MODELLING STOCK RETURN VOLATILITY DYNAMICS IN SELECTED AFRICAN MARKETS

A half-thesis submitted for the partial
fulfilment of the requirements for the degree
of

MASTER OF COMMERCE (FINANCIAL MARKETS)

of

RHODES UNIVERSITY

by

DANIEL JONATHAN KING

December 2012

Supervisor: Mr Ferdi Botha

ABSTRACT

Stock return volatility has been shown to occasionally exhibit discrete structural shifts. These shifts are particularly evident in the transition from ‘normal’ to crisis periods, and tend to be more pronounced in developing markets. This study aims to establish whether accounting for structural changes in the conditional variance process, through the use of Markov-switching models, improves estimates and forecasts of stock return volatility over those of the more conventional single-state (G)ARCH models, within and across selected African markets for the period 2002-2012. In the univariate portion of the study, the performances of various Markov-switching models are tested against a single-state benchmark model through the use of in-sample goodness-of-fit and predictive ability measures. In the multivariate context, the single-state and Markov-switching models are comparatively assessed according to their usefulness in constructing optimal stock portfolios. It is found that, even after accounting for structural breaks in the conditional variance process, conventional GARCH effects remain important to capturing the heteroscedasticity evident in the data. However, those univariate models which include a GARCH term are shown to perform comparatively poorly when used for forecasting purposes. Additionally, in the multivariate study, the use of Markov-switching variance-covariance estimates improves risk-adjusted portfolio returns when compared to portfolios that are constructed using the more conventional single-state models. While there is evidence that the use of some Markov-switching models can result in better forecasts and higher risk-adjusted returns than those models which include GARCH effects, the inability of the simpler Markov-switching models to fully capture the heteroscedasticity in the data remains problematic.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1. Context of the research

Stock market volatility may have several causes. General equilibrium models point to changes in fundamental variables such as consumption, production and preferences (Abel, 1988). Many empirical studies in the efficient markets tradition cite volatility in macroeconomic variables such as inflation, interest rates, the exchange rate, and manufacturing output as influencing stock return variability (Schwert, 1989; Diebold and Yilmaz, 2007; Chinzara, 2011). Aggarwal *et al.* (1999) show that structural breaks in emerging market stock return volatility are related to significant social and political events in addition to economic variables. Alternatively, stock returns are often shown to be more volatile than warranted by a purely 'fundamentals-based' model (Shiller, 1987; Olsen, 1998; Adam *et al.*, 2008).

Whatever the economic cause, a reliable statistical model of stock return volatility is important for the pricing of equity derivative securities and effective hedging of stock market risk (Hamilton and Susmel, 1994; Wang and Theobald, 2008). In addition, changes in the co-movement of stock returns across international markets during high- and low-volatility periods have major implications for diversification strategies (Ramchand and Susmel, 1998; Li, 2009).

Since the seminal studies of Engle (1982) and Bollerslev (1986), the study of financial market (and other) volatility using Autoregressive Conditional Heteroscedasticity (ARCH) and Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models has gained significant popularity. Within ten years of inception, a plethora of studies using some variants of ARCH or GARCH models had been conducted, as evidenced by the survey of Bollerslev *et al.* (1992). Volatility estimates derived from these models are generally found to be more robust and accurate than moving-average or constant volatility estimates. The ease of computation and ability of GARCH models to capture many of the stylised facts associated with financial time series, such as fat tailed distributions and volatility clustering, has ensured their continued popularity throughout the 1990s and the new millennium (Engle, 2004).

A common feature of GARCH-type models using daily financial data is the high level of persistence attributed to shocks, so that the effect of a once off shock to volatility persists for many periods into the future. Many GARCH studies involving financial series have found that the estimated variance is generated by an approximate unit root process (Engle and Bollerslev, 1986; Susmel, 1999). This has led to the development of Engle and Bollerslev's (1986) integrated GARCH (I-GARCH) model.

Despite the robustness of the I-GARCH model across many samples, however, the presence of a unit root in the variance is difficult to justify theoretically (Lamoureux and Lastrapes, 1990; Cai, 1994). In fact, it has been shown that in the presence of structural breaks, GARCH-type models can impose a spuriously high level of persistence of shocks on volatility forecasts (Diebold, 1986; Lamoureux and Lastrapes, 1990; Timmerman, 2000). This finding has led to the parallel development of the Markov-switching ARCH (SWARCH) model, which allows for endogenously identified structural shifts in the volatility generating process (Hamilton and Susmel, 1994; Cai, 1994). The SWARCH specification allows parameters to change according to the particular regime in which the series finds itself, with regime shifts being governed by a hidden Markov process (Hamilton, 1989).

It is generally found that, once having controlled for regime shifts, the persistence of shocks on volatility forecasts is reduced in a statistically significant way (Hamilton and Susmel, 1994; Cai, 1994; Ramchand and Susmel, 1998; Edwards and Susmel, 2003; Marcucci, 2005). Thus, while volatility clustering is still captured, the regime-switching models allow this clustering to be generated by regime changes in addition to within-regime persistence of shocks. That is, where single-regime GARCH models imply a purely time-varying variance, regime-switching models allow for volatility that is both time-varying and state-varying (Ramchund and Susmel, 1998). This specification can thus offer a more intuitively appealing interpretation of the volatility clustering phenomenon than the single-regime GARCH models, as well as improve forecasts due to a higher likelihood of stationarity.

It was initially believed that the regime-switching models would in practice have to be restricted to low order SWARCH due to the recursive nature of GARCH models and the resulting intractability of maximum likelihood estimation for studies with large samples (Hamilton and Susmel, 1994; Cai, 1994). However, this estimation problem has been largely addressed by Gray (1996). The use of Gray's (1996) Markov-switching GARCH (MS-GARCH) procedure or it's extensions (Dueker, 1997; Klaassen, 2002; Haas *et al*., 2004) has allowed richer comparison of parameter estimates across models, as it nests the popular GARCH(1,1) as a special case (Gray, 1996).

In contrast to the growing body of international literature, few studies of African financial market volatility incorporate regime-switching effects. Knedlik and Scheufele (2008) and Duncan and Liu (2009) both address the issue of forecasting South African currency crises within a SWARCH framework. Babikir *et al.* (2010) test the forecasting performance of an MS-GARCH model against a GJR-GARCH, using South African stock market data. As far as could be determined, however, stock market volatility dynamics within and across multiple African markets have not been considered under a regime-switching framework in the literature. Given the importance of accounting for structural shifts in any time series analysis, it is important to fill this gap in the current body of knowledge. The current and future development of derivatives markets in the more financially sophisticated African markets (African Development Bank, 2010) will further increase the potential impact of such a study, as accurate stock return volatility estimates are required for the effective implementation of risk management strategies. The estimated volatility dynamics within and across African markets will also provide an indication of the nature and extent of contagion effects on the continent, with important implications for portfolio managers and policy makers alike.

1.2. Goals of the research

The principal aim of this research is to establish whether accounting for structural changes through the use of Markov-switching models improves estimates and forecasts of stock return volatility within and across selected African countries. As such, the research will aim to address the following questions:

- Do the univariate Markov-switching models of conditional variance provide a superior in-sample fit to the conventional single-state models?
- Do Markov-switching models produce superior forecasts of the conditional variance to the conventional single-state models?
- Do multivariate Markov-switching models of conditional covariance provide a more appropriate and accurate characterisation of stock return volatility and interdependence than the conventional single-state models?

In addition to the above-mentioned primary goals, and as a by-product of the Markov-switching multivariate estimation, the study will aim to establish whether the correlation of returns between African markets changes across volatility states, and the extent to which country-specific volatility states are dependent on one another.

1.3. Methods, procedures, and techniques used

The study will draw significantly on the work of Hamilton and Susmel (1994), Gray (1996), Dueker (1997), Ramchand and Susmel (1998), Marcucci (2005), and Li (2009). Single-state models of the conditional variance and covariance of stock returns will be compared to and tested against various Markov-switching specifications.

The data used consist of daily returns (in U.S. dollars) on the Morgan Stanley Capital International (MSCI) stock market indices for South Africa, Kenya, Mauritius, Morocco and Nigeria from 31 May 2002 until 1 June 2012. The series were chosen based on their availability on *Thomson Reuters Datastream*. Excluded African markets (specifically, Tanzania and Egypt) did not have a large enough sample available, or were closed during much of the Arab Spring. The choice of daily returns is due to the finding that important information regarding volatility is lost at lower frequencies, especially during crisis periods (Edwards, 1998). The period under study will allow the switching models to capture any structural changes in stock return volatility associated with, for example, the high return/low volatility period of the mid-2000s, the financial crisis of 2007, the Arab Spring, and the on-going Eurozone crisis.

In the univariate analysis, the GARCH, simple MS, SWARCH, and MS-GARCH models are compared in order to assess the impact of allowing for changes in regime and differing types of volatility persistence. The in-sample performance of the models will be evaluated based on conventional likelihood ratio tests (LRTs), Davies' (1987) upper-bound test and Ljung-Box Q-tests for autocorrelation in the squared standardised residuals. Out-of-sample forecast performance of the models will be measured through the use of mean squared forecast errors (MSE) and mean absolute forecast errors (MAE). Formal tests of equal predictive ability such as that of Diebold and Mariano (1995) and Clark and West (2007) will be used in addition to the MAE and MSE comparison.

For the multivariate analysis, Bollerslev's (1990) CCC-GARCH, Engle and Kroner's (1995) BEK-GARCH, and the MV-SWARCH model of Hamilton and Lin (1996) and Ramchand and Susmel (1998) will be used in a comparative study of model parameters and performance. Following the rationale of Edwards (1998), Ramchand and Susmel (1998), and Li (2009), South Africa will be used as the originator country in the bivariate analysis due to its prominence as Africa's major economy. Thus, each model will be used to estimate four bivariate variance-covariance matrices (for each of the SA-Other country pairs). In a test of model performance, each model will be used to construct optimal minimum-variance portfolios for the 'typical' risk-averse investor. The portfolio returns, variances, and Sharpe ratios will then be used to assess the associated model's ability to capture accurately the important features of the data.

The results of the MV-SWARCH analysis will also be used to establish whether volatility states are dependent between markets, through the use of LRTs. Consistent with Ramchand and Susmel (1998) and Forbes and Rigobon (2002), inter-country correlation coefficients will then be tested for differences between regimes in order to assess the degree to which contagion is present amongst the African markets studied.

1.4. Structure

The study is structured as follows: Chapter 2 presents the major literature concerning the economic and statistical characterisation of financial volatility and contagion. In addition to a brief overview of the nature of stock return variability, a number of important studies of Markov-switching volatility and cross-correlation of financial series are reviewed. Chapter 3 provides an overview of the data and methods employed in this study, including the relevant models and performance tests used. Chapter 4 presents the empirical results, while Chapter 5 summarises and concludes.

CHAPTER 2: LITERATURE REVIEW

2.1. Introduction

The effective modelling of stock market volatility is important for a number of purposes. Portfolio managers, options traders, and financial policy makers often require an adequate statistical characterisation of stock market volatility in order to perform their duties. Thus, in addition to understanding the underlying causes of stock market volatility, an appropriate reduced-form statistical characterisation of the conditional variance of returns is necessary for accurate volatility estimates and forecasts. Much of the literature pertaining to modelling financial volatility makes use of Bollerslev’s (1986) GARCH model. However, a related strand of literature incorporates endogenous identification of structural breaks within the conditional variance process. Specifically, Markov-switching dynamics have been included within these models of conditional variance in order to capture discrete shifts in financial volatility, and to offer a unique characterisation of volatility persistence (Hamilton and Susmel, 1994).

This chapter reviews the theoretical and empirical developments relating to stock market volatility, contagion, and selected models of conditional variance, with specific emphasis on Markov-switching models. Section 2.2 reviews the underlying economic drivers of stock market volatility and contagion effects. This is in order to supplement the statistical characterisation of stock return volatility which is emphasised in the remainder of the study. Sections 2.3 and 2.4 contain an extensive review of empirical work incorporating univariate and multivariate Markov-switching models of conditional variance. For all the empirical studies considered, the models, data, tests, and major results are summarised and interpreted. Although the focus of this study is on stock returns, it is found that Markov-switching models of conditional variance are well suited to many different asset classes (for example, Gray’s (1996) MS-GARCH has been used to model the conditional variance of factors such as interest rates, exchange rates, stock returns and commodities). Thus, literature relating to interest rates, exchange rates, and stock returns are presented and are considered equally relevant in assessing model performance. Finally, section 2.5 provides a summary of the chapter and concludes.

2.2. Economic causes of stock market volatility and contagion

2.2.1. Volatility

To gain an in-depth understanding of the nature of stock return volatility, it is necessary to review various theoretical developments in this field. Beyond a purely statistical characterisation of the conditional distribution of returns, Hamilton and Lin (1996) argue that a satisfactory explanation for equity market volatility must include underlying economic causes. Indeed, much of the empirical work on the matter aims to measure the effect of real and nominal, micro- and macroeconomic variables on stock return variability.

Black (1976), for example, theorises that changes in stock return volatility are partly caused by the so-called ‘leverage effect’. According to Black (1976:177), a decline in the market value of a firm’s equity, *ceteris paribus,* will through time increase the debt/equity ratio (leverage) of the firm and hence increase its inherent riskiness. The negative relationship between return innovations and future volatility has proven to be robust, and has led to a number of statistical models that incorporate leverage effects, such as the GJR-GARCH model of Glosten, Jagannathan and Runkle (1993). Despite the apparent success of Black’s (1976) theory, Figlewski and Wang (2000) show that the leverage effect may not have much to do with leverage at all. Analysing S&P100 listed firms for the period 1977 until 1996, Figlewski and Wang (2000) illustrate that return volatility tends to increase when stock prices decline, as expected, but shows no such tendency when outstanding debt increases or outstanding shares decline. In light of this evidence, Figlewski and Wang (2000:23) suggest that the leverage effect be more appropriately referred to as a ‘down market effect’.

To the extent that stock returns are related to GDP growth, the leverage effect may be better characterised by studies that incorporate the effect of macroeconomic variables on stock market volatility. For instance, both Schwert (1989) and Hamilton and Lin (1996) find that stock volatility increases during recessions. This could be because the fixed costs associated with rigid labour markets or maintaining physical capital increases the riskiness of investments during periods of low revenue, a phenomenon known as operational leverage (Schwert, 1989:1145). Over and above this effect, however, Schwert (1989:1145) finds that when stock prices drop relative to bond prices, stock market volatility tends to increase, thus maintaining consistency with the financial leverage effect described by Black (1976).

It is not only the level of GDP growth that can affect stock market volatility. Using a variance decomposition technique, Schwert (1989) finds weak evidence that broad macroeconomic volatility is a predictor of U.S. stock market volatility. Diebold and Yilmaz (2007) show that by modifying Schwert's (1989) methodology to include pooled data (incorporating 46 developed and emerging markets), a stronger relationship is observed. In particular, Diebold and Yilmaz (2007) show that for the majority of markets considered, GDP volatility Granger causes stock return volatility.

Extending Schwert's (1989) study, Kearney and Daly (1998) assess the impact of the volatility of domestic macroeconomic variables on stock market volatility in Australia. It is found that the conditional volatilities of domestic interest rates and inflation are positively related to stock market volatility, while the conditional volatilities of industrial production, the money supply and the current account deficit are negatively related to stock market volatility. Although the relationships are statistically significant, no economic theory is offered in support of the evidence. In particular, a negative relationship between the volatilities of industrial production or money supply and stock market volatility is unsupported in the literature. This may point to the instability of the underlying relationship between macroeconomic fundamentals and stock return volatility, and strengthen the behavioural theories of Shiller (1988), among others.

Using stock market data for ten emerging economies from 1985 to 1995, Aggarwal *et al.* (1999) note that stock market volatility tends to exhibit frequent structural shifts associated with not only macroeconomic variables but social and political events as well. Due to the high level of idiosyncratic risks associated with emerging markets, the majority of volatility shifts are found to be domestically generated (Aggarwal *et al.*, 1999). In fact, in their investigation of the sources of global stock return volatility, Aggarwal *et al*. (1999) report that only the U.S. stock market crash of 1987 transmitted higher volatility to multiple markets.

Accounting for the phenomenon of structural shifts described in Aggarwal *et al*. (1999), Chinzara (2011) studies the relationship between a number of macroeconomic variables (including inflation, the prices of oil and gold, the interest rate, and the exchange rate) and the volatility of returns on the Johannesburg Securities Exchange (JSE). Chinzara (2011) finds that volatility in both the domestic short-term interest rate and the exchange rate has a significant positive relationship with stock return volatility. This contrasts with the findings of Kearney and Daly (1998), who report no statistically significant relationship between the conditional volatility of the exchange rate and that of stock returns in Australia. Thus, it may be that the underlying relationships are not only unstable through time but across markets as well.

In a separate but related strand of literature, stock market volatility is explained in terms of changing perceptions of future cash flows. Standard equity pricing models define stock prices as a function of future expected and unexpected dividends and the discount rate. Fama (1990) argues that measuring the total return variation in stock prices due to the aforementioned variables is one way to assess the rationality of the market. Campbell and Ammer (1993) show that a variance decomposition of this sort also allows for the assessment of the proximate causes (i.e. changes in cash flows, the discount rate, and the risk premium) of stock price movements as opposed to fundamental causes (changes in interest rates, economic growth, profitability, and so on). Since the seminal work of Shiller (1981), it is commonly found that stock price variation exceeds its expected magnitude when considering cash flows, interest rates and a constant risk premium (*cf.* Mankiw *et al.*, 1985; Poterba and Summers, 1987; Campbell and Ammer, 1993; Adam *et al.*, 2008). In defence of the efficient markets hypothesis, some authors have explained the discrepancy between actual and expected volatility by introducing a high and variable equity risk premium (Mankiw *et al.*, 1985; Adam *et al*., 2008). However, Poterba and Summers (1987) argue that this adjustment proves unconvincing, as shocks to volatility tend to affect stock returns for short periods at a time.

An alternative interpretation of the excess volatility finding is that psychological factors outside of the rationally defined dividend discount model affect stock prices (Mankiw *et al.*, 1985). Using U.S. stock market data, Campbell and Ammer (1993) find that stock return volatility typically has a standard deviation that is two to three times higher than that justified by expected future divided growth. In the wake of the U.S. stock market crash of October 1987, Cutler *et al*. (1989) find that innovations in macroeconomic fundamentals cannot explain all stock market variations. They posit that psychological factors affecting social consensus may play an important role in moving stock prices (Cutler *et al.*, 1989:5). Referring to the same crisis, Shiller (1988) emphasises the role of the media and communication systems in creating market-moving consensus. Formalising the effect of 'irrationally' formed consensus, Adam *et al.* (2008) are able to replicate many of the empirical regularities in stock return data. Through allowing agents to learn about future stock prices more slowly, a slight modification to the rational expectations-based stock pricing model of Lucas (1978), Adam *et al.* (2008) successfully model stock return overreaction in the short term and mean reversion over longer periods. Thus, behavioural considerations are able to provide a significantly different and more empirically accurate characterisation of stock market volatility than those based on purely economic thinking and rationally formed valuations.

2.2.2. Contagion and volatility spillovers

There has been some disagreement amongst economists on what exactly constitutes financial 'contagion' (Forbes and Rigobon, 2002). Dornbusch *et al.* (2000) distinguish between two components of contagion. The first is co-movement due to common underlying fundamentals affecting different assets. The second component consists of the co-movement between assets that is unexplained by fundamentals. Forbes and Rigobon (2002:2223) define contagion as only those co-movements that are unexplained by normal inter-linkages between markets, defined as excess covariance between returns.

In periods of changing volatility, a simple analysis of the cross-correlation of returns is inadequate for measuring contagion (Forbes and Rigobon, 2002). In fact, once having controlled for heteroscedasticity, Forbes and Rigobon (2002) find no evidence of an increase in cross-correlations associated with the 1987 U.S. stock market crash, the 1994 Mexican crisis, or the 1997 Asian crisis. Hence, a distinction is made between *interdependence* (which is equivalent to the first component of Dornbusch *et al*.'s (2000) definition) and *contagion* (which is equivalent to the second component). The idea that contagion be defined as excess correlation between asset returns has subsequently been adopted by much of the empirical literature on contagion (Bekaert *et al.*, 2003).

Longin and Solnik (1995) investigate whether the cross-correlation of returns between major developed markets is constant over the period 1960 to 1990. Estimating Bollerslev's (1990) CCC-GARCH model over a rolling window sample, Longin and Solnik (1995) find that the cross-correlation of returns is unstable. Specifically, during high volatility periods, the correlation coefficient tends to exhibit a statistically significant increase (Longin and Solnik, 1995). Since Longin and Solnik (1995) model the heteroscedasticity of returns explicitly, the results do not suffer from the bias identified in Forbes and Rigobon (2002)[1], and it is concluded that a contagion effect is present in the data during periods of high volatility.

Having established the existence of the contagion effect, many studies have attempted to explain the underlying causes of the phenomenon. With reference to the observed excess co-movements in international stock markets following the October 1987 crash, King and Wadhwani (1990) model contagion as the outcome of rational agents using imperfect information. In their model, past price changes are used by some market participants as a representation of all available information, and are used to form expectations of future returns. Thus, in an increasingly globalised financial environment, price changes in one market tend to be transmitted into price changes in another (King and Wadhwani, 1990). This "informational freeloading" (Cutler *et al*., 1989:9) effect can cause a self-reinforcing increase in volatility and correlation of returns, particularly during crisis periods when uncertainty is high.

Mendoza and Quadrini (2010) test the assertion that the global financial contagion of 2007/8 is related to the ongoing process of financial globalisation. Simulating the growth in global financial integration, Mendoza and Quadrini (2010) show how a negative shock to banks' equity can, through capital requirements, force banks to liquidate positions in other markets and therefore affect financial asset prices globally. Thus, while King and Wadhwani (1990) emphasise greater informational linkages, Mendoza and Quadrini (2010) emphasise the increasing footprint of international financial institutions in causing contagion.

Referring to contagion across asset classes, Longstaff (2010) shows how, during 2008, the indexed price of subprime asset backed collateralized debt obligations (CDOs) exhibit predictive ability for stock prices. Specifically, CDO prices are shown to affect equity prices with a lag at the height of the crisis. Longstaff (2010:437) interprets this finding as contradicting the correlated information hypothesis, in which new information is fundamentally relevant to a cross-section of securities. If this were the case, the CDO and equity prices would respond contemporaneously (Longstaff, 2010:437). Thus the liquidity channel, through which a shock to one market may force market participants to liquidate positions in other markets over time, is presented as a more satisfactory explanation for the lead-lag contagion effect.

[1] Forbes and Rigobon (2002) show that, in the presence of heteroscedasticity, the cross-correlation of returns between interdependent markets increases during high volatility periods. Since this is merely a statistical artefact, Forbes and Rigobon (2002) argue that it does not constitute evidence of contagion (at least in terms of their definition of 'contagion').

Bekaert *et al.* (2011:2) also refer to the global financial crisis of 2007/8 in their study of contagion across 55 equity markets. A three factor asset pricing model, using U.S. specific variables, global stock price movements, and country-specific macroeconomic indicators, is used to determine what equity market co-movements should be according to the specified fundamental variables. Similarly to Forbes and Rigobon (2002), contagion is defined as any co-movement in excess of what the model predicts (Bekaert *et al.*, 2011:2). In contrast to the 'financial globalisation' hypotheses presented in King and Wadhwani (1990) and Mendoza and Quadrini (2010), Bekaert *et al.* (2011) find that the contagion following the 2007/8 crisis was related more to domestic factors than to global factors. Specifically, countries with low foreign reserves, high current account deficits, and poor sovereign credit ratings experienced a greater degree of contagion than those who were stronger on these fronts, c*eteris paribus*. Bekaert *et al*. (2011:5) explain these results by means of the 'wake-up call' hypothesis, whereby market participants reassess the riskiness of investing in regions with less sound macroeconomic indicators following a market shock.

Although much of the contagion literature does not include African markets, there has been development in this area as well. Through the use of Forbes and Rigobon's (2002) adjusted correlation coefficient, Collins and Biepke (2003) discern whether African equity returns experience contagion from other global and African emerging markets. Of all the African stock markets studied (Egypt, Kenya, Mauritius, Morocco, Namibia, Nigeria, South Africa, and Zimbabwe) only South Africa and Egypt are shown to experience a contagion effect from global emerging market shocks (Collins and Biepke, 2003). In contrast, much of the intra-African contagion effects are related to geographical dispersion. Collins and Biepke (2003:193) interpret this as reflecting the greater trade links of geographically close countries. Thus, while the intra-African contagion effects seem to be driven by changes in fundamental economic variables, the global contagion effects experienced by Egypt and South Africa seem to be driven primarily by behavioural considerations (Collins and Biepke, 2003).

Going beyond conventional measures of contagion, Diebold and Yilmaz (2009) examine cross-market spillovers of both stock returns and volatility. For the 19 developed and developing markets studied (including South Africa), variance decomposition analysis shows significantly different behaviour between return spillovers and volatility spillovers over time (Diebold and Yilmaz, 2009). Since the 1990s, return spillovers tend to show an upward trend with relatively few spikes. According to Diebold and Yilmaz (2009:170), this pattern reflects the gradual integration of global financial markets. By contrast, volatility spillovers display a roughly constant mean, with sharply increasing spikes around crisis periods (Diebold and Yilmaz, 2009). Thus, Diebold and Yilmaz (2009) show that although return and volatility spillovers tend to be of the same magnitude on average, they may be at very different stages of evolution at any given time.

In a similar study, Chinzara and Aziakpono (2009) model the return and volatility linkages between the equity markets of South Africa, Australia, China, Germany, Japan, the UK and the US. Using univariate GARCH and multivariate VAR techniques, both return and volatility linkages between South Africa, Australia, China, and the US are established. The equity markets of Germany, Japan, and the UK do not exhibit any statistically significant linkages with the South African market. Chinzara and Aziakpono (2009:90) speculate that the linkages with South Africa are strong because Australia is a resource based economy, China is an emerging market, and the US market is the largest in the world. Indeed, if the psychological contagion effects described by King and Wadhwani (1990) were dominant, one would expect to observe return and volatility spillovers from Germany, Japan, and the UK as well. Thus, an explanation for return and volatility linkages between stock markets should not exclude the effects of common market fundamentals.

Overall, the underlying causes of volatility and contagion are not yet universally established, and empirical studies offer wide-ranging and sometimes contradictory results. Although much of the work relating macroeconomic fundamentals to volatility and contagion show statistically significant linkages (*cf.* Schwert, 1989; Hamilton and Lin, 1996; Diebold and Yilmaz, 2007), it appears in general that a behavioural theory of investor sentiment can improve our understanding of these phenomena (*cf.* Shiller, 1981; Mankiw *et al.*, 1985; King and Wadhwani, 1990; Adam *et al.*, 2008). The studies of Collins and Biepke (2003), Diebold and Yilmaz (2009), and Chinzara and Aziakpono (2009) further indicate that, in an African context, volatility and contagion effects are best understood with reference to many factors, including global integration, common market exposures, and investor confidence.

Whatever the economic causes of financial volatility, an accurate characterisation of volatility dynamics is essential to various market participants. In light of Aggerwal *et al.*'s (1999) finding that emerging stock markets tend to exhibit frequent structural shifts in the conditional variance of returns, a Markov-switching model (which is able to endogenously identify these shifts) appears to be a promising alternative to the more conventional single-state GARCH models. The following section therefore reviews a number of empirical studies which make use of various Markov-switching specifications. As already noted, since Markov-switching models have been applied with success to various financial series (including interest rates, exchange rates and stock returns, among others), section 2.3 considers each of these studies as equally relevant in assessing the potential benefit from including Markov-switching dynamics in the conditional variance process.

2.3. Empirical studies of Markov-switching volatility

2.3.1. Interest rates

Garcia and Perron (1996) investigate the dynamics of real interest rates under a Markov-switching framework. Given that the null hypothesis of a unit root cannot be rejected in some previous empirical studies of U.S. real interest rates, Garcia and Perron (1996) aim to provide a multiple-regime characterisation of the mean and variance process that allows for structural changes.

The Markov-switching model used by Garcia and Perron (1996) is drawn from Hamilton (1989), with a constant within-regime mean and variance. The authors use monthly and quarterly real interest rate data, obtained from inflation-adjusted 90-day U.S. Treasury Bill rates from 1961 to 1986.

In order to ascertain the appropriate number of regimes, Garcia and Perron (1996) use Davies' (1987) upper-bound test, Gallant's (1977) test, and Davidson and MacKinnon's (1981) *J*-test.[2] For the interest rate series, all the model selection tests favour a two-state model over a single-regime AR(4) specification. Additionally, all the tests favour the three-state specification over the two-state specification. Garcia and Perron (1996:112) thus argue that the inability by previous studies to reject the null of a unit root is due to these studies not accounting for occasional discrete structural shifts.

Dahlquist and Gray (2000) model the dynamics of short-term interest rates in the European Monetary System (EMS) using a Markov-switching framework. Under the Exchange Rate Mechanism, countries within the EMS were obliged to keep their respective currencies within

[2] See Garcia and Perron (1996) for more information on these tests. The Davies and Gallant tests are in-sample tests of model performance in the presence of nuisance parameters. The *J*-test is a forecasting test used for non-nested models.

predefined bands. In the event that central banks lost credibility, speculative attacks would have to be fended off, often by using (among other tools) interest rates as a deterrent (Dahlquist and Gray, 2000:400). In this way, calm periods in which credibility is maintained can result in vastly different behaviour of the short rate to periods in which speculative attacks occur (Dahlquist and Gray, 2000:400).

Dahlquist and Gray (2000) use a discretised version of Cox *et al.*'s (1985) model of short rates, given by:

$$\Delta r_t = \psi_{s_t}\left(\mu_{s_t} - r_{t-1}\right) + \varepsilon_t \tag{1}$$

$$\varepsilon_t | I_{t-1} \sim N(0, \sigma_{s_t}^2 r_{t-1}), \tag{2}$$

where ψ_{st} denotes the state-dependent speed-of-adjustment coefficient, μ_{st} denotes the state-dependent mean interest rate, r_t is the interest rate, I_{t-1} is the information set available at time $t-1$, and $\sigma_{st}^2 r_{t-1}$ is the state-dependent conditional variance of changes in interest rates. The above specification allows the behaviour of the data to change in a number of ways across regimes. Each regime may be associated with a unique long-run mean, a unique speed-of-adjustment parameter, and a unique degree of volatility (Dahlquist and Gray, 2000:405). The model is estimated using constant transition probabilities as well as time-varying transition probabilities.

The data used are one-week Euro-currency rates for Italy, Netherlands, Germany, France, Belgium, and Denmark, observed weekly. The sample period begins on 8 January 1980 for France, Germany, Italy, and Netherlands, and 6 January 1981 and 18 June 1985 for Belgium and Denmark respectively, and ends on 7 July 1998 for all countries (Dahlquist and Gray, 2000:404). For all countries analysed, the periods of weaker credibility are captured by a high volatility, high interest rate regime. The speed-of-adjustment coefficient is higher in this high interest rate regime, and so mean reversion (to a higher long-run mean) is stronger (Dahlquist and Gray, 2000:401). During periods of greater credibility, the short rates tend to be lower and less volatile. The speed-of-adjustment coefficient is also lower in these periods, indicating a weaker degree of mean reversion. The model thus identifies distinct dynamic behaviour between periods of calm and turbulence. Overall, Dahlquist and Gray (2000) argue that the regime-switching model is able to capture economically significant features of interest rate changes that are not adequately captured by single-regime specifications.

Cai (1994) tests the validity a of single regime AR(1)-ARCH(2) model against that of a SWARCH model. The data used are the excess returns of the three-month U.S. Treasury Bill over its one-month counterpart, observed monthly from August 1964 to November 1991. The sample period is selected in order to allow the Markov-switching model to potentially capture the important movements in interest rates associated with the oil crisis of the early seventies and the Fed experiment of 1979-1982 (Cai, 1994).

Rather than employing Hansen's (1992) adjusted LRT or Davies' (1987) upper bound test for more than one regime, Cai (1994) performs a Monte Carlo experiment (using 100 simulations with a sample size of 328 observations) in order to compute critical values for the 90%, 95%, and 99% levels of confidence. This test results in the rejection of the single regime AR(1)-ARCH(2) in favour of a Markov-switching specification. The significance of regime-switching in the data is evidenced by the fact that the null hypotheses of no within-regime ARCH effects cannot be rejected at conventional levels (Cai, 1994). Hence, having accounted for structural shifts in the variance, allowing for time-varying persistence of the ARCH and GARCH type does not significantly improve the fit of the model. Residual diagnostic tests indicate that the single-regime AR(1)-ARCH(2) provides a poor fit of the data, and that both the constant within-regime variance and SWARCH(2,2) specifications provide a better fit, as measured through the squared standardised residuals (Cai, 1994). The ability of the SWARCH model to capture the heteroscedasticity in the data, points to the potential superiority of Markov-switching volatility models over various single-state alternatives.

In a more comprehensive study of interest rate dynamics, Gray (1996) develops the Generalised Regime-Switching (GRS) model to examine one-month U.S. Treasury bill rates, observed weekly from January 1970 through April 1994. Throughout the study, Gray (1996) quotes quasi-LRT statistics (a conventional LRT statistic under the violation of regularity conditions) when comparing single- to multi-regime models. This is due to the computationally burdensome procedure associated with the Hansen (1992) LRT (Gray, 1996). However, the quasi-LRT statistics are often significant enough (under normal circumstances) that a reasonable amount of confidence may be taken in them (Gray, 1996). Besides this shortcoming, Gray (1996) appeals to the economic significance of the results in illustrating model improvements. That is, the estimated parameters imply distinct behaviour within each regime, strengthening the case for a multi-regime model (Gray, 1996).

So as to assess the benefits of each model feature, the major components of the GRS model - regime switches, GARCH effects, level effects, and time-varying transition probabilities - are incrementally added to a single-regime constant volatility model until the full GRS specification is reached. Initially, estimation results for a single-regime constant volatility model show that a poor fit of the data is achieved. Volatility estimates are systematically too high during low volatility periods and too low during high volatility periods. In fact, there is found to be significant autocorrelation in the squared standardised residuals, suggesting heteroscedasticity in the residual series. A quasi-LRT thus rejects the null of a single regime (Gray, 1996). By allowing for regime switches, much of the autocorrelation in the squared standardised residuals disappears. Relaxing the assumption of a constant within-regime variance, conventional GARCH dynamics are added to the model. An LRT indicates that the GARCH terms are important, in that the null hypothesis of no GARCH effects is rejected. Following the inclusion of a GARCH term, Ljung-Box *Q*-statistics show no remaining autocorrelation in the squared standardised residuals. Gray (1996:49) thus states that "allowing GARCH effects in each regime provides a richer characterisation of the conditional variance". In the full GRS model, Gray (1996) incorporates level effects in the variance process and allows for time-varying transition probabilities. Level effects are shown to be statistically significant within the high mean, high volatility regime. In addition, an LRT rejects the null of constant transition probabilities, favouring time-varying probabilities (Gray, 1996). Evidently, at each increment the new model provides a better statistical fit than the previous model (Gray, 1996).

In testing the forecasting accuracy of the various model specifications, Gray (1996) finds that the full GRS model performs the best overall in terms of minimising RMSEs and MAEs, and maximising R^2. The GRS model is thus shown to perform better than existing benchmark models in terms of the in-sample fit (measured by LRTs, quasi-LRTs, and the significance of Ljung-Box *Q*-statistics for standardised residuals and squared standardised residuals), and out-of-sample forecasts (measured by RMSE, MAE, and R^2 between actual and forecast volatility). This indicates that the improved within-sample fit is not simply due to over-fitting the data (Gray, 1996). That is, allowing for two types of volatility persistence (state-varying and time-varying) can capture important features of the conditional variance process in this instance.

2.3.2. Exchange rates

Knedlik and Scheufele (2008) compare three popular methods used for forecasting currency crises, with particular reference to South Africa. These methods include a binary probit model, a signals approach, and a Markov-switching model (Knedlik and Scheufele, 2008:369).

The Markov-switching model used by Knedlik and Scheufele (2008) is a variant of Hamilton's (1989) model, using both a constant within-regime mean and variance, with time-varying transition probabilities (Knedlik and Scheufele, 2008). The data used are monthly observations of a constructed Exchange Market Pressure (EMP) index, which captures changes in not only the exchange rate but interest rates and reserves as well. The sample period used to estimate the models is from 1995 until 2005, and an out-of-sample forecasting test includes the currency crisis of June 2006.

Knedlik and Scheufele (2008) report that the best performing models in the out-of-sample forecasting test are the probit and Markov-switching models. The performance is based on a number of count R^2-type measures, in which the percentage of correct calls is used as a gauge of forecast accuracy (Knedlik and Scheufele, 2008).

It is noted that although the signals approach has a higher level of simplicity, the incorporation of Markov-switching and probit techniques into econometric software packages makes these approaches more viable and simpler to execute (Knedlik and Scheufele, 2008). In addition, the *ad hoc* manner in which the probit model requires identification of crisis periods renders its results possibly unreliable. The Markov-switching model, which uses an endogenous identification technique, does not suffer from this drawback and thus allows for a higher degree of confidence (Knedlik and Scheufele, 2008).

In a related study, Duncan and Liu (2009) attempt to model South African currency crises by identifying structural breaks using Inclan and Tiao's (1994) iterative cumulative sum of squares (ICSS) algorithm. The results from their procedure are then compared to the Markov-switching results of Knedlik and Scheufele (2008), in a test of the performance of each approach. The data used are daily observations of the Rand-Dollar exchange rate, spanning the period 3 January 1994 to 31 March 2009. Duncan and Liu (2009) deliberately choose the sample period to overlap with the period studied in Knedlik and Scheufele (2008).

Using the exchange rate data, Duncan and Liu (2009) initially estimate a standard GARCH(1,1) model, and subsequently compare it to a structural change GARCH (SC-GARCH) model. The SC-GARCH is estimated using dummy variables indicating each crisis period as identified by the ICSS algorithm (Duncan and Liu, 2009). The results of the comparison favour the SC-GARCH over the GARCH(1,1), given that the GARCH(1,1) model implies a non-stationary variance and the SC-GARCH leaves no remaining ARCH effects (Duncan and Liu, 2009). The evidence of structural changes in the data provides the rationale for incorporating Markov-switching dynamics in the model.

By comparing the SC-GARCH model with Knedlik and Scheufele's (2008) Markov-switching model, Duncan and Liu (2009) find that the SC-GARCH is more precise in its identification of crisis periods. As one of the reasons for this, Duncan and Liu (2009) cite the typically low-frequency observations of fundamental economic data. Since Knedlik and Scheufele's (2008) model uses time-varying transition probabilities (which are constructed as functions of fundamental economic data), this Markov-switching approach is restricted to monthly observations. In contrast, Duncan and Liu (2009) use high-frequency data, allowing for more precise measurement of periods.

Despite the strength of Duncan and Liu's (2009) SC-GARCH, it has two inherent weaknesses. Firstly, using extreme exchange rate fluctuations as a proxy for currency crises is restrictive. This is shown by the model's failure to identify accurately the crises of February 1996 and June 2006 (both of which are identified in Knedlik and Scheufele (2008)). Secondly, it is not clear how the SC-GARCH model may be used for forecasting currency crises, as it requires *ex post* identification of structural breaks. This is in contrast to the Markov-switching model, which establishes a probability law with which to forecast (Knedlik and Scheufele, 2008). The Markov-switching model may therefore provide a more useful way to capture the relevant structural breaks, particularly for forecasting purposes.

In another study of developed market exchange rate volatility, Bollen *et al*. (2000) compare the performance of a regime-switching model with that of standard GARCH models in terms of variance forecasts as well as option valuations. The regime-switching specification used is a variant of the Hamilton (1989) model, and the data are assumed to be normally distributed (Bollen *et al*., 2000). An innovation is introduced into this model, however, in that both the mean and variance are allowed to switch between regimes independently, so that there are effectively four primitive states (s_t*), so that:

$s_t* = 1: s_{\mu t} = 1; s_{\sigma t} = 1$ (3)

$s_t* = 2: s_{\mu t} = 2; s_{\sigma t} = 1$ (4)

$s_t* = 3: s_{\mu t} = 1; s_{\sigma t} = 2$ (5)

$s_t* = 4: s_{\mu t} = 2; s_{\sigma t} = 2,$ (6)

where $s_{\mu t}$ and $s_{\sigma t}$ denote the mean and variance regimes, respectively. The economic rationale for modelling the mean and variance regimes as independent is that it allows for stable and unstable periods of appreciation and depreciation (Bollen *et al.*, 2000:245). This four regime specification, as well the traditional single and two-regime specifications are used in comparison with the GARCH(1,1) and AR(1)-GARCH(1,1) (in which returns are allowed to exhibit first order autocorrelation). The data used in Bollen *et al.* (2000) consist of continuously compounded weekly returns on the Deutsche Mark, British Pound, and Japanese Yen, all against the U.S. Dollar from January 1973 until December 1996.

Using Garcia's (1998) LRT, the null hypothesis of a single regime is rejected in favour of the two-regime model at the 99% confidence level for all series (Bollen *et al.*, 2000). Hansen's (1992) adjusted LRT of the four-regime specification against the two-regime specification is not presented due to its numerical complexity; however Bollen *et al.* (2000) appeal to the economic reasoning underlying the four-regime version.

Bollen *et al.*'s (2000) preference for four regimes is validated by a series of Ljung-Box *Q*-tests. Only the four-regime specification is able to eliminate serial correlation in both the standardised and squared standardised residuals. The single regime, two-regime, GARCH(1,1), and AR(1)-GARCH(1,1) models all fail to eliminate serial correlation in one of or both the standardised residuals and squared standardised residuals (Bollen *et al.*, 2000). In order to further test the statistical performance of the models, Bollen *et al.* (2000) use out-of-sample forecasting tests, including the RMSE and MAE tests at 1-, 4-, and 8-week horizons. For the Deutsche Mark and British Pound, the four-regime model outperforms all the other specifications, whereas for the Japanese Yen, the results are mixed. On balance, the regime-switching models provide superior forecasts to the GARCH alternatives.

To illustrate a potential benefit of the regime-switching model, Bollen *et al.* (2000) determine whether valuing options using regime-switching variance forecasts can lead to abnormal profits. The models compared are the simple two-regime model, the Black-Scholes model using lagged implied volatility, and the Black-Scholes model using an estimate of volatility from the previous year's observations (Bollen *et al.*, 2000). For all three currencies, the regime-switching model generates profits that are greater than those of the other two models. Therefore, the results suggest that a relatively simple regime-switching specification of exchange rate volatility can provide not only more accurate forecasts but lead to higher trading profits than its single-regime counterparts (Bollen *et al.*, 2000).

Klaassen (2002) and Haas (2004) develop extensions to Gray's (1996) MS-GARCH model, in which more efficient use of information is made and parameter estimates are more intuitively interpretable. Like Gray (1996) these models allow for two types of volatility persistence, specifically autoregressive persistence and regime persistence.

The Markov-switching specifications in Klaassen's (2002) study are restricted to have two-regimes. Additionally, innovations are assumed to follow the Student's *t* distribution in order to enhance the stability of the regimes (Klaassen, 2002). The comparative study makes use of a single-regime constant variance model, Bollerslev's (1986) GARCH(1,1), a constant within-regime variance model, a SWARCH model in which all the coefficients are allowed to change, Gray's (1996) MS-GARCH, and Klaassen's (2002) MS-GARCH. The data used are daily returns on the British pound, Japanese yen, and German mark exchange rates, all against the U.S. dollar from 3 January 1978 to 23 July 1997 (Klaassen, 2002).

With respect to the estimation results, the GARCH(1,1) is shown to provide a better fit than the single-regime constant variance model, as indicated by an LRT. However, for the GARCH(1,1), the persistence of individual shocks on future volatility is estimated to be very high for all three series. The highly persistent autoregressive dynamics evident in the GARCH models suggest that a regime switching characterisation may be beneficial. Although the presence of a second regime is not formally tested for, the constant within-regime volatility model indicates that the second regime exhibits an unconditional variance which is three to four times greater than the first regime estimate for all series (Klaassen, 2002). In addition, the importance of regime-switching is illustrated by the fact that the degree of autocorrelation present in the squared standardised residuals is greatly reduced (although there is still evidence of some heteroscedasticity) when moving from the constant volatility model to the constant within-regime volatility model (Klaassen, 2002).

The SWARCH models used result in no additional heteroscedasticity for the yen, but fail to account for all ARCH effects for the pound and mark. This indicates the need to include

GARCH terms in the models (Klaassen, 2002). Indeed, the remaining conditional heteroscedasticity for the pound and mark is no longer present once a GARCH term is included. Moreover, an LRT favours the MS-GARCH over the SWARCH models for these two series. In a comparison of Gray's (1996) with Klaassen's (2002) MS-GARCH specifications, Klaassen (2002) notes that the parameter estimates are not significantly different. However, Gray's (1996) specification results in a small amount of residual heteroscedasticity, as well as worse log-likelihood values. Klaassen (2002) speculates that this may be because Gray's (1996) specification makes less efficient use of available information in estimation of parameters.

Given the study's emphasis on volatility forecasts, Klaassen (2002) examines the model performances using three tests of forecast accuracy at 1- and 10-day horizons. Following Anderson and Bollerslev (1998), the *p*-step-ahead measure of observed volatility ($\upsilon_{t,p}$) is given as the mean-adjusted sum of squared daily returns, such that:

$$\upsilon_{t,p} = \sum_{\tau=t}^{t-1+p} (y_\tau - \mu)^2. \tag{7}$$

In the first test, variance forecasts are then regressed against $\upsilon_{t,p}$, so that:

$$\upsilon_{t,p} = \gamma_0 + \gamma_1 E_{t-1}(h_{t+p}) + \varepsilon_t, \tag{8}$$

where $\gamma_0 = 0$ and $\gamma_1 = 1$ if the mean and variance forecasts are conditionally unbiased (Klaassen, 2002). For the GARCH(1,1) model, the null of unbiased forecasts ($\gamma_0 = 0$ and $\gamma_1 = 1$) is rejected in eight out of twelve cases, and it is shown that in each case $\gamma_0 > 0$ and $\gamma_1 < 1$. Klaassen's (2002) MS-GARCH yields better results, in that the null is not rejected in general. This is largely due to the reduction in persistence of shocks on volatility forecasts (Klaassen, 2002).

The other measures of forecast accuracy employed by Klaassen (2002) are the MSE test and an R^2-type statistic similar to that used in Gray (1996:54). According to the MSE measure, the MS-GARCH(1,1) provides more accurate variance forecasts than the single-regime models. Within the multi-regime models, Klaassen's (2002) specification outperforms Gray's (1996) MS-GARCH at the one day horizon (ten day forecasts are not available using Gray's (1996) specification). In eleven out of twelve cases, the MS-GARCH model results in a higher R^2 statistic than the single-regime GARCH. Specifically, the MS-GARCH model outperforms the GARCH model by 22% at the one day horizon and 58% at the ten day horizon (Klaassen, 2002:386). Clearly, the comparative advantage of the Markov-switching specification increases at longer forecast horizons.

In a closely related empirical study, Haas *et al*. (2004) examine continuously compounded daily returns for the Japanese yen, British pound, and Singapore dollar, all against the U.S. dollar from January 1978 to June 2003 for the yen and pound (which roughly encapsulates the period studied by Klaassen (2002)), and January 1981 to June 2003 for the Singapore dollar.

The models used for comparison are the single-regime GARCH(1,1) using normal and Student's *t* distributions, a mixed normal GARCH (MN-GARCH) model in which regimes are drawn from a multinomial distribution, and Haas *et al*.'s (2004) MS-GARCH model. The latter two (mixture) models are estimated using only the normal distribution for the errors and using both two- and three-regime specifications (Haas *et al*., 2004). Although many researchers have illustrated the fat-tailed nature of financial returns series, Haas *et al*. (2004) do not regard the assumption of normality under regime-switching as too restrictive. This is because regime-switching models of volatility provide for other sources of excess kurtosis, such as the switches themselves, which may render within-regime excess kurtosis superfluous (Haas *et al*., 2004).

The estimation of the three competing models and their various specifications yields in-sample results that favour the mixture models over the single-regime models (Haas *et al*., 2004). Comparing the observed autocorrelation of the squared residuals to their estimates, it is shown that the GARCH(1,1) model is unable to accurately capture the ARCH effects, whereas the mixture models provide an improved fit. Comparing the mixture models, the information criteria indicate that the MN-GARCH model provides a better fit for the Japanese yen and British pound, while the Singapore dollar is better modelled using the MMS-GARCH (Haas *et al*., 2004).

In order to further test the performance of the models Haas *et al*. (2004) use one-step-ahead forecast errors, a Value-at-Risk (VaR)-based loss function, and an ARCH-LM test. With respect to the forecasting and VaR tests, the single-regime GARCH(1,1) models are clearly outperformed by the mixture models (Haas *et al*., 2004). Turning to the MN-GARCH and MS-GARCH models, the forecast error and VaR tests both favour the former model over the latter.

An ARCH-LM test of the GARCH(1,1), using the transformed errors, does not indicate any remaining ARCH effects. However, the MS-GARCH models exhibit the lowest values for the ARCH-LM tests, indicating that these models may capture some important features in the data (Haas *et al*., 2004). The results indicate, at the very least, that including more than one state can provide a sufficient fit of the data, as well as improve forecast accuracy over the conventional single-state GARCH models.

2.3.3. Stock returns

Wang and Theobald (2008) use a regime-switching framework to assess the changing behaviour of East Asian stock markets during a period of liberalisation and deregulation of their financial systems. The data used consist of monthly returns on stock market indices for Taiwan, Thailand, the Philippines, Malaysia, Korea, and Indonesia using starting dates ranging from 1971 to 1986 (depending on data availability) and ending in July 2004. The initial goal of the study is to investigate the existence of changing volatility regimes over the sample periods and secondly to assess the behaviour of volatility regimes around the specific dates of financial liberalisation (Wang and Theobald, 2008).

The Markov-switching model used in the study is drawn from that of Hamilton (1989). Specifically, a second order autoregressive model with a constant unconditional mean and a regime-switching variance is used. The focus on regime-switching variance is due to the fact that the authors did not find significant shifts in the mean for any country considered (Wang and Theobald, 2008:268). In this specification, the within-regime conditional variance is assumed constant, and any heteroscedasticity is captured through differences in the variance between regimes (Wang and Theobald, 2008).

In identifying the existence of multiple regimes, Wang and Theobald (2008) use Davies' (1987) upper-bound test. The results of Davies' (1987) test indicate that for all samples considered, the null hypothesis of a single regime is rejected in favour of either two or three volatility regimes. Further LR tests suggest that for Malaysia, Taiwan, and the Philippines, the correct specification consists of two regimes, while for the other countries a three regime specification is best. In order to confirm the validity of these specifications, Wang and Theobald (2008) perform an ARCH test on the standardised residuals from each model. In all cases, no remaining ARCH effects were found. The regime-switching model thus provides an adequate fit of the data, as there appears to be no remaining heteroscedasticity after accounting for the structural breaks in the conditional variance process.

Observing the estimated smoothed probabilities, Wang and Theobald (2008) note that the switching between regimes for all countries tends to follow significant domestic and international events, illustrating a certain degree of interdependence between markets. The effect of financial liberalisation, however, varies widely across the markets, with Thailand, Indonesia, and Korea permanently switching to a higher volatility regime while Taiwan, the Philippines and Malaysia show no such structural shift (Wang and Theobald, 2008). The conclusion drawn is that the effect of financial liberalisation on emerging market stock volatility is country specific.

In another study of stock market regime changes, Chen (2009) investigates the predictability of bear and bull market periods, using both Markov-switching and non-parametric dating techniques. The data used are continuously compounded monthly returns on the S&P 500 stock price index, with a sample period from February 1957 until December 2007.

Chen's (2009) specification is a variant of Hamilton's (1989, 1990) simple Markov-switching model. Comparing a two-regime specification to its single-regime counterpart, Chen (2009) refers to the LR statistic critical values tabulated in Garcia (1998). These critical values are calculated specifically for constant within-regime mean and variance models. Since the calculated LR statistic is far greater than the 99% critical value, Chen (2009) concludes that the Markov-switching model provides a better fit of the data than the single-regime specification. In testing whether a three-regime model outperforms a two-regime model, Chen (2009) notes that appropriately adjusted critical values for the LR statistic are not readily available. Thus, Psaradakis and Spagnolo's (2003) adjusted Akaike information criterion (AIC) and Bayesian information criterion (BIC) figures are calculated for each specification. Providing both a smaller AIC and BIC, the two-regime specification is chosen as the preferred specification in modelling stock returns (Chen, 2009).

In estimation of parameters, the selected two-regime model identifies a high mean, low variance regime and a low mean, high variance regime (Chen, 2009). The high mean, low variance regime is associated with bull markets, and the low mean, high variance regime with bear markets. Both regimes are shown to be highly persistent, with a bull market expected to last more than three times as long as a bear market (Chen, 2009). The distinct behaviour in each regime thus indicates the economic significance of regime-switching effects in the data.

Moving beyond the relatively straightforward model of Hamilton (1989), Hamilton and Susmel (1994) develop the SWARCH model in their study of stock return volatility. The data used in their study are weekly returns of a value-weighted portfolio of stocks traded on the New York Stock Exchange, from 3 July 1962 to 29 December 1987. Comparing a single-regime ARCH (q) model with their SWARCH (K,q) specification, a quasi-LRT statistic strongly rejects the null hypothesis of no regime-switching in the data. Hamilton and Susmel (1994:322) acknowledge the problem of unidentified parameters under the null, but nonetheless regard the quasi-LRT statistic as a "useful descriptive summary of the fit of the alternative models".

Following the establishment of regime switching, numerous Gaussian SWARCH(K,q) models (with $K = 1,\ldots,4$ and $q = 1,\ldots,3$) are fit to the stock market data. Additional specifications include adding a GJR leverage effect term (denoted SWARCH-L(K,q)), and the assumption of Student's t innovations. ARCH coefficients up to $q = 2$ were found to be statistically significant in all specifications. In addition, the leverage effect improved the fit of the model. Both a Wald test and an LRT reject the null of normally distributed innovations in favour of Student's t innovations (Hamilton and Susmel, 1994). The AIC indicates that the SWARCH-L(4,2) model is the best performing model of those estimated, followed by the SWARCH-L(3,2). The Schwartz (1978) criterion (SC) selects the single-regime GARCH-L(1,1) as the best performing model, followed again by the SWARCH-L(3,2) (Hamilton and Susmel, 1994). It is noted, however, that the AIC and SC statistics suffer from the same violation of regularity conditions as the traditional LRT, and that further investigation into model performance should be done (Hamilton and Susmel, 1994).

Because of the precarious nature of the quasi-LRT, AIC, and SC statistics, Hamilton and Susmel (1994) evaluate the performance of the models through forecast errors. It is found that the Student's t SWARCH-L(4,2) performs the best in terms of minimising MSE and MAE, and is the only SWARCH specification that outperforms the constant variance specification. In addition, the SWARCH-L(4,2) performs well in minimising the loss function $\ln(\varepsilon_t^2) - \ln(h_t)$, although not as well as the single regime GARCH-L(1,1). These results are thus mixed overall, but indicate that, in some cases, the SWARCH model may outperform the single-state conditional variance models.

Following the methodology of Hamilton and Susmel (1994), Edwards and Susmel (2001) aim to determine whether stock market volatility increased for a group of Latin American countries and Hong Kong during the 1990s, and whether this increase to a high volatility regime coincides across the selected countries. The specific data used for the study are weekly rates of return on the MSCI stock market indices for Mexico, Brazil, Argentina, Chile and Hong Kong from the last week of August 1989 until the third week of October 1999 (Edwards and Susmel, 2001).

To test the applicability of single-regime models, Edwards and Susmel (2001) initially estimate a GARCH(1,1) for each series. In each case, the persistence of shocks is high. Following Lamoureux and Lastrapes (1990), Edwards and Susmel (2001) interpret this high degree of persistence as a possible indication of structural breaks in the variance. This possibility is thus formally tested using Hansen's (1992) adjusted LRT. In all cases, the null of no switching is rejected in favour of a multiple-regime specification (Edwards and Susmel, 2001).

In their multi-regime study, Edwards and Susmel (2001) use Hamilton and Susmel's (1994) univariate SWARCH specification in order to model the mean and variance of stock market returns for each country. SWARCH models are estimated using two to four regimes and zero to three autoregressive terms. Furthermore, the model is estimated using innovations drawn from Student's *t* distribution, as well as including a leverage effect term (Edwards and Susmel, 2001).

A common feature of SWARCH and MS-GARCH models, and a significant finding of Edwards and Susmel's (2001) study, is that for all series the ARCH parameters are significantly reduced in comparison to the single-regime model. Thus, while the volatility clustering phenomenon is still captured effectively by multiple-regime models, a significant proportion of the clustering is attributed to the persistence of each volatility state (a structural change) rather than individual shocks (Edwards and Susmel, 2001). Despite the finding that for some countries the three-state specification marginally outperforms the two-state model, Edwards and Susmel (2001) choose to focus on the latter due to its tractability when used in a multivariate framework. For each stock market, the scaling parameter associated with the second regime is found to be statistically greater than one, thus distinguishing between a low and high volatility regime (Edwards and Susmel, 2001). In four of the five countries studied, the leverage effect term is found to be statistically insignificant, indicating no asymmetries associated with positive and negative shocks (Edwards and Susmel, 2001). This suggests the possibility that conventional leverage effects may become statistically insignificant when controlling for regime changes in the conditional variance process.

Extending Edwards and Susmel's (2001) study, Canarella and Pollard (2007) apply Hamilton and Susmel's (1994) SWARCH model to a set of Latin American stock markets. The data used are daily returns on stock market indices for Mexico, Peru, Venezuela, Argentina, Chile, and Brazil from 3 January 1994 to 29 April 2005. The study thus extends the sample period used by Edwards and Susmel (2001) to include the Argentinean debt crisis of the early 2000s (Canarella and Pollard, 2007).

For each return series, the mean equation is assumed to follow a single-regime AR(1) specification. Switches between regimes are thus driven entirely through the condtitional variance. Canarella and Pollard (2007) estimate SWARCH(2,4) models using both normal and Student's *t* distributions for the innovations.

In order to test the significance of the second regime, Canarella and Pollard (2007) follow Hamilton and Susmel (1994) and use a quasi-LRT statistic as a "useful descriptive summary". For all markets, and using both distributions for innovations, the quasi-LRT strongly suggests the existence of the second regime. In addition, the scaling parameter for the second regime is significantly different from one in all cases, suggesting economically distinct behaviour of the conditional variance between regimes (Canarella and Pollard, 2007).

Using both distributions, the four ARCH terms are all statistically significant, suggesting that volatility is both time- and state-varying (Canarella and Pollard, 2007). The states are highly persistent for all markets, providing an alternative explanation for the volatility-clustering phenomenon to traditional GARCH models. Overall, the Student's *t* specification is preferred as it results in a significantly higher log-likelihood value (Canarella and Pollard, 2007).

Beyond the simpler Markov-switching and SWARCH models, MS-GARCH models have also been used to model stock market volatility. For example, Dueker (1997) assesses the performance of five Markov-switching GARCH models and a single-regime GARCH(1,1) in modelling daily returns on the S&P 500 index. The sample period spans 6 January 1982 to 31 December 1991, and the out-of-sample test period ends September 1994. Because of the violation of regularity conditions under Markov-switching models, the in-sample fit and forecasting performance of each model are tested using Vlaar and Palm's (1993) test[3] (which is suitable for data that are not identically distributed). For the in-sample test, only the GARCH-DF model (in which the Student's t degrees-of-freedom parameter is allowed to switch between regimes) is shown to fit the data well, and all the models perform poorly in the out-of sample test. The inability of some of the models to deal with time-varying kurtosis is given as the primary reason for their relatively poor performance.

In light of the poor statistical fit of the models, Dueker (1997) further investigates their performance using an economic test. Specifically, the expected difference between the forecasted variance and an option-implied variance (as given by the volatility index (VIX) compiled by the Chicago Board Options Exchange) is used as a measure of forecasting accuracy of a particular model (Dueker, 1997). The VIX is constructed so as to represent the implied volatility on an at-the-money option on the S&P 100 with 22 trading days to expiration. Thus, for each model, the variance estimate used in comparison is the 22-day average forecast, $\bar{\sigma}_t^2$, so that the mean squared error of each model is given by:

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(\bar{\sigma}_t^2 - VIX_t)^2. \tag{9}$$

Using this test, Dueker (1997) shows that only those models that allow for changing kurtosis coefficients outperform the standard GARCH(1,1). In fact, the MSE associated with the GARCH-DF specification is 14% lower than that achieved by the GARCH(1,1). Given the poor performance of the majority of MS-GARCH models relative to the GARCH(1,1), the benefit of including two sources of volatility persistence is seen to be questionable.

A later study which makes use of an MS-GARCH specification for forecasting stock return volatility is that of Marcucci (2005). Marcucci (2005) reiterates the finding by Lamoureux and Lastrapes (1990) that high volatility persistence found in single-regime GARCH models may be due to structural breaks in the volatility process. This is also consistent with the fact that many studies of single-regime GARCH find poor forecasting performances associated with these models (Marcucci, 2005).

In order to examine this finding, the forecasting performance of multiple GARCH models are assessed according to various loss functions. The GARCH models assessed by Marcucci (2005) are the simple GARCH(1,1), Nelson's (1991) EGARCH, Glosten *et al.*'s (1993)

[3] This test is based on the cumulative probability of observing residual values smaller than the actual values observed. See Dueker (1997) or Vlaar and Palm (1993) for a more detailed exposition.

Threshold GARCH, Gray's (1996) MS-GARCH, and Klaassen's (2002) MS-GARCH. Each model is evaluated using innovations from the normal, Student's *t*, and generalised error (G.E.D.) distributions. The extent of forecast errors are then tested according to traditional MSE and MAE measures, R^2-type measures, and Bollerslev and Ghysels's (1996) heteroscedasticity-adjusted MSE (Marcucci, 2005). In addition, directional accuracy measures such as a Success Ratio test and Pesaran and Timmerman's (1992) Directional Accuracy test are employed.[4]

Building on these simple measures of forecast accuracy, Marcucci (2005) tests the models against each other using the Diebold-Mariano (1995) test for equal predictive ability, and both White's (2000) and Hansen's (2001) tests for superior predictive ability.[5] Finally, based on the importance of volatility forecasting in financial risk management (Engle, 2004), Marcucci (2005) tests each model using a VaR-based loss function.

Marcucci (2005) employs data from the daily closing price of the S&P 100 stock market from 1 January 1988 to 15 October 2003. The portion used for out-of-sample testing is from 1 October 2001. According to the equal and superior predictive ability tests, the MS-GARCH models are shown to outperform the single-regime models over short forecast horizons. In particular, the MS-GARCH model using Gaussian innovations performs the best overall (Marcucci, 2005).

In contrast, the VaR based test shows mixed results. At shorter horizons the MS-GARCH model fares best, but at longer horizons the standard GARCH provides superior forecasts (Marcucci, 2005). In this way, the forecasting performance of a particular model can depend critically on the loss function used, as well as the forecast horizon. Although regime-switching models can provide a much better in-sample fit, Marcucci (2005) concludes that it is often the simpler specifications that prove to be more robust when tested according to "economic loss functions" such as VaR.

In the process of developing a Markov-switching E-GARCH model, Henry (2009) studies the mean and conditional variance of U.K. equity returns, and the relationship between these returns and short term interest rates. The specific data used are continuously compounded weekly equity returns on a U.K. total return index, and interbank interest rates for loan

[4] Success Ratio and Directional Accuracy tests measure the extent to which the model correctly predicts the direction of *changes in* the conditional variance (Marcucci, 2005:12).
[5] The Diebold-Mariano (1995) test is a pairwise test of relative predictive ability. The White and Hansen tests are joint tests of relative predictive ability (Marcucci, 2005:12).

maturities ranging from one day to six months. The sample period is from 2 January 1980 until 29 August 2007 (Henry, 2009).

Examining equity returns, a constant mean and variance model is initially compared with a simple two-regime Markov-switching model. A Davies (1987) upper bound test confirms that the Markov-switching model provides a better fit of the data than the single-regime model (Henry, 2009). In addition, a Wald test of equal parameters across regimes indicates that the regimes are statistically distinct at all conventional levels of significance (Henry, 2009). Despite the improvement of the switching model over the single-state model, neither specification fully captures all ARCH effects (Henry, 2009).

In order to account for these ARCH effects, Henry (2009) estimates Nelson's (1991) EGARCH and a Markov-switching EGARCH (MS-EGARCH) model. Within the MS-EGARCH specification all the parameters are allowed to change with a change in regime, as in Gray (1996). A quasi-LRT statistic suggests that the MS-EGARCH model provides a better fit of the data than the single-regime version. Davies' (1987) upper-bound test confirms this finding, suggesting the superiority of the Markov-switching model.

The estimation results of the MS-EGARCH model show distinct behaviours in both regimes. The first regime (regime 0) exhibits high mean returns and low unconditional variance of returns. The second regime (regime 1) is characterised by low mean returns and higher unconditional variance of returns (Henry, 2009). The immediate impact (given by the ARCH term) of shocks is higher in regime 1, while the persistence (given by the GARCH term) of shocks is higher in regime 0. In regime 0 the leverage term is statistically insignificant, but is statistically different from zero and negative in regime 1. Thus in the low return, high variance regime, negative shocks have a greater impact on volatility than do positive shocks (Henry, 2009). Moreover, because the ARCH term in regime 1 is lower in absolute value than the leverage term, positive news decrease volatility during low return, high volatility periods (Henry, 2009).

Extending the MS-EGARCH model to allow for time-varying transition probabilities, Henry (2009) investigates whether short rates affect the probability of switching between regimes. Using Diebold and Mariano (1994) and Filardo's (1994) logistic functions to model probabilities as dependent on interest rates, the null hypothesis of a constant transition probability to the low return, high volatility regime is rejected for all loan maturities except one and twelve months. Henry (2009) states that firms who finance long term assets with short term loans will become increasingly risky as short rates increase, thus explaining the positive relationship between short rates and the probability of switching to regime 1. The MS-EGARCH model is thus shown to capture a number of important features of the conditional variance process, including the changing leverage effect and impact of short rates.

Babikir *et al.* (2010) investigate the presence of structural breaks and asymmetries in the volatility of the Johannesburg Stock Exchange All Share Index (JSE ALSI) from 2 July 1995 until 25 August 2010. Testing for structural breaks, Babikir *et al.* (2010) use both in-sample and out-of-sample tests. For their in-sample test, Inclan *et al.*'s (1994) ICSS algorithm for detecting multiple structural breaks is used. Structural breaks are detected at 10 December 2008 and 15 July 2009. Thus, three sub-periods can be identified. Babikir *et al.* (2010) then estimate three GARCH(1,1) models using the entire sample period and each sub-period as identified through the ICSS algorithm. The parameter estimates across sub-samples are dramatically different (Babikir *et al.*, 2010). During the pre-crisis period (prior to December 2008), shocks are highly persistent in the volatility generating process. Between December 2008 and July 2009, only the intercept term is statistically different from zero, indicating no persistence in the volatility generating process at all. Post July 2009, the persistence of shocks becomes evident once again (Babikir *et al.*, 2010). These results suggest that economically significant structural breaks occur over time in the stock return data considered. Particularly, high-volatility states may be associated with a lower degree of persistence in the conditional variance process than low-volatility states.

Babikir *et al.* (2010) test the robustness of their in-sample finding using out-of-sample forecasting tests. The measures of forecast accuracy used are an MSE test and VaR-based loss function at 1, 20, 60 and 120 week horizons. Three benchmark models – a GARCH(1,1), the RiskMetrics model for daily data, and Baillie *et al.*'s (1996) Fractionally Integrated GARCH (FI-GARCH(1,*d*,1)) model – are compared with five models that are sensitive to structural breaks. The competing models are a GARCH(1,1) using a 50% sample size rolling window, a GARCH(1,1) using a 25% sample size rolling window, a GARCH(1,1) using a weighted maximum likelihood procedure in which more recent observations are given greater weight, a GARCH(1,1) using data since the most recent structural break as identified through the use of the ICSS algorithm, and a simple moving average volatility model using a rolling 250 day window (Babikir *et al.*, 2010). According to the MSE measure, all the competing models outperform the benchmark models. Hansen's (2001) test for superior predictive ability and White's (2000) test for equal predictive ability both indicate that at least one of the competing models provides better forecasts than the benchmarks. Using the VaR-based criterion, Hansen (2001) and White (2000) tests again indicate that at least one of the competing models outperforms the benchmarks (Babikir *et al.*, 2010). Overall, these tests indicate that allowing for structural breaks improves volatility forecasts with respect to the ALSI.

In addition to investigating structural breaks within the data, Babikir *et al.* (2010) seek to determine whether the volatility of the ALSI exhibits leverage effects. In their out-of-sample tests, Babikir *et al.* (2010) include both an MS-GARCH(1,1) and GJR-GARCH(1,1) model to capture possible asymmetries. For both the MSE and VaR-based criteria, the MS-GARCH(1,1) outperforms the GJR-GARCH(1,1) at shorter horizons, whereas the GJR-GARCH(1,1) performs best at longer horizons. As in Marcucci (2005), this illustrates the sensitivity of forecasting tests to the forecast horizons used (in addition to the already established sensitivity to loss-functions used). In general, however, the models of asymmetry fail to outperform the GARCH(1,1), suggesting that the increased flexibility of models such as the MS-GARCH may hinder, rather than improve, forecast accuracy due to possible over parameterisation (Babikir *et al.*, 2010).

2.4. Markov-switching multivariate GARCH models

Markov-switching models have also been used successfully in a multivariate framework. For example, multivariate models of the conditional variance-covariance of returns have been used in order to compute time-varying optimal portfolio weights and hedge ratios or to assess contagion effects (Ramchand and Susmel, 1998; Edwards and Susmel, 2001; Li, 2009).

In one such study, Ramchand and Susmel (1998) employ Hamilton and Susmel's (1994) SWARCH methodology in order to examine the relationship between volatility and correlation of returns across stock markets. Using weekly stock returns from January 1980 to January 1990 for a number of major stock markets, Ramchand and Susmel (1998) use both univariate and bivariate SWARCH techniques to model variances and covariances, and assess the impact of changes in these variables on optimal portfolio weights.

Initially, a univariate SWARCH is estimated using the U.S. stock return series. A quasi-LRT favours the SWARCH(2,1) specification over its single-regime counterpart. The series is then divided into high volatility and low volatility periods using Hamilton and Susmel's (1994) classification that an observation belongs to state s_t if p_{st} (the smoothed probability of being in state s at time t) > 0.5. The correlation of U.S. stock returns with other major markets is then assessed for both high- and low-volatility states separately (Ramchand and Susmel, 1998). The results indicate that when the U.S. is in a high-volatility state, the foreign stock returns are generally more closely correlated to U.S. returns. Specifically, the correlations between U.S. and foreign stock returns are up to 2.63 times higher when the U.S. is in the high-volatility state than when it is in the low-volatility state (Ramchand and Susmel, 1998). If the cross-correlation of returns varies between regimes, the benefits from portfolio diversification will vary as well, with important implications for portfolio management strategies.

However, a shortcoming of the univariate study conducted is that it does not explicitly model the correlation coefficients and thus does not test formally for statistically significant differences in correlations across regimes (Ramchand and Susmel, 1998). Ramchand and Susmel (1998) use a bivariate SWARCH model to address this issue. Using the U.S. as the originator country[6], bivariate SWARCH models are estimated for Japan, the United Kingdom, Germany, and Canada. In all cases, the high volatility scaling parameters are significantly different from one, suggesting that the variance-covariance process is well modelled by a regime-switching system (Ramchand and Susmel, 1998). For the U.S.-U.K. and U.S.-Canada pairs, an LRT rejects the null hypothesis of equal correlation coefficients across regimes. However, for the U.S.-Japan and U.S.-Germany pairs, the test fails to reject the null. Ramchand and Susmel (1998) argue that the failure to reject the null for U.S.-Japan and U.S.-Germany may be due to the low number of observations in the high variance regime, and that the large (albeit statistically insignificant) differences in the correlation coefficients indicate that cross-correlations are in general both time- and state-varying.

To assess the potential benefit from using the MV-SWARCH as opposed to Bollerslev's (1990) CCC-GARCH model, Ramchand and Susmel (1998) compare the Sharpe ratios resulting from minimum variance portfolio weights as calculated using SWARCH and GARCH variances-covariances. The rationale underlying this comparison is that if state-varying variance-covariance estimates improve forecasts, then they should result in improved portfolio performance, as measured through Sharpe ratios (Ramchand and Susmel, 1998). To keep the model tractable, the study is restricted to a bivariate analysis using U.S. stock returns

[6] The 'originator country' is defined as the most likely originator of return and volatility spillovers (Ramchand and Susmel, 1998:406).

and an index consisting of European, Asian, and Far East stock returns (Ramchand and Susmel, 1998).

Ramchand and Susmel's (1998) portfolio optimisation results favour the SWARCH model over the GARCH. The mean Sharpe ratio for the SWARCH-based portfolio is 0.1858 with a standard deviation of 0.0844. In contrast, the mean Sharpe ratio for the GARCH-based portfolio is 0.1789 with a standard deviation of 0.0903. Although the Sharpe ratios are not found to be statistically different from each other, the more volatile portfolio weights associated with the GARCH portfolio would in practice worsen net returns further by increasing transaction costs; strengthening the view that the two models result in significant economic differences (Ramchand and Susmel, 1998). Given that the Sharpe ratios of both SWARCH and GARCH portfolios are similar during low variance periods, Ramchand and Susmel (1998) conclude that the advantage of using a state-varying model of variance-covariances materialises largely during periods of high volatility. During these high volatility periods (when the cross-correlation of returns are known to be higher), the SWARCH model is strongly biased in favour of the less volatile market, whereas the GARCH model is far less so (Ramchand and Susmel, 1998).

Analysing contagion effects between various developing markets, Edwards and Susmel (2001) employ the CCC MV-SWARCH framework of Ramchand and Susmel (1998). That is, they restrict their analysis to bivariate SWARCH models and allow the correlation coefficient to change with changes in the volatility state of the originator country. Using Mexico and Hong Kong as originator countries (see section 2.3.3, page 27 for the details of data used), Edwards and Susmel (2001) investigate whether stock return volatility regime-switches coincide, as well as whether the cross-correlation of returns significantly changes between regimes. In this way they are able to test the idea that following the Mexican and East Asian crises, other developing nations experienced contagion effects through lower returns and higher volatility. Edwards and Susmel's (2001) first test within the multivariate framework is an LRT of the independence of volatility states between countries. The independent states hypothesis is rejected for all the Latin American country pairs, but cannot be rejected for Hong Kong-Brazil and Hong Kong-Chile (Edwards and Susmel, 2001). Edwards and Susmel (2001) suggest that the results for the latter two pairs may reflect the fact that both Brazil and Chile had strong capital account controls for much of the period under study.

For those pairs in which the null hypothesis of independent states is rejected, a stronger test of volatility synchronisation is conducted (Edwards and Susmel, 2001). The aim of this test is to statistically determine whether both countries always experience high volatility states together. In all cases, the null hypothesis of high volatility synchronisation is rejected. Furthermore, the null hypothesis of low volatility synchronisation is rejected for all pairs except Brazil-Argentina (Edwards and Susmel, 2001). Thus, although the volatility states are seen to be somewhat dependent across markets, there is a sufficient degree of independence between markets to reject the full-synchronisation hypotheses.

Analysing cross-correlations of returns, Edwards and Susmel (2001) find that the correlation coefficient between Mexico and other Latin American countries increases by 200% to 400% during Mexico's high volatility periods. Using Hong-Kong as the originator country, cross-correlations are shown to be low, although they too exhibit a significant increase when Hong-Kong enters into a high volatility state (Edwards and Susmel, 2001).

Li (2009) conducts a methodologically similar study to Ramchand and Susmel (1998) in assessing the effect on portfolio performance of using multivariate SWARCH (MV-SWARCH), rather than MV-GARCH estimates of variance-covariances. Li (2009) argues that the results of Ramchand and Susmel's (1998) study may have been distorted since it only considered developed market stock returns, which are highly correlated to begin with. This could result in an understatement of the benefits of diversification. For this reason, Li (2009) uses emerging market (EM) as well as U.S. stock return data in assessing the alternative models. A second diversion from Ramchand and Susmel (1998), and other MV-SWARCH studies such as Edwards and Susmel (2003), is that Li (2009) allows the correlation coefficient within the MV-SWARCH framework to change with all four states, rather than simply with the two 'originator country' states. In this way, a richer characterisation of changes in cross-correlations may be captured.

The objectives of Li's (2009) study include determining the nature of the relationship between volatility regimes and cross-correlations; establishing the extent of differences in diversification benefits between volatility regimes; and evaluating whether the MV-SWARCH-based and MV-GARCH-based minimum variance portfolios result in statistically different Sharpe ratios. The data used are weekly MSCI stock index returns (in U.S. dollars) for the U.S., EM Asia, EM Europe, EM Europe Middle East, EM Far East, and EM Latin America. The sample period covers February 1988 to May 2007.

The specific models compared are Bollerslev's (1990) CCC-GARCH, Engle's (2002) DCC-GARCH, and Ramchand and Susmel's (1998) MV-SWARCH. Bivariate specifications are estimated for U.S.-EM pairs, using each EM series to avoid problems associated with *ad hoc* index selection. For all U.S.-EM combinations, the single-regime MV-GARCH models impart the typical high degree of persistence of shocks on volatility estimates. For the MV-SWARCH model, the second regime scaling parameter is significantly different from one at the 99% confidence level for all U.S.-EM pairs. This provides a high degree of confidence in the existence of a second regime, although formal tests of this are not presented. Under the MV-SWARCH specification, an economically significant result is that the correlation coefficients vary widely across regimes, with the highest degree of correlation occurring in the fourth volatility state (in which both the U.S. and EM series are experiencing high volatility). An LRT rejects the null hypothesis of a constant correlation coefficient in favour of the unrestricted model (Li, 2009).

Due, in part, to the lack of a straightforward statistical test, Li (2009) opts to compare the models based on their performance in constructing minimum variance portfolios. Assuming that investors aim to maximise an exponential utility function, minimum variance portfolio weights can be calculated as a function of the variance-covariance estimates, other variables held constant. Thus, following Ramchand and Susmel (1998), Li (2009) aims to establish the superior model through comparing the mean returns, mean variances and Sharpe ratios of resulting U.S.-EM portfolios.

Across all U.S.-EM pairs, and for all three models, the mean returns on the portfolios are not statistically different. In contrast, the variance of returns is significantly lower for the MV-SWARCH-based portfolio (in all cases) than for the other two portfolios (Li, 2009). In three of the five U.S.-EM pairs, the Sharpe ratio is higher for the MV-SWARCH-based portfolio than the other two portfolios. Thus, the results favour the use of MV-SWARCH models in estimating variance-covariances for portfolio construction, as the potential benefit to Sharpe ratios (through significantly lower variance of returns) tends to outweigh the potential cost (Li, 2009).

In a methodologically distinct multivariate study to those presented above, Lee and Yoder (2007) compare the hedging performance of their Markov-switching BEK-GARCH with that of Engle and Kroner's (1995) single-regime BEK-GARCH. The data used are continuously compounded weekly returns on corn and nickel spot and futures contracts from 2 January 1991 until 29 December 2004 (Lee and Yoder, 2007).

Using estimated variance-covariances, time-varying hedge ratios associated with each model are specified as:

$$\hat{\beta}_t^* = \hat{h}_{sf,t} \big/ \hat{h}_{f,t}^2, \qquad (10)$$

where $\hat{h}_{sf,t}$ is the estimated covariance between spot and futures returns and $\hat{h}_{f,t}^2$ is the variance of futures returns (Lee and Yoder, 2007). The hedging performance of each model is then evaluated based on the variance reduction achieved compared to an un-hedged portfolio. Specifically, White's (2000) Reality Check test is used to assess whether the Markov-switching specification statistically outperforms the single-regime BEK-GARCH (Lee and Yoder, 2007). The results suggest that the null hypothesis of no improvement of the single-regime specification cannot be rejected at conventional significance levels. This may be due to the over parameterisation of the Markov-switching BEK-GARCH: allowing for only two regimes and assuming a constant within-regime mean requires the estimation of 28 parameters (Lee and Yoder, 2007). The estimation difficulties associated with such highly parameterised models thus remains a challenge in the multivariate Markov-switching literature.

2.5. Summary and Conclusions

This chapter presented a review of various theoretical and empirical developments relating to financial volatility, contagion, and selected models of conditional variance, with specific emphasis on Markov-switching models. Since the seminal studies of Engle (1982) and Bollerslev (1986), it is common to model the conditional variance of financial time series as following a single-regime GARCH process. However, the Markov-switching models of Hamilton (1989, 1990), Hamilton and Susmel (1994), and Gray (1996), among others, provide a promising alternative characterisation of the conditional variance of returns. These Markov-switching specifications are particularly suited to modelling sample periods which contain occasional structural shifts in variance dynamics, such as those experienced by emerging markets (Aggerwal *et al*., 1999) or global markets during crisis periods (Ang and Timmerman, 2011).

In addition to the univariate framework, Markov-switching models may be used successfully in a multivariate framework in order to compute time-varying optimal portfolio weights and hedge ratios or to assess contagion effects (Ramchand and Susmel, 1998; Edwards and Susmel, 2001; Li, 2009).

Throughout the literature reviewed, it is commonly found that Markov-switching models can provide a better in-sample fit of the data or more accurate forecasts than the conventional single-state GARCH extensions (Hamilton and Susmel, 1994; Cai, 1994; Gray, 1996; Bollen *et al*., 2000; Haas, 2002; Klaassen, 2004; Wang and Theobald, 2008; Chen 2009). However, these results are on occasion statistically insignificant, and are less clear when assessed through some purely economic loss functions (Marcucci, 2005).

Given the evidence of occasional discrete shifts in the conditional variance process, it is vital to test for the presence of multiple regimes in the conditional variance when a reasonable suspicion exists for structural change. In light of the calm and turbulence of the global and African stock markets during the 2002-2012 period, Markov-switching models may prove to be a more appropriate characterisation of stock return volatility than the popular single-regime GARCH.

In this regard, chapters 3 and 4 are devoted to investigating whether a variety of Markov-switching models are better suited (than single-state GARCH models) to modelling the African stock market data used. In Chapter 3, the relevant models, in-sample goodness-of-fit tests, and out-of-sample forecast tests are presented. Chapter 4 presents the empirical results of the investigation in order to establish the suitability of Markov-switching models, as opposed to single-state GARCH alternatives, when applied to African stock returns.

CHAPTER 3: DATA AND ECONOMETRIC METHODS

3.1. Introduction

Before proceeding with the empirical study of Chapter 4, it is necessary to present and motivate the various data, models and tests used in the current analysis. Thus, drawing on the literature review of Chapter 2, this chapter aims to provide an overview of the data and econometric methods employed. The empirical study is separated into two parts, namely the univariate and multivariate analyses.

As a preliminary analysis, section 3.2 includes a descriptive summary of the stock return data used in this study, noting significant similarities and differences between markets. Section 3.3 details the univariate portion of the study. Sections 3.3.1 to 3.3.2 provide brief overviews of certain single-state and Markov-switching models of conditional variance, respectively. Section 3.3.3 discusses the various methods used to assess model performance, including both in-sample and out-of-sample measures. The multivariate portion of the study is presented in section 3.4. Sections 3.4.1 and 3.4.2, respectively, describe the various single-state and multi-state models used in the empirical analysis. Section 3.4.3 provides a justification for the chosen method of model selection (the 'portfolio optimisation' method). Section 3.5 summarises the chapter.

3.2. Data

In order to accurately compare the properties of each market, it is appropriate to ensure that the return series are compiled according to a standardised method. For this reason, the data used are daily returns on the Morgan Stanley Capital International (MSCI) standard country indices for South Africa, Kenya, Nigeria, Mauritius and Morocco. All the series are obtained from *Thomson Reuters Datastream*. Although available, Egypt was excluded from the analysis in order to avoid the problem of missing data associated with the closing of the Egyptian stock market during the Arab Spring. Tunisia has also been excluded despite availability, as observations for this market only begin in 2004.[7]

[7] The series were selected so as to balance the need for adequate cross-continental representation with the need to analyse the longest possible history of returns. Despite the exclusion of Egypt and Tunisia, the series include a sufficiently diverse set of markets (both in terms of geographical dispersion and GDP growth) for the purposes of this study.

Since the daily return observations for Kenya, Nigeria and Mauritius are available from 03/06/2002, the South African and Moroccan series are also taken from this date. This is primarily because it is mathematically necessary to employ a uniform sample size when conducting the multivariate analysis. Additionally, a consistent sample size will make parameter estimates comparable across markets when conducting the univariate analysis. Thus, the sample for each series spans 03/06/2002 to 01/06/2012, encompassing a total of 2610 daily observations.

Daily returns (r) are calculated as the first difference of the natural log of the index (P_t), multiplied by 100, such that:

$$r_t = 100 \times [\ln(P_t) - \ln(P_{t-1})]. \tag{11}$$

This transformation expresses daily returns in continuously compounded percentage terms. A graphical plot and the summary statistics of each series are presented in Table 1.

	South Africa	**Kenya**	**Nigeria**	**Mauritius**	**Morocco**
Mean	0.0419 (0.2666)	0.0671 (0.0187)	0.0286 (0.3538)	0.0704 (0.0042)	0.0365 (0.0989)
Std. Dev.	1.9256	1.4570	1.5732	1.2578	1.1312
Skewness	-0.3265 (0.0000)	-0.0740 (0.1231)	-0.0055 (0.9093)	0.2557 (0.0000)	-0.3069 (0.0000)
Excess Kurtosis	4.6333 (0.0000)	9.3295 (0.0000)	4.8814 (0.0000)	12.5071 (0.0000)	3.7001 (0.0000)
Jarque-Bera	2380.9871 (0.0000)	9467.8760 (0.0000)	2591.3647 (0.0000)	17040.1071 (0.0000)	1529.8540 (0.0000)
Max	12.3531 {29/10/2008}	9.9037 {31/10/2008}	11.4496 {17/02/2003}	10.8626 {01/04/2009}	5.4950 {13/10/2008}
Min	-13.5659 {16/10/2008}	-11.0232 {01/09/2003}	-10.2154 {04/11/2003}	-8.7257 {10/10/2008}	-7.6987 {05/01/2009}
LB (12)	33.9400 (0.0006)	299.6640 (0.0000)	424.7140 (0.0000)	66.4370 (0.0000)	113.7100 (0.0000)
LB^2 (12)	1824.2290 (0.0000)	863.5790 (0.0000)	605.7260 (0.0000)	540.2450 (0.0000)	498.9970 (0.0000)

Note: Values in round brackets denote p-values, while values in curly brackets denote dates.

Table 1: Summary statistics

A few noteworthy patterns emerge from the summary statistics. The mean daily return figures range between the statistically insignificant 0.03% for Nigeria and 0.07% for Mauritius (corresponding to annual returns of roughly 7.15% and 17.6%, respectively). The markets with third lowest and lowest mean returns, South Africa and Nigeria, also exhibit the largest *ex post* variance of returns. South Africa, Mauritius and Morocco show significant degrees of skewness, whereas Kenya and Nigeria do not. Furthermore, all of the series exhibit statistically significant excess kurtosis, explaining the rejection of the null hypothesis of normally distributed returns in each case (as shown by the Jarque-Bera statistic). This is unsurprising, as the rejection of normally distributed returns is in line with much of the empirical literature using high frequency financial data (e.g.: Dueker, 1997; Canarella and Pollard, 2007).

Considering extreme values, the global financial crisis had a clear impact on all of the markets studied. Excluding Nigeria, every series experiences either a maximum or minimum (or both) observation during the month of October 2008. Another common high-volatility period seems to be in 2003, in which both Kenya and Nigeria exhibit extreme values. According to the Ljung-Box *Q*-statistics, all of the returns series exhibit positive autocorrelation, which is a common finding within studies of emerging and frontier markets (Canarella and Pollard, 2007). Since the focus of this study is explicitly on the conditional variance and covariance of returns, an AR(1) model is used to capture mean returns for all the countries considered. This follows existing research (Lo and MacKinlay, 1990; Hamilton and Susmel, 1994; Ramchand and Susmel, 1998; Canarella and Pollard, 2007), who note that for stock returns, an AR(1) model is usually sufficient. In addition to the autocorrelation found in the returns, each market exhibits positive autocorrelation in the squared returns, suggesting the presence of ARCH effects in the data. Interestingly, the South African market seems to be associated with the weakest autocorrelation in returns but the strongest autocorrelation in squared returns of all the markets analysed.

From the graphical plot of returns in Figure 1 (page 43), it appears that besides the occasional idiosyncratic shock, most series display a prolonged period of low volatility from 2004 to 2007. Following the collapse of Lehman Brothers in 2008, each series exhibits a rapid movement to a high volatility state. Despite the extent of this increase in volatility, many of the markets studied experienced a decline in volatility to pre-crisis levels by 2010. An exception to this decline is South Africa, which shows a more prolonged period in the high volatility state, perhaps due to its relatively extensive trade linkages with the European Union (Edwards *et al.,* 2009).

Given the observed changes in stock return volatility for each market considered, it is necessary to explore appropriate methods for modelling this phenomenon. Section 3.3 presents various univariate techniques for doing so.

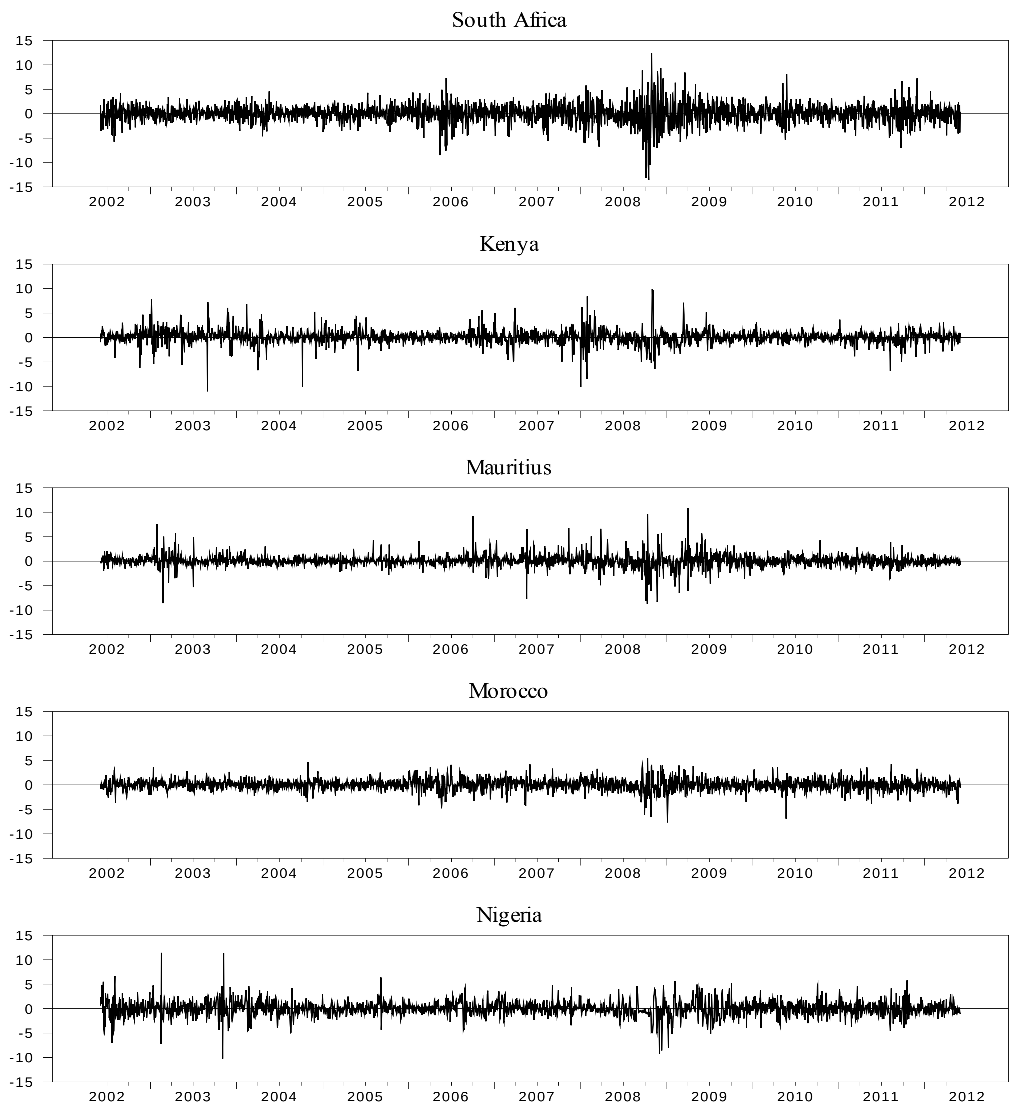

Figure 1: Continuously compounded daily returns (%)

3.3. Univariate models of conditional variance

3.3.1. Single-regime models

While many econometric models operate under the assumption of a constant error term variance, it has been widely recognised (and observed in section 3.1) that financial time series often exhibit significant heteroscedasticity (Engle, 1993; Ang and Timmerman, 2011). Some market practitioners have dealt with this phenomenon through the use of simple moving average estimates of the conditional variance (Engle, 1993:72). Such models of the conditional variance (h_t) of the error term (ε_t) may be characterised by:

$$\varepsilon_t = \upsilon_t \sqrt{h_t}, \tag{12}$$

$$h_t = \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \tag{13}$$

where υ_t is white noise and the weights α_i and lag length (q) are determined on an *ad hoc* basis. While the simplicity of this method may be appealing, the arbitrariness of weightings and lag length are obvious shortcomings.

The central positions that volatility estimates and forecasts take in finance (for example, in portfolio construction, risk-management, and option pricing) require that the best available model of volatility be used at all times. Thus, most research into the measurement of financial market volatility incorporates estimates from the more sophisticated ARCH family of models (Engle, 2004:407).

Engle (1982) originally developed the ARCH model in a test of Friedman's (1977) conjecture that uncertainty about future inflation causes business cycle fluctuations. This conjecture requires that inflation volatility changes over time, a property which the ARCH model is able to capture (Engle, 2003:406). The ARCH model of conditional variance is specified as a constant plus a weighted average of q past squared innovations, such that:

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2. \tag{14}$$

This ARCH (q) specification appears superficially to be equivalent to traditional moving average estimates of volatility. However, the major advance made by Engle (1982) is that the unconditional variance (determined through ω) and weights attached to innovations are determined via maximum likelihood estimation, using information contained in past data (Engle, 1982:990). Furthermore, lag lengths can be chosen using LRTs, residual diagnostics, and relevant information criteria. Thus, the ARCH model is able to capture many important features of financial volatility in a non-arbitrary manner.

In practice, the use of the Engle's (1982) ARCH specification can result in the inclusion of a large number of lags (Bollerslev, 1986:308). For this reason, Bollerslev (1986) proposes a more parsimonious representation of the conditional variance, in which the 'long memory' property of squared returns, for example, can be captured through the inclusion of lagged conditional variance terms (Bollerslev, 1986:309). Specifically, Bollerslev's (1986) GARCH (p,q) models the current conditional variance as a function of q past innovations and p lagged conditional variance terms, such that:

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}, \tag{15}$$

where $\omega \geq 0$, $\alpha_i \geq 0$ and $\beta_j \geq 0$ for all i and j. Intuitively, the GARCH(1,1) forecast of conditional variance at time t is a weighted average of three components: a constant term through which the unconditional variance is determined, the previous period's estimate of the conditional variance, and the new information obtained during the period t-1 (Engle, 2004:407).

There have been a large number of extensions[8] to Bollerslev's (1986) GARCH model (Engle, 2004). However, despite the extensive work done on more flexible specifications, the GARCH(1,1) specification is often an adequate representation of the conditional variance process, and is a common starting point in empirical studies (Engle, 2004). Hansen and Lunde (2001) use the superior predictive ability tests of White (2000) and Hansen (2001) to illustrate that none of the popular volatility models (of 330 models tested) significantly outperform the GARCH(1,1) in predicting exchange rate and stock market volatility. That is, the GARCH(1,1) produces forecasts which are at least as accurate as the more flexible and highly parameterised alternative models. The lack of an obvious outperformer amongst the single-state models tested by Hansen and Lunde (2001) illustrates the robustness of the GARCH(1,1) specification. As such, this study employs the GARCH(1,1) model as a suitable single-state benchmark against which the Markov-switching models are tested.

[8] Popular extensions include Glosten, Jaganathan and Runkle's (1994) Threshold GARCH and Nelson's (1990) Exponential GARCH (commonly used in the presence of leverage effects) and Ding *et al*.'s (1993) more general Power GARCH (which, in addition to capturing leverage effects, is able to model different degrees of nonlinearity in the conditional moments of the data).

3.3.2. Markov-switching models

GARCH-type models, when applied to high-frequency financial series, often impart a high level of persistence of shocks on the future level of conditional variance (Engle and Bollerslev, 1986). As mentioned in the introductory chapter, Engle and Bollerslev (1986) and Nelson (1990) have argued that constraining the sum of the coefficients of a GARCH(1,1) model to equal one can result in a parsimonious and reasonably accurate model of the conditional variance of asset returns. This specification is referred to as the integrated GARCH (I-GARCH) model, and effectively models the conditional variance as following a unit root process, so that the unconditional variance is infinite (Engle and Bollerslev, 1986).

Although the I-GARCH specification has been shown to model a wide range of financial series well, it is difficult to justify the existence of a unit root in the conditional variance process (Lamoureax and Lastrapes, 1990; Cai, 1994). In addition, it is well established that financial markets are prone to sudden and large shocks. This is particularly evident during crisis periods such as the global financial crises of 1987 and 2007 (Ang and Timmerman, 2011). Lamoureux and Lastrapes (1990), among others, show that in the presence of discrete structural breaks in the variance process, traditional GARCH models may impose a spuriously high level of persistence of shocks on future conditional variances. This is analogous to Perron's (1989) finding that structural breaks in the level of a series can result in spuriously high persistence in the mean. Hence the single-regime GARCH models presented in sub-section 2.2.1, with their commonly high degree of autoregressive persistence, are unable to adequately capture the effects of such financial crises on the conditional variance (Lamoureax and Lastrapes, 1990; Schwert, 1990; Nelson, 1991; Marcucci, 2005). When using sample periods covering financial crises or sharp changes in volatility, therefore, it is imperative that structural breaks are accounted for in the data.

One way of accounting for structural changes is to exogenously assign appropriate dummy variables to the data, so that crisis periods may be effectively governed by different ARCH or GARCH processes. However, this simple method may suffer from biases associated with *ad hoc* selection of crisis periods. Another method is to let the data speak for itself, by endogenously identifying the structural breaks using Inclan and Tiao's (1994) ICSS algorithm (*cf.* Duncan and Liu, 2009). Although this is an improvement on the *ad hoc* specification, it is not clear how this method may be used for forecasting. Markov-switching models, on the other hand, provide both endogenous identification of structural breaks as well as an established forecasting method based on a latent first-order Markov chain (Hamilton, 1989).

Hamilton (1989, 1990) developed a constant within-regime mean and variance Markov-switching model (herein referred to as the simple MS model), which has been subsequently used in one form or another within many empirical studies of financial time series (Garcia and Perron, 1996; Dahlquist and Gray, 2000; Bollen *et al.*, 2000; Knedlik and Scheufele, 2008; Wang and Theobald, 2008; Chen, 2009; Duncan and Liu, 2009). The most general version of the simple MS model is:

$$y_t = \mu_{s_t} + \varepsilon_t \tag{16}$$

$$\varepsilon_t \sim D(0, h_{s_t}), \tag{17}$$

so that $y_t \sim D(\mu_{s_t}, h_{s_t})$, where $\mu_{s_t}$is the state-dependent mean of y_t, and D represents the conditional distribution of y_t. As is often the case in practice, the mean equation can be augmented to accommodate autoregressive terms, and the innovations ε_t can be drawn from the normal, Student's *t*, or G.E.D. distributions, for example. Since the focus of this study is on switching conditional variances, the simple MS model used in Chapter 4 entails an AR(1) mean (with no switching) and a switching conditional variance. In this way the estimated, or 'working', conditional variance is a weighted average of the state-specific conditional variances, where the weights are given by the filter probabilities of each state (Hamilton, 1994). Although the simple MS model is comparatively straightforward, it is sometimes able to capture important features of financial time series data, such as heteroscedastic errors (*cf.* Cai, 1994).

For instances in which heteroscedastic errors are not fully captured by the simple MS model, Hamilton and Susmel (1994) and Cai (1994) have developed uniquely specified SWARCH models. Hamilton and Susmel (1994) note that, despite the high persistence of shocks imputed by single-regime GARCH models, their forecasting performance is often poor. For example, Hamilton and Susmel's (1994) GJR-GARCH(1,1) model of stock market volatility implies a decay parameter close to one, indicating highly persistent shocks. This level of persistence is difficult to reconcile with the GJR-GARCH(1,1) model's poor forecasting performance relative to a constant-variance specification (Hamilton and Susmel, 1994). In light of these issues, Hamilton and Susmel (1994) develop a SWARCH model in which a standardised ARCH process is scaled according to the current state, s_t. This allows for discrete jumps in volatility, consistent with many empirical observations of financial time series (Hamilton and Susmel, 1994). Throughout their empirical study, Hamilton and Susmel (1994) use a single-regime, first-order autoregressive specification for the mean equation. The residuals are then modelled as following a SWARCH process. A simple version of this SWARCH(K, q) model for the conditional variance of residuals is:

$$y_t = a_{0s_t} + \sum_{i=1}^{k} a_i y_{t-i} + \varepsilon_t \tag{18}$$

$$\varepsilon_t = \sqrt{\gamma_{s_t}} \times u_t \tag{19}$$

$$u_t = \sqrt{h_t} \times \upsilon_t \tag{20}$$

$$h_t = \omega + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \cdots + \alpha_q u_{t-q}^2 \ , \tag{21}$$

where y_t represents stock returns at time t, ε_t is the observed residual, γ_{st} is the scaling parameter associated with state s_t ($s_t = 1,\ldots,K$), u_t is the de-scaled error term, and υ_t follows a white noise process with unit variance. In maximisation of the log-likelihood function, the scaling parameter γ_{1t} is unidentified and hence set equal to 1. This allows γ_{st} (for $s_t \neq 1$) to be interpreted as the ratio of state s_t variance to state 1 variance (Hamilton and Susmel, 1994).[9]

The model, as specified in equations (18) to (21), has the restriction that the density of y_t can only depend on a finite number of ARCH terms, and hence s_t (Hamilton and Susmel, 1994). This is essentially due to the recursive nature and hence path dependence of a GARCH term: h_t depends directly on h_{t-1}, which in turn depends directly on h_{t-2}, and so on. In this way the conditional variance at time t depends on the entire sequence of states $s_0,\ldots\ldots\ldots,s_t$. Maximum likelihood estimation requires that all possible state paths be integrated out, so that at the tth step there are K^t components of the likelihood function, thereby becoming an extremely large number for most conventionally sized samples (Gray, 1996). As such, Hamilton and Susmel (1994) restrict their analysis to low order ARCH processes within each regime. Although this may seem overly restrictive (the volatility persistence conventionally associated with the GARCH term is lost), it is not clear *a priori* that this will diminish the performance of the model. In particular, the volatility persistence normally associated with

[9] A similar model to Hamilton and Susmel's (1994) SWARCH is presented in Cai (1994). In Cai's (1994) specification, discrete jumps in the conditional variance are realised through differences in the intercept term (ω). This is in contrast to Hamilton and Susmel (1994) who model the discrete jumps in stock returns as changes in the scale of the entire ARCH process. Due to the essential similarity between these models, only Hamilton and Susmel's (1994) specification is used in the following empirical study.

the GARCH term may be better captured through the persistence of regimes (Hamilton and Susmel, 1994).

Despite the apparent estimation difficulties noted above, Gray (1996) developed a widely used technique to avoid the problem of path dependence. Using the fact that, in Gray's (1996) study, the conditional density of changes in interest rates is a mixture of distributions with time-varying mixing parameters (given by the probability of each state occurring at time t) the working conditional variance can be estimated as:

$$h_t = E[\Delta r_t^2 | I_{t-1}] - E[\Delta r_t | I_{t-1}]^2 \tag{22}$$

$$= p_{1,t}(\mu_{1,t}^2 + h_{1,t}) + (1 - p_{1,t})(\mu_{2,t}^2 + h_{2,t}) - [p_{1,t}\mu_{1,t} + (1 - p_{1,t})\mu_{2,t}]^2, \tag{23}$$

and the error term by:

$$\varepsilon_t = \Delta r_t - E[\Delta r_t | I_{t-1}] \tag{24}$$

$$= \Delta r_t - [p_{1,t}\mu_{1,t} + (1 - p_{1,t})\mu_{2,t} , \tag{25}$$

where $\mu_{s,t}$ is the mean for regime s at t and $p_{s,t}$ is the filter probability that the series is in regime s at t. The new conditional variance term h_t is no longer path dependent, as it is calculated through aggregating the conditional variances from each regime at each step (Gray, 1996). The working conditional variance can then be used in constructing both $h_{1,t+1}$ and $h_{2,t+1}$:

$$h_{s_{t+1}} = \omega_{s_{t+1}} + \alpha_{s_{t+1}}\varepsilon_t^2 + \beta_{s_{t+1}}h_t, \tag{26}$$

which can be used as inputs in maximising the log-likelihood function.

Although originally applied to interest rates, Gray's (1996) MS-GARCH has been used in studies of exchange rate movements (*cf.* Klaassen, 2002; Haas *et al.*, 2004) and stock returns (*cf.* Dueker, 1997; Marcucci, 2005; Babikir *et al.*, 2010). However, when applied to the African stock market data used in this study, the maximum likelihood parameter estimates associated with Gray's (1996) MS-GARCH are extremely sensitive to initial guess values, and the maximum likelihood procedure, more often than not, fails to converge. In light of the instability of parameter estimates associated with Gray's (1996) model, Dueker's (1997) extension of Hamilton and Susmel's (1994) SWARCH (which makes extensive use of the above convergence procedure) is preferred in order to capture both regime-persistence and autoregressive persistence in the conditional variance process.

In a study of stock market volatility, Dueker (1997) develops four MS-GARCH models.[10] For each of the MS-GARCH specifications tested, Dueker (1997) assumes a Student's t error distribution. One of the specifications developed simply includes a GARCH term in Hamilton and Susmel's (1994) SWARCH framework (see equations 18-21). The GARCH term is incorporated through the use of the integration technique of Gray (1996), but with more efficient use of available information, as in Klaassen (2002). Effectively, equation (21) is modified by Dueker (1997) to include a GARCH term, such that:

$$y_t = a_{0s_t} + \sum_{i=1}^{k} a_i y_{t-i} + \varepsilon_t \qquad (27)$$

$$\varepsilon_t = \sqrt{\gamma_{s_t}} \times u_t \qquad (28)$$

$$u_t = \sqrt{h_t} \times \upsilon_t \qquad (29)$$

$$h_t = \omega + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \ . \qquad (30)$$

Building, in a natural way, on the SWARCH model of Hamilton and Susmel (1994) and producing more robust and stable parameter estimates than Gray's (1996) MS-GARCH, Dueker's (1997) MS-GARCH (hereafter simply referred to as the MS-GARCH(*K,p,q*) model) is therefore used in the following empirical study.

Due to the high-frequency of the data used in this study, each of the models selected will be restricted to incorporate constant (as opposed to time-varying) transition probabilities. Given that much of the fundamental economic data typically used to express time-varying transition probabilities (such as inflation and GDP growth rates) is available at monthly or higher frequencies, this restriction is considered necessary. Additionally, the restriction is helpful in maintaining tractability and hence convergence of the maximum likelihood procedure.

3.3.3. Performance assessment

3.3.3.1 In-sample performance

[10] These include GARCH extensions to Hamilton and Susmel's (1994) and Cai's (1994) SWARCH models, and two models in which the shape parameter and degree of kurtosis, respectively, are allowed to switch between regimes.

The univariate models to be tested in Chapter 4 include the GARCH(1,1), simple MS model, SWARCH and MS-GARCH. Since many of the conditional variance models presented are nested, it is possible to assess, by comparison, their in-sample fit through the use of LRTs (*cf.* Hamilton and Susmel, 1994; Gray, 1996; Bollen *et al*., 2000; Klaassen, 2002; Edwards and Susmel, 2003; Canarella and Pollard, 2007). In these cases, the LR statistic is:

$$LR_d = -2(L_0 - L_1) \, , \tag{31}$$

where L_i is the maximum log-likelihood value associated with model i and d is the number of restrictions under the null hypothesis (H_0: Model 0 fits the data as well as Model 1). When testing two models with the same number of states, LR_d asymptotically follows the χ^2 distribution.

If the models do not have an equal number of regimes (for example, when comparing a single-state GARCH(1,1) to an MS-GARCH(2,1,1)), the LR statistic is no longer distributed χ^2 (Garcia, 1997; Bollen *et al.*, 2000). This is because the parameters associated with the extra regime are in these cases unidentified under the null hypothesis, thereby violating the regularity conditions required for a χ^2-distributed LR statistic (Bollen *et al.*, 2000). Garcia (1997) derives appropriate critical values for a number of simple Markov-switching models, none of which are included in the empirical section of this study (due to their inclusion of switching means). Hansen (1992) suggests an LRT procedure which entails a number of optimisations over a 'grid' of the relevant unidentified parameters. Because this procedure is extremely computationally burdensome for all but the simplest of Markov-switching models, some studies omit its results in favour of the conventional LRT (a quasi-LRT). Hamilton and Susmel (1994), for example, regard the LR statistic as a useful descriptive summary of model performance.

Despite the usefulness of the conventional LRT in assessing the in-sample fit of models with equal number of states, a more formal test of model performance is necessary for the purposes of this study when comparing models with differing numbers of states. Davies (1987) offers a relatively less burdensome, closed-form LRT. The Davies (1987) upper-bound test adjusts the standard LR significance level to provide an upper-bound on the actual level of significance. Specifically, an upper-bound on the actual level of significance for the LRT statistic (LR) is calculated as:

$$\Pr^*\left[\chi_q^2 > LR\right] = \frac{\Pr[\chi_q^2 > LR] + 2(LR/2)^{q/2}\exp(-\frac{LR}{2})}{\Gamma(\frac{q}{2})}, \tag{32}$$

where q is the number of nuisance parameters and $\Gamma(\cdot)$ is the *gamma* function. Since the adjusted probability $\Pr^*[\chi_q^2 > M]$ is strictly greater than $\Pr[\chi_q^2 > M]$, the Davies (1987) upper-bound test need only be conducted if the null is rejected using the quasi-LRT (Wang and Theobald, 2008). Although Davies' (1987) test provides an upper-bound on the actual significance level, and is therefore somewhat conservative (Garcia, 1996:764), it provides a more formal comparison of one versus two-state and two- versus three-state models. Importantly, greater confidence can be taken in a rejection of the null hypothesis once the *p*-value has been adjusted in this way.

Since the focus of this study is on the nature of stock return volatility, it is imperative that the models used are able to capture fully the heteroscedasticity present in the data. Much of the work examining models of conditional variance make use of the Ljung-Box *Q*-test in this regard (*cf.* Hamilton and Susmel, 1994; Gray, 1996; Susmel, 1999; Bollen *et al.*, 2000; Klaassen, 2002; Edwards and Susmel, 2003; Canarella and Pollard, 2007). In particular, the squared standardised residuals from each model are tested for any remaining autocorrelation, the existence of which signals the presence of excess heteroscedasticity. Thus, following Canarella and Pollard (2007), in this study the models are tested by specifically making use of the *Q*-statistic at 12 lags (autocorrelation found only at higher lags is assumed to be spurious). The results of the LRTs, quasi-LRTs, Davies' (1987) tests, and the Ljung-Box *Q*-tests are then taken together as an indication of the in-sample goodness-of-fit of each model considered.

3.3.3.2 Forecasting performance

In addition to the in-sample tests presented in the preceding sub-section, forecasting tests are widely used in assessing the overall performance of univariate Markov-switching models of conditional variance (*cf.* Hamilton and Susmel, 1994; Gray, 1996; Susmel, 1999; Klaassen, 2002; Haas, 2004; Marcucci, 2005). This stems from the difficulty in comparing the in-sample performance of models that are not strictly nested, and the violation of regularity conditions when comparing, for example, switching to single-state models.

In order to conduct a forecast analysis for the purposes of this study, the sample was shortened by 115 observations.[11] A rolling estimation of each model was then conducted using a constant in-sample window size of 2 495 observations. At each estimation step, 1- to 15-day forecasts of the conditional variance were estimated. This procedure resulted in 100 sets of 1- to 15-day forecasts being obtained. These were then compared to the observed daily squared residuals, according to the particular loss functions chosen. Although it is accepted that the daily squared residuals can be a very noisy estimate of actual volatility (Anderson and Bollerslev, 1998), the data are not available at a higher frequency.[12] Furthermore, although noisy, the daily squared residuals remain an unbiased estimate of the actual volatility observed, and are thus adequate for comparative purposes.

Some studies of forecasting performance are restricted to a simple comparison of forecast errors, represented by selected loss functions (Hamilton and Susmel, 1994; Susmel, 1999). In a preliminary analysis of forecasts conducted in the following chapter, mean absolute forecast errors (MAE) and mean squared forecast errors (MSE) are used as loss functions in order to gauge the suitability of the various model specifications. The MSE and MAE figures are calculated as:

$$MSE = \frac{1}{T}\sum_{t=1}^{T}\{\hat{\varepsilon}_t^2 - h_t\}^2 \quad (33)$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|\hat{\varepsilon}_t^2 - h_t| \quad (34)$$

where T is the number of observations in the forecasting window, $\hat{\varepsilon}_t^2$ is the observed squared residual, and h_t is the forecasted conditional variance. The MSE loss function is included because of its prevalence in the forecasting performance literature (Clark and West, 2007:3). Although the MSE measure can sometimes be used in isolation when forecasting data in levels, it is not clear that the fourth moment of the data in question always exists. Hence, rather than take the square of a square (as in the case of the MSE) as the only measure of forecasting performance, this study includes the MAE, which does not exhibit the same problem. The MAE measure has the added advantage of neutrality (or equal weighting) between larger and smaller errors, whereas the MSE measure adds more weight to larger errors – an assumption better left to the practitioner.

[11] The relatively high computational burden associated with the rolling estimation of SWARCH and MS-GARCH models restricted the plausible size of the forecasting window.
[12] This would be required in order to construct Anderson and Bollerslev's (1998) measure of 'realised' volatility.

More formal tests of forecasting performance are conducted by, among others, Klaasen (2002), Haas (2004) and Marcucci (2005). To formally measure the adequacy of forecasts, Marcucci (2005) makes use of the Diebold and Mariano (1995) test for equal predictive ability, as well as White's (2000) and Hansen's (2001) tests for superior predictive ability. The latter two tests are preferred by Marcucci (2005) on the basis that they test for the existence of a superior model out of a set of alternatives rather than the more tedious pairwise test of Diebold and Mariano (1995). While this point is valid when comparing a larger set of alternative models, the forecasting tests of Chapter 4 are confined to four alternatives, rendering the Diebold-Mariano (1995) test adequate for the current study. The Diebold-Mariano (1995) test statistic is obtained via a two-step procedure. Initially, the mean loss (*d*) is calculated as:

$$d = \frac{1}{T}\sum_{t=1}^{T}\{f(e_t^1) - f(e_t^2)\}, \tag{35}$$

where e_t^1 is the forecast error for model 1 at time *t*, and $f(x)$ is the relevant loss function. The Diebold-Mariano (1995) test statistic is then defined as:

$$DM = d/\sqrt{\sigma_d^2/(T-1)}, \tag{36}$$

where σ_d^2 is the sample variance of *d*. This DM statistic, under certain conditions, follows a *t*-distribution. When comparing the forecasting performance of the GARCH(1,1) model to those of the simple MS and SWARCH models in the following chapter, the Diebold-Mariano (1995) test, utilising absolute forecast errors as the relevant loss function, is used.

While the Diebold-Mariano (1995) test, as presented so far, is suitable for comparisons between the non-nested models, the statistic no longer has a *t*-distribution when the competing models are nested (Clark and McCracken, 2001). Under these conditions, a more suitable test of forecasting performance is provided by Clark and West (2007). Clark and West's (2007) test is thus used to compare the forecasting performance of the GARCH(1,1), simple MS and SWARCH(2,1) against the MS-GARCH(2,1,1) (which nests the other three models). Moreover, the test is also used to compare the performance of the simple MS model against that of the SWARCH(2,1). The Clark and West (2007) test adjusts the squared forecast error of the larger model due to any inherent parameter estimation error. Specifically, a new series (q_t) can be constructed as:

$$q_t = (e_t^1)^2 - \{(e_t^2)^2 - (h_t^1 - h_t^2)^2\}, \tag{37}$$

where e_t^i is the forecast error and h_t^i denotes the actual forecasts (model 1 is nested within model 2). In order to obtain a test statistic, the q_t series is then regressed on a constant. If the *t*-statistic associated with the estimated constant is significantly greater than zero, the null hypothesis of equal predictive ability can be rejected in favour of the larger model (Clark and West, 2007).

Taken together, the in-sample and forecasting tests of sub-sections 3.3.3.1 and 3.3.3.2 are used to gauge the relative performance of each model, and are thus able to provide an indication as to whether the Markov-switching univariate models are capable of improving on the conventional single-state GARCH.

3.4. Multivariate models of conditional variance

3.4.1. Single-regime MV-GARCH models

It is widely observed that the volatility of many financial series move together across markets (Edwards and Susmel, 2001; Bauwens *et al.*, 2006). This apparent interdependence has resulted in the development of the multivariate GARCH (MV-GARCH) framework, under which a variety of interactions between conditional variances and covariances may be specified. The variance-covariance matrix (an output of MV-GARCH models) can be used as an important input in assessing contagion effects, computing time-varying hedge ratios, pricing assets through the Capital Asset Pricing Model, or constructing minimum variance portfolios (Bauwens *et al.*, 2006).

In an attempt to capture dynamic linkages between series, Bollerslev (1990) developed the constant conditional correlation (CCC) model. In this formulation, the conditional variance of each series is estimated as in the univariate case. The covariances between the series are then estimated as a constant proportion (ρ_{ij}) of the corresponding product of conditional standard deviations (Bauwens, 2006). Considering a bivariate CCC model, the variance-covariance matrix is:

$$H_t = \begin{bmatrix} h_{11t} & \rho_{12}(h_{11t}h_{22t})^{0.5} \\ \rho_{12}(h_{11t}h_{22t})^{0.5} & h_{22t} \end{bmatrix}. \tag{38}$$

The CCC specification significantly reduces the number of parameters to be estimated – in the bivariate case, only one extra parameter (ρ_{ij}) is required. Through reducing greatly the number of parameters to be estimated, the CCC model offers a greater possibility of convergence during the maximisation procedure (Bauwens, 2006). The ease of estimation and the flexibility of the variance structure have thus ensured that the CCC and related models remain popular in high-dimensional empirical work (Engle, 2004).

While the ease of estimation is an attractive feature of the CCC model, Tsui and Yu (1999) have shown that the null hypothesis of a constant correlation coefficient can be rejected for many financial series. A more flexible specification of market co-movements therefore may be necessary in order to capture this feature of the data.[13] Engle and Kroner (1995) propose a specification in which all the coefficients enter the model in quadratic form. Specifically, the BEK(1,1,1) bivariate model is given as:

$$H_t = W'W + A'\tilde{\varepsilon}_{t-1}\tilde{\varepsilon}'_{t-1}A + B'H_{t-1}B, \tag{39}$$

$$W = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \tag{40}$$

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \tag{41}$$

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \tag{42}$$

where W, A and B are matrices of coefficients to be estimated, $\tilde{\varepsilon}_t$ is a vector of error terms, and H_t is a positive definite matrix of variance-covariances.[14] Compared to the CCC model, the BEK specification contains a relatively large number of parameters for any study involving more than two series. However, the ability of the BEK model to capture a range of market interdependencies (for example, the lagged conditional variances of each series enter into the covariance equations, allowing for the volatility of one series to directly impact its co-movement with another), makes it an attractive alternative, especially for low-dimensional analyses.

[13] In this respect, Engle's (2002) DCC-GARCH model (in which the correlation coefficient is explicitly time-varying) was also considered for estimation. However, the maximum likelihood procedure failed to converge for most of the countries considered. The DCC model is thus excluded from this analysis.

[14] It is the positive definiteness of H_t that results in the BEK model being preferred over other models with flexible covariance structures, such as Bollerslev *et al.*'s (1988) VEC and diagonal VEC models.

Although the time-varying nature of cross-correlation of returns can be captured by a number of multivariate models, this study makes use of the CCC and BEK models in particular. The CCC model serves as a benchmark with which to test the null hypothesis of no time-variation whatsoever in the cross-correlation of returns. The BEK model is used to provide a comparison between a more flexible single-state model and the regime-switching alternative. These models thus serve as single-state representatives for the purposes of this study.

3.4.2. Markov-switching MV-GARCH models

Hamilton and Lin (1996) develop the multivariate SWARCH (MV-SWARCH) model in their analysis of the relationship between output growth and stock return volatility. Following the work of Hamilton and Lin (1996), MV-SWARCH models have been typically restricted to fit certain assumptions. Specifically, given that MV-SWARCH models tend to be computationally burdensome, it is often assumed that the individual series experience only two volatility regimes, and the studies are commonly restricted to bivariate analysis (Ramchand and Susmel, 1998; Edwards and Susmel, 2003; Li, 2009). This simplified specification thus includes four 'primitive' states (s_t*) such that:

$s_t* = 1$: Series $i – s_t = 1$; Series $j – s_t = 1$ (43)

$s_t* = 2$: Series $i – s_t = 1$; Series $j – s_t = 2$ (44)

$s_t* = 3$: Series $i – s_t = 2$; Series $j – s_t = 1$ (45)

$s_t* = 4$: Series $i – s_t = 2$; Series $j – s_t = 2$. (46)

In this case, the number of primitive states (K*) is four, and the variance-covariance matrix is hence generated by a four-state latent Markov chain. Inclusion of more states or a higher dimensional analysis results in rapid growth in parameters to be estimated, rendering maximum likelihood convergence unlikely.

Ramchand and Susmel (1998) develop a MV-SWARCH model in which the mean equation is constant across regimes. In vector form, the system is represented by:

$$\tilde{y}_t = A + B\tilde{y}_{t-1} + \tilde{\varepsilon}_t \quad (47)$$

$$\tilde{\varepsilon}_t \sim BD(0, H_t) \text{ ,} \quad (48)$$

where $\tilde{\varepsilon}_t$ follows a bivariate distribution (*BD*) and the time-varying conditional covariance matrix H_t is characterised by the elements:

$$\frac{h_{xx,t}}{\gamma_{x,s_t}} = \omega_x + \sum_{k=1}^{q} \alpha_{x,k} \frac{\varepsilon_{x,t-k}^2}{\gamma_{x,s_{t-k}}}; \qquad x = \{i,j\} \tag{49}$$

$$h_{ij,t} = \rho_{i,s_t}(h_{ii,t}h_{jj,t})^{0.5} \tag{50}$$

The above model is thus a regime-switching generalisation of Bollerslev's (1990) CCC model. It is common practice within the MV-SWARCH framework to assume that the correlations (ρ_{xst}) in $s_t* = 1$ and $s_t* = 2$ are equal, and likewise for $s_t* = 3$ and $s_t* = 4$ (Ramchand and Susmel, 1998; Edwards and Susmel, 2003; Li, 2009). This restricts the correlation coefficient to vary with the state of the 'originator' country series (*Series i*) and facilitates convergence during the maximisation procedure through maintaining tractability of the system (Ramchand and Susmel, 1998). Although derived from and somewhat similar to the CCC specification, the MV-SWARCH model is not as restrictive as its single-state counterpart. Rather than assuming a constant correlation coefficient, the potential for time-variation in the cross-correlation between series is captured through changing filter probabilities through time. In this way, a regime-switching characterisation of the conditional covariance of returns may also be superior to other single-state alternatives such as the DCC-GARCH or BEK-GARCH models.

In addition to the flexibility offered in the correlation coefficient, the MV-SWARCH allows for statistical testing of different assumptions regarding the interaction of variance states between countries. For example, if the variance states of series *I* and *J* are independent, then $p_{13}^* = p_{12}^i p_{11}^j$. Alternatively, if the variance states are fully dependent, and thus always equal, $p_{13}^* = 0$ (Ramchand and Susmel, 1998). The MV-SWARCH thus offers a significant deal of flexibility and opportunities to test various hypotheses regarding the nature of variance-covariance states and the cross-correlation of returns.

Beyond the restricted MV-SWARCH model of Hamilton and Lin (1996), there have been a number of extensions to the multivariate Markov-switching conditional variance-covaraince framework. Lee and Yoder (2007) and Haas and Mittnik (2008) develop Markov-switching versions the BEK and VEC models, respectively. Although the Markov-switching BEK and VEC models allow for both autoregressive and regime persistence in the covariance of returns, they are very highly parameterised and computationally intensive. For example, Lee and Yoder's (2007) Markov-switching BEK model requires (in the bivariate case, assuming a constant within regime mean) the estimation of at least 28 parameters. Because of the computational burden associated with the switching BEK and the resulting difficulty in obtaining convergence, the empirical study of Chapter 4 utilizes only the MV-SWARCH model as a Markov-switching alternative to the single-state CCC-GARCH and BEK-GARCH specifications. As noted in section 3.3.2, it is not clear *a priori* that the persistence associated with regimes is an inferior characterisation of the data than the traditional autoregressive persistence of some single-state models.

3.4.3. Performance assessment: Portfolio optimisation

Because the multivariate models used in this study are not nested, LRTs are an inappropriate test of comparative performance. Following previous research (Ramchand and Susmel, 1998; Lee and Yoder, 2007; Li, 2009), these models are thus assessed according to their ability to aid in the construction of optimal portfolios. The logic for such a test is that the multivariate models that provide the most accurate variance-covariance estimates should improve the performance of portfolios formed based on these estimates.

Consistent with Ramchand and Susmel (1998) and Li (2009), the bivariate models are each used to construct optimal 2-element portfolios for a 'typical' risk-averse investor. The assumed utility function of this investor is the familiar exponential utility function:

$$U(W) = -\exp\{-\lambda\left(\frac{W}{W_0}\right)\}, \qquad (51)$$

where W is expected wealth, W_0 is current wealth, and λ is the coefficient of absolute risk aversion (assumed to be 3).[15] For a given set of portfolio weights, w_t, associated with a vector of mean returns, μ_t, and a covariance matrix, Σ_t, this implies an expected utility of:

$$E[U(w_t)] = -\exp\{-\lambda(w_t\mu_t - \frac{\lambda w_t{}'\Sigma w_t}{2}) \qquad (52)$$

Optimal portfolio weights, represented by the 2x1 vector w_t^*, for this particular investor can then be obtained through:

$$\mu_t = \lambda w_t^{*\prime}\Sigma_t \qquad (53)$$

[15] The relative performance of each model is insensitive to this figure (Ramchand and Susmel, 1998; Li, 2009).

which, when rearranged, yields:

$$w_t^* = \left(\frac{1}{\lambda}\right)[\Sigma_t^{-1}\{\mu_t - (i'\Sigma_t^{-1}i)^{-1}(\mu_t'\Sigma_t^{-1}i - \lambda)i\}], \tag{54}$$

where i is a 2x1 vector of ones (Ramchand and Susmel, 1998).

Using the calculated weights associated with each model, the portfolio (and hence the underlying model) that exhibits the highest returns and lowest variances are preferred over the alternatives. To test for differences between portfolio returns and variances, Li (2009) devises two test statistics, z_1 and z_2, which are used in the multivariate analysis of Chapter 4:

$$z_1 = \sqrt{T} \times \frac{(\frac{1}{T})\times\sum_{t=1}^{T} r_t^{i-j}}{SD(r_t^{i-j})} \tag{55}$$

$$z_2 = \sqrt{T} \times \frac{(\frac{1}{T})\times\sum_{t=1}^{T}(r_t^{i-j}\times r_t^{i+j})}{SD(r_t^{i-j}\times r_t^{i+j})}, \tag{56}$$

where:

$$r_t^{i-j} = r_t^i - r_t^j \tag{57}$$

$$r_t^{i+j} = r_t^i + r_t^j, \tag{58}$$

and r_t^i is the observed return at time t for portfolio i. The test statistic z_1 tests whether the portfolio returns are significantly different, and z_2 tests for differences between the variances. In each case, the test statistic asymptotically follows the standard normal distribution. For z_1, if the calculated z-statistic is greater than the appropriate critical value, model i produces a significantly higher return than model j. Likewise, for z_2, if the calculated z-statistic is greater than the appropriate critical value, model i produces a significantly higher variance than model j.

Because many investors are not concerned with portfolio returns or variances in isolation, the Sharpe ratios, which provide a measure of risk-adjusted returns, for each portfolio are used to provide an overall assessment of the adequacy of each model (Ramchand and Susmel, 1998; Li, 2009). Assuming a zero risk-free rate, the Sharpe ratio for portfolio i can be defined as:

$$SR^i = \frac{1/_T \sum_{t=1}^{T} r_t^i}{SD(r_t^i)} \tag{59}$$

Those portfolios, and hence underlying models, that exhibit the highest Sharpe ratios are preferred over the alternatives.

3.5. Summary

To account for the changing volatility evident in all the markets analysed, various econometric models of conditional variance were considered. In the univariate portion of this chapter, a variety of single-state models are presented. Due to its robustness in many empirical studies, the GARCH(1,1) is taken as a suitable single-state benchmark against which the simple MS, SWARCH and MS-GARCH multi-state models will be compared. The models will be tested based on their in-sample fit (measured through the use of LRTs, quasi-LRTs, Davies' (1987) tests and residual diagnostics) and their out-of-sample forecasting performance (measured by MAEs, MSEs, the Diebold-Mariano (1995) test, and Clark and West's (2007) test).

In the multivariate section of the chapter, a number of popular single-state models are considered. The CCC-GARCH and BEK-GARCH models are selected as representative single-state models against which the MV-SWARCH will be tested. The MV-SWARCH model allows for a state-varying correlation coefficient, unlike the CCC-GARCH (constant correlation coefficient) and BEK-GARCH (time-varying correlation coefficient). The portfolio optimisation test, in which the variance-covariance estimates from each model are used to construct optimal minimum-variance portfolios, is presented as the primary method by which the models will be assessed. Using this test, the model producing the highest portfolio Sharpe ratios will be considered superior to the others and, by extension, a more appropriate characterisation of market inter-linkages.

CHAPTER 4: EMPIRICAL RESULTS

4.1. Introduction

Given the clear historical periods of calm and turbulence in stock markets, a Markov-switching model, in which stock market volatility may pass through more than one state, is a plausible alternative characterisation of the conditional variance to the conventional single-state GARCH models. The rapid increase in stock market volatility associated with the global financial crisis, and the tendency for developing markets to exhibit discrete structural shifts (Aggarwal *et al*., 1999), creates a particularly strong case for including endogenously identified volatility states in a model of African stock market variability over the past decade. If a multi-state characterisation of stock market variability is appropriate, Markov-switching models should provide better in-sample and out-of sample estimates than those which are restricted to one regime. In this chapter, therefore, a number of univariate and multivariate Markov-switching models are presented and tested against the selected benchmark single-state models.

In section 4.2, univariate models of volatility such as the GARCH(1,1), simple MS, SWARCH and MS-GARCH models are compared, with the aim of identifying the most appropriate characterisation of the conditional variance process for the African markets selected. These models, and some extensions of them, are evaluated based on the economic implications of parameter estimates, LRTs, the Davies (1987) upper-bound test (where appropriate), residual diagnostics, and forecast errors.

In section 4.3, three multivariate models are evaluated. Specifically, estimates of the bivariate MV-SWARCH, CCC-GARCH, and BEK-GARCH models are presented. Each model is then compared based on the economic implications of parameter estimates and a portfolio optimisation test. Section 4.4 provides a brief summary of the chapter and concludes.

4.2. Univariate models

4.2.1. The time-path of volatility states

Figures 2-6 show, for each country examined, the continuously compounded daily returns series and the smoothed probability series, p_{1t}, obtained through the estimation of the SWARCH(2,1) model with Student's *t* errors. Since there are only two states in this instance, p_{2t} can be derived by subtracting p_{1t} from 1. Graphically, p_{2t} can be seen as the smoothed probability figure, turned upside-down. Although the smoothed probabilities are derived from the SWARCH(2,1) estimates specifically, they are qualitatively very similar to those obtained using a number of other Markov-switching models (such as the simple MS and more highly parameterised SWARCH models), and hence provide a sufficient overview of the general result presented below.

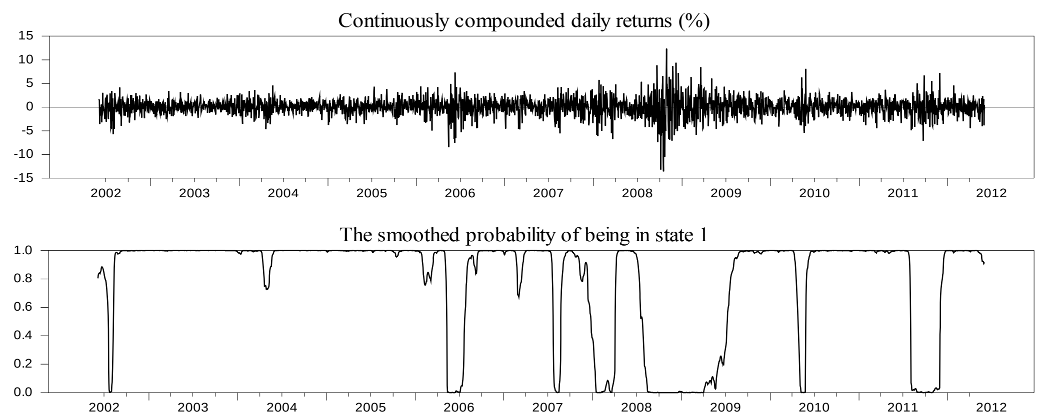

Figure 2: Time-path of volatility states (South Africa)

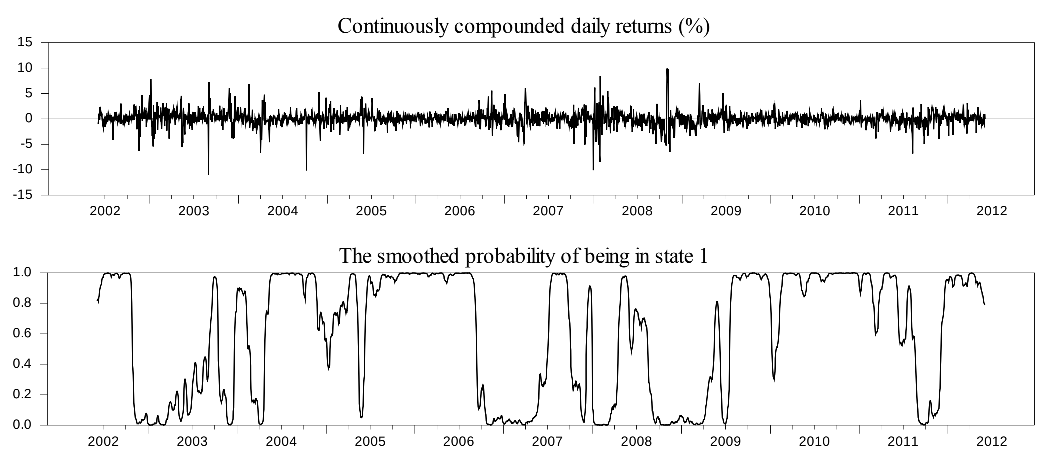

Figure 3: Time-path of volatility states (Kenya)

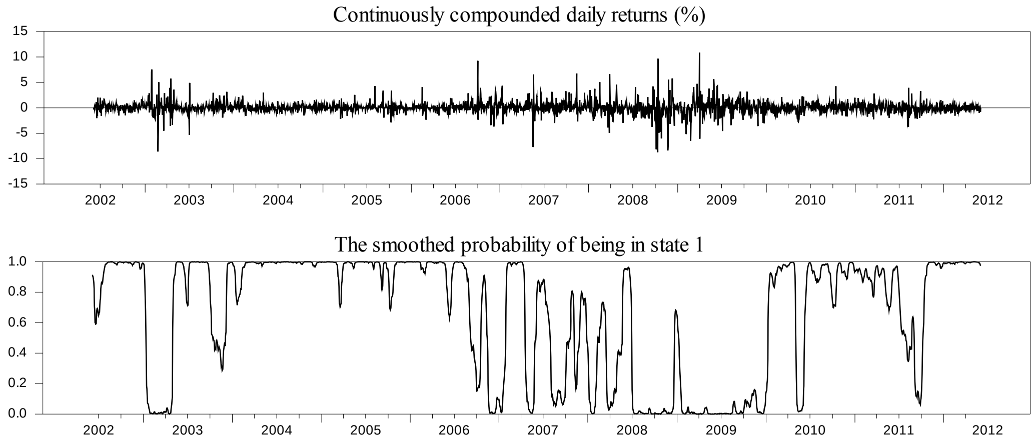

Figure 4: Time-path of volatility states (Mauritius)

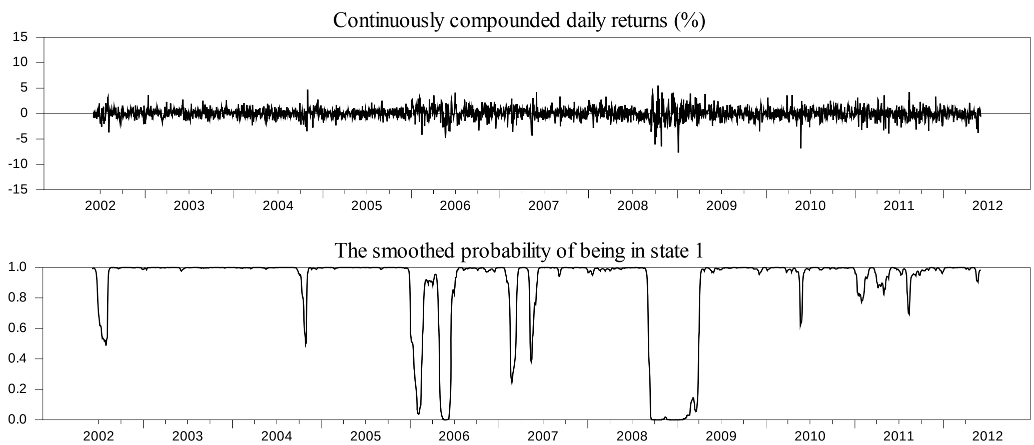

Figure 5: Time-path of volatility states (Morocco)

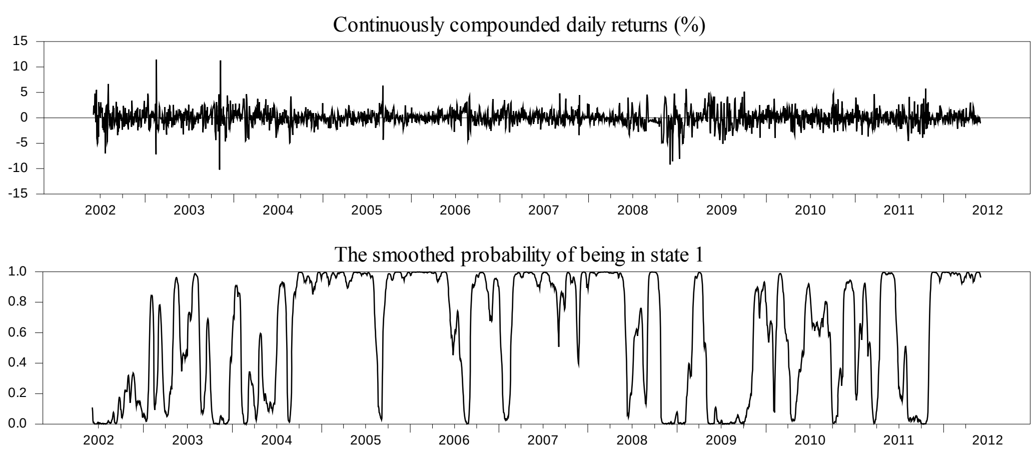

Figure 6: Time-path of volatility states (Nigeria)

Using the rule-of-thumb proposed by Hamilton (1989), an observation at time t is classified as belonging to state 1 if $p_{1t} > 0.5$. As seen in figures 2-6, South Africa and Morocco begin the sample period in a prolonged period of low volatility until the period 2006 to 2007, in which the smoothed probability switches a number of times between the low- and high-volatility states. Contrastingly, Kenya, Mauritius, and Nigeria begin the sample period in a somewhat unsettled state, switching often between low- and high-volatility states between 2002 and the end of 2004. Following this uncertain period, these three countries enter into the low-volatility state for the majority of the period until 2007. As a result of the contagion experienced in the wake of the developed market financial crisis, each series enters into a high-volatility state between the third and fourth quarter of 2008. South Africa, Kenya, and Morocco exit the high-volatility state after roughly two quarters, remaining in the low-volatility state for much of the following two years. Mauritius and Nigeria experience more prolonged periods in the high-volatility state, post-crisis, and Nigeria in particular remains in an uncertain state for the remainder of the sample. While Morocco remains in the low-volatility state towards the end of the sample, all of the other countries exhibit a sharp switch into the high-volatility state during the third quarter of 2011, accompanied by a sharp exit out of this state after roughly one quarter.

The joint pattern that emerges is one of generally low volatility leading up to the global financial crisis, followed by two clear yet relatively short-lived periods in the high-volatility state.[16] This pattern is consistent with the preliminary data analysis in Chapter 3. The lower persistence associated with the periods of high volatility is also consistent with much of the empirical literature, and is one of the features of stock market data which supports the use of conditional variance models with more than one state (Susmel, 1999).

4.2.2. Error distribution and leverage effects

For all the models estimated, specifications incorporating Student's t errors are strictly preferred to those assuming normally distributed errors. As illustrated in Tables B1-B4 of Appendix B, LRTs are used to test the null of Gaussian errors against the alternative of Student's t distributed errors for all countries examined. In each case, the null of normally distributed errors is rejected at the 1% level of significance for the GARCH(1,1), simple MS,

[16] This pattern is followed, to a lesser extent, by Kenya and Nigeria, who display a number of idiosyncratic state transitions as well.

SWARCH(2,1), and MS-GARCH(2,1,1) models. This is a common finding in studies using high-frequency financial data, and is due largely to the excess kurtosis found in daily stock returns (Dueker, 1997). For this reason, the remainder of the empirical study focuses exclusively on those specifications using Student's *t* errors, unless otherwise stated.

A somewhat unique feature of the African stock market used is the lack of a clear 'leverage' effect (in which volatility increases more in response to negative news than positive news). As shown in Appendix B, Tables B6 and B10, including a leverage term in both the GARCH(1,1) and SWARCH(2,1) models does not, in general, improve their fit. Specifically, for all the countries considered except South Africa, the null hypothesis of no asymmetry cannot be rejected according to an LRT. In light of this result, the models considered in the remainder of this chapter do not include a leverage effect term.

4.2.3. Univariate parameter estimates and their economic meaning

As noted in Chapter 1, Lamoureaux and Lastrapes (1990), among others, show that when applied to financial data, single-state GARCH models typically exhibit a high degree of persistence in the conditional variance. They have argued that the reason for the commonly high persistence in these models may be the existence of discrete structural breaks in the conditional variance series.

Using the African stock market data available, this pattern of high persistence has been repeated. Referring to the GARCH(1,1) coefficient estimates (Table A9, Appendix A), the degree of persistence is measured by the sum $\alpha_1 + \beta_1$. For all countries, the sum of these coefficients is relatively high, the lowest being 0.9463 for Nigeria and the highest 1.086 for Mauritius. In fact, for Kenya and Mauritius, an LRT indicates a failure to reject the null hypothesis that the conditional variance of returns follows an I-GARCH process. For South Africa and Nigeria, the null can be rejected at the 5% level, while it can be rejected at the 1% level for Morocco. Given the lack of a strong theoretical reason for such high autoregressive persistence in the conditional variance, the preceding result provides sufficient motivation to investigate the existence of more than one volatility state.

The most basic Markov-switching model of the conditional variance presented in this study is the simple MS model (Table A1, Appendix A). This model characterises the conditional variance as alternating between two constants, associated with a low-volatility state and a high-volatility state respectively. The simple MS estimation yields similar results across the various countries. The high-volatility regime typically has a conditional variance that is between five times (Morocco and South Africa) and ten times (Kenya) that of the low-volatility regime, suggesting that the two regimes exhibit economically distinct behaviour. More formal tests of the existence of the second regime (and higher) are deferred to the next sub-section.

Based on the estimated transition probabilities, both states are highly persistent for most countries. South Africa experiences the most persistent states[17] (the low-volatility state is expected to last for about 213 trading days on average, and the high-volatility state about 54 trading days), with Kenya showing the least persistent states (29 days in low-volatility and 8 days in high-volatility). The level of persistence for Mauritius, Morocco and Nigeria lie in between these extremes, but generally closer to the Kenyan case than that of South Africa.

Referring to Table A2 in Appendix A, the SWARCH(2,1) model parameters tell a somewhat similar story to that of the simple MS model. The SWARCH(2,1) model characterises the conditional variance as following an ARCH process that is scaled upwards by a scaling parameter, γ_2, when the series enters into the high-volatility state. For Kenya, Mauritius, Morocco, and Nigeria, the ARCH term is significant at the 1% level, while for South Africa, it is statistically insignificant. The scaling parameters range between 4.27 for Kenya and 6.10 for Mauritius and, according to a Wald test, are all significantly different from 1, thus strongly suggesting (although not formally proving) the existence of a second regime for all countries. Unsurprisingly, the low-volatility regime is more persistent than the high-volatility regime for each country. Based on the transition probabilities, the low-volatility regimes are expected to last between 263 (Morocco) and 48 (Nigeria) trading days, while the high-volatility regimes are expected to last between 57 (South Africa) and 31 (Morocco) trading days. The ergodic (or unconditional) probability of being in state 1 ranges between 59.73% (Nigeria) and 89.30% (Morocco), while the minority of time is spent in state 2. This is consistent with the differing level of persistence observed and the pattern of smoothed probabilities illustrated in Figures 2-6 above.

Including a GARCH term in the Markov-switching framework allows for an alternative source of volatility persistence. Estimation of Dueker's (1997) MS-GARCH(2,1,1) thus leads

[17] This could be due to the greater market capitalisation, liquidity, or relative institutional efficiency of the Johannesburg Securities Exchange.

to a different set of conclusions with respect to some of the countries studied (Table A3, Appendix A). For instance, although the scaling parameters are similar in magnitude to those of the SWARCH(2,1) model, a *t*-test shows that for South Africa and Morocco, they are not significantly different from 1. Clearly, the inclusion of a GARCH term has, for these two countries, negated the explanatory power of regime persistence in favour of autoregressive persistence. In fact, the GARCH term is significant at the 1% level for all the countries analysed. For Kenya, Mauritius and Nigeria, however, the scaling parameters remain significantly greater than 1, in spite of an alternative source of persistence. Strengthening the case for a one-state model, some of the transition probabilities estimated are not significant at conventional levels. Specifically, for South Africa and Nigeria, the transition probability *P(1,2)* is not significant, suggesting that the second regime is not well defined.

These results indicate that a two-state characterisation of the conditional variance of returns may be useful in the case of Kenya, Mauritius, and perhaps Nigeria, but may be inappropriate for South Africa and Morocco (although, once again, this is not a formal test of the existence of a second regime, which is deferred to the next sub-section).

4.2.4 Model selection

4.2.4.1. In-sample goodness of fit

In this sub-section, various univariate models are tested based on their likelihood ratios, as well as their ability to capture fully the existing heteroscedasticity in the data. The actual test statistics and/or their associated *p*-values are presented in Appendix B.

As a preliminary test for the existence of more than one regime, a basic AR(1) with a constant variance is compared to the alternative simple MS model. In this case, the only structural difference between the two models is that the latter expresses the 'working' conditional variance as a weighted average of two state-specific conditional variance estimates.

Since the second state-specific conditional variance is unidentified under the null, the standard LRT statistic is no longer distributed χ^2. Nonetheless, a quasi-LRT indicates that, for all countries, the AR(1) can be rejected at the 1% level of significance. Additionally, a Davies

(1987) upper-bound test confirms the findings of the quasi-LRT, and provides more conclusive evidence in favour of the regime-switching specification.

Having established the superior fit of the simple MS over the AR(1) model, it is necessary to test whether the simple MS model fully captures the observed heteroscedasticity in the data. Taking the squared standardised residuals for each country, a Ljung-Box Q test indicates some residual heteroscedasticity in most cases. Specifically, Q-statistics at 12 lags are found to be significant at the 1% level for Nigeria, Mauritius and South Africa, and at the 5% level for Morocco. Notably, there is no residual heteroscedasticity found in the case of Kenya, indicating that a simple MS model may adequately capture Kenya's stock market volatility process.

Given that the simple MS model fails to fully capture the heteroscedasticity for most of the countries analysed, the following step is to examine whether a model that allows for some degree of autoregressive behaviour in the squared errors is able to do so. The SWARCH(2,1) and MS-GARCH(2,1,1) models provide natural extensions to the simple MS model, allowing for the existence of both regime-switching and conventional (G)ARCH dynamics. Since the persistence attributable to the lagged conditional variance term in a GARCH(1,1) model may be more appropriately captured by the persistence of different states in a Markov-switching framework, the next test considers the SWARCH(2,1) against the simple MS model. Following this, the existence of GARCH effects (over and above regime persistence) is tested by comparing both the simple MS and SWARCH(2,1) models with the MS-GARCH(2,1,1).

Considering the inclusion of one lagged squared residual term, an LRT indicates that for Kenya, Mauritius, and Morocco, the simple MS model is rejected in favour of the SWARCH(2,1) at the 1% level of significance, while for Nigeria it is rejected at the 5% level. The SWARCH(2,1) does not provide a statistically significant improvement in the log-likelihood value for South Africa, however, as the LRT results in a failure to reject the null. Analysing the squared standardised residuals, the SWARCH(2,1) model is able to more fully capture the heteroscedasticity of returns than the simple MS model. Specifically, Q-statistics at 12 lags are insignificant at conventional levels for Kenya, Mauritius and Morocco. For South Africa and Nigeria, however, there is remaining autocorrelation in the squared standardised residuals.

Since none of the models thus far have been able to account across all countries for the total extent of heteroscedasticity, the MS-GARCH(2,1,1) model is tested as a plausible alternative characterisation of the conditional variance. An LRT indicates that the MS-GARCH(2,1,1) is strongly favoured against both the simple MS and the SWARCH(2,1) models. In each comparison, the LR statistic is significant at the 1% level for all the countries. In addition, estimation of the MS-GARCH(2,1,1) model results in no remaining heteroscedasticity for all the series. Specifically, the *Q*-statistic at 12 lags (for the squared standardised residuals) is not significant at conventional levels for any of the countries considered, indicating that the MS-GARCH(2,1,1) fits the data well in all cases.

Because the strongest result obtained thus far seems to be associated with the inclusion of a GARCH term, it is appropriate to test the in-sample performance of the simpler, single-regime GARCH(1,1) model against the alternative MS-GARCH(2,1,1). Both the GARCH(1,1) and MS-GARCH(2,1,1) models are able to account adequately for the autocorrelation in the squared returns. Analysing the squared standardised residuals resulting from each model, both models perform well (as indicated by the insignificant *Q*-statistics at 12 lags of all the series tested). A more telling distinction is made between the models when referring to the LRT results.

In this case, a Davies (1987) upper-bound test is necessary in addition to the quasi-LRT, because the parameters associated with the second state are unidentified under the null hypothesis of a GARCH(1,1) conditional variance. Ignoring the problem of the nuisance parameters, a quasi-LRT leads to the rejection of the null hypothesis at the 1% level of significance for Kenya, Mauritius, Morocco and Nigeria. Thus, for these countries, the MS-GARCH(2,1,1) provides a better in-sample fit than the GARCH(1,1) model. In contrast, for South Africa the test indicates a failure to reject the null hypothesis. In addition, the Davies (1987) test confirms these preliminary findings. Specifically, according to the Davies (1987) test, the adjusted *p*-value is 0.7389 in the case of South Africa, and less than 0.01 for all other countries. Although the Davies (1987) test provides an upper-bound for the actual *p*-value of the test statistic and is hence somewhat conservative, the relatively extreme values obtained here provide a reasonable level of confidence in the results. Accordingly, with the exception of South Africa, the inclusion of more than one state in the conditional variance process improves the in-sample fit of the models.

Despite the outperformance of the MS-GARCH(2,1,1) when compared to the GARCH(1,1), simple MS model, and SWARCH(2,1), it remains plausible that more flexible specifications of these models may not be outperformed to the same extent, if at all. For example, adding an additional lag to the SWARCH(2,1) specification (SWARCH(2,2)) improves the log-likelihood value at the 1% level for all the countries.[18] Moreover, when comparing the SWARCH(2,2) to a SWARCH(2,4) specification, an LRT results in the rejection of the SWARCH(2,2) at the 1% significance level for South Africa, Kenya and Nigeria. Although including additional lags generally improves the SWARCH(2,1) model's in-sample fit, the MS-GARCH(2,1,1) remains the better fitting model, at least in terms of LR statistics.

Including additional lags is not the only way to increase flexibility. Adding another state to the SWARCH(2,1) model (SWARCH(3,1)) also provides a better in-sample fit. A quasi-LRT and Davies (1987) upper-bound test indicates that the two-state version can be rejected in favour of three states at the 1% level for all the countries considered.

It is clear that due to the multitude of potential model specifications available, it is not a straightforward task to select the single best performer. In fact, much of the literature addressing Markov-switching volatility models go beyond in-sample tests in order to assess overall performance. Because of the computational difficulty associated with comparing models with different states, forecasting performance can be adopted as an additional and more formal measure of model adequacy (Bollen *et al.*, 2000:253). As such, the four basic models assessed so far (GARCH(1,1), simple MS model, SWARCH(2,1), and MS-GARCH(2,1,1)) will be subject to forecast accuracy tests in the following section.

4.2.4.2. Forecasting performance

In order to test the forecasting performance of each model for a particular country, this study employs the use of both MAE and MSE measures.[19] Tables C1-C5 in Appendix C contain MAE and MSE estimates for each country at each forecast horizon.

For South Africa, the results are strikingly consistent. Using both the MSE and MAE measures, the simple MS model provides the most accurate forecasts at all horizons, followed closely by the SWARCH(2,1) model. The MS-GARCH(2,1,1) and GARCH(1,1) provide the least accurate forecasts, with the former performing the worst. In terms of improving the MAE measure, the simple MS model performs 8.48% better than the MS-GARCH(2,1,1) at the 1-day horizon, and 5.58% better at the 3-week horizon.

---

[18] See Appendix B, Tables B1-B16, for the actual LR statistics.
[19] See section 3.3.3.2 for the specific methodology used.

With respect to Kenya, the results are more mixed. At the 1-day horizon, the MAE indicates that the SWARCH(2,1) model performs the best, followed by the GARCH(1,1) and MS-GARCH(2,1,1) models, with the simple MS faring the worst. By contrast, the MSE suggests that the simple MS model performs the best at the 1-day horizon, with the GARCH(1,1) and SWARCH(2,1) performing slightly worse, and the MS-GARCH(2,1,1) performing worst overall. At the 1-week horizon both measures indicate that the simple MS model performs the best, followed by the SWARCH(2,1), GARCH(1,1), and MS-GARCH(2,1,1). At the 3-week horizon, the top two performers remain the same, while the MS-GARCH(2,1,1) marginally outperforms the GARCH(1,1). When measured by the MAE, the simple MS model performs 2.84% worse than the GARCH(1,1) at the 1-day horizon. However, at the 1- and 3-week horizons, the simple MS model improves the MAE by 19.10% and 48.91%, respectively, over the GARCH(1,1) forecasts. For Kenya, therefore, the simple MS specification is thus clearly the preferred model by this measure.

In the case of Mauritius, the simple MS and SWARCH(2,1) models continue to dominate the GARCH(1,1) and MS-GARCH(2,1,1). At the 1-day horizon, the SWARCH(2,1) model performs the best overall. The MAE improvement over the worst performing model (the MS-GARCH) is 27.59%. At the 1- and 3-week horizons, the simple MS model performs the best, followed by the SWARCH(2,1), the GARCH(1,1), and the MS-GARCH(2,1,1) models. At the 3-week horizon, the simple MS model improves the MAE by 80.63% over the GARCH(1,1) and 84.69% over the MS-GARCH(2,1,1).

Considering Morocco, the GARCH(1,1) model outperforms all other models at the 1-day and 1-week horizons, using both the MAE and MSE measures. The other three models produce similar results at these horizons, with the GARCH(1,1) improving the MAE by only 6.06% over the worst model at 1-week. At the 3-week horizon, the GARCH(1,1), simple MS, and SWARCH(2,1) models produce very similar results by both measures, with the MS-GARCH(2,1,1) clearly performing the worst. On balance, the GARCH(1,1) model is relatively more successful using the Moroccan series than for the other countries examined.

For Nigeria, the MS-GARCH(2,1,1) model outperforms all the others using both measures of forecast accuracy at each horizon. This constitutes the best relative performance of the MS-GARCH(2,1,1) model for all of the countries studied. The GARCH(1,1) model performs second-best at the 1-day horizon, but falls to worst at the 1- and 3-week horizons. Conversely, when measured by the MAE, the simple MS model performs worst at the 1-day horizon but rises to second-best at the longer horizons. The MS-GARCH(2,1,1) improves the MAE over the worst performing model by 8.41% at 1 day, 10.24% at 1 week, and 21.64% at 3 weeks.

Overall the forecasting results are somewhat mixed, but tend to favour the models without a GARCH term, especially at the longer forecast horizons. The GARCH(1,1) model shows a comparative advantage at the 1-day horizon, but performs better than the simple MS and SWARCH(2,1) models only for Morocco. Except for Nigeria, the MS-GARCH(2,1,1) model is the poorest performer. This may be due to over-parameterisation, as the MS-GARCH(2,1,1) is the only model tested that allows for two sources of persistence. For South Africa, Kenya, and Mauritius, the simple MS model performs the best at the two longer forecast horizons. In these cases, it is also the SWARCH(2,1) that consistently performs the second-best out of all the models tested. The forecasting outperformance of the simpler Markov-switching models over the GARCH and MS-GARCH models points to a potential benefit to be derived from a different characterisation of volatility persistence.

Although the evidence in favour of the simpler Markov-switching models appears strong, no formal statistical test of forecasting performance has been conducted. A simple *F*-test for differences in the MSE figures is inappropriate if the condition of contemporaneously uncorrelated forecast errors, among other conditions, is violated. Since this is highly likely to be the case,[20] the Diebold and Mariano (1995) test of equal predictive ability is deemed more appropriate.

In order to avoid the problems of *ad hoc* assumption-making, the Diebold-Mariano (1995) test makes use of absolute forecast errors as the relevant loss function. Under this test, the simple MS and SWARCH(2,1) models can be formally compared to the single-state GARCH(1,1), which acts as the benchmark. Under the null hypothesis, the pair of models tested has equal predictive ability. A highly positive statistic indicates that the Markov-switching model produces better forecasts, whereas a highly negative statistic indicates that the GARCH(1,1) is the better forecaster. The test statistics are presented in Tables C6 and C7 in Appendix C.

According to the Diebold-Mariano (1995) test, the simple MS model, which performed well under the preliminary analysis, continues to do so relative to the GARCH(1,1), particularly at

[20] Since many of the models are nested, their forecasts (and the changes in their forecasts through time) tend to be similar.

the longer forecast horizons. For South Africa, the simple MS is shown to produce better 1-day and 1-week forecasts at the 1% level of significance. The 3-week forecasts are also shown to be superior, but only at the 10% level. Referring to both the Kenyan and Mauritian data, the null hypothesis of equal predictive ability for the 1-day horizon cannot be rejected at any conventional significance level. However, as is consistent with the analysis so far, the 1- and 3-week forecasts of the simple MS model are superior to the GARCH(1,1) at the 1% significance level. In the Moroccan case, the GARCH(1,1) model outperforms the simple MS model at all horizons (this is the only instance in which the GARCH(1,1) is clearly preferred). For Nigeria, at the 1-day horizon the GARCH(1,1) produces better forecasts at the 5% level of significance, whereas at the longer horizons the simple MS model performs better at the 1% level.

The SWARCH(2,1) model, because of its similarity to the simple MS model, produces a comparable forecasting performance. For South Africa and Kenya, the SWARCH(2,1) provides superior forecasts to the GARCH(1,1) at two of the three forecast horizons. In the South African case, the 3-week forecasts are not statistically different, while for Kenya it is at the 1-day horizon that the null of equal predictive ability cannot be rejected. With respect to Mauritius, the SWARCH(2,1) is superior to the GARCH(1,1) at all the forecast horizons used. For Morocco, the GARCH(1,1) outperforms the SWARCH(2,1) at the 1-day and 1-week horizons. At the 3-week horizon, the null hypothesis cannot be rejected at any conventional level of significance. Using the Nigerian data, the 1-day and 1-week horizons show no significant differences between the models' forecasts. At the 3-week horizon, however, the SWARCH(2,1) shows superior predictive ability at the 1% level of significance.

According to Clark and McCracken (2001), the Diebold-Mariano (1995) test statistic no longer has a *t*-distribution if the pair of models used is nested. For this reason, the MS-GARCH(2,1,1) model, which nests all the others, is tested using the adjustment procedure suggested by Clark and West (2007). Additionally, the Clark and West (2007) procedure can be used to compare the forecasting performance of the SWARCH(2,1) and simple MS models. Tables C8-C11 (Appendix C) contain the *p*-values associated with the Clark and West (2007) *t*-statistics. For each country, the MS-GARCH(2,1,1) is tested against the remaining three models, and the SWARCH(2,1) is tested against the simple MS.

Under the Clark and West (2007) test, the MS-GARCH(2,1,1) model performs relatively poorly. For South Africa, the MS-GARCH(2,1,1) outperforms the GARCH(1,1), simple MS, and SWARCH(2,1) models at the 3-, 1-, and 1-week horizons, respectively. For Nigeria, the MS-GARCH(2,1,1) also outperforms the simple MS model at the 1-week horizon. Besides these isolated cases, for South Africa, Kenya, Morocco and Nigeria, the alternative models produce better forecasts at every horizon. In stark contrast to the other countries used, Mauritius produces a somewhat surprising result, given the relatively poor MSE and MAE figures associated with the MS-GARCH(2,1,1). In this case, the MS-GARCH(2,1,1) is shown to outperform the GARCH(1,1) and SWARCH(2,1) models at every horizon, at the 1% level of significance. It also outperforms the simple MS model at the 1- and 3-week horizons.

Testing the simple MS model against the SWARCH(2,1) produces results that are consistent with the preliminary comparisons of MSE and MAE figures. In general, that is, the SWARCH(2,1) fails to significantly outperform the simple MS model. Exceptions are for Kenya and Mauritius at the 3-week horizon, and Mauritius and Nigeria at the 1-week horizon.

In summary, the forecast tests generally favour the simple MS model and the SWARCH(2,1). These models exhibit little or no autoregressive persistence, in contrast to the GARCH(1,1) and MS-GARCH(2,1,1). In order to gain a better understanding of why the simpler Markov-switching models outperform those with lagged conditional variance terms, it is helpful to investigate any systematic differences in forecasts that may exist. As an illustrative example, Figure 7 (page 76) shows the 1-day, 1-week and 3-week difference between the simple MS and GARCH(1,1) forecasts for Kenya.

Consistent with the analysis so far, there appears to be no systematic difference between the simple MS and GARCH(1,1) forecasts at the 1-day horizon. However, at the 1- and 3-week horizons, the simple MS model produces systematically lower conditional variance estimates than those of the GARCH(1,1). This needs to be understood within the context of the volatility regime towards the end of the rolling estimation sample, and the volatility regime after the sample. Removing 115 observations (as was done in order to conduct the forecast tests) re-adjusts the sample so that the last observation occurs roughly at the end of 2011. At that point, each country (with the exception of Morocco) transitions from the high-volatility state to the low-volatility state, and remains there for the first five months of 2012.

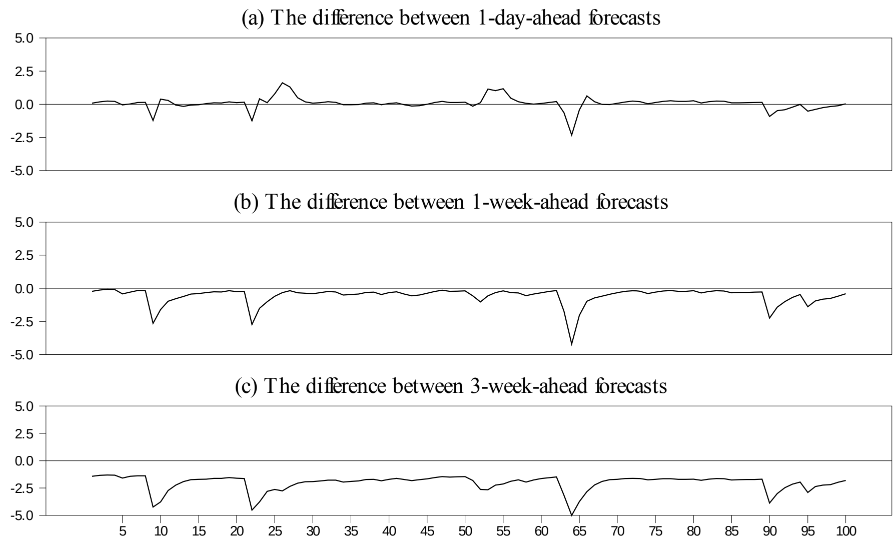

Figure 7: Forecast differences

Since the GARCH(1,1) model does not identify volatility states, the high-volatility shocks occurring towards the end of 2011 result, via the autoregressive mechanism, in systematically higher forecasts than those of the regime-switching models. The simpler regime-switching models (those with little or no autoregressive dynamics in the conditional variance), immediately identify the transition to a low-volatility state and adjust their forecasts accordingly. In this way, the simpler Markov-switching models can provide superior forecasts to the GARCH(1,1) or MS-GARCH(2,1,1) models, especially over periods in which a series moves from one state to another.

4.3. Multivariate models

To more fully capture volatility dynamics in an increasingly linked world, many researchers are finding it is necessary to explicitly model the interdependence between different markets (Bauwens *et al.*, 2006:79). Multivariate GARCH models remain a conventional method of illustrating the phenomena of volatility spillover and contagion, for example (Bauwens *et al.*, 2006:79). It is thus necessary to assess the performance of a regime-switching model against the already established single-state models in the multivariate space.

For the multivariate study, bivariate MV-SWARCH(2,1), CCC-GARCH and BEK-GARCH models are estimated. In all cases, models were estimated assuming both normal and

Student's t distributed errors. As in the univariate case, however, LRTs indicate that those specifications which assume Student's t distributed errors provide a better fit of the data. Hence, unless otherwise stated, the analysis is restricted to these models.

Following Ramchand and Susmel (1998) and Edwards and Susmel (2001), this study restricts the Markov-switching model to a bivariate extension of the SWARCH(2,1) specification. This is due to the large number of parameters that typically have to be estimated in multivariate models: Restricting each country to two states results in a four-state bivariate model, as illustrated in Chapter 3. For this relatively simple Markov-switching model, 25 parameters need to be estimated. Inclusion of further flexibility in the model substantially decreases the probability of convergence to a global maximum during the quasi-maximum likelihood estimation procedure.

4.3.1. The time-path of volatility states

As an illustrative example, the return series for South Africa and Mauritius are presented in Figure 8 (page 78), coupled with the smoothed probabilities of each primitive state, which have been derived through estimation of the South Africa-Mauritius MV-SWARCH model. As can be seen in Figure 8, much of the sample period is spent within primitive states 1 and 4 (in which both South Africa and the Mauritius are in state 1 and 2, respectively). It is thus worthwhile to test, for each country, whether the model can be restricted so that both markets are always in equivalent states, without losing significant explanatory power. This 'dependent states' hypothesis is tested by restricting the model to two primitive states, in which both markets are either in state 1 or 2, and subsequently conducting an LRT against the unrestricted model. Since there are 10 restrictions imposed on the dependent states model, the 1% critical χ^2 value is 23.21. For all the market pairs used, the LR statistic exceeds the 1% critical value. Thus, the null hypothesis of fully dependent states is rejected in favour of the more flexible model, in which Kenya, Mauritius, Morocco, and Nigeria transition at least partially independently of South Africa between volatility states.

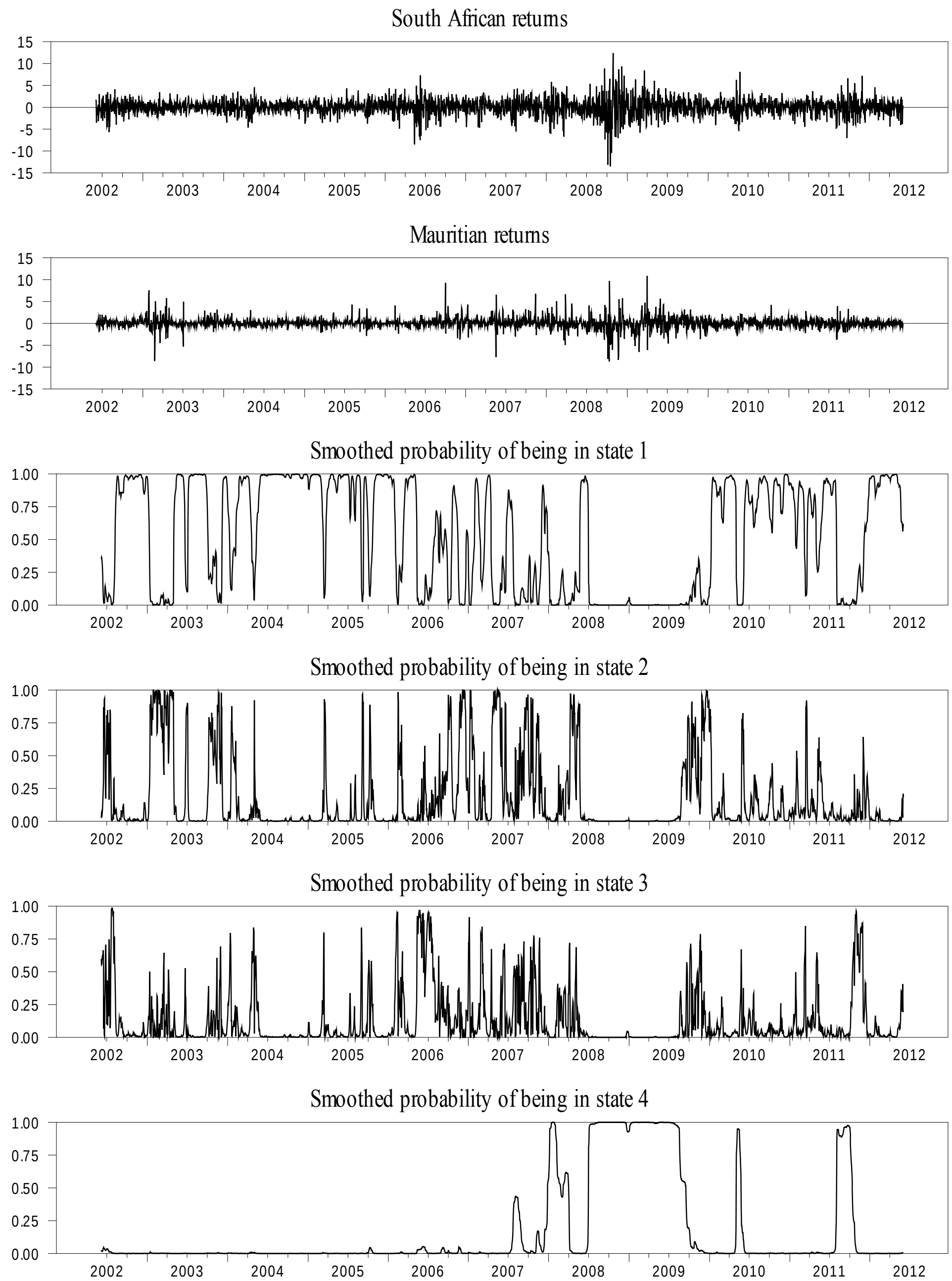

Figure 8: South Africa-Mauritius volatility states

4.3.2. Bivariate parameter estimates and their economic meaning

The focus of this sub-section is on the statistical significance and economic implications of the covariance parameter estimates for each of the three models presented (see Appendix A, Tables A12-A14). The inclusion of a covariance structure is one of the advantages of moving from a univariate to bivariate setting, and is primarily what allows for the study of inter-linkages between markets.

The CCC-GARCH model is the most restrictive multivariate model presented, as it assumes that the correlation coefficient, which directly affects the conditional covariance between two series, remains constant. Nonetheless, the relative ease with which it can be estimated and the simplicity of the innovation provides a good starting point for any multivariate analysis of volatility.

For all the 'South Africa-Other' pairs used, the correlation coefficients estimated under the CCC-GARCH specification are positive. This indicates some degree of common risk sources between the markets analysed. For example, if the countries were all net exporters of commodities, one would expect the returns and volatility of these markets to be related in a similar way to the level or volatility of a commodities price index, and hence to each other indirectly (Diebold and Yilmaz, 2008). Alternatively, if all the markets examined are highly integrated with the global financial economy, the swings of international capital may affect them similarly during optimistic and pessimistic periods of investor sentiment (King and Wadhwani, 1990). For the South Africa-Kenya, South Africa-Mauritius, and South Africa-Nigeria pairs, the correlation coefficient is significant and ranges between about 0.04 and 0.05. By contrast, the South Africa-Morocco correlation coefficient is a relatively high at 0.25 and is statistically significant, suggesting a much greater degree of interdependence between these two markets than in the other cases.

Engle and Kroner's (1995) BEK-GARCH allows for a greater degree of flexibility in the conditional covariance. In addition to a constant, the conditional covariance term (and each conditional variance term) in a bivariate framework is modelled as linearly dependent on the lagged squared residuals from both series, the lagged residuals multiplied by each other, and the lagged variance and covariance terms. This allows for time variation in the conditional covariance, in an equivalent manner to a univariate GARCH model.

Referring to the South Africa-Kenya and South Africa-Morocco pairs, Tables A14 (i) and A14 (iii) in Appendix A illustrate that no off-diagonal matrix elements are statistically significant. Due to the particular structure of the BEK model, this means that the constant, lagged squared errors, and lagged conditional variance terms fall away from the conditional covariance equation. The lagged conditional covariance terms and the multiplied residuals, however, retain their explanatory power, allowing for persistence and variation in the conditional covariance series. The statistical insignificance of the lagged conditional variance coefficients indicates that the volatility experienced in either market (in both cases) plays little role in determining the extent of interdependence.

In the case of South Africa-Mauritius, the only significant off-diagonal elements are α_{12} and β_{12}. What this indicates is that, for the conditional covariance equation, the constant term, lagged squared residuals and conditional variance terms for Mauritius fall away. Contrastingly, the equivalent South African terms retain significant explanatory power. That is, the co-variation in returns between South Africa and Mauritius depends directly on the extent of return volatility in the South African market, but not on that of the Mauritian market.

For South Africa-Nigeria, the only off-diagonal element which is statistically significant is α_{12}, at the 5% level. This can be interpreted to mean that although the volatility of the South African market directly affects the extent of covariance between South Africa and Nigeria, this effect is not persistent enough to warrant the inclusion of the South African lagged conditional variance term.

Having established a direct relationship between South African volatility and the degree of co-variation between South Africa and other markets (especially those of Mauritius and Nigeria), it may be beneficial to incorporate this property into a Markov-switching framework. This would allow the volatility of the South African market to directly affect the cross-correlation of returns, which may be interpreted as evidence of a contagion effect during high-volatility periods. Tables A12 (i)-A12 (iv) in Appendix A present the parameter estimates for the bivariate MV-SWARCH model. In estimating these parameters, South Africa has been established as the *originator* country, so that the correlation coefficient can take on one of two values, depending on the volatility state in which South Africa finds itself. This assumption is deemed appropriate given the finding, under the BEK framework, that the conditional covariance of returns between South Africa and other African markets can in some instances depend directly on the conditional variance of returns in South Africa.

Under the MV-SWARCH framework, the South Africa-Kenya and South Africa-Mauritius market pairs exhibit similar behaviour in the correlation coefficient. For both pairs, the correlation coefficient is statistically insignificant when South Africa is in the low-volatility state, but jumps to 0.11 and 0.08 (both significant at the 1% level) respectively, when South Africa is in the high-volatility state. This may indicate the presence a contagion effect, in which high volatility periods in South Africa, such as those during financial crises, are accompanied by increased inter-linkages between international markets.

For the South Africa-Morocco pair, the relatively strong inter-linkages established under the single-state models are maintained. When South Africa is in the low-volatility state, $\rho_1 = 0.22$ and is significant at the 1% level. When South Africa moves to the high-volatility state, the correlation coefficient increases to 0.37 and remains significant. In the South Africa-Nigerian case, ρ is insignificant for both South African states.

In a formal test of whether the correlation coefficient remains constant between states, the MV-SWARCH models are re-estimated with the restriction $\rho_1 = \rho_2$. An LRT is then conducted, using the null hypothesis of one correlation coefficient against the alternative of two. For each pair except that of South Africa-Morocco, the null hypothesis is rejected at the 1% level of significance in favour of the more flexible correlation coefficient. This suggests that contagion effects, in which there is excess cross-correlation of returns during high-volatility periods, are important for all country pairs except South Africa-Morocco.

4.3.3. Risk reduction through diversification

Given the overall result that the correlation coefficients appear to increase during South African high-volatility periods, it is expected that the benefit from diversification between the market pairs will diminish relative to low-volatility periods. Table D3 in Appendix D illustrates the extent of risk reduction from diversification for each originator country state. Specifically, using portfolio weights of 0.5 (half of wealth invested in each market),[21] the risk reduction figure illustrates the percentage decrease in the standard deviation of returns

[21] This is consistent with Li (2009). The general result – that the benefits of diversification tend to diminish during high volatility periods – is insensitive to the exact figure used.

resulting from diversification when South Africa is in a low-volatility state, and when South Africa is in a high-volatility state. The risk reduction percentage is calculated as:

$$Risk\ reduction\ \% = 100[\sigma_p - (1/2 \times \sigma_{SA} + 1/2 \times \sigma_{OTH})]/(1/2 \times \sigma_{SA} + 1/2 \times \sigma_{OTH}), \qquad (60)$$

where σ_p is the standard deviation of the diversified portfolio, and σ_{SA} and σ_{OTH} are the standard deviations of the South African and other country market indices, respectively (Li, 2009). The test is executed by separating each sample into a low-volatility and high-volatility period (based on South Africa only), where an observation belongs to state 1 if $p_1 \geq 0.5$. Standard deviation figures are then applied to each sample separately, and the benefit from diversification is calculated for each South African volatility regime.

In all cases except South Africa-Mauritius, the benefit from diversification is greater when South Africa is in the low-volatility state. For South Africa-Kenya, the risk reduction benefits of diversification decrease by 22.98% when South Africa is in the high volatility state. For South Africa-Morocco, the extent of the decrease in risk reduction benefit is 34.23%, while this decrease is only 3.01% for South Africa-Nigeria. Overall, it appears that the benefits from pan-African diversification are more strongly felt when South Africa is in the low volatility state, than otherwise. This is consistent with the observed increase in cross-correlation coefficients when South Africa is in the high-volatility state, as well as other existing research (*cf.* Ramchand and Susmel, 1998; Susmel, 2000; Edwards and Susmel, 2001).

4.3.4. Model selection: Portfolio optimisation

A noticeable feature of the CCC-GARCH and MV-SWARCH correlation coefficients is that the former appear to be weighted averages of the latter. In the case of South Africa-Kenya, the CCC-GARCH produces the estimate $\rho = 0.03$. The MV-SWARCH estimates are $\rho_1 = 0$ and $\rho_2 = 0.14$. Considering the fact that, from the univariate analysis, the ergodic probability that South Africa is in the high-volatility state is 21%, ρ can be calculated as: (1-0.21)0 + (0.21)0.14 = 0.03 (rounded).

Since the correlation coefficient is more flexible under the regime-switching framework than the single-state models, it is plausible that the MV-SWARCH produces better in-sample estimates and out-of-sample forecasts than its single-state counterparts. For example, if two return series exhibit a low or insignificant degree of co-variation for 70% of a sample, but a high and significant degree of co-variation for the rest, a model that restricts the correlation coefficient to remain constant will overestimate (underestimate) the conditional correlation of returns during the low-volatility (high-volatility) state. Likewise, if the persistence in the covariance of returns is better described by the persistence of different states rather than an autoregressive dynamic, the MV-SWARCH model should provide more reliable estimates of the conditional variances and covariance than the BEK-GARCH model.

As the models considered here are not nested within each other, LRTs prove to be an inappropriate measure of their comparative performance. Hence, following Ramchand and Susmel (1998) and Li (2009), the models presented will be subjected to a portfolio optimisation test. In particular, the variance-covariance estimates obtained from each bivariate model are used to construct utility optimising, 2-element portfolios for a 'typical' risk-averse investor. Each model is therefore associated with unique portfolio weights (w_1 and w_2). The portfolios (and hence models) which exhibit the highest risk-adjusted returns (measured by the Sharpe ratio under the assumption of a zero risk-free rate) are then favoured over the others.

The results of the abovementioned test are presented in Table D1 in Appendix D. In terms of mean returns, the MV-SWARCH outperforms both the BEK-GARCH and CCC-GARCH models. With respect to portfolio variances, the MV-SWARCH model performs the worst, in general. According to the risk-adjusted returns, the benefit of higher returns generally outweighs the cost of excess volatility, as the MV-SWARCH model produces the highest Sharpe ratios. This preliminary analysis of the portfolio optimisation results, while indicative of the MV-SWARCH model's superiority, does not provide conclusive evidence in its favour. In order to establish statistical differences in the mean and variances obtained, a simple z-test is used (*cf.* Li, 2009). If the differences are significant, the assertion that the MV-SWARCH model outperforms the other two models will gain additional credibility.

For South Africa-Kenya and South Africa-Nigeria, the MV-SWARCH produces higher returns than the CCC-GARCH model at the 5% and 1% levels of significance, respectively. Mean returns do not differ significantly across these two models for the South Africa-Mauritius and South Africa-Morocco pairs, however. With respect to the volatility of returns, the MV-SWARCH produces higher portfolio variances for South Africa-Kenya and South Africa-Mauritius at the 1% and 5% levels of significance, respectively. For South Africa-

Nigeria, the difference in portfolio variances is statistically insignificant, while for South Africa-Morocco the MV-SWARCH portfolio shows lower volatility than the CCC-GARCH portfolio at the 1% significance level.

On the basis of further *z*-tests, the MV-SWARCH and BEK-GARCH portfolios produce similar results. With respect to the South Africa-Mauritius and South Africa-Nigeria market pairs, neither the mean returns nor the variance of returns from either portfolio exhibit statistically significant differences. The strongest result in favour of the MV-SWARCH is for South Africa-Morocco, in which both the mean return and variance are significantly better (higher returns and lower volatility) than those of the BEK-GARCH portfolio at the 1% level. For the South Africa-Kenya portfolios, both the mean and variance of returns from the MV-SWARCH portfolio are significantly higher than that of the BEK -GARCH portfolio.

In Figures 9 and 10 (page 85), the weight differentials between the MV-SWARCH/CCC-GARCH and MV-SWARCH/BEK-GARCH portfolios are presented for each country. For example, the SWARCH-CCC series are derived by subtracting the proportion of wealth stored in the South African market for the CCC-GARCH portfolio ($w_{1,CCC}$) from the proportion of wealth stored in the South African market for the MV-SWARCH portfolio ($w_{1,SW}$). This weight differential then gives an indication of how the MV-SWARCH achieves its advantage in terms of risk-adjusted returns.

In Figure 9, the MV-SWARCH portfolio is shown to underweight the South African market until 2008, following which it tends to overweight South Africa. This pattern is clear for all the country pairs, excluding South Africa-Mauritius, for which the MV-SWARCH tends to underweight the South African market throughout the sample.

Comparing the MV-SWARCH with the BEK-GARCH portfolios (Figure 10), a roughly equal weighting is evident in the South African market for South Africa-Mauritius. For the South Africa-Morocco and South Africa-Kenya portfolios, the pattern is similar to those of Figure 9, in that South Africa is underweighted prior to the global financial crisis, after which South Africa tends to be overweighted by the MV-SWARCH model. In the case of South Africa-Nigeria, the MV-SWARCH model favours the South African market throughout the sample period, with an increase in favour around the end of 2008.

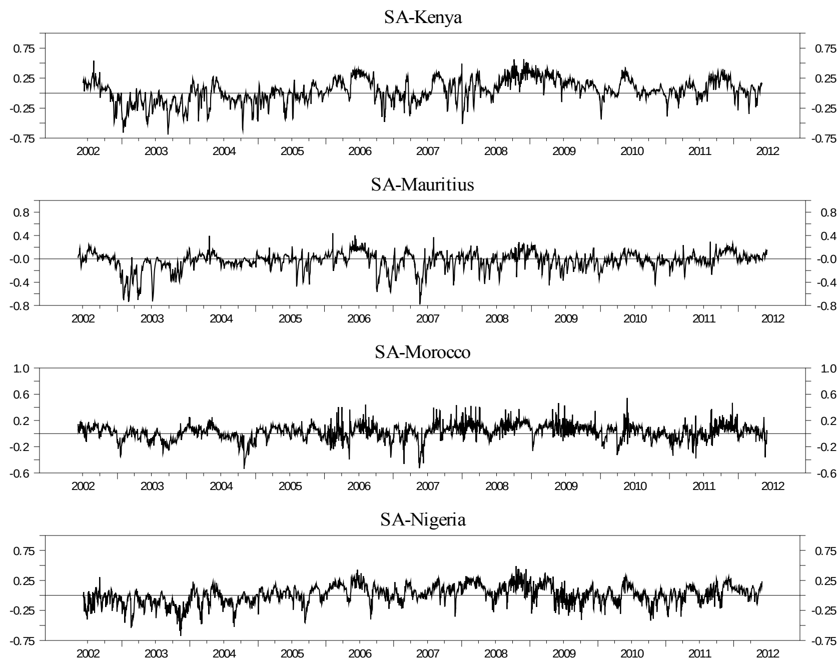

Figure 9: SWARCH-CCC weight differentials

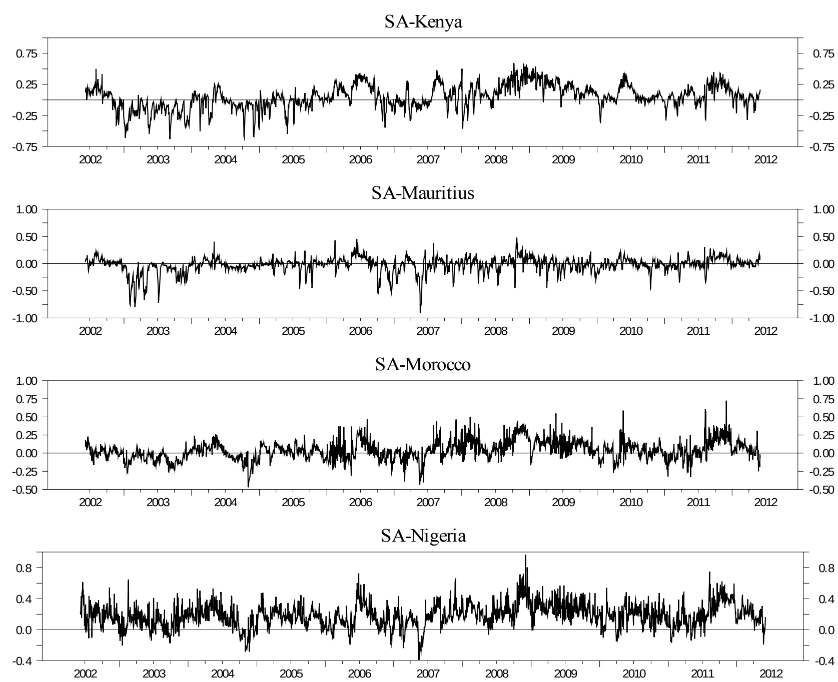

Figure 10: SWARCH-BEK weight differentials

Overall, the MV-SWARCH portfolios have generated excess returns through increased exposure to the less developed and frontier markets during the low volatility periods of pre-2008, while increasing the proportion of wealth held in South Africa during, and in the wake of, the global financial crisis.

4.4. Summary and Conclusions

A preliminary analysis of the univariate parameter estimates indicates that a two-state characterisation of the conditional variance of returns may be useful in the case of Kenya, Mauritius, and perhaps Nigeria, but may be inappropriate for South Africa and Morocco.

A set of LRTs, quasi-LRTs, and Davies (1987) upper-bound tests indicate that the best in-sample fit of the data is achieved by the GARCH(1,1) and MS-GARCH(2,1,1) models. In addition, the simple MS and SWARCH(2,1) specifications are unable to fully account for the heteroscedasticity in the data. However, it is clear that due to the multitude of potential model specifications available, it is not a straightforward task to select the single best performer based on in-sample measures alone.

In order to strengthen the comparison between model specifications, a number of forecasting accuracy tests are used. The forecast tests generally favour the simple MS model and the SWARCH(2,1). These models exhibit little or no autoregressive persistence, in contrast to the GARCH(1,1) and MS-GARCH(2,1,1). A plausible hypothesis for this outperformance is that the simpler regime-switching models immediately identify the transition to a low-volatility state and adjust their forecasts accordingly. By contrast, the models which include stronger autoregressive dynamics exhibit a much higher degree of persistence in the conditional variance. In this way, the simpler Markov-switching models can provide superior forecasts to the GARCH(1,1) or MS-GARCH(2,1,1) models, especially over periods in which a series moves from one state to another.

For the multivariate study, bivariate MV-SWARCH(2,1), CCC-GARCH and BEK-GARCH models are estimated. In a preliminary test of regime-switching, an LRT is conducted by restricting the correlation coefficient in the MV-SWARCH model to remain constant. For each market pair except that of South Africa-Morocco, the null hypothesis of no switching is rejected. As a further illustration of the importance of regimes, the risk-reduction benefits of diversification are contrasted between the low- and high-volatility states. In all cases except

South Africa-Mauritius, the benefit from diversification is greater when South Africa is in the low-volatility state.

To test formally which of the three multivariate models performs best, a portfolio optimisation test is used. In this test, the MV-SWARCH model produces the most favourable risk-adjusted returns. While this difference appears generally statistically significant compared to the CCC-GARCH model, the outperformance is less clear relative to the BEK-GARCH model. According to portfolio weight differentials, it appears that the way MV-SWARCH portfolios generate excess returns is by overweighting the less developed and frontier markets during the high-return/low-volatility period leading up to the global crisis, while tilting more towards South Africa during, and in the wake of, the global financial crisis.

On balance, it is not unambiguously clear whether the Markov-switching models outperform the single-state alternatives. In the univariate analysis, the GARCH term appears vital in order to fully capture the heteroscedasticity in the data. This suggests that, in the African context, models of conditional variance should at least include some degree of autoregressive persistence. However, the forecasting tests point to an almost exactly opposite conclusion. It is the simpler Markov-switching models that outperform the models with higher degrees of autoregressive persistence.

In the multivariate analysis, the flexibility of the MV-SWARCH model is clearly able to capture some important features of the data, such as contagion effects across the African markets. However, the excess benefit (in terms of portfolio risk and return) to be derived from the use of the MV-SWARCH is not clear when compared to a flexible single-state model such as the BEK-GARCH. There is evidently room for further comparative research in this area.

CHAPTER 5: SUMMARY AND CONCLUSIONS

5.1. Summary

The primary goal of this study was to establish whether accounting for structural changes through the use of Markov-switching models would improve estimates and forecasts of stock return volatility within and across selected African countries. Chapters 2 to 4 presented the core literature, data, econometric methods, and empirical findings relating to this study. An executive summary of these chapters thus follows.

Chapter 2 entailed a broad review of the literature concerning the economic and statistical nature of financial volatility and market inter-linkages. Aimed at a thorough understanding of the nature of stock market volatility, a number of studies regarding its underlying economic causes were analysed. Broadly speaking, these studies can be categorised into three schools. The first school attempts to relate macroeconomic variables, including volatility in these variables, to stock market volatility through various econometric techniques. The second school relates changes in stock market volatility to changes in expected dividends, discount rates and a risk premium. The third school argues that stock prices tend to move more than what is justified by the 'fundamentals-based', rational expectations models of schools one and two. It was found that although much of the literature under schools one and two provide statistically significant and useful results, the behavioural models of school three appear capable of explaining a non-negligible portion of stock return volatility, thus providing a more empirically accurate characterisation of stock return dynamics.

The study of market inter-linkages pointed to a number of plausible underlying causes of the contagion effect. Where contagion is defined as co-movement between markets, a number of studies cite common fundamental risk exposures as a reason for the positive covariance of returns. However, when using the more accepted definition of contagion as 'excess' correlation between returns, a range of different hypotheses is proposed. These hypotheses relate to the use of imperfect information by investors, liquidity shocks forcing massive sell-offs, financial globalisation increasing inter-linkages through time, and 'wake-up' calls in which markets with poor fundamentals get sold off in the wake of seemingly unrelated shocks. The existence of more than one plausible hypothesis suggested a complex process underlying the phenomenon of contagion, in which many of the aforementioned factors may play a part.

Having established an economic basis for the study, a review of various Markov-switching models of financial market conditional variance and covariance was conducted. It was found that Markov-switching models of conditional variance are widely applied to various financial series, including interest rates, exchange rates, and stock returns, among others. Within each market category, it was revealed that while some univariate studies restrict their analyses to the simple MS models, there are many that employ the more complex SWARCH or MS-GARCH specifications. Typically, the performance of these models is compared to a set of single-state benchmarks. In the univariate studies reviewed, the performance of each model is commonly assessed through the use of LRTs, adjusted-LRTs, tests for residual heteroscedasticity, and predictive ability tests. Overall, the studies predominantly conclude that the Markov-switching models, through their ability to capture discrete structural breaks in the conditional variance, provide a better fit of the data than many of the conventional single-state alternatives. In addition, Markov-switching models are capable of producing superior forecasts to various GARCH specifications, although this result is somewhat sensitive to the loss function used.

Studies of multivariate Markov-switching models were also examined. In these studies, the performance of bivariate Markov-switching models is compared, through various means, to a number of single-state models such as the CCC-GARCH, DCC-GARCH, and BEK-GARCH. In most of the studies presented, various hypotheses regarding the extent of dependence between country-specific volatility states and the correlation of returns are tested. Typically, it is found that the volatility states display some degree of dependence across series, but are at least partially independent. Furthermore, the correlation coefficient is shown in many studies to vary between states, consistent with much of the contagion literature. To assess model performance comparatively, some of the studies make use of hedging or portfolio optimisation tests. These tests generally indicate a moderate degree of outperformance by the Markov-switching models over the single-state alternatives, although the differences in performance are statistically insignificant in certain cases.

Chapters 3 and 4 presented the data, empirical methods, and core results of the analysis conducted. In Chapter 3, the African stock market data used in the study were described. The series selected were the daily returns on MSCI standard country indices of South Africa, Kenya, Mauritius, Morocco, and Nigeria. A preliminary analysis of returns identified a general pattern of low volatility leading up to the global financial crisis of 2008, followed by a sharp entry into a high-volatility state. Most of the series analysed then exited the high-volatility state after two quarters, entering the high-volatility state again for only one quarter towards the end of the sample period. The two distinct periods of high volatility, during 2008/2009 and 2011, appeared far more short-lived than the low volatility periods. The varying degrees of persistence evident in the data provided a strong case for the existence of more than one volatility state. The remainder of Chapter 3 was thus devoted to the selection of single-state and Markov-switching models of variance and co-variance to be estimated and tested against one another. For the univariate study, the GARCH(1,1) was found to be a suitable single-state benchmark model. The simple MS, SWARCH, and MS-GARCH models were identified as plausible alternative characterisations of the conditional variance. In the multivariate study, the CCC-GARCH and BEK-GARCH were identified as appropriate single-state benchmark models. The MV-SWARCH model was selected as the Markov-switching alternative, due to its relatively parsimonious representation of the covariance of returns.

In Chapter 4, the core results of the study were presented and interpreted. In the univariate portion of this chapter, a number of Markov-switching models of conditional variance were tested against the benchmark single-state GARCH(1,1) model. The parameter estimates of the simple MS and SWARCH(2,1) models strongly suggested the existence of a second regime. Furthermore, the estimated transition probabilities confirmed the lower persistence associated with the high-volatility state. LRTs indicated the superior fit of the simple MS over the standard AR(1) model, and the SWARCH(2,1) over the simple MS model. The simple MS model failed to entirely capture the heteroscedasticity in all the series except that of Kenya. Similarly, the SWARCH(2,1) failed to account for the heteroscedasticity in the South African and Nigerian data. To address these issues, Dueker's (1997) MS-GARCH(2,1,1) and the single-state GARCH(1,1) models were considered.

Including a GARCH term in the Markov-switching framework weakened the evidence in favour of a regime-switching conditional variance. For South Africa and Morocco, the scaling parameter was no longer significantly different from unity. In addition, the MS-GARCH(2,1,1) model was able to fully account for the heteroscedasticity in the data for all countries considered. A Davies (1987) test further indicated that the MS-GARCH(2,1,1) provided a superior in-sample fit to the GARCH(1,1) model for all countries except South Africa.

In order to complement the in-sample tests of model performance, a number of forecasting tests were used. MAE and MSE functions suggested that the simple MS and SWARCH(2,1) models provided more accurate forecasts than the MS-GARCH(2,1,1) and GARCH(1,1). The more formal Diebold-Mariano (1995) and Clark and West (2007) tests generally confirmed the MAE and MSE results. Overall, the forecasting tests suggested that the models with little or no autoregressive persistence tended to outperform the others. The switch from a high- to low-volatility state immediately prior to the forecast window was cited as a plausible reason for this outperformance. Specifically, the GARCH(1,1) and MS-GARCH(2,1,1) models would produce forecasts that are systematically too high.

To conduct the multivariate analysis, the CCC-GARCH, BEK-GARCH, and MV-SWARCH models were used. Using South Africa as the originator country, bivariate parameter estimates were generated for Kenya, Mauritius, Morocco and Nigeria. For the CCC-GARCH model, the correlation coefficients for all country pairs were significant and ranged mostly between 0.04 and 0.05. Morocco was the outlier in this sense, with a relatively high correlation coefficient of 0.25. Estimation of the BEK-GARCH suggested that in some cases, notably South Africa-Mauritius and South Africa-Nigeria, the South African lagged conditional variance estimate directly and positively affects the conditional covariance of returns. Given this result, the MV-SWARCH model was characterised by a correlation coefficient that switches between one of two values, depending on the volatility state of South Africa (the volatility originator). For all country pairs, the correlation coefficient was higher when South Africa is in the high-volatility state. Restricting the correlation coefficients to be equal across regimes significantly reduced the explanatory power of the model for all pairs considered, as was shown through the use of LRTs. Restricting the number of primitive states to two also significantly reduced the explanatory power of the model for each pair considered, suggesting that the countries move at least partially independently between states.

The degree of risk reduction from portfolio diversification was tested for each country pair. It was shown that when South Africa is in the high volatility state, the benefits from portfolio diversification are diminished, except in the case of South Africa-Mauritius.

As a test of model performance, each of the three multivariate models was used to construct optimal portfolios for a typical risk-averse investor. The MV-SWARCH portfolio produced the most favourable risk-adjusted returns in each case, measured by Sharpe ratios. Based on tests for differences in the mean return and variance of returns, this outperformance was clear when compared to the CCC-GARCH portfolios, but less so when compared to the BEK-GARCH portfolios.

5.2. Conclusions

A number of existing studies of developed and developing financial markets show that Markov-switching models can produce superior in-sample fit and better forecasts of the conditional variance than many single-state models. In this study, the empirical results are less clear. In the univariate part of the study, in-sample performance tests favour the single-state GARCH and MS-GARCH models over the simple MS and SWARCH models. Unlike the former two models, the latter two fail to completely capture the heteroscedasticity present in all the data series. This result points to the importance of conventional GARCH effects in capturing time-varying volatility, and weakens the case for a dominant Markov-switching element in the conditional variance of African stock returns.

In contrast, tests for equal predictive ability overwhelmingly favour the simple MS and SWARCH models over the GARCH and MS-GARCH. A plausible hypothesis for this result is that, immediately prior to the forecast window, most of the series analysed exited the high-volatility state. In this context, those models that included a highly persistent GARCH term tended to overestimate future conditional variance. Thus, although the simple MS and SWARCH models appear to be underspecified, they can provide more accurate and robust forecasts than models including a GARCH term, especially during periods of transition between states.

The bivariate analysis somewhat favours the Markov-switching specification over the two single-state alternatives. The higher Sharpe ratios produced by the MV-SWARCH portfolio suggest a potential benefit from making use of this model's variance-covariance estimates in practice. While the outperformance of the MV-SWARCH relative to the CCC-GARCH is clear, it is less so against the BEK-GARCH. However, the diminished benefit from diversification when South Africa is in a high-volatility state suggests that a state-varying characterisation of the cross-correlation of returns is appropriate.

In sum, the use of Markov-switching variance and covariance models can prove beneficial in some instances. The superior forecasts and risk-adjusted returns obtained during the performance tests bear testament to this finding. However, the inability of the simple MS and SWARCH models to fully capture the evident heteroscedasticity, and the lack of a clear outperformer in the multivariate analysis, suggests that there is scope for further comparative research in this field. In particular, more flexible Markov-switching specifications should perhaps be considered, and these should be tested against a wider range of alternative single-state models. Moreover, the inconsistency of in-sample and forecasting performance should be further investigated, as this may point to problems with over-fitting when conducting conventional GARCH analyses. Under the multivariate framework, the Markov-switching multivariate models presented cannot plausibly be applied to high-dimensional systems. In this respect, there is a clear need to restrict the number of parameters to be estimated in a way that does not hinder the ability of the models to capture the important features of the data.

REFERENCES

ADAM, K., MARCET, A. and NICOLINI, J.P., 2008. Stock market volatility and learning. Working Paper No. 862. European Central Bank Working Paper Series.

ANG, A., and TIMMERMAN, T., 2011. Regime changes and financial markets. NBER Working Paper No. 17182. National Bureau of Economic Research.

ABEL, A., 1988. Stock prices under time-varying dividend risk: An exact solution in an infinite-horizon general equilibrium model. *Journal of Monetary Economics.* 22: 375-392.

AFRICAN DEVELOPMENT BANK, 2010. *African fixed income and derivatives guidebook*. Tunis-Belvedère, Tunisia: African Development Bank.

AGGARWAL, R., INCLAN, C. and LEAL, R., 1999. Volatility in emerging stock markets. *Journal of Financial and Quantitative Analysis.* 34 (1): 33-55.

BABIKIR, A., GUPTA, R., MWABUTWA, C. and OWUSEKYERE, E., 2010. Structural breaks and GARCH models of stock return volatility. Working Paper No. 201030. University of Pretoria Department of Economics

BAUWENS, L., LAURENT, S. and ROMBOUTS, J.V.K., 2006. Multivariate GARCH models: A survey. *Journal of Applied Econometrics*. 21: 79-109.

BAILLIE, R.T., BOLLERSLEV, T. and MIKKELSEN, H.O., 1996. Fractionally integrated generalised autoregressive conditional heteroscedasticity. *Journal of Econometrics*. 74: 3-30.

BEKAERT, G., EHRMANN, M., FRATZCHER, M. and MEHL, A.J., 2011. Global crises and equity market contagion. NBER Working Paper No. 17121. National Bureau of Economic Research.

BEKAERT, G., HARVEY, C.R. and NG, A., 2003. Market integration and contagion. NBER Working Paper No. 9510. National Bureau of Economic Research.

BLACK, F., 1976. Studies of stock market volatility changes. Proceedings of the American Statistical Association, Business and Economic Statistics Section. pp. 177-181.

BOLLEN, N.P.B., GRAY, S.F. and WHALEY, R.E., 2000. Regime switching in foreign exchange rates: Evidence from currency option prices. *Journal of Econometrics*. 94: 239-276.

BOLLERSLEV, T., 1986. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics.* 31: 307-327.

BOLLERSLEV, T. and ENGLE, R.F., 1986. Modelling the persistence of conditional variances. *Journal of Econometric Reviews*. 5 (1): 1-50.

BOLLERSLEV, T., 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalised ARCH approach. *Review of Economic Statistics*. 72: 498-505.

BOLLERSLEV, T., CHOU, R. and KRONER, K., 1992. ARCH modelling in finance: A review of the theory and empirical evidence. *Journal of Econometrics.* 52: 5-59.

BOLLERSLEV, T. and GHYSELS, E., 1996. Periodic autoregressive conditional heteroscedasticity. *Journal of Business and Economic Statistics*. 14: 139-157.

CAI, J., 1994. A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics.* 12 (3): 309-316.

CANARELLA, G. and POLLARD, S.K., 2007. A switching ARCH (SWARCH) model of stock market volatility: Some evidence from Latin America. *International Review of Economics*. 54: 445-462.

CAMPBELL, J.Y. and AMMER, J., 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns. *Journal of Finance*. 48 (1): 3-37.

CHEN, S., 2009. Predicting the bear stock market: Macroeconomic variables as leading indicators. *Journal of Banking and Finance*. 33: 211-223.

CHINN, M. and ITO, H., 2008. A new measure of financial openness. *Journal of Comparative Policy Analysis.* 10 (3): 309-322.

CHINZARA, Z., 2011. Macroeconomic uncertainty and conditional stock market volatility in South Africa. *South African Journal of Economics.* 79 (1): 27-49.

CHINZARA, Z. and AZIAKPONO, M.J., 2009. Dynamic returns linkages and volatility transmission between South African and world major stock markets. *Journal for Studies in Economics and Econometrics*. 33 (3): 69-94.

CLARK, T.E., and MCCRACKEN, M.W., 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*. 105 (1): 85-110.

CLARK T.E. and WEST, K.D., 2005. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics.* 138: 291-311.

COLLINS, D. and BIEPKE, N., 2003. Contagion: A fear for African equity markets? *Journal of Economics and Business*. 55: 285-297.

COX, J.C., INGERSOLL, J.E. and ROSS, S.A., 1985. A theory of the term structure of interest rates. *Econometrica*. 53: 385-407.

CUTLER, D.M., POTERBA, J.M. and SUMMERS, L.H., 1989. What moves stock prices? *Journal of Portfolio Management*. 15 (3): 4-12.

DAHLQUIST, M. and GRAY, S.F., 2000. Regime-switching and interest rates in the European monetary system. *Journal of International Economics*. 50: 399-419.

DAVIDSON, R. and MACKINNON, J.G., 1981. Several tests for model specification in the presence of alternative hypotheses. *Econometrica*. 49: 781-793.

DAVIES, R.B., 1987. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*. 74 (1): 33-43.

DIEBOLD, F., 1986. Modelling the persistence of conditional variance: A comment. *Econometric Reviews.* 5: 51-56.

DIEBOLD, F., LEE, J. and WEINBACH, G., 1994. Regime switching with time-varying transition probabilities. In: Hargreaves, C. (ed.). *Nonstationary Time Series Analysis and Cointegration*. Oxford: Oxford University Press, pp. 283-302.

DIEBOLD, F. and MARIANO, R.S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics*. 13 (3): 253-263.

DIEBOLD, F. and YILMAZ, K., 2007. Measuring financial asset return and volatility spillovers, with application to global equity markets. *Economic Journal*. 119: 158-171.

DIEBOLD, F. and YILMAZ, K., 2008. Macroeconomic volatility and stock market volatility, worldwide. NBER Working Paper No. 14269. National Bureau of Economic Research.

DING, Z., GRANGER, C.W.J. and ENGLE, R.F., 1993. A long-memory property of stock market returns and a new model. *Journal of Empirical Finance*. 1: 83-106.

DORNBUSCH, R., PARK, Y.C. and CLAESSENS, S., 2000. Contagion: Understanding how it spreads. *World Bank Research Observer*. 15: 177-197.

DUEKER, M.J., 1997. Markov switching in GARCH processes and mean-reverting stock-market volatility. *Journal of Business and Economic Statistics*. 15 (1): 26-34.

DUNCAN, A. and LIU, G., 2009. Modelling South African currency crises as structural changes in the volatility of the rand. *South African Journal of Economics.* 77 (3): 364-379.

EDWARDS, S., 1998. Interest rate volatility, capital controls and contagion. NBER Working Paper No. 6756. National Bureau of Economic Research.

EDWARDS, S. and SUSMEL, R., 2001. Volatility dependence and contagion in emerging equity markets. NBER Working Paper No. 8506. National Bureau of Economic Research.

EDWARDS, S. and SUSMEL, R., 2003. Interest rate volatility in emerging markets. *Review of Economics and Statistics.* 85 (2): 328-348.

EDWARDS, L., CASSIM, R. and VAN SEVENTER, D., 2009. Trade policy in South Africa. In: Aaron, J., Kahn, B. and Kingdon, G. (eds.). *South African Economic Policy Under Democracy.* Oxford: Oxford University Press, pp. 151-181.

ENGLE, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica.* 50 (4): 987-1007.

ENGLE, R.F., 1993. Statistical models for financial volatility. *Financial Analysts Journal*. 49 (1): 72-78.

ENGLE, R.F., 2004. Risk and volatility: Econometric models and financial practice. *American Economic Review.* 94 (3): 405-420.

ENGLE, R.F. and KRONER, K.F., 1995. Multivariate simultaneous generalised ARCH. *Journal of Econometric Theory*. 11: 122-150.

FAMA, E., 1990. Stock returns, expected returns, and real activity. *Journal of Finance*. 45: 1089-1108.

FIGLEWSKI, S. and WANG, X., 2000. Is the 'leverage effect' a leverage effect? NYU Working Paper No. S-00-37. New York University.

FILARDO, A.J., 1994. Business cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics*. 12: 299-308.

FORBES, K.J. and RIGOBON, R., 2002. No contagion, only interdependence: Measuring stock market comovements. *Journal of Finance.* 57 (5): 2223-2261.

FRIEDMAN, M., 1977. Nobel lecture: Inflation and unemployment. *Journal of Political Economy*. 85: 451-472.

GALLANT, A.R., 1977. Testing a nonlinear regression specification: A non-regular case. *Journal of the American Statistical Association*. 72: 523-530.

GARCIA, R., 1998. Asymptotic null distribution of the likelihood ratio test in Markov switching models. *International Economic Review*. 39 (3): 763-788.

GARCIA, R. and PERRON, P., 1996. An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics*. 78: 111-125.

GLOSTEN, L.R., JAGANNATHAN, R. and RUNKLE, D., 1993. Relationships between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*. 48: 1779-1801.

GRAY, S.F., 1996. Modelling the conditional distribution of interest rate as a regime-switching process. *Journal of Financial Economics.* 42: 27-62.

HAAS, M., MITTNIK, S. and PAOLELLA, M., 2004. A new approach to Markov-switching GARCH models. *Journal of Financial Econometrics.* 2 (4): 493-530.

HANSEN, B.E., 1992. The likelihood ratio test under nonstandard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics*. 7: S61-S82.

HANSEN, P.R., 2001. An unbiased test for superior predictive ability. Working paper No. 2001-06, Brown University.

HANSEN, P.R. and LUNDE, A., 2001. A comparison of volatility models: Does anything beat a GARCH(1,1)?. Unpublished manuscript. Department of Economics, Brown University.

HAMILTON, J., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica.* 57 (2): 357-384.

HAMILTON, J., 1990. Analysis of time series subject to changes in regime. *Journal of Econometrics.* 45: 39-70.

HAMILTON, J. and SUSMEL, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics.* 64: 307-333.

HAMILTON, J. and LIN, G., 1996. Stock market volatility and the business cycle. *Journal of Applied Econometrics*. 11: 573-593.

HENRY, O.T., 2009. Regime switching in the relationship between equity returns and short-term interest rates in the UK. *Journal of Banking and Finance*. 33: 405-414.

INCLAN, C. and TIAO, G.C., 1994. Use of cumulative-sum-of-squares for retrospective detection of changes in variance. *Journal of the American Statistical Association*. 89 (427): 913-923.

KEARNEY, C. and DALY, K., 1998. The causes of stock market volatility in Australia. *Applied Financial Economics*. 8: 597-605.

KING, M.A. and WADHWANI, S., 1990. Transmission of volatility between stock markets. *Review of Financial Studies*. 3 (1): 5-33.

KLAASSEN, F., 2002. Improving GARCH volatility forecasts with regime-switching GARCH. *Journal of Empirical Economics.* 27: 363-394.

KNEDLIK, T. and SCHEUFELE, R., 2008. Forecasting currency crises: Which methods signalled the South African crisis of June 1996? *South African Journal of Economics*. 76 (3): 367-383.

KOEDIJK, K., NISSEN, F., SCHOTMAN, P. and WOLFF, C., 1997. The dynamics of short-term interest rate volatility reconsidered. *European Finance Review.* 1: 105-130.

LAMOUREUX, C. and LASTRAPES, W., 1990. Persistence in variance, structural change, and the GARCH model. *Journal of Business and Economic Statistics.* 8 (2): 225-234.

LEE, H. and YODER, J., 2007. A bivariate Markov regime switching GARCH approach to estimate time-varying minimum hedge ratios. *Applied Economics*. 39 (10): 1253-1265.

LI, M.L., 2009. Change in volatility regimes and diversification in emerging stock markets. *South African Journal of Economics*. 77 (1): 59-80.

LO, A.W., and MACKINLAY, C., 1990. Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies*. 3 (3): 431-468.

LONGSTAFF, F.A., 2010. The subprime credit crisis and contagion in financial markets. *Journal of Financial Economics*. 97: 436-450.

LONGIN, F. and SOLNIK, B., 1995. Is the correlation in equity returns constant: 1960-1990? *Journal of International Money and Finance.* 14 (1): 3-26.

LUCAS, R.E., 1978. Asset prices in an exchange economy. *Econometrica*. 46: 1426-1445.

MANKIW, N.G., ROMER, D. and SHAPIRO, M., 1985. An unbiased re-examination of stock market volatility. *Journal of Finance*. 40 (3): 677-687.

MARCUCCI, J., 2005. Forecasting stock market volatility with regime-switching GARCH models. *Studies in Nonlinear Dynamics and Econometrics.* 9 (4): 1-42.

MENDOZA, E.G. and QUADRINI, V., 2010. Financial globalisation, financial crises and contagion. *Journal of Monetary Economics*. 57: 24-39.

NELSON, D.B., 1990. Stationarity and persistence in the GARCH(1,1) model. *Journal of Econometric Theory*. 6: 318-334.

NELSON, D.B., 1991. Conditional heteroscedasticity in asset returns: A new approach. *Econometrica*. 59: 347-370.

OLSEN, R.A., 1998. Behavioural finance and its implications for stock price volatility. *Financial Analysts Journal.* 54 (2): 10-18.

PERRON, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*. 57: 1361-1401.

PESARAN, M.H. and TIMMERMAN, A., 1992. A simple non-parametric test of predictive performance. *Journal of Business and Economic Statistics*. 10 (4): 461-465.

POTERBA, J.M. and SUMMERS, L.H., 1987. The persistence of volatility and stock market fluctuations. *American Economic Review*. 76: 1146-1151

PSARADAKIS, Z. and SPAGNOLO, N., 2003. On the determination of the number of regimes in a Markov-switching autoregressive model. *Journal of Time Series Analysis*. 24 (2): 237-252.

RAMCHUND, L. and SUSMEL, R., 1998. Volatility and cross correlation across major stock markets. *Journal of Empirical Finance.* 5: 397-416.

SCHWERT, G.W., 1989. Why does stock market volatility change over time? *Journal of Finance.* 44 (5): 1115-1153.

SCHWERT, G.W., 1990. Stock market volatility. *Financial Analysts Journal.* 46 (3): 23-34.

SHILLER, R., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*. 71: 421-436.

SHILLER, R., 1987. The volatility of stock market prices. *Science.* 235: 33-37.

SHILLER, R. 1988. Causes of changing financial market volatility. In: Federal Reserve Bank of Kansas City (ed.). *Financial Market Volatility*. Kansas: Federal Reserve Bank of Kansas City. pp. 1-22.

SUSMEL, R., 1999. Switching volatility in international equity markets. Unpublished manuscript. Department of Finance, University of Houston.

TIMMERMAN, A., 2000. Moments of Markov-switching models. *Journal of Econometrics*. 96: 75-111.

VLAAR, P.J. and PALM, F.C., 1993. The message in weekly exchange rates in the European Monetary System: Mean reversions, conditional heteroscedasticity and jumps. *Journal of Business and Economic Statistics*. 11: 351-360.

WANG, P. and THEOBALD, M., 2008. Regime-switching volatility of six East Asian emerging markets. *Research in International Business and Finance*. 22: 267-283.

WHITE, H., 2000. A reality check for data snooping. *Econometrica*. 68 (5): 1097-1126.

APPENDIX A: MODEL PARAMETER ESTIMATES

Tables A1-A14 present the parameter estimates of selected univariate and multivariate models. Figures in round brackets represent p-values associated with each parameter estimate. Figures in square brackets are normally distributed Wald test statistics associated with the null hypotheses $\gamma_2 = 1$ (when displayed below an estimate of γ_2) or $\gamma_3 = \hat{\gamma}_2$(when displayed below an estimate of γ_3).

1. Univariate models

1.1. Markov-switching univariate models

(A1) Simple MS model

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1061 (0.0002)	0.0535 (0.0018)	0.0374 (0.0036)	0.0434 (0.0000)	0.0045 (0.8239)
a_1	0.0233 (0.1969)	0.2410 (0.0000)	0.06950 (0.0001)	0.1420 (0.0000)	0.3505 (0.0000)
ω_1	2.0169 (0.0000)	0.6717 (0.0000)	0.4653 (0.0000)	0.8065 (0.0000)	0.8642 (0.0000)
ω_2	10.0133 (0.0000)	6.8034 (0.0000)	4.1535 (0.0000)	3.8725 (0.0000)	4.6894 (0.0000)
v	11.1383 (0.0000)	5.5828 (0.0000)	3.5848 (0.0000)	7.0964 (0.0000)	8.1843 (0.0000)
P(1,1), P(1,2)	0.9953, 0.0183	0.9652, 0.1277	0.9716, 0.0614	0.9896, 0.0600	0.9749, 0.0518
Log-likelihood	-5081.4914	-3940.2773	-3475.7438	-3746.0095	-4321.8238

(A2) SWARCH(2,1)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1084 (0.0002)	0.0517 (0.0014)	0.0299 (0.0215)	0.0431 (0.0172)	-0.0022 (0.9139)
a_1	0.0243 (0.2133)	0.2553 (0.0000)	0.0729 (0.0001)	0.1365 (0.0000)	0.3588 (0.0000)
ω	1.9396 (0.0000)	0.5910 (0.0000)	0.4434 (0.0000)	0.7878 (0.0000)	0.6910 (0.0000)
α_1	0.0334 (0.1302)	0.3691 (0.0000)	0.3650 (0.0000)	0.1375 (0.0002)	0.2284 (0.0000)
v	11.3358 (0.0000)	3.7721 (0.0000)	3.2385 (0.0000)	6.1539 (0.0000)	7.7246 (0.0000)
γ_2	4.8858 [7.9847]	4.2719 [6.1012]	6.0967 [5.1802]	4.7575 [4.7463]	4.2808 [5.4509]
P(1,1), P(1,2)	0.9953, 0.0176	0.9876, 0.0192	0.9885, 0.0230	0.9962, 0.0317	0.9793, 0.0307
Log-likelihood	-5080.2226	-3912.0988	-3437.9768	-3735.0540	-4294.0691

(A3) MS-GARCH(2,1,1)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1186 (0.0000)	0.0522 (0.0012)	0.0303 (0.0134)	0.0495 (0.0058)	-0.0066 (0.7466)
a_1	0.0261 (0.2036)	0.2517 (0.0000)	0.0781 (0.0000)	0.1404 (0.0000)	0.3604 (0.0000)
ω	0.0430 (0.0011)	0.2815 (0.0000)	0.2013 (0.0004)	0.0163 (0.0302)	0.1704 (0.0000)
α_1	0.0699 (0.0000)	0.3335 (0.0000)	0.4431 (0.0000)	0.0739 (0.0001)	0.2511 (0.0000)
β_1	0.9162 (0.0000)	0.4401 (0.0000)	0.3850 (0.0004)	0.9035 (0.0000)	0.6152 (0.0000)
γ_2	4.6765 [1.2583]	3.0918 [2.9688]	3.9458 [2.5668]	3.6569 [1.5742]	2.5805 [3.1006]
v	12.0058 (0.0000)	3.6358 (0.0000)	2.9173 (0.0000)	15.2990 (0.0138)	6.3384 (0.0000)
P(1,1), P(1,2)	0.9987, 0.0294	0.9939, 0.0089	0.9966, 0.0032	0.8940, 0.3094	0.9981, 0.0017
Log-likelihood	-5034.9246	-3896.8155	-3424.3784	-3720.7971	-4277.7507

(A4) SWARCH(2,1) with asymmetry

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1036 (0.0008)	0.0491 (0.0028)	0.0301 (0.0235)	0.0400 (0.0195)	0.0018 (0.9301)
a_1	0.0268 (0.2098)	0.2569 (0.0000)	0.0729 (0.0002)	0.1380 (0.0000)	0.3587 (0.0000)
ω	1.9492 (0.0000)	0.5779 (0.0000)	0.4430 (0.0000)	0.7860 (0.0000)	0.7015 (0.0000)
α_1	-0.0128 (0.6054)	0.3301 (0.0001)	0.3731 (0.0000)	0.1033 (0.0164)	0.2575 (0.0000)
λ	0.0847 (0.0298)	0.0806 (0.4231)	-0.0166 (0.8723)	0.0737 (0.2065)	-0.0630 (0.3479)
v	11.2177 (0.0000)	3.7804 (0.0000)	3.2386 (0.0000)	6.1632 (0.0000)	7.6934 (0.0000)
γ_2	4.9064 [8.0679]	4.2953 [5.8769]	6.0973 [5.1678]	4.6487 [4.6916]	4.3018 [5.5830]
P(1,1), P(1,2)	0.9955, 0.0171	0.9877, 0.0193	0.9886, 0.0230	0.9962, 0.0313	0.9800, 0.0311
Log-likelihood	-5077.8030	-3911.7642	-3437.9651	-3734.3141	-4293.6458

(A5) SWARCH(2,2)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1107 (0.0001)	0.0502 (0.0036)	0.0265 (0.0494)	0.0439 (0.0091)	-0.0062 (0.7526)
a_1	0.02851 (0.1622)	0.2590 (0.0000)	0.0788 (0.0000)	0.1418 (0.0000)	0.3666 (0.0000)
ω	1.6946 (0.0000)	0.5476 (0.0000)	0.3485 (0.0000)	0.7614 (0.0000)	0.5789 (0.0000)
α_1	0.0398 (0.0723)	0.3543 (0.0000)	0.4419 (0.0000)	0.1477 (0.0000)	0.2533 (0.0000)
α_2	0.1242 (0.0002)	0.1404 (0.0001)	0.1793 (0.0003)	0.0692 (0.0100)	0.0957 (0.0000)
v	12.5015 (0.0000)	3.6586 (0.0000)	2.9592 (0.0000)	5.7938 (0.0000)	6.9651 (0.0000)
γ_2	4.6604 [7.0726]	3.5800 [4.8143]	4.1942 [5.0640]	4.8564 [4.0852]	3.6243 [4.7850]
P(1,1), P(1,2)	0.9955, 0.0167	0.9915, 0.0126	0.9951, 0.0043	0.9982, 0.0199	0.9850, 0.0144
Log-likelihood	-5064.5990	-3900.4402	-3424.5046	-3728.9450	-4287.5418

(A6) SWARCH(2,4)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1174 (0.0000)	0.0533 (0.0009)	0.0292 (0.0248)	0.0495 (0.0059)	-0.0093 (0.6101)
a_1	0.0325 (0.0894)	0.2495 (0.0000)	0.0790 (0.0000)	0.1435 (0.0000)	0.3601 (0.0000)
ω	1.4124 (0.0000)	0.5215 (0.0000)	0.3281 (0.0000)	0.3787 (0.0468)	0.4877 (0.0000)
α_1	0.0340 (0.0949)	0.3569 (0.0000)	0.4404 (0.0000)	0.1962 (0.0004)	0.2900 (0.0000)
α_2	0.1326 (0.0000)	0.1274 (0.0003)	0.1685 (0.0000)	0.1260 (0.0003)	0.0916 (0.0006)
α_3	0.0851 (0.0013)	0.0895 (0.0033)	0.0358 (0.2212)	0.0797 (0.0126)	0.1341 (0.0000)
α_4	0.0836 (0.0009)	-0.0095 (0.0000)	0.0301 (0.2434)	0.1047 (0.0001)	0.0833 (0.0041)
v	12.7926 (0.0000)	3.6602 (0.0000)	2.9506 (0.0000)	7.2066 (0.0021)	6.1362 (0.0000)
γ_2	4.3207 [6.0742]	3.2061 [4.2333]	4.0899 [4.8403]	2.3788 [1.1219]	2.7563 [6.3411]
P(1,1), P(1,2)	0.9966, 0.0125	0.9931, 0.0098	0.9955, 0.0038	0.5440, 0.5272	0.9980, 0.0018
Log-likelihood	-5049.4649	-3891.3230	-3421.7818	-3738.9363	-4270.5448

(A7) SWARCH(3,1)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1179 (0.0000)	0.0544 (0.0016)	0.0288 (0.0157)	0.0507 (0.0056)	-0.0038 (0.8529)
a_1	0.0247 (0.2216)	0.2519 (0.0000)	0.0717 (0.0001)	0.1370 (0.0000)	0.3557 (0.0000)
ω	1.0397 (0.0000)	0.4873 (0.0000)	0.3367 (0.0000)	0.5516 (0.0000)	0.5782 (0.0000)
α_1	0.0122 (0.5543)	0.1801 (0.0003)	0.3992 (0.0000)	0.1131 (0.0002)	0.1951 (0.0000)
v	15.7817 (0.0000)	4.8455 (0.0000)	3.0490 (0.0000)	7.0255 (0.0000)	8.6315 (0.0000)
γ_2	2.3932 [8.0569]	3.0484 [5.1728]	3.1538 [5.3169]	1.8448 [6.7621]	2.5614 [5.5090]
γ_3	10.6663 [7.6744]	21.3264 [4.0682]	11.6310 [4.4165]	7.5027 [4.9245]	8.7216 [5.4242]
P(1,1), P(1,2), P(1,3)	0.9818, 0.0063, 0.0000	0.9843, 0.0153, 0.0221	0.9943, 0.0035, 0.0041	0.9962, 0.0000, 0.0142	0.9880, 0.0060, 0.0124
P(2,1), P(2,2), P(2,3)	0.0181, 0.9881, 0.0206	0.0099, 0.9703, 0.0822	0.0035, 0.9950, 0.0052	0.0038, 0.9947, 0.0196	0.0087, 0.9853, 0.0198
Log-likelihood	-5053.9262	-3891.5744	-3412.6213	-3717.7185	-4274.4309

1.2. Single-state univariate models

(A8) AR(1)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1005 (0.0027)	0.0559 (0.0024)	0.0630 (0.0088)	0.0360 (0.0514)	0.0054 (0.8194)
a_1	0.0200 (0.2297)	0.2206 (0.0000)	0.1061 (0.0000)	0.1586 (0.0000)	0.4048 (0.0000)
ω	3.8268 (0.0000)	4.4823 (0.0121)	1.5644 (0.0000)	1.3037 (0.0000)	2.2809 (0.0000)
v	4.0279 (0.0000)	2.2629 (0.0000)	2.0025 (0.0000)	3.9575 (0.0000)	3.5493 (0.0000)
Log-likelihood	-5236.4721	-4091.7263	-3641.7143	-3822.2070	-4469.8039

(A9) GARCH(1,1)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1196 (0.0000)	0.0537 (0.0016)	0.0362 (0.0073)	0.04852 (0.0070)	-0.0010 (0.9616)
a_1	0.0255 (0.2030)	0.2477 (0.0000)	0.0779 (0.0002)	0.1443 (0.0000)	0.3532 (0.0000)
ω	0.0499 (0.0018)	0.2764 (0.0000)	0.1156 (0.0002)	0.0564 (0.0028)	0.1628 (0.0000)
α_1	0.0802 (0.0000)	0.3905 (0.0000)	0.3910 (0.0000)	0.0963 (0.0000)	0.2607 (0.0000)
β_1	0.9058 (0.0000)	0.5646 (0.0000)	0.6950 (0.0000)	0.8605 (0.0000)	0.6856 (0.0000)
v	10.7346 (0.0000)	3.2818 (0.0000)	2.7730 (0.0000)	5.6115 (0.0000)	5.8099 (0.0000)
Log-likelihood	-5035.8510	-3910.3631	-3448.9863	-3730.0275	-4289.9170

(A10) GARCH(1,1) with asymmetry

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.0803 (0.0067)	0.0540 (0.0024)	0.0350 (0.0106)	0.0445 (0.0152)	0.0003 (0.9892)
a_1	0.0295 (0.1365)	0.2475 (0.0000)	0.0779 (0.0000)	0.1453 (0.0000)	0.3534 (0.0000)
ω	0.0657 (0.0001)	0.2759 (0.0000)	0.1145 (0.0003)	0.0603 (0.0063)	0.1630 (0.0000)
α_1	0.0146 (0.1839)	0.3928 (0.0000)	0.3609 (0.0000)	0.0817 (0.0002)	0.2658 (0.0000)
β_1	0.9065 (0.0000)	0.5652 (0.0000)	0.6976 (0.0000)	0.8554 (0.0000)	0.6854 (0.0000)
λ	0.1112 (0.0000)	-0.0053 (0.9410)	0.0541 (0.4354)	0.0331 (0.1670)	-0.0105 (0.8071)
v	12.6148 (0.0000)	3.2816 (0.0000)	2.7771 (0.0000)	5.5936 (0.0000)	5.8146 (0.0000)
Log-likelihood	-5017.8875	-3910.3602	-3448.6737	-3728.8737	-4289.8855

(A11) I-GARCH(1,1)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
a_0	0.1212 (0.0000)	0.0539 (0.0008)	0.0372 (0.0044)	0.0501 (0.0039)	-0.0042 (0.8393)
a_1	0.0247 (0.2217)	0.2470 (0.0000)	0.0793 (0.0001)	0.1427 (0.0000)	0.3520 (0.0000)
ω	0.0271 (0.0028)	0.2808 (0.0000)	0.1088 (0.0000)	0.0262 (0.0068)	0.1361 (0.0000)
α_1	0.0877 (0.0000)	0.4336 (0.0000)	0.3042 (0.0000)	0.1113 (0.0000)	0.3090 (0.0000)
β_1	0.9123 (0.0000)	0.5664 (0.0000)	0.6958 (0.0000)	0.8887 (0.0000)	0.6910 (0.0000)
v	9.5117 (0.0000)	3.1294 (0.0000)	3.0416 (0.0000)	4.7909 (0.0000)	5.1105 (0.0000)
Log-likelihood	-5038.5826	-3910.7623	-3450.5195	-3736.9924	-4292.6583

2. Multivariate models

2.1. Markov-switching multivariate models

(A12) MV-SWARCH(2,1)

(i) South Africa-Kenya

	SA	**Kenya**
a_0	0.1271 (0.0000)	0.0509 (0.0016)
a_1	0.0137 (0.4844)	0.2521 (0.0000)
ω	1.9464 (0.0000)	0.6289 (0.0000)
α_1	0.0362 (0.1379)	0.2177 (0.0000)
γ_2	4.5129	9.1885
ρ_1	0.0155 (0.5795)	
ρ_2	0.1110 (0.0023)	
v	7.6170 (0.0000)	
P(1,1), P(1,2), P(1,3), P(1,4)	0.9825, 0.0000, 0.0101, 0.1305	
P(2,1), P(2,2), P(2,3), P(2,4)	0.0123, 0.8494, 0.0326, 0.0000	
P(3,1), P(3,2), P(3,3), P(3,4)	0.0052, 0.0000, 0.9573, 0.0552	
Log-likelihood	-9045.8158	

(ii) South Africa-Mauritius

	SA	Mauritius
a_0	0.1224 (0.0001)	0.0314 (0.0337)
a_1	0.0140 (0.5152)	0.0813 (0.0001)
ω	1.8433 (0.0000)	0.2997 (0.0000)
α_1	0.0377 (0.1597)	0.2228 (0.0000)
γ_2	4.3696	8.0604
ρ_1	0.0265 (0.3579)	
ρ_2	0.0855 (0.0471)	
v	6.5755 (0.0000)	
P(1,1), P(1,2), P(1,3), P(1,4)	0.9728, 0.0634, 0.0227, 0.0000	
P(2,1), P(2,2), P(2,3), P(2,4)	0.0000, 0.8341, 0.2263, 0.0000	
P(3,1), P(3,2), P(3,3), P(3,4)	0.0232, 0.1025, 0.7510, 0.0128	
Log-likelihood	-8551.9244	

(iii) South Africa-Morocco[22]

	SA	Morocco
a_0	0.1178 (0.0001)	0.0551 (0.0031)
a_1	0.0170 (0.3372)	0.1449 (0.0000)
ω	1.8507 (0.0000)	0.6733 (0.0000)
α_1	0.0398 (0.0482)	0.1712 (0.0000)
γ_2	5.3834	3.9406
ρ_1	0.2229 (0.0000)	
ρ_2	0.3736 (0.0000)	
v	9.6737 (0.0000)	
P(1,1), P(1,2), P(1,3), P(1,4)	0.9921, 0.0749, 0.0000, 0.0000	

[22] Continued on the following page

P(2,1), P(2,2), P(2,3), P(2,4)	0.0000, 0.0435, 0.5393, 0.0000
P(3,1), P(3,2), P(3,3), P(3,4)	0.0072, 0.8680, 0.4607, 0.0186
Log-likelihood	-8701.6057

(iv) South Africa-Nigeria

	SA	**Nigeria**
a_0	0.1169 (0.0000)	0.0104 (0.6268)
a_1	0.0233 (0.2048)	0.3462 (0.0000)
ω	1.7369 (0.0000)	0.9314 (0.0000)
α_1	0.0707 (0.0086)	0.2689 (0.0000)
γ_2	4.6223	4.7827
ρ_1	0.0360 (0.1718)	
ρ_2	0.0278 (0.5374)	
v	10.3663 (0.0000)	
P(1,1), P(1,2), P(1,3), P(1,4)	0.9904, 0.0428, 0.0214, 0.0000	
P(2,1), P(2,2), P(2,3), P(2,4)	0.0000, 0.0893, 0.0116, 0.5112	
P(3,1), P(3,2), P(3,3), P(3,4)	0.0049, 0.0000, 0.9670, 0.0185	
Log-likelihood	-9418.5263	

2.2. Single-state multivariate models

(A13) CCC-GARCH(1,1)

(i) South Africa-Kenya

	SA	**Kenya**
a_0	0.1372 (0.0000)	0.0564 (0.0021)
a_1	0.0136 (0.4927)	0.2549 (0.0000)
ω	0.0500 (0.0088)	0.1980 (0.0000)
α_1	0.0897 (0.0000)	0.2863 (0.0000)
β_1	0.9095 (0.0000)	0.6019 (0.0000)
ρ	0.0401 (0.0689)	
v	5.4131 (0.0000)	
Log-likelihood	-8995.0931	

(ii) South Africa-Mauritius

	SA	**Mauritius**
a_0	0.1355 (0.0000)	0.0399 (0.0091)
a_1	0.0140 (0.4824)	0.0911 (0.0001)
ω	0.0579 (0.0027)	0.0668 (0.0000)
α_1	0.0880 (0.0000)	0.2323 (0.0000)
β_1	0.9116 (0.0000)	0.7259 (0.0000)
ρ	0.0472 (0.0359)	
v	4.8564 (0.0000)	
Log-likelihood	-8552.9897	

(iii) South Africa-Morocco

	SA	**Morocco**
a_0	0.1204 (0.0001)	0.0512 (0.0053)
a_1	0.0193 (0.3495)	0.1406 (0.0000)
ω	0.0616 (0.0012)	0.0571 (0.0017)
α_1	0.0832 (0.0000)	0.0908 (0.0000)
β_1	0.9024 (0.0000)	0.8584 (0.0000)
ρ	0.2490 (0.0000)	
v	7.9629 (0.0000)	
Log-likelihood	-8688.7928	

(iv) South Africa-Nigeria

	SA	**Nigeria**
a_0	0.1221 (0.0000)	0.0065 (0.7616)
a_1	0.0192 (0.3422)	0.3451 (0.0000)
ω	0.0535 (0.0027)	0.1477 (0.0000)
α_1	0.0847 (0.0000)	0.2396 (0.0000)
β_1	0.9028 (0.0000)	0.7001 (0.0000)
ρ	0.0363 (0.0841)	
v	9.2714 (0.0000)	
Log-likelihood	-9343.8190	

(A14) BEK-GARCH(1,1)

(i) South Africa-Kenya

	SA	**Kenya**
a_0	0.1370 (0.0000)	0.0544 (0.0025)
a_1	0.0042 (0.8325)	0.2567 (0.0000)
ω_{11}	0.1960 (0.0000)	
ω_{21}	0.0190 (0.8617)	
ω_{22}	0.4818 (0.0000)	
α_{11}	0.2335 (0.0000)	
α_{12}	-0.0007 (0.9766)	
α_{21}	0.0003 (0.9898)	
α_{22}	0.5322 (0.0000)	
β_{11}	0.9707 (0.0000)	
β_{12}	0.0046 (0.5735)	
β_{21}	-0.0033 (0.8507)	
β_{22}	0.7642 (0.0000)	
v	5.2410 (0.0000)	
Log-likelihood	-9019.2003	

(ii) South Africa-Mauritius

	SA	**Mauritius**
a_0	0.1325 (0.0000)	0.0442 (0.0018)
a_1	-0.0012 (0.9546)	0.0933 (0.0000)
ω_{11}	0.1959 (0.0000)	
ω_{21}	0.0891 (0.2551)	
ω_{22}	0.2775 (0.0000)	
α_{11}	0.2210 (0.0000)	
α_{12}	0.0624 (0.0053)	
α_{21}	0.0131 (0.6622)	
α_{22}	0.4757 (0.0000)	
β_{11}	0.9736 (0.0000)	
β_{12}	-0.0178 (0.0039)	
β_{21}	0.0110 (0.5261)	
β_{22}	0.8369 (0.0000)	
v	4.7743 (0.0000)	
Log-likelihood	-8574.1463	

(iii) South Africa-Morocco

	SA	**Morocco**
a_0	0.1231 (0.0001)	0.0479 (0.0137)
a_1	0.0247 (0.2029)	0.1411 (0.0000)
ω_{11}	-0.2131 (0.0000)	
ω_{21}	0.0255 (0.7137)	
ω_{22}	0.2449 (0.0000)	
α_{11}	0.2573 (0.0000)	
α_{12}	0.0227 (0.3316)	
α_{21}	-0.0691 (0.1013)	
α_{22}	0.2637 (0.0000)	
β_{11}	0.9567 (0.0000)	
β_{12}	-0.0041 (0.6348)	
β_{21}	0.0399 (0.1112)	
β_{22}	0.9345 (0.0000)	
v	7.8460 (0.0000)	
Log-likelihood	-8691.7959	

(iv) South Africa-Nigeria

	SA	**Nigeria**
a_0	0.1181 (0.0001)	0.0141 (0.5274)
a_1	0.0081 (0.7064)	0.3481 (0.0000)
ω_{11}	0.1892 (0.0000)	
ω_{21}	-0.0821 (0.3676)	
ω_{22}	0.4078 (0.0000)	
α_{11}	0.2252 (0.0000)	
α_{12}	-0.0310 (0.0401)	
α_{21}	-0.0178 (0.4468)	
α_{22}	0.4839 (0.0000)	
β_{11}	0.9689 (0.0000)	
β_{12}	0.0076 (0.2191)	
β_{21}	0.0218 (0.1695)	
β_{22}	0.8318 (0.0000)	
v	8.8148 (0.0000)	
Log-likelihood	-9366.1990	

APPENDIX B: IN-SAMPLE PERFORMANCE TESTS

1. Likelihood ratio tests and Davies (1987) upper-bound tests

Tables B1-B16 present the relevant likelihood ratio (LR) statistics. Only those models which are nested can be compared using likelihood ratio tests.

When comparing single state to multi-state models, the LRT statistic is no longer distributed χ^2 (hence the term "quasi-LRT"). To obtain a more reliable *p*-value, Davies' (1987) upper-bound test is used. For tables which include the Davies (1987) test, *q* indicates the number of parameters present only under the alternative (analogous to degrees of freedom (df) for the conventional LRT).

For models which have an equal number of states, *** indicates a significant LR statistic at the 1% level, and ** indicates significance at the 5% level. For models with differing numbers of states, an upper-bound on the actual level of significance is given (Davies, 1987).

1.1. Normal vs. Student's *t* errors

(B1) Null: GARCH(1,1) Normal, Alt: GARCH(1,1) Student's t (1 df)

	SA	Kenya	Mauritius	Morocco	Nigeria
LR	35.4962***	600.9794***	626.551***	144.6888***	124.7242***

(B2) Null: Simple MS Normal, Alt: Simple MS Student's t (1 df)

	SA	Kenya	Mauritius	Morocco	Nigeria
LR	27.8202***	70.7264***	148.7704***	20.4786***	47.685***

(B3) Null: SWARCH(2,1) Normal, Alt SWARCH(2,1) Student's t (1 df)

	SA	Kenya	Mauritius	Morocco	Nigeria
LR	26.1300***	83.5708***	152.6250***	35.2784***	45.1106***

(B4) Null: MS-GARCH(2,1,1) Normal, Alt: MS-GARCH(2,1,1) Student's t (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	11.9340***	78.0072***	56.2530***	5.4460**	24.8694***

1.2. The GARCH model

(B5) Null: I-GARCH(1,1), Alt: GARCH(1,1) (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	5.4632**	0.7984	3.0664	13.9298***	5.4826**

(B6) Null: GARCH(1,1), Alt: GARCH(1,1) with asymmetry (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	35.9270***	0.0058	0.6252	2.3076	0.0630

1.3. The simple MS model

(B7) Null: AR(1), Alt: Simple MS) (q = 3)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	309.9614***	302.898***	331.941***	152.395***	295.9602***
Davies	0.0000	0.0000	0.0000	0.0000	0.0000

(B8) Null: Simple MS, Alt: SWARCH(2,1) (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	2.5376	56.357***	75.534***	21.911***	55.5094**

(B9) Null: Simple MS, Alt: MS-GARCH (2,1,1) (2 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	93.1336***	86.9236***	102.7308***	50.4248***	88.1462***

1.4. The SWARCH and MS-GARCH models

(B10) Null: SWARCH(2,1), Alt: SWARCH(2,1) with asymmetry (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	4.8392**	0.6692	0.0234	1.4798	0.8466

(B11) Null:SWARCH(2,1), Alt: SWARCH(2,2) (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	31.2472***	23.3172***	26.9444***	12.218***	13.0546***

(B12) Null:SWARCH (2,2), Alt: SWARCH(2,4) (2 df)

	SA	**Kenya**	**Mauritius**	**Morocco**[23]	**Nigeria**
LR	30.2682***	18.2344***	5.4456	-19.9826	33.994***

(B13) Null: SWARCH(2,1), Alt: MS-GARCH (2,1,1) (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	90.5960***	30.5666***	27.1968***	28.5138***	32.6368***

(B14) Null: SWARCH(2,2), Alt: MS-GARCH(2,1,1) (1 df)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	59.3488***	7.2494***	0.2524	16.2958***	19.5822***

(B15) Null: GARCH(1,1), Alt:MS-GARCH(2,1,1) Student (q = 3)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	1.8528	27.0952***	49.2158***	18.4608***	24.3326***
Davies	0.7389	0.0001	0.0000	0.0033	0.0003

[23] Convergence to a global maximum value could not be achieved for the SWARCH (2,4) when applied to Morocco.

(B16) Null: SWARCH(2,1), Alt: SWARCH(3,1) (q=3)

	SA	**Kenya**	**Mauritius**	**Morocco**	**Nigeria**
LR	52.5928***	41.0488***	50.711***	34.671***	39.2764***
Davies	0.0000	0.0000	0.0000	0.0000	0.0000

2. Ljung-Box *Q*-tests

The figures presented in Tables B17-B22 are Ljung-Box *Q*-stats at 12 lags for the squared standardised residuals (LB^2 *(12)*). The symbol *** indicates significance at the 1% level, and ** indicates significance at the 5% level.

(B17) South Africa

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
LB^2 *(12)*	8.9857	90.9840***	86.8657***	7.6559

(B18) Kenya

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
LB^2 *(12)*	2.32239	12.4464	6.3500	1.7387

(B19) Mauritius

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
LB^2 *(12)*	7.5714	40.2419***	11.7549	12.2728

(B20) Morocco

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
LB^2 *(12)*	17.9991	24.9757**	6.8069	12.8947

(B21) Nigeria

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
LB^2 *(12)*	15.2056	143.7105***	43.2551***	9.4645

APPENDIX C: FORECASTING TESTS

1. Forecast errors

Tables C1-C5 present mean absolute forecast errors (MAE) and mean squared forecast errors (MSE) at the 1 day, 1 week (5 trading days) and 3 week (15 trading days) horizons. Each cell contains two figures separated by a semi-colon. The figure on the left is the MAE and the figure on the right is the MSE.

(C1) South Africa

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
1 day	2.21; 11.91	2.05; 11.12	2.08; 11.34	2.24; 12.01
1 week	2.27; 12.40	2.15; 11.68	2.16; 11.75	2.32; 12.56
3 weeks	2.58; 16.37	2.54; 15.95	2.55; 16.03	2.69; 16.67

(C2) Kenya

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
1 day	1.71; 7.75	1.76; 7.72	1.61; 7.83	1.74; 7.92
1 week	2.42; 10.03	1.96; 7.55	2.17; 8.34	2.47; 10.30
3 weeks	3.66; 15.38	1.87; 5.96	2.31; 7.95	3.36; 14.29

(C3) Mauritius

	GARCH(1,1)	**Simple MS**	**SWARCH(2,1)**	**MS-GARCH(2,1,1)**
1 day	0.68; 0.62	0.70; 0.72	0.63; 0.59	0.87; 1.09
1 week	1.54; 2.74	0.91; 0.99	1.17; 1.79	2.60; 8.54
3 weeks	6.30; 41.63	1.22; 1.61	1.90; 4.09	7.97; 90.55

(C4) Morocco

	GARCH(1,1)	Simple MS	SWARCH(2,1)	MS-GARCH(2,1,1)
1 day	0.84; 1.46	0.89; 1.52	0.88; 1.48	0.86; 1.46
1 week	0.93; 1.93	0.97; 1.90	0.97; 1.89	0.99; 2.02
3 weeks	1.16; 4.10	1.19; 4.09	1.17; 4.08	1.33; 4.41

(C5) Nigeria

	GARCH(1,1)	Simple MS	SWARCH(2,1)	MS-GARCH(2,1,1)
1 day	1.03; 1.63	1.07; 1.65	1.04; 1.69	0.98; 1.58
1 week	1.27; 2.13	1.21; 1.93	1.27; 2.11	1.14; 1.88
3 weeks	1.71; 3.31	1.43; 2.43	1.52; 2.71	1.34; 2.33

2. The Diebold-Mariano (1995) test

Tables C6 and C7 present *t*-statistics associated with the Diebold-Mariano (1995) test statistic. Positive *t*-statistics indicate that Model *Y* produces the superior forecasts, whereas a negative *t*-statistic indicates that Model *X* produces the superior forecasts (where the table headings are denoted Model *X*/Model *Y*). Since *p*-values are not calculated directly, *** indicates that the statistic is significant at the 1% level, ** indicates that the statistic is significant at the 5% level, and * indicates that the statistic is significant at the 10% level.

(C6) GARCH/Simple MS

	SA	Kenya	Mauritius	Morocco	Nigeria
1 day	4.5711***	-1.1208	-0.9825	-3.3181***	-1.8583**
1 week	3.8233***	6.6449***	19.0403***	-3.1939***	3.8999***
3 weeks	1.4160*	15.5154***	39.6069***	-2.0714**	12.8162***

(C7) GARCH/SWARCH

	SA	Kenya	Mauritius	Morocco	Nigeria
1 day	3.5872***	1.2882	2.5015***	-1.5952*	-0.2554
1 week	3.6772***	4.4948***	11.9898***	-2.7572***	0.2233
3 weeks	0.9640	14.0176***	47.1087***	-0.9155	11.5176***

3. The Clarke and West (2007) test

The figures presented in the Tables C8-C11 are *p*-values associated with the Clarke and West (2007) slope coefficient. If the *p*-value is less than 0.1, the larger of the two models produces the more accurate forecasts. The table headings are denoted Model *X*/Model *Y*, where model *Y* is nested by model *X*.

(C8) MS-GARCH/GARCH

	SA	Kenya	Mauritius	Morocco	Nigeria
1 day	0.1896	0.8258	0.0039	0.1762	0.1372
1 week	0.1259	0.2818	0.0000	0.3570	0.4435
3 weeks	0.0868	0.6515	0.0000	0.2905	0.1226

(C9) MS-GARCH/Simple MS

	SA	Kenya	Mauritius	Morocco	Nigeria
1 day	0.1531	0.8081	0.2502	0.4988	0.2043
1 week	0.0992	0.9336	0.0000	0.1395	0.0658
3 weeks	0.9992	0.0561	0.0000	0.3199	0.1568

(C10) MS-GARCH/SWARCH

	SA	Kenya	Mauritius	Morocco	Nigeria
1 day	0.1614	0.5467	0.0092	0.5906	0.6494
1 week	0.0973	0.9765	0.0000	0.1315	0.3043
3 weeks	0.9305	0.2769	0.0000	0.2379	0.6987

(C11) SWARCH/Simple MS

	SA	Kenya	Mauritius	Morocco	Nigeria
1 day	0.4325	0.6828	0.3873	0.6048	0.2835
1 week	0.2631	0.8670	0.0593	0.2494	0.0082
3 weeks	0.1537	0.0726	0.0004	0.5523	0.0000

APPENDIX D: MULTIVARIATE MODELS

1. Portfolio optimisation test *(Table D1)*

	SA-Kenya			**SA-Mauritius**			**SA-Morocco**			**SA-Nigeria**		
Model	*MV-SWARCH*	*CCC-GARCH*	*BEK-GARCH*	*MV-SWARCH*	*CCC-GARCH*	*BEK-GARCH*	*MV-SWARCH*	*CCC-GARCH*	*BEK-GARCH*	*MV-SWARCH*	*CCC-GARCH*	*BEK-GARCH*
Mean return	0.0590	0.0373	0.0367	0.0643	0.0590	0.0601	0.0508	0.0346	0.0348	0.0502	0.0435	0.0427
Variance	1.4785	1.2016	1.2192	1.1868	1.1340	1.1613	1.0212	1.1144	1.1201	1.3421	1.3009	1.3064
Sharpe Ratio	0.0484	0.0340	0.0332	0.0590	0.0554	0.0558	0.0502	0.0328	0.0329	0.0433	0.0382	0.0373

2. Li's (2009) test statistics *(Table D2)*

A positive z_1 (z_2) statistic indicates that the MV-SWARCH portfolio exhibits the higher mean (variance), whereas a negative statistic indicates a lower mean (variance).*** indicates that the test statistic is significant at the 1% level, ** at the 5% level, and * at the 10% level.

	SA-Kenya		**SA-Mauritius**		**SA-Morocco**		**SA-Nigeria**	
Model Pair	*SWARCH/CCC*	*SWARCH/BEK*	*SWARCH/CCC*	*SWARCH/BEK*	*SWARCH/CCC*	*SWARCH/BEK*	*SWARCH/CCC*	*SWARCH/BEK*
z_1	1.9068**	1.9583**	0.7336	0.5788	2.9843***	2.3929***	0.6841	0.8299
z_2	5.1817***	4.8292***	1.5636*	0.6906	-3.6121***	-14.6771***	1.0512	0.9005

3. Risk reduction (%) *(Table D3)*

	SA-Kenya	**SA-Mauritius**	**SA-Morocco**	**SA-Nigeria**
SA in LV state	30.0604	36.8566	22.0605	28.3835
SA in HV state	23.1527	38.0389	14.5083	27.5283

Note: *LV* denotes "low volatility", while *HV* denotes "high volatility".