2021

Modelling stochastic volatility using hidden Markov models: a case study of the Kenyan securities Market.

Bosire, Matilda Bosibori
Strathmore Institute of Mathematical Sciences
Strathmore University

Strathmore UNIVERSITY

MODELLING STOCHASTIC VOLATILITY USING HIDDEN MARKOV MODELS: A CASE STUDY OF THE KENYAN SECURITIES MARKET

BOSIRE MATILDA BOSIBORI

Research Thesis Submitted to Strathmore University in Fulfilment of the Requirements for the Master of Science in Mathematical Finance.

2021

Abstract

The biased parameter estimates generated by the Black-Scholes model have been attributed to the failure of the normality and constant volatility assumption to hold. Results of the improvement of the Black-Scholes-Merton model include time-varying volatility models which capture certain stylized facts of stock returns, with their use expected to improve the ability to price assets beyond the benchmarks provided by Black-Scholes. This thesis models stochastic volatility using Hidden Markov Models in Kenya. The univariate Stochastic volatility Model is calibrated to the Nairobi Securities Exchange 20 share index daily data for the period January 2012 to February 2021. The hidden Markov model (HMM) is employed in establishing volatility regimes while the Expected Maximization (EM) algorithm is employed in parameter estimation. Markov chain Monte Carlo (MCMC) and Sequential Monte Carlo (SMC) techniques are employed in filtering out noisy observations in parameter estimation. The 4-state model, which divides the economy into periods of very high, high, low, and very low volatility, is established to be optimal. Under each regime and filtering technique, different parameter estimates for single and multiple state models suggest a more dynamic framework for modeling the volatility process. The research findings contribute to theoretical literature on volatility-backed financial asset valuation and risk management in the context of regime switches.

Keywords: Hidden Markov Models, Stochastic Volatility, NSE20, Volatility Regimes

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Bosire Matilda Bosibori

Signature: ______________________

Date: 10th September 2021

Approval

The thesis of Bosire Matilda Bosibori was reviewed and approved by the following:

Dr. Samuel Chege,

Senior Lecturer, Strathmore Institute of Mathematical Sciences,

Strathmore University

Dr. Godfrey Achono Madigu,

Dean, Strathmore Institute of Mathematical Sciences,

Strathmore University

Dr. Bernard Shibwabo,

Director of Graduate Studies,

Strathmore University

Acknowledgement

I thank the Almighty God for the gift of life and for making this research a success thus far. I acknowledge the support I have received from both the lecturers and my fellow students at the Strathmore Institute of Mathematical Sciences during my study. I extend my heartfelt gratitude to Dr. Samuel Chege, my supervisor, for introducing me to the world of Hidden Markov Models and for taking time to guide, direct and assist me whenever I encountered challenges. His support made this work a reality. Finally, I am thankful to my parents, siblings and friends for their moral support and motivation during my studies.

God bless you all.

Dedication

To my beloved parents and family, classmates and friends.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC: Arkaike Information Criterion

AR: AutoRegressive

BIC: (Schwartz) Bayesian Information Criterion

CBK: Central Bank of Kenya

CIR: Cox, Ingersoll and Ross (Model)

ESS: Effective Sample Size

EM: Expectation-Maximisation (Algorithm)

GARCH: Generalized AutoRegressive Conditional Heteroskedasticity

HMM: Hidden Markov Model

MCMC: Markov Chain Monte Carlo

MLE: Maximum Likelihood Estimation

MSM: Markov Switching Model

NSE: Nairobi Securities Exchange

PDF: Probability Density Function

S&P 500: Standard & Poor's 500 Index

SDE: Stochastic Differential Equation

SMC: Sequential Monte Carlo

SV: Stochastic Volatility

Chapter 1

Introduction

1.1 Background of the Study

The probability that stock prices will rise or decline is an increasing function of volatility, which in turn leads to an increase in the value of options. The use of volatility as a proxy to risk has resulted in an increased need to accurately model and forecast volatility which is vital for a range of applications including financial asset pricing, hedging strategies, portfolio selection and asset management. The Black-Scholes model, as a forerunner to the option pricing framework, is still widely used in the financial market. It assumes that continuously compounded log spot asset prices are normally distributed with a constant mean and variance. However, empirical studies have shown that this is not always the case, as market prices have shown peakedness and fat tails when plotted, and the constant variance assumption does not fit market-realistic models particularly for volatility-backed financial assets.

Stirred by this major shortcoming of the Black Scholes pricing model, researchers in recent years have come up with asset pricing models that allow for the volatility to be heteroskedastic despite some conditioning it on some prespecified criteria. Allowing for time-varying variance has proven to be a significant improvement to the modeling dynamics of various financial assets and macroeconomic factors, with applications present in the modeling of inflation (Chan, 2013), foreign exchange (Ahlip, 2008), volatility (Maqsood, Safdar, Shafi, & Lelit, 2017) and macroeconomic forecasting (Clark & Ravazzolo, 2015).

Stochastic volatility models specify a stochastic differential equation driven by its own risk process for each of the stock price and volatility equation. While local volatility models are a valuable simplification of the time-varying volatility, they tend to produce a flattened forward implied volatility and have generally shown an inability to price complex structured products close to their market prices. Various studies have suggested the use of hidden Markov models (HMMs) in the estimation of standard and non-standard stochastic volatility models. With regard to capturing regime changes as unobservable Markov chains, their tractability and good econometric properties have promoted for their widespread application in finance.

Hidden Markov models describe the underlying state of the economy by encoding all information in the financial markets into a single (state) process. They have their applications in the determination of financial time series data regimes and amplification of signals for momentum indicators aimed at improving investment decision-making. A hidden Markov model is a statistical Markov chain observed in noise, where the system being modelled is assumed to be a Markov process with unobservable or hidden states and transition probabilities. The observation process is assumed to be strictly Markovian and its behaviour depends only on the finite or infinite state process, which is only partially observable through the observation process.

Hidden Markov models have further applications in econometrics, signal processing, pattern recognition, computational finance and bioinformatics. Signal processing techniques were traditionally exclusively applied in analysing electric signals in engineering, with their applications proving to be general enough to be applied to any field. Where a signal is defined as any sequence in numerical data that varies with respect to an underlying independent variable (mostly time), financial signal processing techniques have been majorly employed in technical analysis by quantitative analysts in robust asset selection and high frequency trading to gain from short random asset market price fluctuations. Essentially, signal processing in finance entails analysis of key behaviours and patterns within financial markets over time for prediction purposes.

The HMM has been employed in diverse fields due to its versatility. In Kenya, for instance, HMMs have been employed in modelling daily rainfall occurrence (Nyongesa, Zeng, & Ongoma, 2020), estimation of malaria symptom data set (Mbete, Nyongesa, & Rotich, 2019), credit scoring for the poor and unbanked financial consumers (Bundi, Patrick, et al., 2016) and calibration and state estimation for the Vasicek term-structure model to the evolution of interest rates (Chelimo, 2017).

This study models stochastic volatility and identifies possible volatility regimes in the Kenyan securities market. The hidden Markov model (HMM) is employed in establishing volatility regimes while the Expected Maximization (EM) algorithm is employed in parameter estimation. Markov chain Monte Carlo (MCMC) and sequential Monte Carlo (SMC) techniques are employed in filtering out noisy observations in parameter estimation. The Nairobi Securities Exchange 20 (NSE20) share index historical price data, for the period January 2012 - February 2021, is employed in calibrating the univariate stochastic volatility (SV) model in this analysis.

1.2 Statement of the Problem

The use of volatility as a proxy for risk has led to an increased emphasis on obtaining good estimates for conditional variance. Although still widely used, market evidence shows that the Black-Scholes model (Black & Scholes, 1973) fails to fully capture the real market dynamics due to some of its overly simplifying assumptions. Of particular mention is the constant volatility assumption in the underlying asset dynamics of which, empirical data points to the existence of non-constant volatility commonly referred to as the volatility smile. One improvement to the Black-Scholes model is stochastic volatility models where volatility is allowed to be heteroskedastic. In addition, there is a general need to capture any underlying (economic) processes that drive the volatility in the economy. To this end, Hidden Markov Models are employed in determining volatility regimes as they are considered simple enough to perform parameter estimation yet rich enough for their applications to be extended to the real world.

1.3 Objectives

1.3.1 General Objective

To model stochastic volatility in the Kenyan securities market using the Hidden Markov Model.

1.3.2 Specific Objectives

1. To extract hidden volatility states for the NSE20 share index under the HMM framework.

2. To perform parameter estimation of the univariate Stochastic Volatility model in the hidden Markov modelling framework.

3. To evaluate the effects of change of states/regimes on parameter estimates.

4. To forecast volatility using the estimated transitional probabilities.

1.4 Rationale of the Study

The purpose of this paper is to model stochastic volatility using a hidden Markov model in the Kenyan economy. The question arises as to how to model volatility in a way that is consistent with the market-observed variation of non-constant volatility together with identifying hidden states that drive the volatility process. This stems from the Black-Scholes pricing framework's assumption that return volatility is constant.

We consider the univariate stochastic volatility model when modelling volatility, as it is simple but flexible enough to incorporate volatility stylized facts together with allowing for non-linear modelling of the latent state space. The HMM is employed in identifying hidden factors and possible volatility regimes in the volatility process. The Markov property of the HMM allows for computationally viable algorithms as well as a general enough structure that can model complex behavior.

In the long run, volatility is a necessary prediction tool for investors in asset pricing and portfolio management. This study seeks to create a theoretical foundation for the application of HMMs in stochastic volatility modelling to the growing financial sector in Kenya and other frontier markets. Possible fields of application include asset allocation, financial valuation, portfolio and risk management.

The rest of this thesis is structured as follows: we review relevant literature on volatility modelling, stochastic volatility models, hidden Markov models, parameter estimation and filtering techniques in the hidden Markov modelling framework in Chapter 2. Chapter 3 provides a mathematical definition of the univariate stochastic volatility model, hidden Markov model, the expected maximization algorithm, Markov chain Monte Carlo and sequential Monte Carlo filtering techniques. Results are presented and interpreted in Chapter 4, as well as a discussion of volatility forecasting, the effects of macroeconomic variables on volatility regime switches, and an analysis of the effects of change of states/regimes on parameter estimates under Markov chain Monte Carlo and sequential Monte Carlo filtering techniques. Chapter 5 provides a conclusion to the study and suggestions for further research.

Chapter 2

Literature Review

2.1 Introduction

Empirical work has evaluated the asymmetry (leverage effects) and volatility clustering in asset returns, first noted by Mandelbrot (1963). Financial time series are mainly characterized by high persistence in the autocorrelation of squared observations and leptokurtosis (Haas, 2009), volatility clustering (Mandelbrot, 1963), non-constant and persistent volatility which eventually exhibits mean-reversion in the long run, and leverage effects (volatility asymmetry). There is a general belief among researchers that alternatives or extensions to the basic Black Scholes model can better accommodate the volatility clustering and fat-tailed distributions of returns.

Equity return volatility in Kenya has been found to be highly persistent with non-significant leverage effects, while the impact of news on volatility is not significantly asymmetric (Nyamongo, Misati, et al., 2010). According to Mekoya (2013), Kenyan stock market volatility was persistent with long memory from 1998 to 2002, although it was unclear whether negative external shocks were more persistent than the positive resulting asymmetric volatility trend. The presence of leverage effects was established by Kirui, Wawire, and Onono (2014), who also discovered that bad news had a significant impact on share price volatility relative to good news. Ndei, Muchina, and Waweru (2019) observed that by taking structural breaks into account, volatility persistence was reduced while leverage effects were found to lead to explosive volatility.

2.2 Stochastic Volatility Models

Stochastic volatility is a concept that deals with the prevalent time-varying volatility as seen in financial markets. Introduced by Taylor (1982), stochastic volatility models provide economical and flexible ways of modelling time-varying volatility of financial asset returns closer to the real world volatility dynamics. Stochastic Volatility models specify a constant drift in the mean asset value for the volatility SDE while the variance is non-constant and reverts to its mean value after a disturbance.

Despite their success in modelling volatility stylized facts, the higher-order multiple integrals of stochastic volatility models have proven to be analytically intractable particularly where volatility is introduced as a latent process. A result of this has been the need for complex computational and inferential modelling techniques. Recent literature has proposed a variety of estimation methods such as the Generalized Method of Moments (Andersen & Sørensen, 1996), Quasi-Maximum Likelihood (Harvey & Shephard, 1996), Simulated Maximum Likelihood (Danielsson, 1994), Efficient Method of Moments (Gallant, Hsieh, & Tauchen, 1997) and approximate Maximum Likelihood (Fridman & Harris, 1998).

Stochastic volatility models, which have extensive applications in option pricing, have also been used to model volatility regimes, with more research being done in this area. Their applications have been extended into modelling both linear and non-linear state space volatility models in various studies (Genon-Catalot, Jeantheau, Larédo, et al., 2000; Rossi & Gallo, 2006; Shimada & Tsukuda, 2005; Pitt, Malik, & Doucet, 2014; Heimbürger, 2016; Malakorn & Iamtan, 2018). Langrock, MacDonald, and Zucchini (2012) applied a numerical maximization on an approximation of the likelihood function, a method that is often efficiently evaluated using recursive techniques applied on hidden Markov models (HMMs).

Goutte, Ismail, and Pham (2017) proposed a semi-analytical expression for the calibration and estimation of S&P 500 and VIX option prices, and later proved that discretely observed SV models could be viewed as HMMs, and that SV models provided a material example of HMMs with a continuous state space. They developed and implemented a MLE extension for a regime-switching SV model by considering a mean-reverting C.I.R. process for the instantaneous variance. They found that their extension allowed for efficient calibration taking advantage of the piecewise affine structure of the volatility process and the good inertia of each small volatility shift rate.

2.3 Hidden Markov Models

HMMs are well-suited for modelling, filtering, classification and prediction of time sequences in a range of partially observable stochastic environments, with a larger range of application in signal processing domains. According to Heimbürger (2016), the two properties that have led to the success of the HMM are the Markov property which allows for computationally viable algorithms and a general enough structure that allows for complex behaviour to be modelled. By studying deposits and withdrawals in modeling credit scores for Kenya's poor and unbanked, Bundi et al. (2016) described HMMs as quick, flexible, and dynamic to changing market conditions.

Thrun and Langford (1998) brought out some deficiencies in earlier HMM algorithms, which assumed discrete state and observation processes, supported compact feature-based state representations and did not provide mechanisms for adapting computational requirements to available resources such as offline computation domains. They presented Monte Carlo HMMs for approximating a large, non-parametric class of density functions, which employed continuous state and observation processes, and which could be used in any time-fashion once trained. Genon-Catalot et al. (2000) proposed explicit contrast functions to replace the intractable likelihood function leading to consistent and asymptotically Gaussian estimators, while acknowledging one area of difficulty in their study; that transitional probabilities of the hidden chain they exhibited were not explicitly known.

Rossi and Gallo (2006) performed their estimation using the GARCH models where they modelled the dynamic regimes in the conditional variance of asset returns across discrete levels as driven by a latent Markov chain. Their study confirmed volatility persistence with their data strongly supporting asymmetric effects and leptokurtic innovations. Moreover, they showed that the Markov chain framework compared relatively well with other approaches in modelling short term volatility as it provided a non-linear framework that could accommodate different treatment of innovations with regard to the size and sign. This translated the notion that remarkable signals in returns are momentary and less persistent relative to small-sized changes.

Langrock et al. (2012) employed a structured HMM to estimate non-standard stochastic volatility models while considering conditionally gamma-distributed volatility for an SV model, where they used dependent and independent mixtures of auto-regressive components as the log-volatility process. Abanto-Valle, Lachos, and Dey (2015) analysed daily stock returns data on the S&P 500 index in their study that saw them approximate a Bayesian estimation method for SV models with scalar mixtures of normal distributions (SMN) based on the proposal of Langrock et al. (2012). They employed the HMM forward algorithm to obtain forecast distributions and based their estimation method on the fact that powerful HMM machinery could be applied in accurate approximation of the likelihood functions of SV model with SMN distributions.

HMMs, as a class of Regime-Switching Models where an unobservable Markov chain drives the switching process, have been reported to understand structural changes in data, therefore allowing for relatively complex modelling techniques and better application to simplistic statistical models (Jha, 2019). A common approach involves having a model's parameters dependent on a Markov Chain, with many but finite states, which encode all information present in the market into a single process that explains the state of the economy. With applications to regime-switching, Jha (2019) concluded that the model was able to capture major uptrends on his five-year $S\&P500$ data. He proposed better features to train the HMM on in order to capture short-term (less than a year) rapid regime switches and momentum trading signals.

Krishnamurthy, Leoff, and Sass (2018) considered both a HMM and a Markov Switching Model (MSM) in continuous time which allowed for derivation of optimal states explicitly while the Markov chain was relevant in controlling the time-varying volatility. HMMs consider constant volatility where volatility jumps with the Markov chain for the MSM, which leads to a difference in behaviour of information between the two models. They established that the MSM had relatively better econometric properties than the HMM which failed to account for some stylized facts such as volatility clustering. However, the MSM failed to achieve continuous-time discretization consistency, which was achieved by the HMM, whose continuous-time optimal strategies were comparable to its discrete-time optimal strategies.

Rossi and Gallo (2006) noted that a commonly encountered problem in Markov-switching models parameter estimation was that the number of parameters became unmanageable when the number of states in the Markov chain increased. They proposed simple parameterization together with considering a small number of states. The presence of regimes in their study was an advantage consistent with the stylized facts of persistence in the behaviour of the time–varying variance of returns.

Various authors have argued in favour of the HMM in SV parameter estimation noting that the approximation was especially convenient for fitting experimental variants of SV models. This is given that simple formulae for HMM forecast distributions, evaluating well-defined residuals, and approximating process volatility are available. Conversely, due to the high dimensional multiple integral of the likelihood, these authors did not hesitate to state that likelihood-based parameter estimation in the SV modeling framework was challenging.

2.3.1 Parameter Estimation

Parameter estimation in HMM is carried out depending on the nature of the states and observations in one's data set, that is, whether discrete or continuous. To achieve practical sub-optimal estimators for discrete states and observations in HMM parameter estimation, Elliott, Moore, and Aggoun (1995) note that recursive estimation is used to update parameters and predictor errors for the data set, resulting in better parameter estimates. Filter and smoother based estimations are used for the state process when only the observation process is continuous, with data smoothing performed recursively to update smoothed estimates.

Given some deficiencies of the likelihood function in obtaining parameter estimates, several iterative methods were proposed. The Expected Maximization Algorithm, used to compute the MLE for latent data models, maximizes a conditional pseudo log-likelihood, whose increase guarantees an increase in the likelihood. The algorithm, which computes estimates from incomplete data with applications to clustered, censored, or truncated data, variance component estimation, and iteratively re-weighted least squares, was introduced by Dempster, Laird, and Rubin (1977).

The EM method generates sequences of estimates that converge towards the true parameter for stochastic state space models and is widely known for its computational ease and numerical stability. Where closed form expressions are available for the M-step besides other optimisation methods, numerical approximations for the EM algorithm result in an analytically intractable E-step beyond cases of linear Gaussian or finite state-space HMM, which necessitates approximations.

Notwithstanding the EM algorithm's advantages, Wu (1983) detected erroneous results from Dempster's study and investigated the convergence properties of the EM algorithm. Wu studied whether the algorithm converged to a local maximum or stationary value of the likelihood function, or whether the sequence of parameter estimates generated converged. He demonstrated that under more practical conditions, such as when the likelihood function was unimodal and a certain differentiability condition was met, multiple but unique convergence results could be obtained. This is supported by Rossi and Gallo (2006) who found a single maximum when starting the maximization procedure from several points of the parameter space.

2.3.2 Filtering Techniques

In HMM, filtering techniques are employed to cancel out certain frequencies while preserving others, with an aim to reduce noise from signals. In state inference, optimal filtering involves estimating hidden states up until a particular time given observations up till the given time, while smoothing involves using future observations to distributions at a certain time. Smoothing has been proven to be relatively computationally challenging but generally leads to smoother trajectory estimates. State inferences are generally made with regard to whether a model is linear or non-linear.

The Kalman Filter and Rauch-Tung-Striebel smoothers were developed in the early

60's, as exact methods to linear state estimation, while the nature of real-world financial time series data to be non-linear led to the development of non-linear Gaussian Filters, Sequential Monte Carlo (SMC) methods and numerical approximations to exact methods such as the Extended, Unscented and Ensemble Kalman Filters. Kalman and Bucy (1961) introduced the Kalman filter, whose solution completely specified the optimal filter for either finite or infinite smoothing intervals, in both stationary and non-stationary statistics.

The Kalman filter turned out to be an exception to the norm as generally, the likelihood, filter and smoothing distributions are almost always analytically intractable. To add to this, linearity was a limiting assumption for the Kalman Filter as in most real-world applications, the state space model is often non-linear. For linearization of certain non-linear models, the Extended Kalman filter was introduced, where dynamics are linearized around the mean and variance. However, a common shortcoming of this approach is introduced where dynamics are severely non-linear. Additionally, to ensure mathematical tractability, the Extended Kalman filter fails to take into account all salient statistical features of processes under consideration resulting in poor results (Douc, Garivier, Moulines, Olsson, et al., 2011).

Wan and Van Der Merwe (2000) proposed the unscented Kalman filter which employs a deterministic approach to choose a subset of carefully selected sigma points around the mean. The sample points are then propagated through non-linear functions, yielding more precise estimates of the true mean and covariance of Gaussian random variables. Julier and Uhlmann (2004) demonstrated substantial performance improvements in terms of accuracy and ease of implementation of the unscented Kalman filter in the context of state estimation of nonlinear control, while Wan and Van Der Merwe (2000) noted that the unscented Kalman filter and extended Kalman filter had the same computational complexities. To add to this, generally, the unscented Kalman filter may lead to serious errors where the state distribution is non-Gaussian (Yang, Xing, Wang, & Tsui, 2016).

While theoretical results did not hold limit for their study, Thrun and Langford

(1998) note that while practical applications are only limited to finite sets, Monte Carlo HMMs are better suited for many real-world applications. This is due to their support of representations of continuous state and observation processes, their built-in mechanisms for model selection and their support of any-time computation, which allows them to be extremely compliant in time-critical applications with bounded computational resources. The MCMC and SMC methods are both classes under Bayesian inference in general state space models which has its framework conditioned on observed data (Andersson, 2019).

SMC methods have proven to be relatively more flexible, easy to implement, parallelisable and more applicable in general settings with general improvements in applied statistics, engineering and probability accounting (Douc et al., 2011). SMC algorithms infer state variables through a general representation of probability distributions by performing *importance sampling* to generate a sequence of weighted random numbers or particle samples, whose distribution converges asymptotically to the sample posterior distribution. Applying the *re-sampling step* in the filtering process is finally performed in order to eliminate samples with relatively smaller (near zero) weights. Heimbürger (2016) suggests using SMC methods to approximate the smoothing distribution and proposes averaging where errors associated with the smoothed sufficient statistics are present.

Extensions of the SMC in higher-order HMM were carried out in nonlinear state space models by Allaya, Coulibaly, Mouhamadou, Sene, et al. (2019), who mimicked the prediction step and update step in derivation of the filter distribution, while noting that growth of the SMC method has been limited by its computational intensiveness, which limit them to HMMs of unit order. Malakorn and Iamtan (2018) were also keen to mention this drawback of the particle filters as being time-consuming particularly due to the re-sampling step.

Where the SMC method is an inference method that efficiently estimates time-dependent hidden variables and involves a collection of multiple samples approximately distributed according to the posterior, MCMC iteratively transform a single sample thus representing inference using posterior draws over time. MCMC sampling for HMMs involves latent states underlying a noisy observation process from which the MCMC method then generates posterior samples for the HMM parameters given observed data (Turek, de Valpine, & Paciorek, 2016). With their application in Bayesian statistics, computational physics, biology and linguistics, MCMC algorithms have been employed in sampling from probability distributions by recording states from the Markov chain. MCMC methods have been established to have a profound influence on computational statistics, more so in the Bayesian paradigm.

A MCMC sampler alternates between sampling the parameters conditional on the observed data and the hidden Markov chain, and updating the hidden chain conditional on the data and parameters (Rydén et al., 2008). Endo, van Leeuwen, and Baguelin (2019) remarked that the MCMC is widely utilized to efficiently sample from a high dimensional parameter space particularly where the likelihood of the observed data was available in a functional form. Andersson (2019) investigated convergence of the MCMC to the target posterior distribution, noting a guarantee of convergence as the simulation draws approached infinity, a weak guarantee of convergence for finite simulation draws and divergence where the parameter space was hard to explore. Endo et al. also noted that high correlation between hidden variables in time series data substantially slowed down the mixing of the MCMC.

Andersson further established slow convergence where the models were complicated, attributing it to the random walk behaviour and conditional sampling of the algorithms. He suggested less complex models or more complicated algorithms such as the Hamiltonian Monte Carlo (HMC) as solutions to the problem of slow convergence. The Gibbs sampling and Metropolis-Hastings are known and easily implemented MCMC algorithms based on conditional sampling. Scott (2002) noted that the statistical efficiency of the MCMC output could be improved through incorporating established recursive algorithms, the most important being the forward-backward algorithm, and went ahead to prove that the Gibbs sampler using the forward-backward recursions mixed rapidly relative to other samplers used for HMM.

Chapter 3

Methodology

3.1 Research Design

This is a quantitative research as statistical inference will be used to draw conclusions on the study. The research is aimed at establishing possible states and regimes in stochastic volatility in the Kenyan Securities market, where secondary data from the NSE20 share index is employed to calibrate the univariate SV model. The underlying states are the independent variables while volatility is the dependent variable. Parameter estimation is performed using the iterative Expectation Maximization algorithm. Sequential Monte Carlo (SMC) and Markov Chain Monte Carlo (MCMC) are employed in filtering out noisy data in parameter estimation. Under the established state models, an analysis of the impact of changes in states on the parameter estimates of the SV Model is performed.

3.2 The Stochastic Volatility Model

Volatility can be modelled probabilistically through state-space models where the logarithm of squared volatilities (latent states) follows an AR (1) process (Andersen, Davis, Kreiß, & Mikosch, 2009; Kastner, 2016). The log return of asset price, P_t, is defined by $r_t = log\frac{P_t}{P_{t-1}}$ and admits the univariate SV model:

$$r_t = \sigma_t \epsilon_t, \tag{1}$$

where ϵ_t is a Gaussian white noise process with zero mean and unit variance, and $\sigma_t > 0$ is a stochastic process representing the volatility of r_t. Based on empirical results, σ_t^2 follows a log-normal distribution. As such, there exists a random variable $z_t = \log \sigma_t^2$, which is normally distributed, and which reduces equation (1) to:

$$r_t = \exp(\frac{z_t}{2})\epsilon_t. \tag{2}$$

Traditionally, z_t is assumed to follow a first order AR(1) process with Gaussian innovation white noise:

$$z_t = \phi z_{t-1} + c + \omega_t, \tag{3}$$

where ϕ and c are constants, while $\omega_t \sim N(0, Q)$ and ϵ_t are mutually independent. Additionally, if $|\phi| < 1$, the above process is wide-sense stationary.

When a scaling factor, β, is be introduced to equation (2) to remove the constant term c in equation (3), we have the canonical SV model for the log return as:

$$z_t = \phi z_{t-1} + \omega_t, \tag{4}$$

$$r_t = \beta \exp(\frac{z_t}{2})\epsilon_t, \tag{5}$$

with initial state z_0.

Equations (4) and (5) are the state model and the measurement/observation model, respectively. They are a particular type of stochastic non-linear state space models, where z_t is the unobserved state variable commonly interpreted as the latent time-varying volatility process, r_t is the output of the model which is the return process in this case, ω_t and ϵ_t are considered as the process noise and the measurement noise respectively, and $\theta = \{\phi, Q, \beta\}$ are the model parameters.

Squaring equation (5) and introducing a logarithm reduces the equation to:

$$x_t = \alpha + z_t + v_t, \tag{6}$$

where:

$$x_t = \log(r_t^2),$$

$$\alpha = \log(\beta^2) + E[\log(\epsilon_t^2)],$$

$$v_t = \log(\epsilon^2 - E[\log(\epsilon_t^2)]) \sim \log(\chi_1^2) - E[\log(\chi_1^2)].$$

This leaves us with the univariate SV model:

$$z_t = \phi z_{t-1} + \omega_t, \tag{7}$$

$$x_t = \alpha + z_t + v_t, \tag{8}$$

where $w_t \sim N(0, Q)$ and $v_t \sim \log(\chi_1^2) - E[\log(\chi_1^2)]$.

The vector of parameters is $\theta = \{\alpha, \phi, Q\}$ representing the level of log-variance, persistence of log-variance and volatility of log-variance respectively. To get initial parameter estimates, the method of moments as suggested by (Kim, 2005) is employed. It generates consistent estimates for a linear system:

$$\alpha^{(0)} = \frac{\sum_{t=0}^{n} x_t}{n},$$

$$\phi^{(0)} = \frac{\sum_{t=3}^{n}(x_t - \overline{x}_t)(x_{t-2} - \overline{x}_{t-2})}{\sum_{t=3}^{n}(x_{t-1} - \overline{x}_{t-1})(x_{t-2} - \overline{x}_{t-2})},$$

$$Q^{(0)} = \frac{\sum_{t=2}^{n}(x_t - \overline{x}_t) - \phi(x_{t-1} - \overline{x}_{t-1})}{n} - \hat{var}(v_t) - \hat{\phi}^2 \hat{var}(v_{t-1}).$$

3.3 Hidden Markov Model

We consider a bivariate probabilistic HMM that consists of a state process z_k, which is discrete and takes its values from some finite set $\mathbf{Z}$ equipped with a countably generated σ-algebra Ω, and the observation process x_k from another finite space $\mathbf{X}$ with corresponding σ-algebra Λ, such that ($\mathbf{Z}$,Ω) and ($\mathbf{X}$, Λ) are measurable spaces. z_k are considered to be latent/hidden variables governed by $Pr(x_k|z_k)$ and $Pr(z_k|z_{k-1})$.

The joint distribution of x_n and z_n is such that:

$$Pr(x_1, ..., x_n, z_1, ..., z_n) = Pr(z_1)Pr(x_1 \mid z_1) \prod_{k=2}^{n} Pr(z_k \mid z_{k-1})\, Pr(x_k \mid z_k),$$

with stationary transitional probabilities:

$$T_{(i,j)} = Pr(z_{k+1} = j \mid z_k = i) \qquad \forall i, j \epsilon (1, ..., n),$$

emission probability:

$$\xi_i(x) = Pr(x_k \mid z_k = i) \qquad \forall i \epsilon (1, ..., n),$$

and an initial distribution:

$$\pi(i) = Pr(z_1 = i) \qquad \forall i \epsilon (1, ..., n).$$

These reduce the joint distribution to:

$$Pr(x_1, ..., x_n, z_1, ..., z_n) = \pi(i) \xi_{z_1}(x_1) \prod_{k=2}^{n} T(z_{k-1}, z_k) \xi_{z_k}(x_k).$$

Following the formal definition of (Bickel, Ritov, Ryden, et al., 1998), the stochastic process $(z_n, n \geq 1)$ with a state space $(\mathbf{Z}, \Omega)$ is a hidden Markov model if the following hold:

1. We are given, but do not observe, a strictly stationary Markov Chain $z_1, ..., z_n$ with state space $(\mathbf{Z}, \Omega)$.

2. For all n given $(z_1, ..., z_n)$, the $(x_i, i = 1, ..., n)$ are conditionally independent and the conditional distribution of x_i depends only on z_i.

3. The conditional distribution of x_i given z_i does not depend on i.

We assume that the process $(x_i, i \geq 1)$ is strictly stationary, and if the hidden Markov chain $(z_i, i \geq 1)$ is ergodic, then $(x_i, i \geq 1)$ is also ergodic.

Expectation - Maximization Algorithm

The EM algorithm is employed in finding the (local) maximum likelihood estimates for parameters where the model depends on unobserved latent variables and where equations cannot be solved directly. The EM iteration alternates between an expectation (E) step, which creates a function for the expectation of the log-likelihood of the complete data over the smoothing distribution, which is evaluated using the current parameter estimates, and a maximization (M) step, which re-estimates parameters by maximizing the expected marginal log-likelihood found on the E-step. These parameter estimates are then used to determine the distribution of the latent variables in the next E-step.

Given the complete data-set, $\{x, z\}$, of observed data and unobserved data respectively, and a vector of unknown parameters, θ, along with a likelihood function, $L(\theta; x, z) = p(x, z \mid \theta)$, then:

E – Step:

The E-step of the EM algorithm computes the expected value of $L(\theta; x, z)$ given the observed data, x, and the current parameter estimate, θ_t. In particular, we define:

$$\begin{aligned} Q(\theta; \theta_t) &:= E[L(\theta; x, z) \mid x, \theta_t], \\ &= \int L(\theta; x, z) p(z \mid x, \theta_t)\, dz, \end{aligned}$$

where $p\,(z \mid x, \theta_t)$ is the conditional density of z given the observed data, x, and assuming $\theta = \theta_t$.

M – Step:

The M-step consists of maximizing over θ the expectation step computed above. That is, we set:

$$\theta_{t+1} = arg_{\theta}\, max(\theta\, |\theta_t).$$

We then set $\theta_t = \theta_{t+1}$. This is repeated up until θ_t converges, with convergence to a local maximum guaranteed. The idea is to first initialize the parameters θ to some (random) variables, compute the conditional probability $Pr(z|\theta)$, then use the computed values of z to compute better estimates for the parameters θ. This is performed iteratively until convergence, which is achieved when parameter estimates become stable and no further improvements can be made to the likelihood value. The ML estimate of θ is then taken as the best estimate of the local maxima.

Filtering Techniques

Markov Chain Monte Carlo

Given a collection of observations $x_{1:t} := (x_1, ..., x_t)$, inference is made with regard to the parameter set, θ, and the states, $z_{1:t} := (z_1, ..., z_t)$. In the Bayesian framework, inference relies on the posterior density:

$$p(\theta, z_{1:t}|x_{1:t}) = p(\theta|x_{1:t})p_\theta(z_{1:t}|x_{1:t}),$$

where,

$$p(\theta|x_{1:t}) = \frac{p_\theta(x_{1:t})p(\theta)}{p(x_{1:t})}, \qquad p_\theta(z_{1:t}|x_{1:t}) = \frac{p(z_{1:t}, x_{1:t})}{p_\theta(x_{1:t})},$$

with,

$$p_\theta(z_{1:t}, x_{1:t}) = \varphi_\theta(z_1) \prod_{n=2}^{t} f_\theta(z_n|z_{n-1}) \prod_{n=1}^{t} g_\theta(x_n|z_n).$$

It is feasible to design efficient MCMC algorithms for linear Gaussian and finite state space models where it is possible to sample from $p_\theta(z_{1:t}|x_{1:t})$ and compute $p_\theta(x_{1:t})$. Common practice is to build MCMC kernel updating subblocks as:

$$p_\theta(z_{n:n+K-1}|x_{n:n+K-1}, z_{n:n+K-1}) \propto \prod_{m=n}^{n+K} f_\theta(z_m|z_{m-1}) \prod_{m=n}^{t} g_\theta(x_m|z_m).$$

The prior distribution of the parameter vector is specified by choosing independent components of each parameter, that is, $p(\theta) = p(\alpha)p(\phi)p(Q)$. $\alpha \in \mathbb{R}$ is equipped with a normal and uninformative prior $\alpha \sim N(b_\alpha, B_\alpha)$ where it is common practice to set

$b_\alpha = 0$ and $B_\alpha \geq 100$ for daily log-returns. For $\phi \in [-1, 1]$, $(\phi + 1)/2 \in \mathbb{B}(a_0, b_0)$ where a_0 and b_0 are hyperparameters and $\mathbb{B}(x, y)$ denotes the beta function. As for the volatility of the variance process, $Q \in \mathbb{R}^+ \sim \mathbb{B}_Q * \Xi_1^2 = \mathbb{G}(1/2, 1/2\mathbb{B}_Q)$ or an equivalent centered normal distribution $\pm\sqrt{Q} \sim N(0, \mathbb{B}_Q)$ (Frühwirth-Schnatter & Wagner, 2010).

Joint (batched) sampling of all instantaneous volatilities through the "all without a loop" (AWOL) is a key feature that enables for a significant reduction in the correlation of the draws. Complex models such as volatility models are such that it is possible to sample from the prior only as a result of which the question arises on how to efficiently evaluate a MCMC sampler point-wise on these models. The SMC methods provide approximations for $p_\theta(z_{1:t}|x_{1:t})$ and $p_\theta(x_{1:t})$ sequentially, and are considered for general state HMMs.

Sequential Monte Carlo

SMC filters aim to estimate, recursively in time, the posterior distributions of hidden states of a Markov Process given some noisy and partial observations. They employ a set of particles (samples) to represent the posterior distribution and often assume the states, z_k, and the observations, x_k, can be modelled in the form:

- The state process, $\{z_o, z_1, ...\}$, is modelled as a Markov process on $\mathbb{R}^{d_z}$ (for some $d_z \geqslant 1$), with initial distribution $Pr(z_0)$ that evolves according to the transition probability density, $Pr(z_k|z_{k-1})$.

- The observations, $\{x_0, x_1, ...\}$, take values in some state space on $\mathbb{R}^{d_x}$ (for some $d_x \geqslant 1$) and are conditionally independent provided that $z_0, z_1, \cdots$ are known.

Under Bayes' Rule for conditional probability, we have the non-linear filtering equation:

$$Pr(z_0, ..., z_k|x_0, ..., x_k) = \frac{Pr(x_0, ..., x_k|z_0, ..., z_k)Pr(z_0, ...z_k)}{Pr(x_0, ..., x_k)},$$

where:

$$Pr(x_0, ..., x_k) = \int Pr(x_0, ..., x_k | z_0, ..., z_k) Pr(z_o, ..., z_k) dz_0 ... dz_k,$$

$$Pr(x_0, ..., x_k | z_0, ..., z_k) = \prod_{i=0}^{k} Pr(z_i | z_{i-1}),$$

$$Pr(z_0, ..., z_k) = Pr_0(z_0) \prod_{i=0}^{k} Pr(z_i | z_{i-1}).$$

The non-linear filtering problem involves computing the conditional distribution, $Pr(z_k | x_0, ..., x_{k-1})$, sequentially and the filtering equation is given by the recursion:

$$\xrightarrow{\text{updating}} Pr(z_k | x_0, ..., x_k) = \frac{Pr(x_k | z_k) Pr(z_k | x_0, ..., x_{k-1})}{\int Pr(x_k | z_k') Pr(z_k' | x_0, ..., x_{k-1}) dz_k'},$$

$$\xrightarrow{\text{prediction}} Pr(z_{k+1} | x_0, ..., x_k) = \int Pr(z_{k+1} | z_k Pr(z_k | x_o, ..., x_k) dz_k$$

We assume $Pr(x_k | z_k)$ is continuous with the convention $Pr(z_0 | x_0, ..., x_k) = Pr(z_0)$ for $k = 0$.

Chapter 4

Data Analysis and Results

4.1 Data Description

This study uses secondary data, recorded daily, for the period January 2012 to February 2021[1] from the Nairobi Securities Exchange 20 (NSE20) Share index. The NSE20 is the long-standing benchmark index in the exchange that tracks the performance of twenty best performing (large) companies listed on the Nairobi Securities Exchange and represents the geometric mean of the constituent companies' share prices. These firms are listed firms based on quantitative financial merit and represent 80% of the market capitalization and are therefore a good proxy for measuring the study variables. The study period should be able to capture the influence of various macroeconomic and political factors on regime changes on the share prices and return volatility. Figure 1 gives a plot of the daily recorded prices.

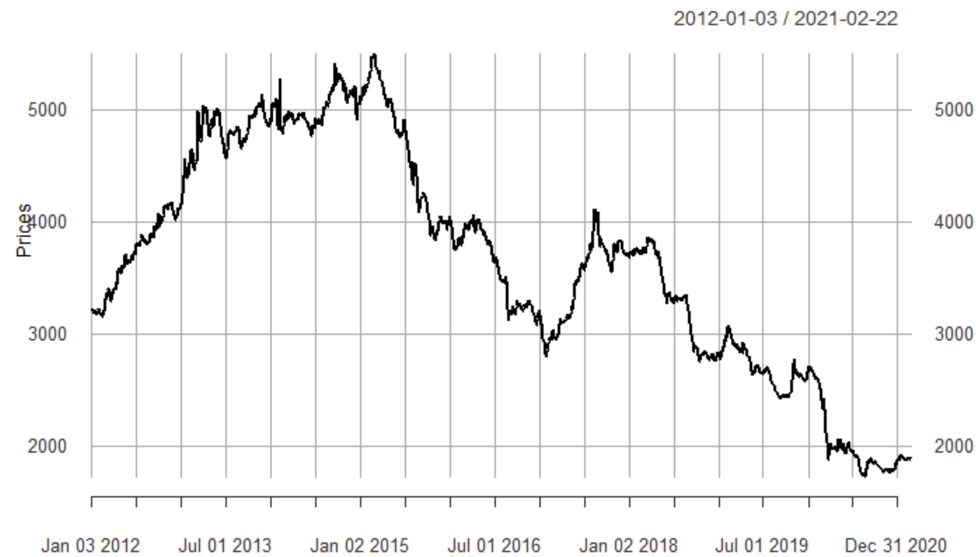


Figure 1: NSE20 daily share prices for the period January 2012 to February 2021

The figure shows that the daily prices for the NSE20 share index fluctuated through years. This is an indication of non-constant stock market performance in Kenya with the trend showing an increase from January 2012 to the March 2015. Beyond that the

[1]Secondary data obtained from `https://www.investing.com/`

prices have been decreasing with a one time increase from the first quarter of 2017 to the third quarter of the same year. For the study period, the highest price recorded was Ksh. 5499.64 and the lowest price Ksh. 1723.96.

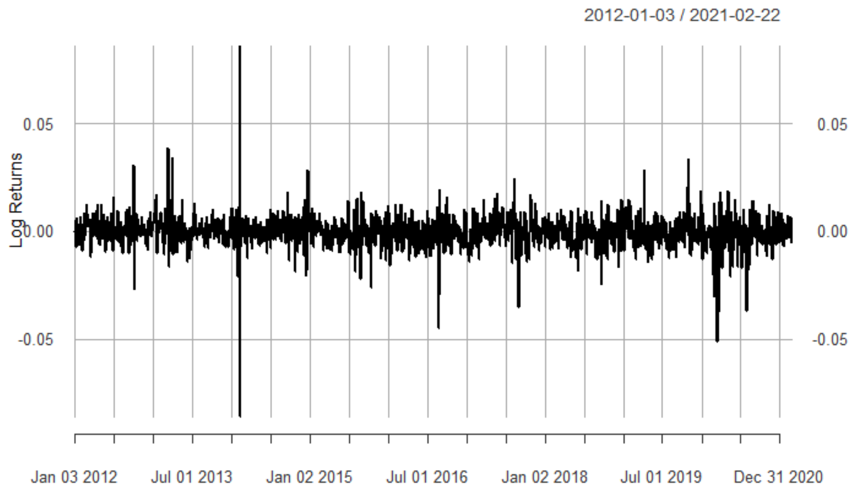


Figure 2: Log returns for NSE20 share index daily price data for the period January 2012 to February 2021

The demeaned log returns are seen to vary around a mean value of averagely 0.00, with periods of high and low variance as shown in Figure 2. The volatility plots in Figure 3 and Figure 4 show periods of high volatility and periods of low volatility. The greatest variance can be observed in early 2014 in the plot of log returns, a trend that is also visible in the rolling volatility and latent volatilities quantile plots below.

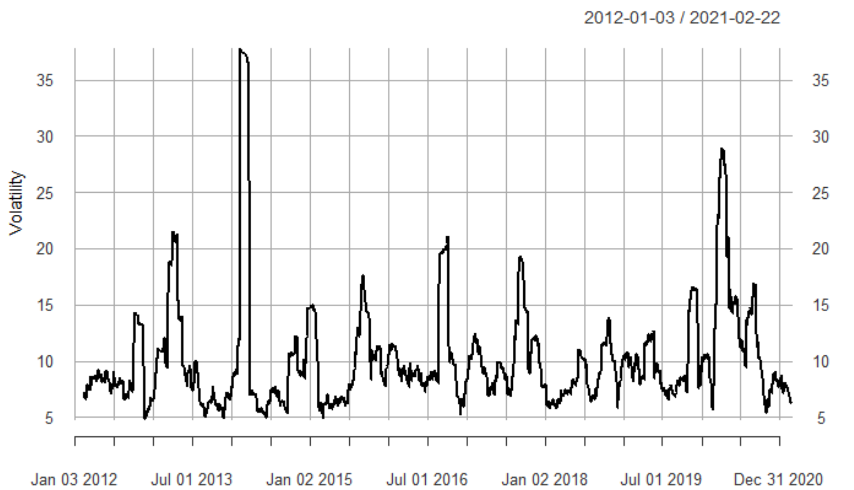


Figure 3: Rolling volatility for daily price data for the period January 2012 to February 2021

The volatility evolves continuously. In Figure 4, the volatility estimates are approximated using initial parameter estimates $\alpha = -0.0002, \phi = 0.75900$ and $Q/\sigma = 0.1$. The volatility estimates are based on .05, .5 and .95 quantiles of latent volatility.

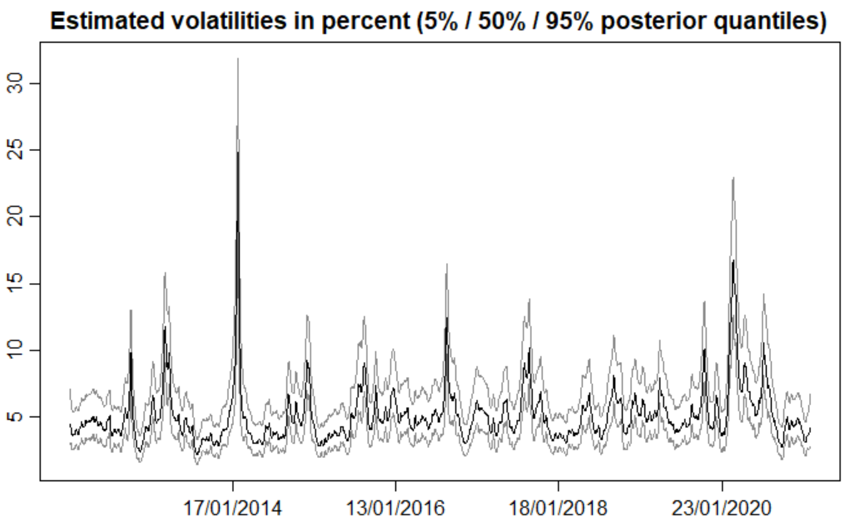


Figure 4: Quantile plot of latent volatilities for daily price data for the period January 2012 to February 2021

4.2 Determination of Regimes

The HMM is fitted on rolling volatility data. Results for the implementation of a 2-state, 3-state and 4-state HMM filter using the R-library *depmixS4()*, whose codes are available under Appendix IV, are discussed below.

4.2.1 2-Regime Model

The two-state HMM allows for high and low volatility without any additional differentiation. Based on the findings, the 2-state process assumes that the log-volatility process starts at state 1 where state 2 represents high volatilities while state 1 represents low volatilities. The 2-state HMM converges at 25 iterations with 7 degrees of freedom, a log-likelihood of -4742.109, AIC value 9498.217 and BIC value 9538.23.

Table 1: Transition probability matrix for the 2-state HMM

	to state 1	to state 2
from state 1	0.989	0.011
from state 2	0.027	0.973

The transitional matrix for a 2-state fitted HMM model, as seen in Table 1, depicts stable states, that is, a higher probability of remaining in the same state given by 0.989 and 0.973 for state 1 and state 2 respectively. There is a higher probability of staying in state 1 relative to state 2. There is a higher transitional probability from state 2 to state 1, 0.027, compared to the probability of moving from state 1 to state 2, 0.011.

Table 2: Descriptive statistics for the 2-state HMM

	Re1.(intercept)	Re1.sd
State 1	6.926	1.261
State 2	13.591	5.323

The descriptive statistics for parameter estimates of the Gaussian response variables are presented in Table 2. State 1 has a lower mean (intercept) value of 6.926 and a lower standard deviation 1.261 where state 2 has a higher mean value 13.591 and high standard deviation value 5.323. This shows that the resulting model has two well-separated states, where state 2 has fast responses while state 1 has slow responses. Figure 5 is a graphical representation of state changes for the 2-state HMM model.

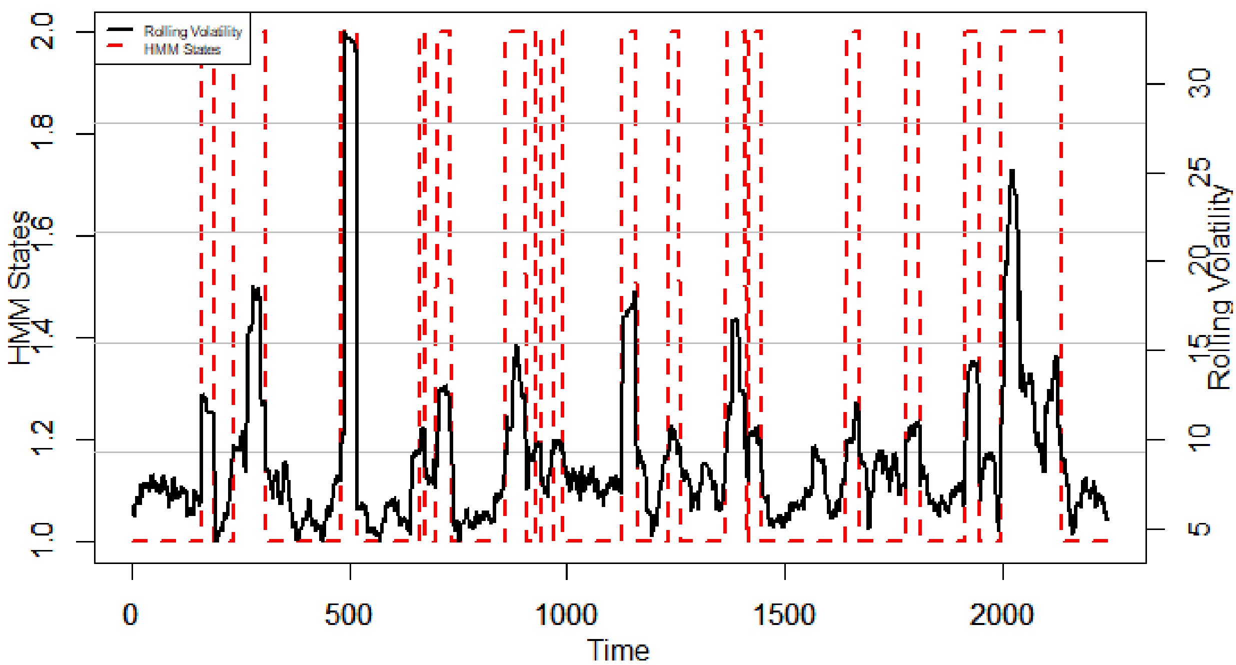


Figure 5: State changes for a 2-state HMM applied to rolling volatility data for the period January 2012 to February 2021

4.2.2 3-Regime Model

The three-state HMM provides for a high volatility regime, a low volatility regime and a transitional regime bridge between the high and low volatility regimes. The 3-state process assumes that the volatility process begins at state 1, where state 1 represents low volatilities, state 2 represents high volatilities and state 3 represents the transitional regime bridge between state 1 and state 2. The 3-state HMM converges at 52 iterations with 14 degrees of freedom, a log-likelihood of -3876.486, AIC value 7780.972 and BIC value 7860.997.

Table 3: Transition probability matrix for the 3-state HMM

	to state 1	to state 2	to state 3
from state 1	0.978	0.000	0.022
from state 2	0.003	0.971	0.025
from state 3	0.019	0.011	0.970

Given Table 3, there is a higher likelihood of remaining in all states, implying stable states. The probability of staying in state 1, 0.978, is the highest whereas there is no probability of moving from state 1 to state 2 with an equally lower probability, 0.003, of moving from state 2 to state 1.

Table 4: Descriptive statistics for the 3-state HMM

	Re1.(intercept)	Re1.sd
State 1	5.999	0.739
State 2	15.808	5.665
State 3	8.527	1.057

Table 4 gives the descriptive statistics for the Gaussian response variables. State 2 has the highest response rate given that it has the highest response parameter estimates, while state 1 has the lowest response rate. The graphical representation of state changes for the 3-state HMM model are as shown in Figure 6 below:

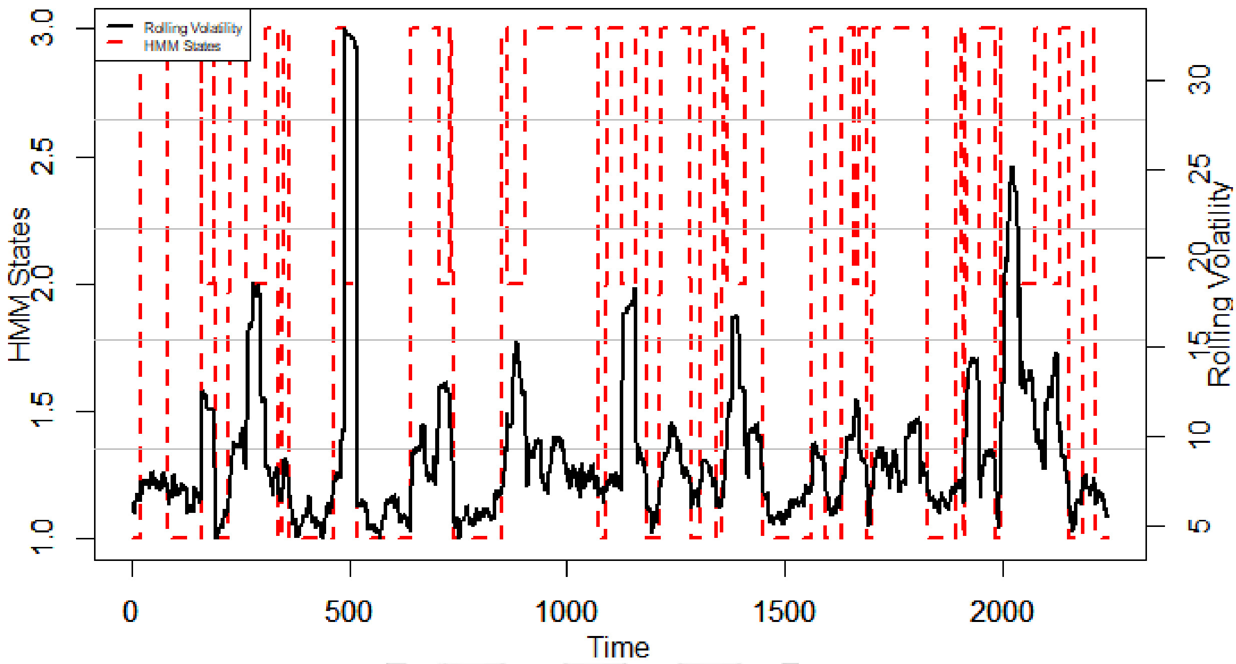


Figure 6: State changes for a 3-state HMM applied to rolling volatility data for the period January 2012 to February 2021

4.2.3 4-Regime Model

The four-state HMM provides for very high and very low volatility regimes, and two intermediate states representing high and low volatility regimes respectively. Given the findings, the 4-state process assumes that the log-volatility process starts at state 1. State 1 represents very low volatilities while state 4 represents low volatilities. State 2 represents high volatilities while state 3 represents very high volatilities. The 4-state HMM converges at 67 iterations with 23 degrees of freedom, a log-likelihood of -3223.031, AIC value 6492.063 and BIC value 6623.531. The transition matrix is given in Table 5.

Table 5: Transition probability matrix for the 4-state HMM

	to state 1	to state 2	to state 3	to state 4
from state 1	0.973	0.000	0.000	0.027
from state 2	0.002	0.943	0.019	0.036
from state 3	0.005	0.021	0.971	0.003
from state 4	0.019	0.029	0.002	0.950

There is a higher likelihood of staying in the same state. The highest probability is in staying in state 1 which represents very low volatility while the probability of moving from state 1 to state 2 or state 3 is zero. The parameter estimates of the Gaussian response variables have their descriptive statistics as shown in Table 6.

Table 6: Descriptive statistics for the 4-state HMM

	Re1.(intercept)	Re1.sd
State 1	5.653	0.575
State 2	9.424	0.746
State 3	16.031	5.688
State 4	7.336	0.521

State 3 has the highest response rate given its high intercept and standard deviation values, while state has the lowest response rate. The 4-state HMM model's state changes are represented graphically by Figure 7.

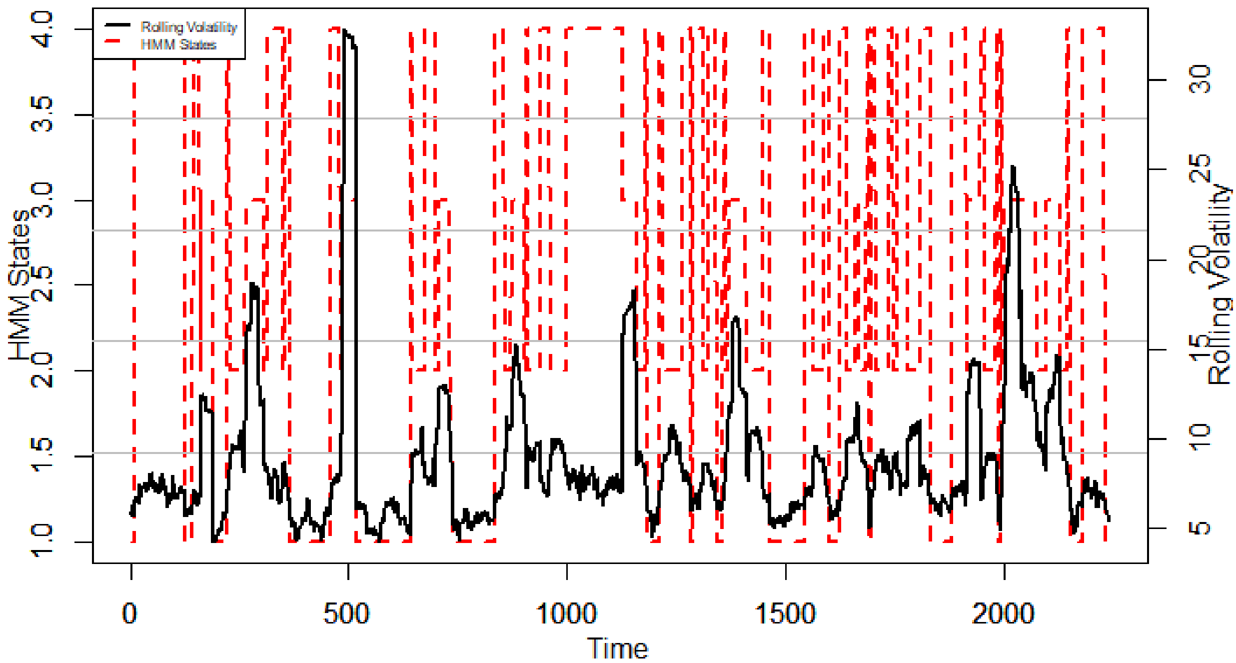


Figure 7: State changes for a 4-state HMM applied to rolling volatility data for the period January 2012 to February 2021

It has been empirically observed that the log-likelihood increases with increased number of states, which is a result of increased number of parameters to be estimated. The same is evidenced in this study, given the resultant log-likelihood values -4742.109, -3876.486 and -3223.031 for the 2-state, 3-state and 4-state model respectively. Due to the shortcomings of the log-likelihood in state estimation, we rely on the information criteria to determine the optimal states. Information criteria reward goodness of fit but impose a penalty that is an increasing function of the number of estimated parameters and which discourages model over-fitting. Based on the BIC, which imposes a greater penalty for additional parameters, the 4-state economy is optimal with the lowest BIC value of 623.531, relative to the 9538.23 and 7860.997 values for the 2-state and 3-state respectively.

4.3 Parameter Estimation

4.3.1 Markov Chain Monte Carlo

The single regime model assumes that there are no regimes in the economy. As such, the model's input is all the data set used in the study, that is, a numeric vector of squared log returns without any missing values. Given the definitions of Kim (2005) the initial parameter estimates are as: $\theta = \{\ \alpha \sim N(mean(x_t), var(x_t))$, $(\phi+1)/2 \sim \mathbb{B}(a_0, b_0)$, $Q = 0.1\}$[2]. With implementation codes available under Appendix IV, the parameter estimates obtained are:

Table 7: Parameter estimates for the single state SV model under MCMC

Parameter	Mean	Standard Deviation	ESS
μ	-0.00037	0.03568	9301
ϕ	0.99947	0.00031	2290
σ	0.24462	0.02407	105
Total (n)	2273		
Iterations per second	626		
Time Elapsed (in seconds)	17.58		

There is low deviation from the mean values of the parameter estimates as shown by the standard deviation values. The effective sample size (ESS) is a metric for determining how well a converged Markov chain traversed the posterior space. Intuitively, it's the number of independently distributed and identically distributed draws, with higher values indicating better mixing. Figure 8 summarizes the resultant simulation

[2] μ and α are used interchangeably and represent the level of log-variance while both Q and σ represent the volatility of the log variance process.

density plots of the Markov chains of the parameters where the grey-dashed lines and black-solid lines represent prior and posterior densities respectively.

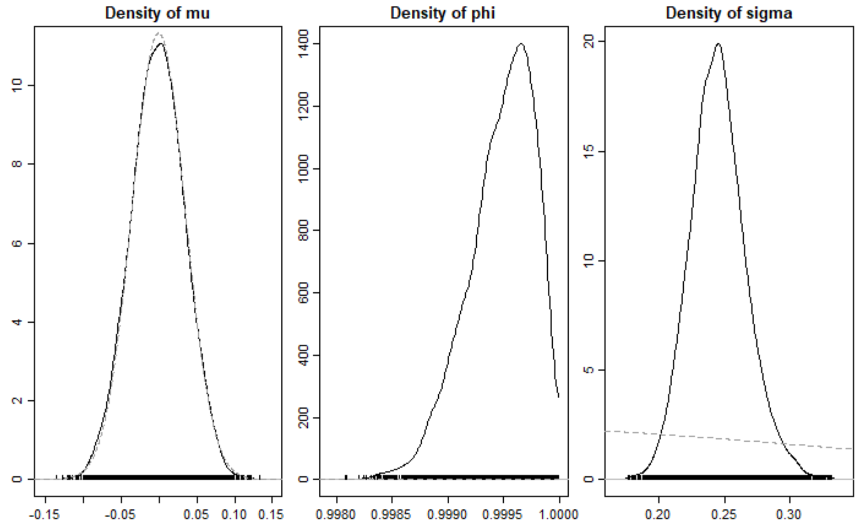


Figure 8: MCMC density plots of parameter posteriors for the single state univariate stochastic volatility model

The multi-regime SV model assumes 2-state, 3-state and 4-state volatility regimes with parameter estimates under each regime presented and discussed below. The simulation density plots of the Markov chains of each parameter for state 2, state 3 and state 4 are summarized in Figure 11, Figure 12 and Figure 13 under Appendix I respectively.

2-State Model

The 2-State model assumes that there are two regimes in the economy, where state 1 represents low volatility and state 2 represents high volatility as discussed under the determination of regimes section. Parameter estimates for the 2-state model under the MCMC filtering technique are presented in Table 8.

Table 8: Parameter estimates for the 2-state SV model under MCMC

Parameter	State 1	State 2
μ	-0.0526	-0.0004
ϕ	1.0000	0.9978
σ	0.4047	0.4796
Total (n)	1640	632

State 1 has a lower mean value, μ, and higher variance of the volatility process, σ, relative to state 2 in the 2-state HMM. The ϕ values for the two states have low variance from each other and are as $|\phi| < 1$ which shows that both processes are stationary in the weak sense.

3-State Model

The 3-state process assumes that state 1 models low volatilities, state 2 represents high volatilities and state 3 represents the transitional regime bridge between state 1 and state 2. Parameter estimates for each state process in the 3-state model are presented in Table 9.

Table 9: Parameter estimates for the 3-state SV model under MCMC

Parameter	State 1	State 2	State 3
μ	-0.0002	-0.0007	-0.0162
ϕ	0.9999	1.0000	1.0000
σ	0.0325	0.6199	0.3679
Total (n)	908	396	968

State 1 has the highest mean value while state 3 has the lowest mean value. On the contrary, state 1 has the highest variance of the volatility value while state 2 has the lowest variance of the volatility process value. The ϕ values for the different states have low variance from each other and are as $|\phi| < 1$ indicating that all three processes are stationary.

4-State Model

Given the findings under regime determination, the 4-state process assumes that state 1 represents very low volatilities, state 4 represents low volatilities, state 2 represents high volatilities while state 3 represents very high volatilities. Parameter estimates for each of the 4 states are presented in Table (10).

Table 10: Parameter estimates for the 4-state SV model under MCMC

Parameter	State 1	State 2	State 3	State 4
μ	-0.0559	-0.0001	-0.0006	-0.0606
ϕ	1.0000	1.0000	0.9996	1.0000
σ	0.3248	0.2482	0.7441	0.4195
Total (n)	660	512	377	723

State 2 has the highest mean value while state 4 has the lowest mean value. State 3 has the lowest variance of the volatility value while state 2 has the highest volatility value. All ϕ values for the 4 states have low variance from each other and are as $|\phi| < 1$ which implies that all four processes are stationary.

4.3.2 Sequential Monte Carlo/Particle Filters

Initial parameters for the single-regime model are as $\alpha \sim N(mean(x_t), var(x_t))$, $(\phi + 1)/2 \sim \mathbb{B}(a_0, b_0)$ and $Q = 0.1$. The algorithm takes time to converge for larger data sets. In theory, it converges where the particle size and number of iterations is sufficiently large. However, in practice, it is not feasible to use infinitely large particle sizes and/or number of iterations. As a result, it is of importance to select an appropriate particle size and number of iterations to implement the particle filters algorithm. The parameter estimates are as shown in Table 11 with implementation codes, as available under Appendix IV, evaluated per maximum number of iterations $M = length(observations)$, and the number of particles $= length(observations)$.

Table 11: Parameter estimates for the single state SV model under SMC

Parameter	Parameter Estimates
μ	-0.00021
ϕ	0.9966
σ	0.2528
Total (n)	2273
ESS	2046.693

The single state process takes on all the observations in the study. The ϕ value, 0.9966 is less than one which shows that the process is stationary. Figure 9 shows sequential summaries of the parameter posteriors, with each panel plotting the (.05, .50, .95) posterior quantiles for the given parameter and the true parameter values are represented by the mid line. The density plots of the parameter estimates derived from the posterior distribution are shown in Figure 10.

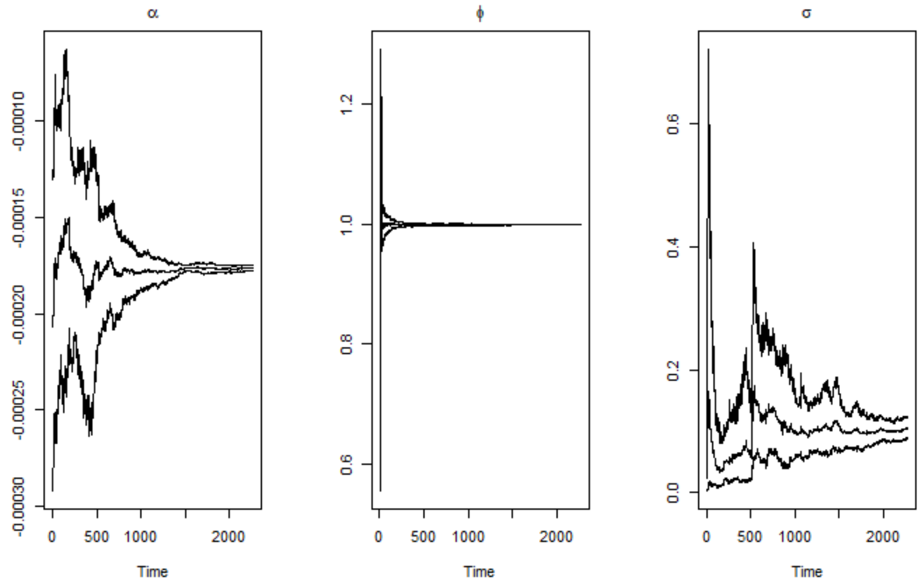


Figure 9: SMC convergence of parameter estimates for the single-state univariate stochastic volatility model

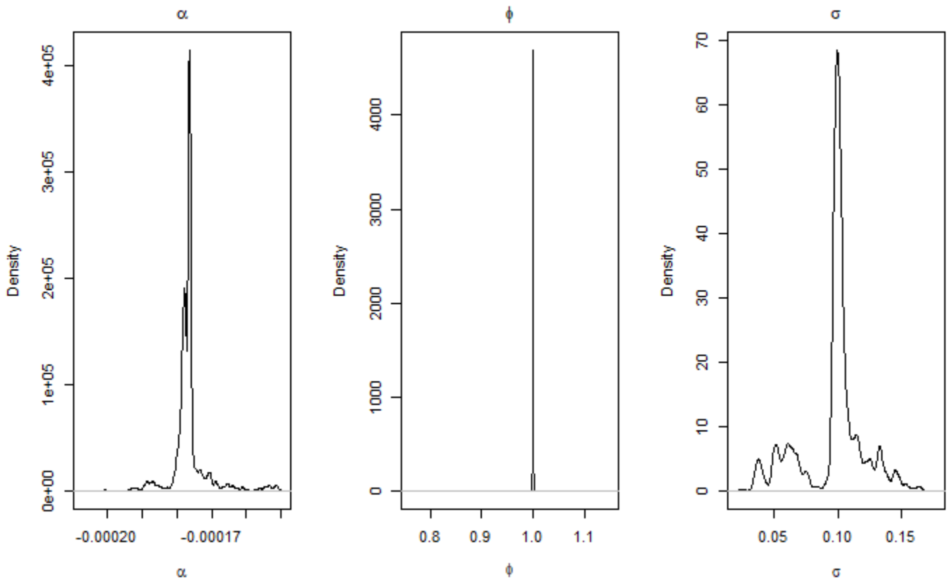


Figure 10: SMC density plots of posterior parameter densities for the single state univariate stochastic volatility model

The multi-regime SV model assumes 2-state, 3-state and 4-state volatility regimes with parameter estimates under each regime presented and discussed below. Figure 10, Figure 15, Figure 16 and Figure 17 in Appendix II present sequential summaries of the parameter posteriors for state 2, state 3 and state 4 respectively. Each panel plots the (.05, .50, .95) posterior quantiles and density functions for each parameter.

2-State Model

Parameter estimates for the 2-state model under SMC are presented in Table (12).

Table 12: Parameter estimates for the 2-state SV model under SMC

Parameter	State 1	State 2
α	-0.0001	-0.0004
ϕ	0.9946	0.9910
σ	0.1856	0.0138
Total (n)	1640	632
ESS	1517.641	629.8072

State 1 has a lower mean and variance of the volatility process value. Both processes are stationary as their ϕ values are less than 1.

3-State Model

The parameter estimates for the 3-state model under SMC, presented in Table (13), indicate that the highest mean value is in state 3, followed by state 1 and then state 2.

Table 13: Parameter estimates for the 3-state SV model under SMC

Parameter	State 1	State 2	State 3
α	-0.0002	-0.0010	0.0229
ϕ	0.9981	0.9172	0.9810
σ	0.1809	0.2218	0.3728
Total (n)	908	396	968
ESS	836.2122	363.8864	847.4931

State 1 has the highest variance of the volatility process. The values of ϕ indicate that all three processes are stationary.

4-State Model

The four-state HMM provides for very high and very low volatility regimes, and two intermediate states representing high and low volatility regimes. Parameter estimates for each of these state processes are presented in Table (14).

Table 14: Parameter estimates for the 4-state SV model under SMC

Parameter	State 1	State 2	State 3	State 4
α	-0.0368	-0.0001	-0.0008	0.0460
ϕ	0.9663	0.7166	0.7784	0.9699
σ	4.8551	14.4947	8.2678	0.0933
Total (n)	660	512	377	723
ESS	330.8861	211.0747	182.6429	693.7399

State 4 has the highest mean value while state 2 has the highest variance of the volatility process. The values of ϕ are as $|\phi| < 1$ showing all four processes are stationary.

4.4 Forecasting Volatility

The HMM filter on rolling volatility indicates well separated volatility states for the 2-state, 3-state and 4-state regimes, and characterizes the transition process underlying the data. Given that volatility is not directly observable in the market, we rely on the transition matrices, given by Table 1, Table 3 and Table 5, for predictive purposes when it comes to forecasting volatility. The matrices are achieved upon fitting the HMM filter as shown under Appendix IV. The transition matrices for all three models indicate stable states given higher probabilities of staying within a state relative to moving to other states. This supports the volatility clustering stylized fact, as periods of high volatility are usually followed by periods of high volatility while periods of low volatility are followed by periods of low volatility.

The 3-state model assumes that state 1 represents low volatility, state 2 represents high volatility while state 3 is the bridge between low and high volatility. According to the empirical results, there is no probability of moving from state 1 to state 2, and the probability of moving from state 2 to state 1 is low. This indicates that the volatility process has a low likelihood of jumping from low to high or from high to low levels. The same holds for the 4-state model where the transition probability from state 1 (very low volatility) to state 2 (high volatility) and to state 3 (very high volatility) is 0. Transition probabilities from state 2 to state 1, state 3 to state 1 and state 4 (low volatility), and from state 4 to state 3, are also relatively low.

4.5 Discussion

Performance of equity market depends on the nature of the economic environment. As a frontier market, Kenya has seen a steady rise in its gross domestic product averaging 5% per year over the last decade, as shown in Figure 17. Studies in Kenya have analyzed the effects of various macro-economic variables on stock returns, the most significant being money supply, inflation, exchange rates and interest rates. Various authors have established existence of leverage effects in the Kenyan economy as discussed under the subsection 2.1 Introduction.

Figure 1 shows a growing trend in NSE20 share prices from the beginning of the modeling period until the year 2015. This may be attributed to higher economic growth due to lower inflation rates in comparison to 2011, as seen in Figure 18, which saw a shift away from consumption and toward investment. The high inflation in 2011 was attributed to a reduced Central Bank Rate meant to revive lending and stimulate the economy through increased consumption. The NSE20's share prices have been steadily declining since mid 2015, owing to the listed companies' poor financial results as a result of a tough economic environment. Furthermore, high exchange rates, as shown in Figure 19, may have resulted in a drop in foreign investments in listed companies over the same time span.

Latent Markov models have been employed in modelling regime switches in economics. Following the first specific objective of this study, the HMM filter provides a clear separation of states where the input data is rolling volatility over a 30-day window. The 4-state economy is optimal given that it has a relatively low AIC and BIC value. Under the 4-state regime, the HMM filter classifies the economy into four volatility regimes: very high, high, low and very low volatility. It is worth noting that experimental results show that all three state models start at state 1 which represents low volatility for the 2-state and 3-state models, and very low volatility for the 4-state model.

Given the 4-state optimal model, very low volatility characterizes the beginning of the modelling period which could be attributed to moderate economic growth but of significant increase from the year 2011. Averagely low volatility can be seen in early 2016, 2018 and 2019. These periods are characterized by fairly stable exchange and inflation rates. In early 2016, a relatively stable currency and lower inflation rates were witnessed. As seen in Figure 20, this was due to the Central Bank of Kenya (CBK) reducing the benchmark interest rate from 11.5% to 10.5%, resulting in lower credit costs. Political risk dominated the years 2012 - 2014 as a result of the International Criminal Court (ICC) hearings, resulting in some price volatility during the same period.

Periods of high volatility during the modelling period could be attributed to the general elections in March 2013 and September 2017 with a consequent nullification in the last quarter of 2017. Given tensions as a result of a hostile political environment encountered in previous elections, this may have resulted in inflation and political risk. Political risk also causes financial market uncertainty, with unpredictable price movements seen during election-related information releases. The CBK's decision to lift its base rate in July 2015 resulted in volatile 91-day Treasury Bill (T-bill) prices in late 2015, as shown in Figure 20. Further to that, in mid-2016, a Banking Amendment Act that capped interest rates levied by financial lending institutions caused some market volatility.

The International Criminal Court's decision in January 2014 to postpone the start of the prosecutor's trial until February 2014 (ICC, 2014), as well as many other critical information releases on the same case, may have contributed to the very high volatility observed in early 2014. The election business cycle may have amplified this, resulting in uncertainty about monetary policy and fiscal spending transmission or policy risk, which often leads to delayed recovery from political events (Born & Pfeifer, 2014). Charges of crimes against humanity against Kenya's top political figures may have disincentivised foreign investment in Kenyan companies. High inflation and a low exchange rate of the US dollar against the Kenyan shilling resulted in a depreciated currency in early 2014. The corona virus pandemic, which occurred in early 2020, was the most recent cause of extremely high volatility. There was very high market volatility given the first reported case of COVID-19 in March 2020, evidenced by Figure 3, with relatively high volatility experienced around key informational releases on policies regarding the pandemic in the year 2020. The NSE20 share prices were on a constant dip from the year 2015 and the pandemic only further exacerbated the situation. The pandemic has had its impact on both the domestic and global fronts resulting in a contracted quarterly gross domestic product contrary to the growth that had been witnessed in the last decade.

The univariate stochastic volatility model takes the volatility of the variance process into account and is a general state-space model. The AR(1) process, whose one-time shock impacts values of the evolving variable infinitely far into the future, captures the persistent nature of volatility. This may be the case particularly where there is no market resolution following a macroeconomic or political event. The EM algorithm is particularly employed in parameter estimation as it achieves maximum likelihood parameter estimates without having to deal with the likelihood which cannot be presented in a closed form for the SV model, and where the data is not fully observable. Under this model, a persistence of log-variance (ϕ) parameter value that is less than unit value indicates stationarity.

The MCMC and SMC rely on Bayesian estimation of parameters which are distributed approximately with regard to the posterior. In Bayesian inference, an assumption about the prior distributions of the parameter set is made and parameter estimates are then simulated using the assumed priors. In this research, prior distributions for the univariate SV model's parameters are as normal for the level of log-variance (μ/α), beta for persistence of log-variance (ϕ) and chi-square for the volatility of log-variance (σ). The posterior distribution is obtained by evaluating the likelihood of each parameter estimate, $Pr(\theta)$, given observation information, $Pr(\theta \mid x_t)$.

Filtering techniques are employed to extract latent state variables from noisy data to be used in parameter estimation. MCMC utilizes batched and offline samplers while SMC employs sequential and online techniques for parameter estimation. For the second specific objective of the study, parameter estimates for the single state models under MCMC (Table 7) and SMC (Table 11) differ but have a low deviation from each other. For both filters, the σ values exhibit the most deviation in the various state models. Under SMC, σ values have the most deviation under the 4-state model returning the highest values for states 1, 2 and 4. The μ and ϕ values remain relatively within the same range, depending on whether the particular state models low, moderate or high volatility. For MCMC, parameter estimates for the level of log-variance are close to each other in the 2-state model but their range widens in the 3-state and 4-state models. Resultant posterior density plots under MCMC in Appendix I and SMC in Appendix II depict a relatively even distribution for the μ parameter, a negative (left) skew for the ϕ parameter and positive (right) skew for the σ parameter.

The Markov Chain is a stochastic model that describes a series of possible events where the probability of each event is solely determined by the state obtained in the previous event. A transition probability matrix occurs in a first-order Markov chain process, which specifies the likelihood of transitioning from one state to another in successive time periods. As volatility is not directly observable in the market, we rely on transition probabilities to forecast volatility. In this study, the transitional probability matrices are given by Table 1, Table 3 and Table 5. The transition matrices for all three models imply to stable states which supports the stylized fact of volatility clustering.

Chapter 5

Conclusions and Recommendations

Hidden Markov Models (HMMs) have their applications in state inference based on observations. Secondary price data for the NSE20 share index is used to estimate rolling volatility with a goal of estimating volatility regimes by employing a HMM filter. The univariate stochastic volatility model established by Taylor (1982) and acknowledged by Andersen et al. (2009) provides a straightforward yet versatile framework for modeling time-varying volatility, as evidenced by empirical research.

The use of the Schwartz Bayesian Information Criterion (BIC) is motivated by a well-known drawback of the log-likelihood in state estimation, which increases as the number of states increases. The BIC allows us to select an optimal model, that is, the model with the lowest BIC value. Estimation results under determination of regimes indicate that the 4-state model, which divides the economy into periods of very high, high, low and very low volatility, is optimal. It has the lowest BIC value of 6623.531 relative to 9538.23 and 7860.997 for the 2-state and 3-state models respectively. From the discussion in Chapter 4, it is evident that different macroeconomic variables, such as the election business cycle, variations in inflation rates, interest rates and exchange rates, drive the volatility process at different times during the modeling period.

Markov chain Monte Carlo (MCMC) and Sequential Monte Carlo (SMC) filtering techniques are employed in parameter estimation under a single regime and multiple regimes economy for the volatility process. These methods generate approximations to filtering distributions and are commonly used in non-linear and/or non-Gaussian state space models. Given the different parameter estimates for the state processes under each modelling regime and under each filtering technique (Table 7 to Table 14), a more versatile framework for modeling the volatility process is implemented. As a result, when analyzing volatility dynamics in the pricing of various volatility-backed financial assets, regime switching should be factored. Higher same state transition probabilities in all regimes are notable from Table 1, Table 3 and Table 5, which implies volatility clustering in this analysis.

A useful extension to this research, particularly in the Kenyan financial market can include applying a multivariate stochastic volatility model as well as other already established stochastic volatility models. In addition, a comparative analysis of the said models with the univariate SV model can offer insights on the ability of the models to capture stylized facts of the volatility process under regime changes. Furthermore, while Kalman filters and Monte Carlo methods are foundational in the study of state space models, there are a broad range of filters employed in state-space estimation. Given the nature of financial time series data, it would be important to consider nonlinear, non-Gaussian models such as the point mass filter (PMF), rejection sampling (RS), importance resampling (IR), Metropolis-Hastings independence sampling (MH), Markov chain sequential Monte Carlo (MCSMC) where Markov chain Monte Carlo is employed within a Sequential Monte Carlo algorithm to update parameter values, and extensions to particle filters such as the auxiliary particle filters (APF) and unscented particle filters (UPF). This may provide useful extensions to this research in the context of Kenyan financial markets, as well as comparative analyses between different filtering techniques.

The research focused on the NSE20 stock index, which has been criticized for only including blue chip firms. The NSE25 and the All Share Index (NASI) also act as measures of market performance and can be used as secondary data sources for modelling stochastic volatility in the context of regime switches. The research can also be extended to modelling regimes in other financial instruments in the Kenyan financial market, such as bonds, exchange rates and commodities. Finally, while this study focused on daily data, tick-by-tick data could be used in modeling market returns and subsequent volatility for better risk management and control, particularly in highly liquid markets. Other risk controls, such as the Value-at-Risk (VaR) and risk measures for shorter time periods, could improve from this.

References

Abanto-Valle, C., Lachos, V., & Dey, D. K. (2015). Bayesian estimation of a skew-student-t stochastic volatility model. *Methodology and Computing in Applied Probability*, *17*(3), 721–738.

Ahlip, R. (2008). Foreign exchange options under stochastic volatility and stochastic interest rates. *International Journal of Theoretical and Applied Finance*, *11*(03), 277–294.

Allaya, M. M., Coulibaly, A., Mouhamadou, M. K., Sene, B., et al. (2019). On some extensions of the sequential monte carlo methods in high-order hidden markov models. *Afrika Statistika*, *14*(2), 1977–1998.

Andersen, T. G., Davis, R. A., Kreiß, J.-P., & Mikosch, T. V. (2009). *Handbook of financial time series*. Springer Science & Business Media.

Andersen, T. G., & Sørensen, B. E. (1996). Gmm estimation of a stochastic volatility model: A monte carlo study. *Journal of Business & Economic Statistics*, *14*(3), 328–352.

Andersson, L. (2019). *An application of bayesian hidden markov models to explore traffic flow conditions in an urban area.*

Bickel, P. J., Ritov, Y., Ryden, T., et al. (1998). Asymptotic normality of the maximum-likelihood estimator for general hidden markov models. *The Annals of Statistics*, *26*(4), 1614–1635.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. In *World scientific reference on contingent claims analysis in corporate finance: Volume 1: Foundations of cca and equity valuation* (pp. 3–21). World Scientific.

Born, B., & Pfeifer, J. (2014). Policy risk and the business cycle. *Journal of Monetary Economics*, *68*, 68–85.

Bundi, D., Patrick, W., et al. (2016). Credit scoring for m-shwari using hidden markov model. *European Scientific Journal*, *12*(15).

Chan, J. C. (2013). Moving average stochastic volatility models with application to inflation forecast. *Journal of Econometrics*, *176*(2), 162–172.

Chelimo, J. K. (2017). *Calibration of vasicek model in a hidden markov context: the case of kenya* (Unpublished doctoral dissertation). Strathmore University.

Clark, T. E., & Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, *30*(4), 551–575.

Danielsson, J. (1994). Stochastic volatility in asset prices estimation with simulated maximum likelihood. *Journal of Econometrics*, *64*(1-2), 375–400.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1–22.

Douc, R., Garivier, A., Moulines, E., Olsson, J., et al. (2011). Sequential monte carlo smoothing for general state space hidden markov models. *Annals of Applied Probability*, *21*(6), 2109–2145.

Elliott, R. J., Moore, J. B., & Aggoun, L. (1995). Practical recursive filters. *Hidden Markov Models: Estimation and Control*, 163–194.

Endo, A., van Leeuwen, E., & Baguelin, M. (2019). Introduction to particle markov-chain monte carlo for disease dynamics modellers. *Epidemics*, *29*, 100363.

Fridman, M., & Harris, L. (1998). A maximum likelihood approach for non-gaussian stochastic volatility models. *Journal of Business & Economic Statistics*, *16*(3), 284–291.

Frühwirth-Schnatter, S., & Wagner, H. (2010). Stochastic model specification search for gaussian and partial non-gaussian state space models. *Journal of Econometrics*, *154*(1), 85–100.

Gallant, A. R., Hsieh, D., & Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of econometrics*, *81*(1), 159–192.

Genon-Catalot, V., Jeantheau, T., Larédo, C., et al. (2000). Stochastic volatility models as hidden markov models and statistical applications. *Bernoulli*, *6*(6), 1051–1079.

Goutte, S., Ismail, A., & Pham, H. (2017). Regime-switching stochastic volatility model: estimation and calibration to vix options. *Applied Mathematical Finance*, *24*(1), 38–75.

Haas, M. (2009). Persistence in volatility, conditional kurtosis, and the taylor property in absolute value garch processes. *Statistics & probability letters*, *79*(15), 1674–1683.

Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics*, *14*(4), 429–434.

Heimbürger, H. (2016). *Modelling of stochastic volatility using partially observed markov models.*

ICC. (2014). *Situation in the republic of kenya in the case of the prosecutor v. uhuru muigai kenyatta.* Retrieved from `https://www.icc-cpi.int/CourtRecords/CR2014_02248.PDF`

Jha, O. (2019). When to 'buy the dip'. medium.

Julier, S. J., & Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, *92*(3), 401–422.

Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Journal of Fluids Engineering*.

Kastner, G. (2016). Efficient bayesian interference for stochastic volatility.

Kim, J. (2005). *Parameter estimation in stochastic volatility models with missing data using particle methods and the emalgorithm* (Doctoral dissertation, University of Pittsburgh). Retrieved from `http://d-scholarship.pitt.edu/8265/`

Kirui, E., Wawire, N. H., & Onono, P. O. (2014). Macroeconomic variables, volatility and stock market returns: a case of nairobi securities exchange, kenya. *International Journal of Economics and Finance*, *6*(8), 214–228.

Krishnamurthy, V., Leoff, E., & Sass, J. (2018). Filterbased stochastic volatility in continuous-time hidden markov models. *Econometrics and statistics*, *6*, 1–21.

Langrock, R., MacDonald, I. L., & Zucchini, W. (2012). Some nonstandard stochastic volatility models and their estimation using structured hidden markov models. *Journal of Empirical Finance*, *19*(1), 147–161.

Malakorn, T., & Iamtan, T. (2018). Estimating parameters of a stochastic volatility model using the expectation-maximization algorithm coupled with a gaussian particle filter. *Asia-Pacific Journal of Science and Technology*, *23*(4).

Mandelbrot, B. (1963). New methods in statistical economics. *Journal of political economy*, *71*(5), 421–440.

Maqsood, A., Safdar, S., Shafi, R., & Lelit, N. J. (2017). Modeling stock market volatility using garch models: A case study of nairobi securities exchange (nse). *Open Journal of Statistics*, *7*(2), 369–381.

Mbete, D., Nyongesa, K., & Rotich, J. (2019). Estimation of malaria symptom data set using hidden markov model. *Asian Journal of Probability and Statistics*, 1–29.

Mekoya, M. W. (2013). *Modeling and forecasting stock market volatility at nairobi securities exchange* (Unpublished doctoral dissertation). University of Nairobi.

Ndei, C. M., Muchina, S., & Waweru, K. (2019). Modeling stock market return volatility in the presence of structural breaks: Evidence from nairobi securities exchange, kenya. *International Journal of Research in Business and Social Science (2147-4478)*, *8*(5), 156–171.

Nyamongo, M. E., Misati, R., et al. (2010). Modelling the time-varying volatility of equities returns in kenya. *African Journal of Economic and management studies*, *1*(2), 183–196.

Nyongesa, A. M., Zeng, G., & Ongoma, V. (2020). Non-homogeneous hidden markov model for downscaling of short rains occurrence in kenya. *Theoretical and Applied Climatology*, *139*(3), 1333–1347.

Pitt, M. K., Malik, S., & Doucet, A. (2014). Simulated likelihood inference for stochastic volatility models using continuous particle filtering. *Annals of the Institute of Statistical Mathematics*, *66*(3), 527–552.

Rossi, A., & Gallo, G. M. (2006). Volatility estimation via hidden markov models. *Journal of Empirical Finance*, *13*(2), 203–230.

Rydén, T., et al. (2008). Em versus markov chain monte carlo for estimation of hidden markov models: A computational perspective. *Bayesian Analysis*, *3*(4), 659–688.

Scott, S. L. (2002). Bayesian methods for hidden markov models: Recursive computing in the 21st century. *Journal of the American statistical Association*, *97*(457), 337–351.

Shimada, J., & Tsukuda, Y. (2005). Estimation of stochastic volatility models: an approximation to the nonlinear state space representation. *Communications in Statistics—Simulation and Computation®*, *34*(2), 429–450.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. *Time series analysis: theory and practice*, *1*, 203–226.

Thrun, S., & Langford, J. (1998). *Monte carlo hidden markov models* (Tech. Rep.). Carnegie-Mellon Univ Pittsburg PA School of Computer Science.

Turek, D., de Valpine, P., & Paciorek, C. J. (2016). Efficient markov chain monte carlo sampling for hierarchical hidden markov models. *Environmental and Ecological Statistics*, *23*(4), 549–564.

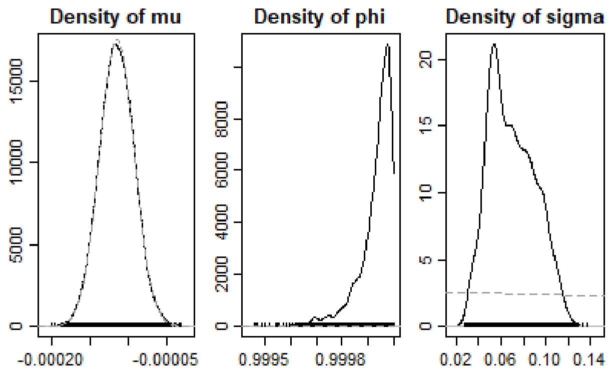
Wan, E. A., & Van Der Merwe, R. (2000). The unscented kalman filter for nonlinear estimation. In *Proceedings of the ieee 2000 adaptive systems for signal processing, communications, and control symposium (cat. no. 00ex373)* (pp. 153–158).

Wu, C. J. (1983). On the convergence properties of the em algorithm. *The Annals of statistics*, 95–103.

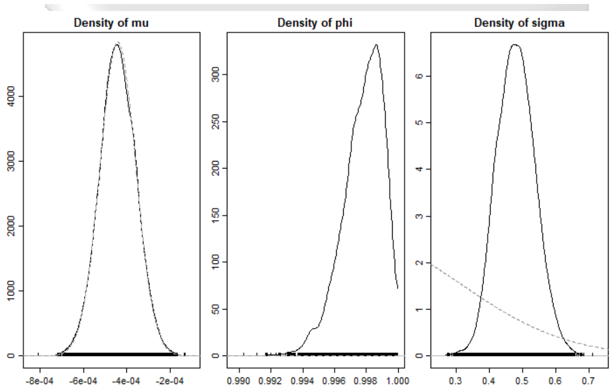
Yang, F., Xing, Y., Wang, D., & Tsui, K.-L. (2016). A comparative study of three model-based algorithms for estimating state-of-charge of lithium-ion batteries under a new combined dynamic loading profile. *Applied energy*, *164*, 387–399.

Appendix I

Markov Chain Monte Carlo Two State Model



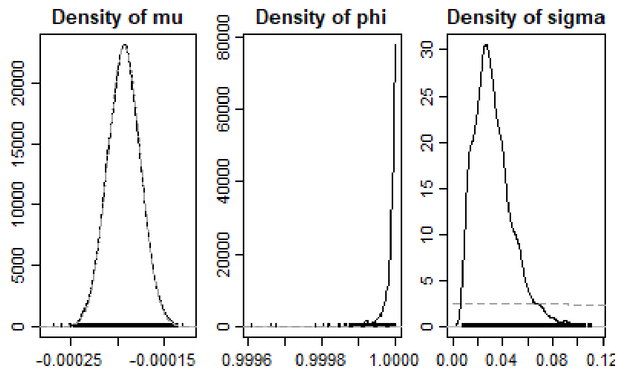
(a) State 1 density plots of MCMC draws



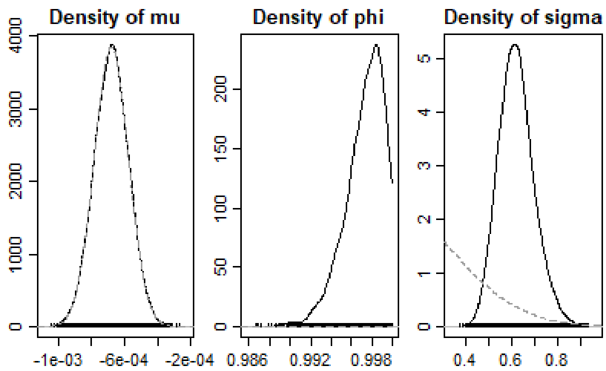
(b) State 2 density plots of MCMC draws

Figure 11: Density plots of MCMC draws from parameter posteriors for the 2-State Model

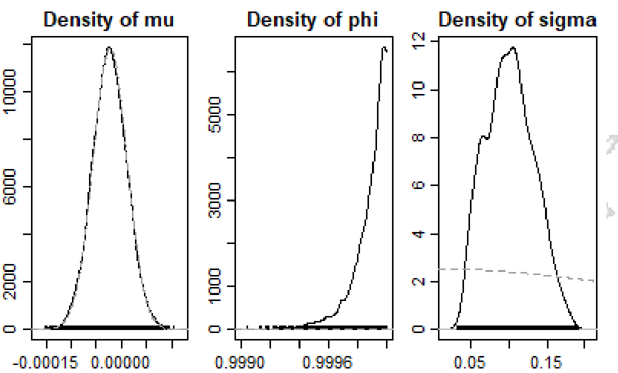
Three State Model



(a) State 1 density plots of MCMC draws



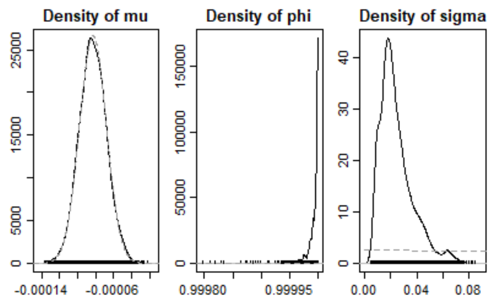
(b) State 2 density plots of MCMC draws



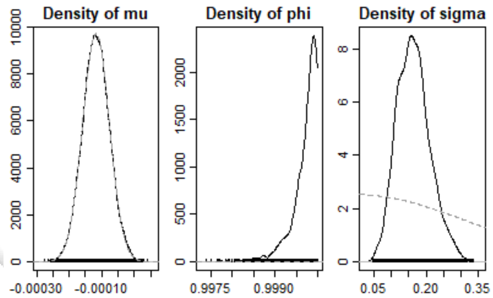
(c) State 3 density plots of MCMC draws

Figure 12: Density plots of MCMC draws from parameter posteriors for the 3-State Model

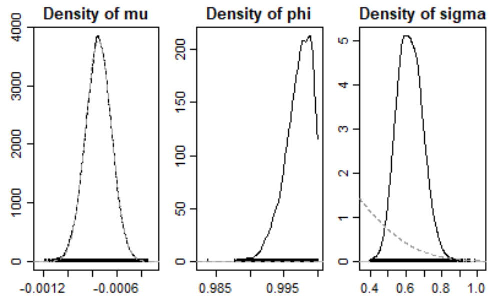
Four State Model



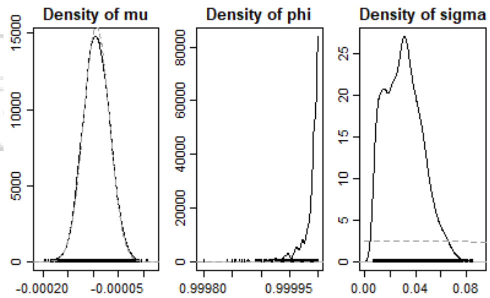
(a) State 1 density plots of MCMC draws



(b) State 2 density plots of MCMC draws



(c) State 3 density plots of MCMC draws



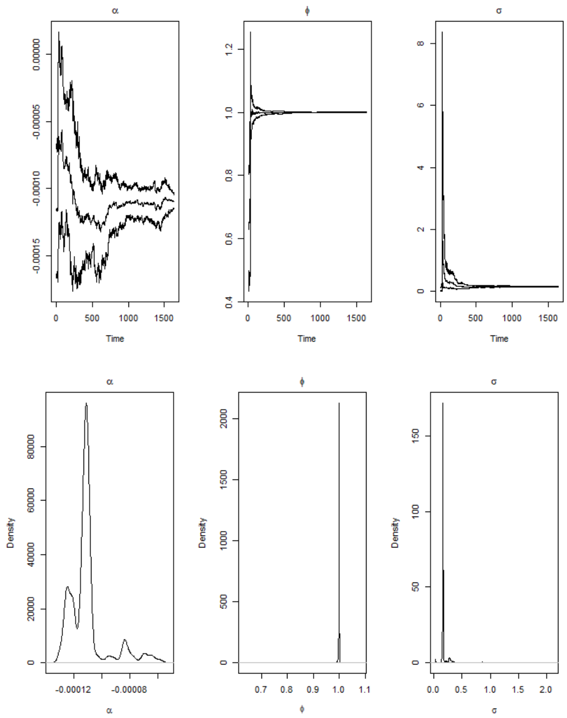
(d) State 4 density plots of MCMC draws

Figure 13: Density plots of MCMC draws from parameter posteriors for the 4-State Model

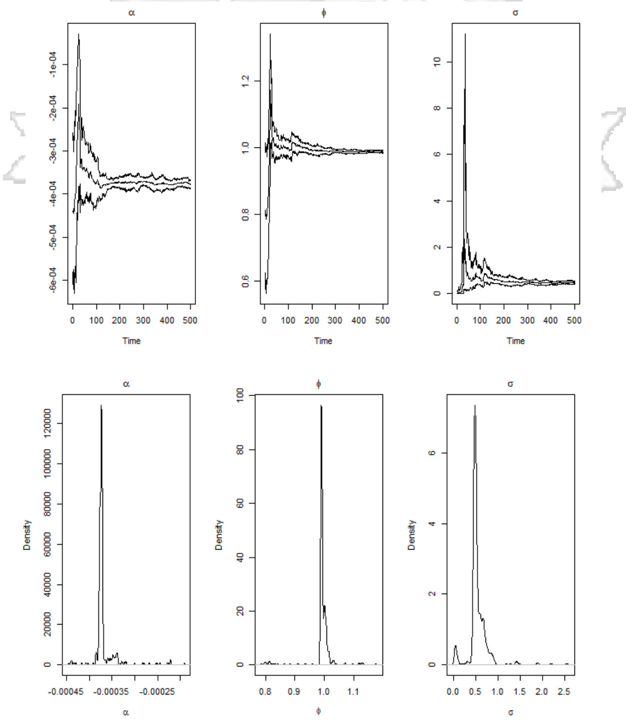
Appendix II

Sequential Monte Carlo

Two State Model



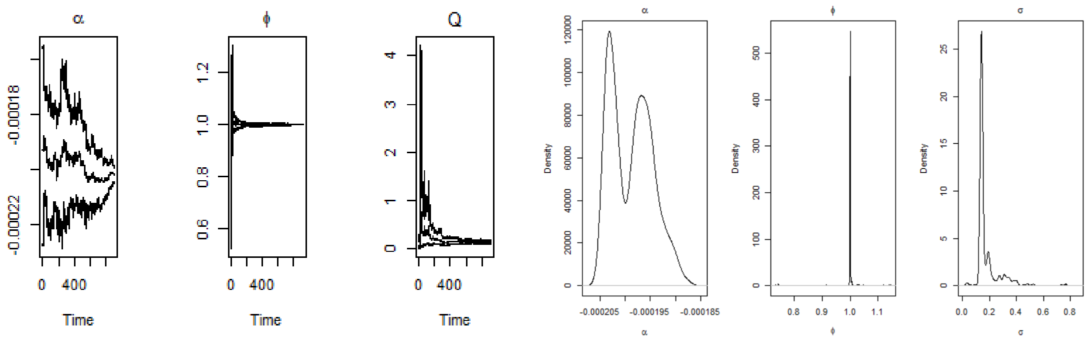
(a)State 1 parameter estimates and density plots



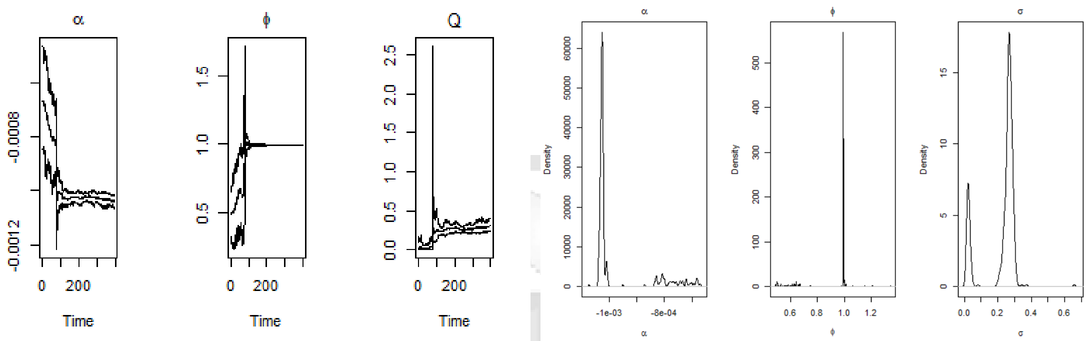
(b) State 2 parameter estimates and density plots

Figure 14: SMC convergence of parameter estimates and resultant density plots for the 2-State Model

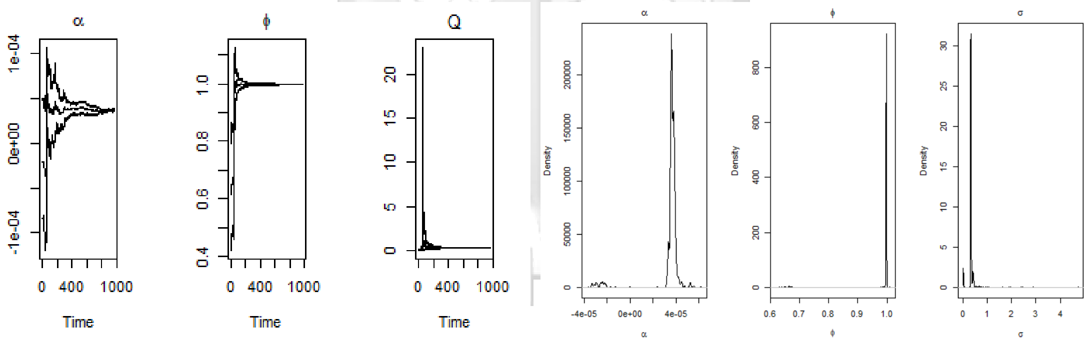
Three State Model



(a)State 1 parameter estimates and density plots



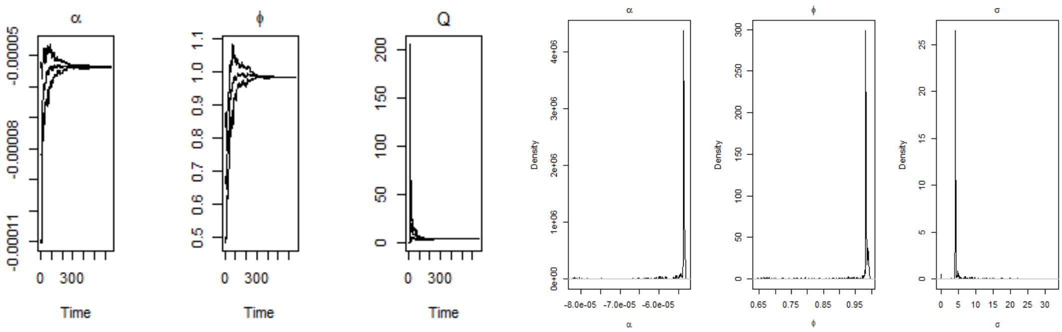
(b)State 2 parameter estimates and density plots



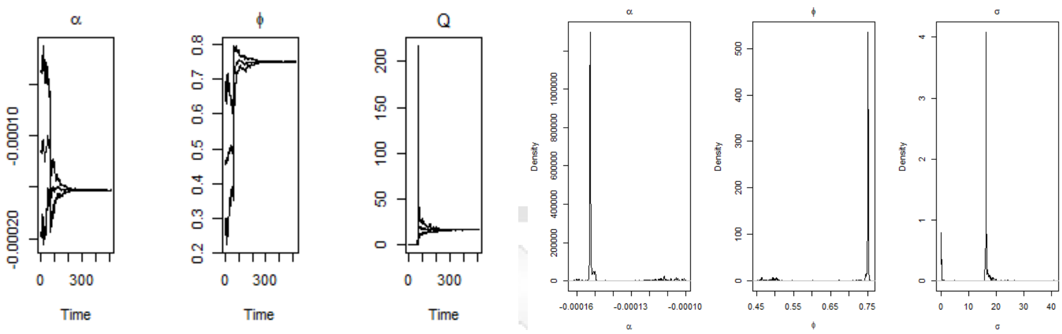
(c)State 3 parameter estimates and density plotsl

Figure 15: SMC convergence of parameter estimates and resultant density plots

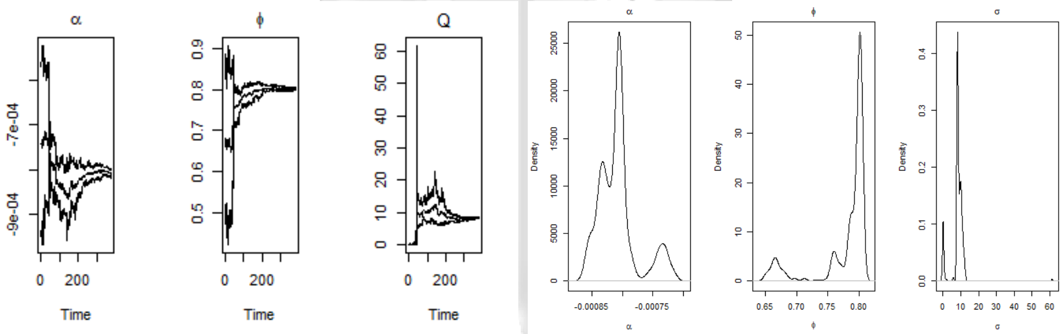
Four State Model



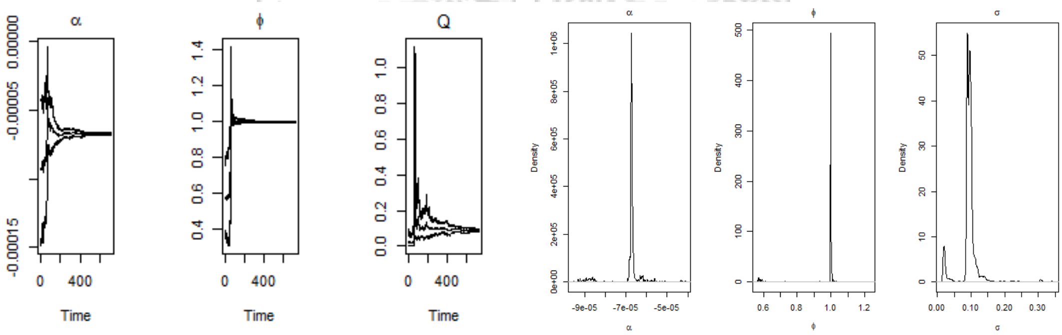
(a)State 1 parameter estimates and density plots



(b)State 2 parameter estimates and density plots



(c)State 3 parameter estimates and density plots



(d)State 4 parameter estimates and density plots

Figure 16: SMC convergence of parameter estimates and resultant density plots for the 4-State Model

Appendix III

Macroeconomic Variables

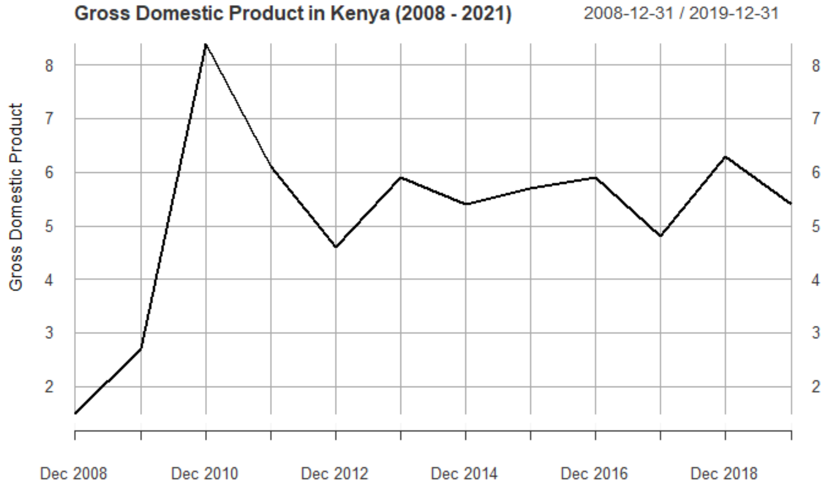


Figure 17: Percentage annual Gross Domestic Product in Kenya for the Period 2008 - 2021

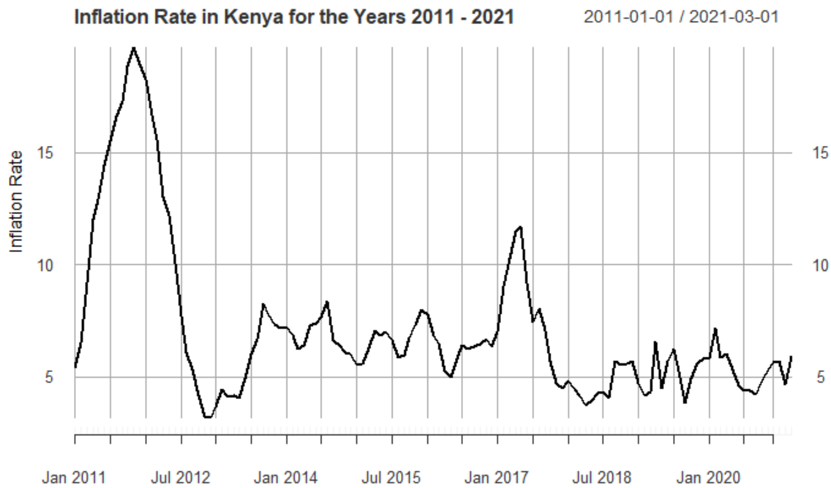


Figure 18: Monthly Inflation rate in Kenya for the Period 2011 - 2021

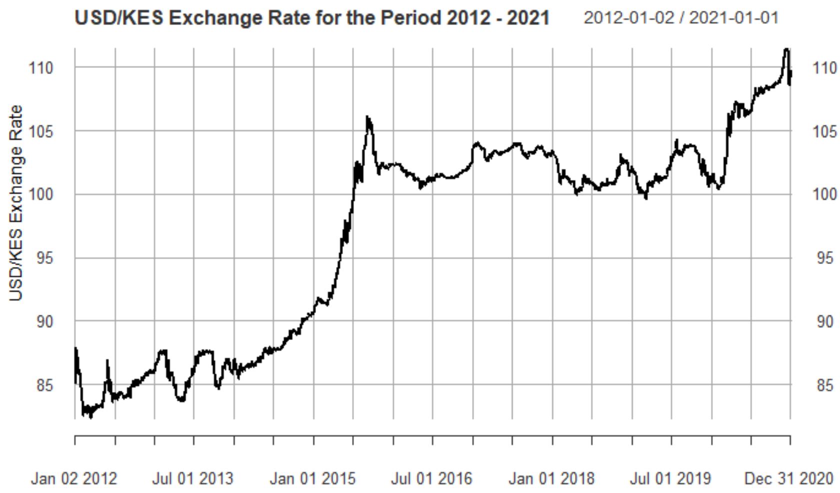


Figure 19: Daily USD/KES Exchange Rates for the period 2012 - 2021

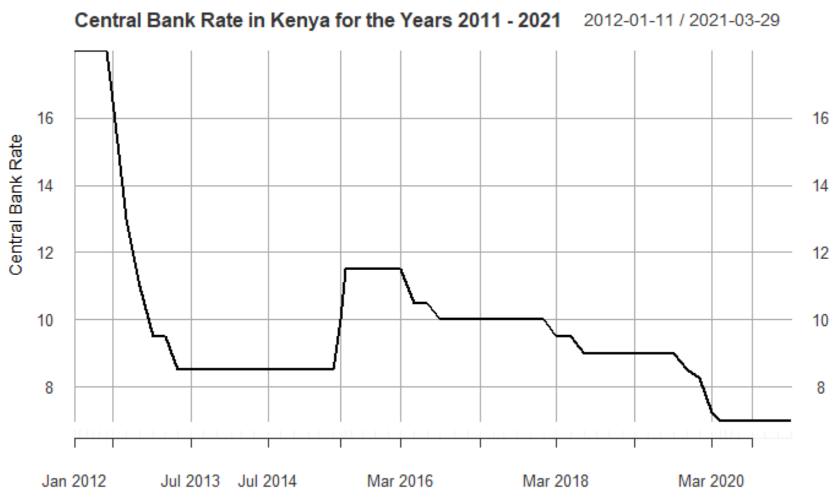


Figure 20: Monthly Central Bank Rate in Kenya for the period 2012 - 2021

Appendix IV

R Codes

Fitting a Hidden Markov Model

```
library(depmixS4)
ModelData <- data.frame(na.omit(finvol))
colnames(ModelData) <- c("finvol")
#finvol is rolling volatility data

#Fitting HMM Filter
HMM <- depmix(finvol~1, data=ModelData, nstates = n)
#n = number of states to be modelled
HMMfit <- fit(HMM, verbose=FALSE)
#Summary statistics for state model
print(HMMfit)
summary(HMMfit)
HMMpost <- depmixS4::posterior(HMMfit)

# Graphical analysis
plot.ts(HMMpost$state, axes=F, col = 'red', type = 'l', lty = 2, lwd
    = 2, ylab = "", xlab = "")
axis(2, col='black',lwd=2)
mtext(2, text= 'HMM States',line=2)
par(new = TRUE)
plot.ts(ModelData$finvol, axes = F, ylab = "", xlab = "", lwd = 2,
   type = 'l')
axis(4, col='black',lwd=2)
mtext(4, text= 'Rolling Vol', line=1.3)
axis(1,col = 'black', lwd=2)
mtext(1, text = "Time", col='black',line=2)
legend("topleft", legend=c("Rolling Volatility", "HMM States"),lty=c
   (1,2), col = c("black", "red"), lwd = c(2,2), cex = 0.5)
```

Markov Chain Monte Carlo Filter

```
library(stochvol)

#Input data (y) is logarithm of squared returns
#Initial parameters estimates \cite{kim2005parameter}
initalpha = mean(y)
y2 = y[1:(length(y)-2)]
y1 = y[2:(length(y)-1)]
y0 = y[3:length(y)]
A = cov(y0, y2)
B = cov(y1, y2)
initphi = min(A/B,0.95)
initQ = max(mean(( (y0-mean(y0))-initphi*(y1-mean(y1)))^2) - 5 -5*
   initphi^2,0.1)

## Fit AR(1) - SVL model to NSE20 share index log returns data
ress <- svsample(y*10, priormu = c(-0.0002, 0.03518331), priorphi =
   c(initphi, 1.5), priorsigma = initQ, designmatrix = "ar1")
volplot(ress, forecast = 100, dates = Kenya_NSE20$Date[-1])
summary(ress, showlatent = FALSE)

# Graphical analysis
#par(mfrow = c(3, 1))
#paratraceplot(ress)
par(mfrow = c(1, 3))
paradensplot(ress)
```

Sequential Monte Carlo Filter

```
#Initializing prior, posterior, likelihood and quantile functions
pri        = function(x,m0,sC0){dnorm(x,m0,sC0)}
likelihood = function(y,x){dnorm(y,0,exp(x/2))}
rlike      = function(x){rnorm(1,0,exp(x/2))}
post       = function(y,x,m0,C0){pri(x,m0,sC0)*likelihood(y,x)}
quant025   = function(x){quantile(x,0.025)}
quant975   = function(x){quantile(x,0.975)}
# Data and prior hyperparameters, input data log returns(y) and log
   variance (x)
initalpha = mean(y)
y2 = y[1:(length(y)-2)]
y1 = y[2:(length(y)-1)]
y0 = y[3:length(y)]
A = cov(y0, y2)
B = cov(y1,y2)
initphi = min(A/B,0.95)
initQ = max(mean(( (y0-mean(y0))-initphi*(y1-mean(y1)))^2) - 5 -5*
   initphi^2,0.1)
n     =  length(y)
#Initial parameters same as MCMC initial parameters
alpha = initalpha
beta   =  initphi
tau2   =  initQ/10
alpha.true = alpha
beta.true   = beta
tau2.true   = tau2
m0       = 0.0
C0       = 0.1
sC0      = sqrt(C0)
ealpha   = alpha
valpha   = 0.01
ephi     = beta
vphi     = 0.01
nu       = 3
lambda   = tau2
```

```
N       =  length(y)
xs      = rnorm(N,m0,sC0)
pars    = cbind(rnorm(N,ealpha,var(y)),rnorm(N,ephi, sqrt(vphi)),log
   (1/rgamma(N,nu/2,nu*lambda/2)))
delta   = 0.75
h2      = 1-((3*delta-1)/(2*delta))^2
a       = sqrt(1-h2)
parss   = array(0,c(N,3,n))
xss     = NULL
ws      = NULL
ESS     = NULL

#Fitting the SMC Filter
par(mfrow = c(1,1))
for (t in 1:n){
  # Resampling
  mpar         = apply(pars,2,mean)
  vpar         = var(pars)
  ms           = a*pars+(1-a)*matrix(mpar,N,3,byrow=T)
  mus          = pars[,1]+pars[,2]*xs
  weight       = likelihood(y[t],mus)
  k            = sample(1:N,size=N,replace=T,prob=weight)
  # Propagating
  if (delta<1){
    ms1 = ms[k,]+matrix(rnorm(3*N),N,3)%*%chol(h2*vpar)
  }else{
    ms1 = ms
  }
  xt           = rnorm(N,ms1[,1]+ms1[,2]*xs[k],exp(ms1[,3]/2))
  w            = likelihood(y[t],xt)/likelihood(y[t],mus[k])
  w            = w/sum(w)
  ind          = sample(1:N,size=N,replace=T,prob=w)
  xs           = xt[ind]
  pars         = ms1[ind,]
  xss          = rbind(xss,xs)
  parss[,,t]   = pars
```

```
  ws          = rbind(ws,w)
  cv2         = var(w)/(mean(w)^2)
  ESS         = c(ESS,N/(1+cv2))
  #ts.plot(ESS,xlim = c(1,n))
}

# Posterior summary
malpha = apply(parss[,1,],2,mean)
lalpha = apply(parss[,1,],2,quant025)
ualpha = apply(parss[,1,],2,quant975)
mbeta  = apply(parss[,2,],2,mean)
lbeta  = apply(parss[,2,],2,quant025)
ubeta  = apply(parss[,2,],2,quant975)
mtau2  = apply(exp(parss[,3,]),2,mean)
ltau2  = apply(exp(parss[,3,]),2,quant025)
utau2  = apply(exp(parss[,3,]),2,quant975)
# Graphical analysis
par(mfrow=c(1,3))
ts.plot(malpha, ylim = range(lalpha, ualpha), ylab = "", main=
   expression(alpha))
lines(lalpha, lwd = 1.25)
lines(ualpha, lwd = 1.25)
ts.plot(mbeta,ylim = range(lbeta, ubeta), ylab = "", main=expression
   (phi))
lines(lbeta, lwd = 1.25)
lines(ubeta, lwd = 1.25)
ts.plot(mtau2, ylim = range(ltau2,utau2), ylab = "", main=expression
   (sigma))
lines(ltau2, lwd = 1.25)
lines(utau2, lwd = 1.25)
par(mfrow = c(1,3))
plot(density(malpha), xlab = expression(alpha), main = expression(
   alpha))
plot(density(mbeta), xlab = expression(phi), main = expression(phi))
plot(density(mtau2), xlab = expression(sigma), main = expression(
   sigma))
```