that the linear approximation described in Equation 7.2.4 is reasonable over the observed range of $EQDY$. There are a number of ways to set α and β , for example, minimising the squared or absolute error over the observed range of values or minimising the expected error using the empirical distribution of values. However, the results of the EMH test using Equation 7.2.8 discussed below appear to be relatively robust to the method used to estimate α and β .

The minimum and maximum values of $EQDY_t$ in the period 1960 to 1993 are 0.88 and 2.15 respectively. Setting α and β so that Equation 7.2.4 holds exactly for $EQDY$ equal to 0.88 and 2.15 gives $\alpha=0.045$ and $\beta=-0.016$. This approximation results in negative errors over the entire range of approximation. Setting α and β so that Equation 7.2.4 holds exactly for $EQDY$ equal to the mean of 1.634 and minimises the total squared error results in all errors being positive, and gives $\alpha=0.035$ and $\beta=-0.010$. These two extremes provide a range for α of (0.035, 0.045). From Thomson's AR(1) model for $EQDY$, $\phi_1=0.810$ so φ ranges from -0.226 to -0.218 , depending on the values chosen for α and β in the approximation described by Equation 7.2.4.

It should be noted that the higher the value of ϕ_1 , the closer φ is to zero. Hence, as a result of using the standard (but inappropriate) time series null hypothesis for transfer function models, i.e. $H_0: \varphi=0$, EMH transfer function models will often be rejected, since φ is often not significantly different from zero. This problem is exacerbated when the process for $EQDY$ is close to non-stationary. However, by using the correct null hypothesis, which incorporates the EMH, such errors can be avoided.

Using Thomson's (1996) data from 1960 to 1993, the estimated coefficient for φ in Equation 7.2.8 is $\varphi = -0.051$. When tested against the zero null hypothesis, this is not significantly different from zero, so the standard (but inappropriate) time series null hypothesis cannot be rejected and the $EQDY_{t-1}$ term in the transfer function model for $EQDG_t$ would normally be dropped. However, when tested against the null hypothesis of an efficient market, the EMH cannot be rejected: a Wald test of the more extreme

restriction $H_0: \varphi = -0.226$ gives a chi-squared value of 3.270 with a p-value of 0.071 while the same test with $H_0: \varphi = -0.218$ gives a chi-squared value of 2.961 with a p-value of 0.085 indicating that the EMH cannot be rejected at the 5% level. This suggests that the $EQDY_{t-1}$ term in the equation for $EQDG_t$ should be included in a transfer function model of dividend yields and dividend growth rates, and that with $\alpha=0.035$ and $\phi_1=0.81$, the parameter φ should be set equal to about -0.218. (It should be noted that the Wilkie model does in fact include an equity dividend yield term in the equation for equity dividends; however, the null hypothesis of an efficient market remains to be tested in that model using the approach proposed in this section.)

Although the Wald test of the restriction $H_0: \varphi = -0.210$ cannot be rejected at the 5% level, it can be rejected at the 10% level. As discussed, this does not imply a rejection of the EMH since what we are actually testing is the joint hypothesis that the market is efficient and that the model describing returns is appropriate. Clearly, alternative models for $EQDG$ and $EQDY$ should be investigated. Such alternative models are beyond the scope of this thesis.

It should also be noted that the value of ϕ_1 used in the above test for the EMH is the maximum likelihood estimate based on the sample data. It has a standard deviation of 0.122, suggesting that its true value could be higher or lower than 0.81 (assuming an AR(1) model is correct in the first place). With $\varphi = -0.051$ and a lower estimated value of ϕ_1 in the AR(1) model for $EQDY$, the EMH is more likely to be rejected, although the $EQDY$ data supporting such an estimate for ϕ_1 would also affect the estimate of φ in the transfer function for $EQDG$ with input $EQDY$. The problem here is that the same $EQDY$ data that is used to estimate ϕ_1 (and hence the efficient market null hypothesis) is used to estimate φ in the transfer function for $EQDG$ with input $EQDY$. It is not clear at this stage how to resolve the problems surrounding this data-driven hypothesis test, and this is left for future research.

3 CONCLUSION

The objective of this thesis has been to provide improved models and techniques for actuarial use in the quantification and management of financial risks relating to the assets and liabilities of financial institutions and individuals in South Africa.

The models developed in Chapters 3-6 of this thesis have each aimed at providing a better description of the major economic and financial factors that might impact such assets and liabilities. They are essentially descriptive models, although the models developed in Chapters 4-6 also consider important theoretical features that are thought to be desirable in long-term stochastic models for actuarial use. In particular, they reference the two key financial economic concepts of market efficiency and the principle of no-arbitrage.

In addition, the framework presented in the previous section of this chapter addresses the theoretical concern of market efficiency voiced about existing models of dividend yields and dividend growth rates. It has been shown to be a tractable framework for the testing of the Efficient Market Hypothesis and implementation of market efficiency in such models. It is suggested that this framework be employed in the development of new models containing these variables.

The models of Chapters 4-6 provide a useful set of tools and techniques for modelling and managing major financial risks relating to ALM in South Africa.

In the specific case that certain liabilities take the form of a set of known nominal liability cash flows, the immunization techniques presented in Chapter 6 provide an optimal framework for managing the relevant assets in relation to those liabilities. However, in practice, such cash flows are often based on expected values, and the decision-maker should ensure that the likely distribution of cash flows suggests sufficient certainty to implement an immunization process. It may also be the case in practice that:

- there is a large surplus or deficit in the fund that must be managed;
- these liabilities form a subset of the total liabilities under consideration; or,
- short selling constraints prevent the implementation of a full immunization strategy.

Whatever the case, the decision-maker may wish to consider a partial immunization strategy that immunizes only a subset of the liability cash flows. For the remainder of the liabilities, alternative more risky strategies may be considered. In the absence of sufficient evidence for market inefficiency, these strategies are best analyzed via Monte Carlo simulation using a stochastic model such as the MMS model presented in Chapter 5.

It is not intended that decision-makers using the MMS model for projection purposes rely uncritically on the parameters estimated; indeed they should reconsider certain parameters in the light of market information available at the time. However, market information is limited and does not provide information on certain parameters. For example, the autoregressive parameters, which are critical to the dynamics of certain variables, are not available from market information. The MMS model removes much of the bias found in corresponding parameter estimates from linear time series models. This, together with the fact that these parameters appear to be stable across regimes suggests that they are more likely to provide stable estimates for forecasting purposes.

Factors that should be considered and that are not available in current market information include:

- the autoregressive parameters for long- and short-term interest rates, and inflation;
- the extrapolated forward rate curve and the implied level of future long-term rates;
- contemporaneous correlations; and,
- the size of the equity risk premium

While such information is not available from current market information, it is not suggested that all corresponding empirical estimates are appropriate for projection purposes. In particular, users should consider carefully the appropriate size of equity risk premium to use for projection purposes.

Users should also consider exogenous factors that may impact the MMS transition probabilities such as:

- the current framework of inflation targeting at the time,
- the probability of future inflation shocks, and
- the various monetary policy responses to such shocks.

Users may also wish to consider a sensitivity analysis to the possibility of alternative monetary policy regimes and risk premiums.

In setting various parameters for projection purposes, users may also wish to consider the anticipated market prices of risk for different asset categories. The MMS model is not an equilibrium model, although for projection purposes it is possible to set its parameters to incorporate market equilibrium using portfolio theory.

Portfolio theory provides a framework for selecting investments so as to maximize an investor's expected utility (see Elton & Gruber, 1991). It is commonly assumed that the expected utility can be measured in terms of the mean and standard deviation of returns. In this case, investors will only include a security in their portfolio if its inclusion either increases the expected return or decreases the standard deviation of the portfolio returns. Hence, for inclusion, a security with a lower expected return than the portfolio should either have a sufficiently lower expected standard deviation or a correlation with the portfolio that is sufficiently less than one to adequately reduce the standard deviation of the portfolio. The aggregation of each investor's optimum portfolio must result in the market portfolio. Assuming investors have homogeneous expectations, it is then possible to set equilibrium parameters such that this constraint is satisfied.

In reality, different measures of risk are used by different investors, and investors have different objectives and liabilities. These factors complicate the equilibrium market prices of risk for different asset categories. Nonetheless, they are an important consideration in setting parameters for projection purposes and decision making. A full discussion on the parameterization of models to currently available market information and the identification of equilibrium market prices of risk for different asset categories is beyond the scope of this thesis.

References

- Agiakloglou, C. (1996), 'A note on the unit root test based on the sample autocorrelations,' *The University of Piraeus Journal of Economics, Business, Statistics and Operations Research*, **46**, 97-106.
- Anderson, N., Breedon, F., Deacon, M., Derry, A. & Murphy, G. (1996) *Estimating and interpreting the yield curve*, John Wiley & Sons.
- Ang, A. & Moore, D. (1994) "The Australian yield curve," *unpublished working paper*, Macquarie University, Sydney, Australia.
- Aron, J. & Muellbauer, J. (2000), 'Inflation and Output Forecasting for South Africa: Monetary Transmission Implications,' *Working paper Series*, 2000-23, Nuffield College, Oxford University.
- Asher, A. (2007) *The design of retirement schemes: possibilities and imperatives*, Unpublished PhD thesis. Department of Statistics & Actuarial Science, University of the Witwatersrand, Johannesburg, South Africa.
- Barber, J.R. & Copper, M.L. (1996), 'Immunization using principal components analysis,' *The Journal of Portfolio Management*, **22** (3), 99–105
- Begg, D.K.H. (1982), *The Rational Expectations Revolution in Macro-Economics*, Johns Hopkins University, Princeton.
- Bierens, J.H. (1993), 'Higher-order sample autocorrelations and the unit root test,' *Journal of Econometrics*, **57**, 137-160.
- Bollerslev, T. (1986), 'Generalized Autoregressive Conditional Heteroscedasticity,' *Journal of Econometrics*, **31**, 307-327.
- Bollen, N.P.B. (1998). 'Valuing options in regime switching models,' *Journal of Derivatives*, **6**, 38-49.
- Box, G.E.P. & Jenkins, G.M. (1970), *Time series analysis, forecasting and control*, Holden Day, San Francisco.
- Box, G.E.P. & Tiao, G.C. (1975), 'Intervention analysis with applications to economic and environmental problems,' *Journal of the American Statistical Association*, **70**, 70-79.
- Bollerslev, T. (1986), "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, **31**, 307-327.

- Breusch, T.S. & Pagan, A.R. (1978), 'A simple test for heteroscedasticity and random coefficient variation,' *Econometrica*, **46**, 1287-1294.
- Boyle, P.P. (1978), 'Immunization under stochastic models of the term structure,' *J.I.A.*, **105**, 177-187.
- Brennan, M.J. & Schwartz, E.S. (1982), 'An equilibrium model of bond pricing and a test of market efficiency,' *Journal of Financial and Quantitative Analysis*, **17**, 75-100.
- Bühler, A. & Zimmermann, H. (1996), 'A statistical analysis of the term structure of interest rates in Switzerland and Germany' *The Journal of Fixed Income*, **6**, 55-67
- Carter, J. (1991), 'The derivation and application of an Australian stochastic investment model,' *Transactions of the Institute of Actuaries of Australia*, **1**, 315-428.
- Cattell, R. B. (1965), 'Factor analysis: an introduction to fundamentals,' *Biometrics*, **21**, 190-215.
- Chaplin, G. B. (1998), 'A review of term structure models and their applications,' *B.A.J.*, **4**, 323-349.
- Chow, G. (1960), 'Tests for equality between sets of coefficients in two linear regressions,' *Econometrica*, **28**, 591-605.
- Christiano, L.J. (1992), 'Searching for break points in GNP,' *Journal of Business and Economic Statistics*, **10**, 237-250.
- Claasen, M.S. (1993), 'A systematic approach to the financial management and risk control of employee benefit funds,' *3rd AFIR International Colloquium*.
- Claasen, M.S. & Huber, P.P. (1992), 'Integrating pension fund assets and liabilities: techniques for the 1990's,' *Transactions of the Actuarial Society of South Africa*, **9**, 209-224.
- Cosslett, Stephen R., & Lung-Fei Lee (1985), 'Serial Correlation in Discrete Variable Models,' *Journal of Econometrics*, **27**, 79-97.
- Cox, J.C., Ingersoll, J.E. & Ross, S.A. (1985), 'A theory of the term structure of interest rates,' *Econometrica*, **53**, 385-407.
- D'Ecclesia, R.L. & Zenios, S.A. (1994), 'Risk factor analysis and portfolio immunization in the Italian bond market,' *The Journal of Fixed Income*, **4**, 51-9
- Darroch, J. N. (1965), 'An optimal property of principal components,' *Annals of Mathematical Statistics*, **36**, 1579.

- Davies, R. (1977) 'Hypothesis testing when a nuisance parameter is present only under the alternative' *Biometrika*, **64**, 247-254.
- 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.
- Dickey, D.A. & Fuller, W.A. (1981), 'Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root,' *Econometrica*, **49**, 1057-1072.
- Diebold, F.X. & Nerlove, M. (1990), 'Unit roots in economic time series: a selective survey,' in Rhodes, G.F. and Fomby, T.B. (eds), *Advances in Econometrics*, Vol. 8. Greenwich, CT: JAI Press, 3-69.
- Diebold, F.X., Lee, J-H. & Weinbach, G.C. (1994), 'Regime switching with time-varying transition probabilities,' in C.Hargreaves (ed.), *Nonstationary Time Series Analysis and Cointegration*. Oxford University Press.
- Dolado, J.J., Jenkinson, T. & Sosvilla-Rivero, S. (1990), 'Cointegration and unit roots,' *Journal of Economic Surveys*, **4**, 249-273.
- Draper, N.R. & Smith, H. (1981) *Applied Regression Analysis*. Second edition, Wiley.
- Durland, J.M. & McCurdy, T.H. (1994), "Duration Dependent Transitions in a Markov Model of U.S. GNP Growth," *Journal of Business and Economic Statistics*, **12**, 279-88.
- Elton, E.J. & Gruber, M.J. (1991), *Modern Portfolio Theory and Investment Analysis*. John Wiley & Sons Inc.
- Engle, R.F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation,' *Econometrica*, **50**, 987-1007.
- Engle, R.F. (1995), *ARCH Selected Readings*, Oxford University Press.
- EViews (1994), *User's Guide*, Quantitative Micro Software, Irvine, California, USA.
- Falkena, H.B., Fourie, L.J. & Kok, W.J. (1984) *The Mechanics of the South African Financial System*, MacMillan South Africa, 2nd edition.
- Fama, E.F., Fisher, L., Jensen, M.C. & Roll, R. (1969), 'The Adjustment of Stock Prices to New Information,' *International Economic Review*, **10**, 1, 1-21.
- Fama, E.F. & French, K.R. (1988), 'Dividend Yields and Expected Stock Returns,' *Journal of Financial Economics*, **22**, 3-25.

- Fama (1991), 'Efficient Markets: II,' *Journal of Finance*, **46**, 1575-1617.
- Fama, E.F. (1998), 'Market efficiency, long-term returns, and behavioral finance,' *Journal of Financial Economics*, **49**, 283-306.
- Feldman, K.S., Bergman, B., Cairns, A.J.G., Chaplin, G.B., Gwilt, G.D., Lockyer, P.R. & Turley, F.B. (1998), 'Report of the fixed interest working group,' *BAJ*, 4, 213–383
- Fisher, L. & Weil, R.L. (1971), 'Coping with the risk of interest rate fluctuations: returns to bond holders from a naïve and optimal strategy,' *Journal of Business*, 44 (3), 408–31
- Garcia, R. & Perron, P. (1996), 'An analysis of the real interest rate under regime shifts,' *The Review of Economics and Statistics*, **X**, 111-125.
- Garcia, Rene (1998), 'Asymptotic Null Distribution of the Likelihood Ratio Test in Markov Switching Models,' *International Economic Review*, **39**, 763-788.
- Godfrey, L.G. (1978), 'Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables,' *Econometrica*, **46**, 1293-1302.
- Gott, D.V. (2009) *Stochastic modelling of asset returns with applications in pricing life-insurance embedded derivatives*. Unpublished M.Sc. thesis. University of the Witwatersrand, Johannesburg, South Africa.
- Granger, C.W.J. & Newbold, P. (1974), 'Spurious regressions in Economics' *Journal of Econometrics*, **2**, 111-120.
- Granger, C.W.J. & Timmermann, A.G. (2002), 'Efficient market hypothesis and forecasting,' *CEPR Discussion Paper*, No. 3593.
- Goldfeld, Stephen M., & Richard E. Quandt (1973), "A Markov Model for Switching Regressions," *Journal of Econometrics*, **1**, 3-16.
- Hamilton, J.D. (1989), 'A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle,' *Econometrica*, **57**, 357-384.
- Hamilton, J.D. (1990), 'Analysis of time series subject to changes in regime,' *Journal of Econometrics*, **45**, 39-70.
- Hamilton, J.D. (1993), 'Estimation, inference and forecasting of time series subject to changes in regime,' in G.S.Maddala, C.R.Rao & H.D.Vinod (eds), *Handbook of Statistics*, Vol. 11. New York: North-Holland.

- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Hardy, M.R. (2001). 'A regime-switching model of long-term stock returns,' *North American Actuarial Journal*, **52**, 41-53
- Harris, G. (1994) 'On Australian stochastic share return models for actuarial use,' *The Institute of Actuaries of Australia Quarterly Journal*, September, 34-54.
- Harris, G. (1996) "Market phases and cycles? A regime switching approach." *The Institute of Actuaries of Australia Quarterly Journal*, December, 28-44.
- Harris, G. (1999), 'Regime switching vector autoregressions: a Bayesian Markov Chain Monte Carlo approach,' *ASTIN*, **29**, 1, May 1999.
- Head, S.J., Adkins, D.R., Cairns, A.J.G., Corvesor, D.O., Cule, D.O., Exeley, C.J., Johnson, I.S., Spain, J.G. & Wise, A.J. (2000), 'Pension fund valuations and market values,' *B.A.J.*, **6**, 55–143.
- Heath, D., Jarrow, R. & Morton, A. (1992), 'Bond pricing and the term structure of interest rates; a new methodology,' *Econometrica*, **60**, 77-105.
- Heston S.L. (1993) A closed-form solution for options with stochastic volatility with applications to bond & currency options. *Review of Financial Studies* 6 (2), 327-343.
- Ho, T.S.Y. (1992), 'Key rate durations: measures of interest rate risk,' *The Journal of Fixed Income*, 2, 29–44
- Holden, D. & Perman, R. (1994), 'Unit roots and cointegration for the economist,' in Rao, B.B. (ed.), *Cointegration for the Applied Economist*, MacMillan, London.
- Homer, S. & Sylla, R. (2005), *A History of Interest Rates*. Fourth Edition, Wiley.
- Howie, R.J. (2007) *A theory based stochastic investment model for actuarial use*. Unpublished M.Bus.Sc. thesis. University of Cape Town, South Africa.
- Howson, C. (2000) *Hume's Problem : Induction and the justification of Belief*. Clarendon Press, Oxford.
- Hua, G. (1994), 'Modelling equity returns using a simple ARCH model,' *5th AFIR International Colloquium*, **2**, 537-565.
- Huang, C.-F. & Litzenberger, R.H. (1988), *Foundations for financial economics*, Prentice-Hall, Englewood Cliffs, NJ
- Huber, P.P. (1997), 'A review of Wilkie's stochastic asset model,' *B.A.J.*, **3**, 181-210.

- Huber, P.P. & Verrall, R.J. (1999), 'The need for theory in actuarial economic models,' *B.A.J.*, **5**, II, 377-395.
- Hume, D. (1910) [1748]. *An Enquiry concerning Human Understanding*. P.F. Collier & Son, (<http://18th.eserver.org/hume-enquiry.html>).
- INET Graphics and Database Service (1998). Intelligent Network (Pty) Ltd.
- INET Graphics and Database Service (2006). Intelligent Network (Pty) Ltd.
- Jarque, C.M. & Bera, A.K. (1981), 'An efficient large sample test for normality of observations and regression residuals,' *Working papers in Economics and Econometrics*, No. 47, Australian National University.
- Johansen, S. & Juselius, K. (1992), 'Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK,' *Journal of Econometrics*, **53**, 211-244.
- Johnson, N.L. & Kotz, S. (1972). *Distributions in statistics : continuous multivariate distributions*. John Wiley & Sons, Inc.
- JSE Information Products Sales Division (30 October 2005) "FTSE / JSE End of Day Indices Dissemination User Manual."
- Juselius, K. (1995), 'Do PPP and UIP hold in the long-run,' *Journal of Econometrics*, **69**, 11-240.
- Kendall, M., Stewart, A. & Ord, J.K. (1983), *The Advanced Theory of Statistics*, Vol. 3, Charles Griffen
- Kendall, M. & Ord, J.K. (1990). *Time Series. (Third edition.)* A Charles Griffin Title, Edward Arnold.
- Krolzig, Hans-Martin (1997), *Markov-Switching Vector Autoregressions : Modelling, Statistical Inference, and Application to Business Cycle Analysis*, Berlin: Springer.
- Kwiatkowski, Denis, Peter C. B. Phillips, Peter Schmidt & Yongcheol Shin (1992). 'Testing the Null Hypothesis of Stationary against the Alternative of a Unit Root,' *Journal of Econometrics*, **54**, 159-178.
- Lane, G.B. (1995), 'A practical guide to stochastic asset modelling,' *Transactions of the Actuarial Society of South Africa*, **14**, 209-224.
- Lee, E.M. (1986), *An introduction to pension schemes*, Institute of Actuaries, London

- Litterman, R. & Scheinkman, J. (1991), 'Common Factors Affecting Bond Returns,' *Journal of Fixed Income*, 1 (1), 54–61
- McLeod, H. D. (1990), 'The development of a market yield curve: the South African solution,' *1st AFIR International Colloquium*, 2, 197-212.
- Macaulay, F.R. (1938), 'Some theoretical problems suggested by the movement of interest rates, bond yields, and stock prices in U.S. since 1856,' National Bureau of Economic Research, New York
- MacKinnon, J.G. (1991), 'Critical values for cointegration tests,' in Engle, R.F. & Granger, C.W.J. (eds), *Long-Run Economic Relationships: Readings in Cointegration*, Oxford University Press.
- Maitland, A.J. (1996). 'A Review of Thomson's Stochastic Investment Model.' *Transactions of the Actuarial Society of South Africa*, 15, published conference proceedings.
- Maitland, A.J. (1997). 'Non-Stationarity in some South African Financial and Economic Series,' *Proceedings of the 7th International AFIR Colloquium*, 7, 631-46, published conference proceedings.
- Maitland, A.J. (2000), 'Interpolating the South African Yield Curve,' *Proceedings of the 10th AFIR Colloquium*, 469–82, published conference proceedings.
- Maitland, A.J. (2001). An Empirical Approach to Immunization in South Africa. *South African Actuarial Journal*, 1, 114-33.
- Maitland, A.J. (2002). Interpolating the South African Yield Curve using Principal Components Analysis: A Descriptive Approach. *South African Actuarial Journal*, 2, 129-45.
- Maitland, A.J. (*forthcoming*), 'A Multiple Markov Switching Model for Actuarial Use in South Africa,' *South African Actuarial Journal*.
- Malkiel, B.G. (1992). Efficient Market Hypothesis. In: Newman et al. (Eds.), *New Palgrave Dictionary of Money and Finance*, London: Macmillan.
- Malkiel, B.G. (2003), 'The Efficient Market Hypothesis and Its Critics,' *CEPS Working Paper*, 91.
- Maturity Guarantees Working Party of the Institute of Actuaries (MGWP) (1980). *Report of the MGWP*. J.I.A., 107, 2, no. 435.

- Mboweni, T.T. (2000), "A new monetary policy framework," Statement of the Monetary Policy Committee of South Africa, 6 April 2000.
- Mills, T.C. (1994), *The econometric modelling of financial time series*. Cambridge University Press.
- Pemberton, J.M. (1998), 'The methodology of Actuarial Science,' *B.A.J.*, **5**, 115-195.
- Perron, P. (1988), 'Trends and random walks in macroeconomic time series: further evidence from a new approach,' *Journal of Economic Dynamics and Control*, **12**, 297-332.
- Perron, P. (1989), 'The Great Crash, the Oil Price Shock and the Unit Root Hypothesis,' *Econometrica*, **57**, 1361-1401.
- Perron, P. (1991), 'Test consistency with varying sampling frequency,' *Econometric Theory*, **7**, 341-368.
- Perron, P. (1993), 'The Great Crash, the Oil Price Shock and the Unit Root Hypothesis: erratum,' *Econometrica*, **61**, 248-249.
- Perron, P. (1994), 'Trend, unit root and structural change in macroeconomic time series,' in Rao, B.B. (ed.), *Cointegration for the Applied Economist*, MacMillan, London.
- Pesaran, M.H. & Shin, Y. (1996), 'Cointegration and speed of convergence to equilibrium,' *Journal of Econometrics*, **71**, 117-143.
- PGN 201 (2003) "RETIREMENT FUNDS - ACTUARIAL VALUATION REPORTS AND RELATED TOPICS," Actuarial Society of South Africa.
- Phillips, P.C.B. (1987), 'Time series regression with a unit root,' *Econometrica*, **55**, 227-301.
- Phillips, P.C.B. & Perron, P. (1988), 'Testing for a unit root in time series regression,' *Biometrika*, **75**, 335-346.
- Ramsey, F.P. (1930), 'On a problem in formal logic,' *Proc. London Math. Soc.*, Vol. 30, pp. 264-286
- Ranne, A. (1998) "The Finish Stochastic investment model," Transactions of the 26th International Congress of Actuaries, 7, 213-238.
- Redington, F.M. (1952), 'Review of the principles of life office valuations,' *JIA*, **78**, 286-315

- Redington, F.M. (1986), *A ramble through the actuarial countryside. Collected papers and speeches*. Ed G. Chamberlain. Institute of Actuaries Student Society.
- Reitano, R.R. (1991), 'Multivariate immunization theory,' *TSA*, **XLIII**, 392–438
- Reitano, R.R. (1992), 'Non-parallel yield-curve shifts and immunization,' *Journal of Portfolio Management*, **18**, 36–43
- Said, S.E. & Dickey, D.A. (1984), 'Testing for unit roots in autoregressive-moving average models of unknown order,' *Biometrika*, **71**, 599-607.
- SAS/ETS (1993), *User's Guide. Version , second edition*. Copyright 1993 by SAS Institute Inc., Cary, NC, USA.
- Schwert, G.W. (1987), 'Effects of model specification on tests for unit roots in macroeconomic data,' *Journal of Monetary Economics*, **20**, 73-105.
- Seber, G.A.F. (1984). *Multivariate Observations*, John Wiley & Sons.
- Sherris, M. (1995), 'Interest rate risk factors in the Australian bond market,' *5th AFIR Colloquium*, 859–70
- Sherris, M., Tedesco, L. & Zehnwirth, B. (1996), 'Stochastic investment models: unit roots, cointegration, state space and GARCH models for Australian data,' *Research Paper Series, Centre for Actuarial Studies, University of Melbourne*, **33**.
- Sherris, M., Tedesco, L. & Zehnwirth, B. (1999) "Investment returns and inflation models: some Australian evidence," *British Actuarial Journal*, **5**, 1, 237-267.
- Smith, A.D. (1996) , 'How actuaries can use financial economics,' *B.A.J.*, **2**, 1057-1193.
- Statistical Research Division (2000). *X-12-ARIMA Reference Manual*. U.S Census Bureau, Washington DC, 20233-9100.
- Stock, J.H. & Watson, M. (1988), 'Testing for common trends,' *Journal of the American Statistical Association*, **83**, 1097-1107.
- Stock, J.H. (1993), 'Unit roots and trend breaks,' in Engle, R.F. & McFadden, D. (eds) *Handbook of Econometrics*, Amsterdam: North-Holland, Vol. 4.
- Thomson, R.J. (1994), 'A stochastic investment model for actuarial use in South Africa,' *Transactions of the Actuarial Society of South Africa*, **13**
- Thomson, R.J. (1995), 'A methodology for the modelling of interest rates and other economic variables with reference to the money and capital markets of South Africa,' *Proceedings of the 5th AFIR Colloquium*

- Thomson, R.J. (1996) "Stochastic investment modelling - the case of South Africa," *British Actuarial Journal*, **2**, 765-802.
- Thomson, R.J. (2004) *The development and application of models in actuarial science*, Unpublished PhD thesis. Department of Statistics & Actuarial Science, University of the Witwatersrand, Johannesburg, South Africa.
- Thomson, R.J. (2006) "A typology of models used in Actuarial Science," *South African Actuarial Journal*, **6**, 19-36.
- Thomson, R.J. (2008). "Modelling the market in a risk-averse world: The case of South Africa," Convention of the Actuarial Society of South Africa, unpublished.
- Thomson, R.J. & Gott, D.V. (2006) "Stochastic models for actuarial use: the equilibrium modelling of local markets," Convention of the Actuarial Society of South Africa.
- Thomson, R.J. & Gott, D.V. (2009) "Stochastic models for actuarial use: the equilibrium modelling of local markets," *ASTIN*, **39**, 1, 339-370.
- Thorlacius, A.E. (2000), 'Arbitrage in asset modelling for integrated risk management,' *Proceedings of the 10th AFIR Colloquium*.
- Tilley, J.A. (1992), 'An actuarial layman's guide to building stochastic interest rate generators,' *Transactions of the Society of Actuaries*, **44**, 509-564.
- Tucker, L. R., Koopman, R. F. & Linn, R. L. (1969), 'Evaluation of factor analytic research procedures by means of simulated correlation matrices,' *Psychometrica*, **34**, 421-459.
- Vasicek, O. (1977), 'An equilibrium characterisation of the term structure,' *Journal of Financial Economics*, **5**, 177-188.
- Wilkie, A.D. (1984) 'Steps towards a comprehensive stochastic investment model,' *Occasional Actuarial Research Discussion Paper*, 36, Institute of Actuaries, London, U.K.
- Wilkie, A.D. (1986a) 'A stochastic investment model for actuarial use,' *Transactions of the Faculty of Actuaries*, **39**, 341-403.
- Wilkie, A.D. (1986b), 'Some applications of stochastic investment models,' *J.S.S.*, **29**, 25-51
- Wilkie, A.D. (1995). 'More on a stochastic asset model for actuarial use,' *British Actuarial Journal*, **1**, 777-964.