2018

Efficiency of the markov regime switching GARCH Model in modelling volatility for tea prices

Mathew Kiplimo Maiyo
Strathmore Institute of Mathematical Sciences (SIMs)
Strathmore University

Efficiency Of The Markov Regime Switching GARCH Model In Modelling Volatility For Tea Prices

Mathew Kiplimo Maiyo

Master of Science in Mathematical Finance

2018

Efficiency Of The Markov Regime Switching GARCH Model In Modelling Volatility For Tea Prices

Mathew Kiplimo Maiyo

Submitted in partial fulfillment of the requirements for the Degree of Master of Science in Mathematical Finance at Strathmore University

Institute of Mathematical Sciences
Strathmore University
Nairobi, Kenya

June, 2018

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.

....
....
....

The thesis of Mathew Kiplimo Maiyo was reviewed and approved by the following.

Dr. Samuel Chege Maina
Lecturer, Strathmore Institute of Mathematical Sciences
Strathmore University

Ferdinand Othieno
Dean, Strathmore Institute of Mathematical Sciences
Strathmore University

Prof. Ruth Kiraka
Dean, School of Graduate Studies
Strathmore University

Abstract

This study examines the ability of the Markov Regime Switching GARCH model, in comparison with the univariete GARCH models, in modelling and forecasting price volatility of the tea traded at the Mombasa Tea Auction, within some time horizon. The study uses weekly data, from 2010 to 2017, to analysis regime switching in volatility and provides an in-sample and out-of-sample forecast. Volatility regime switching is first modelled with a Markov switching framework. In-sample and out-of-sample forecasts of volatility using competing MRS-GARCH models and the single regimes GARCH models are then provided. Comparison of in-sample forecast is done on the basis of goodness-of-fit and the comparison of the out-of-sample forecasts is done on the basis of forecast accuracy, using the statistical loss function. The results show that the MRS-GARCH models can remove the high persistence of GARCH models. This shows the priority of MRS-GARCH models and provides evidence of regime clustering. In out-of-sample forecast perfomance, the MRS-GARCH models were better than the single regime GARCH model. However, this superioirity fades for longer time horizon.

Keywords: Volatility, Markov Regime Switching GARCH, GARCH, exponential GARCH, GJR-GARCH, Persistency, In-sample forecast, Out-of-sample forecast.

Contents

List of Tables

List of Figures

Acknowledgement

I hereby acknowledge my supervisor, Dr Chege, and the faculty at the Institute of Mathematical Sciences in Strathmore University who have been instrumental in the course of my study.

To

Lucy, Ella and Liam

Chapter 1

Introduction

Greater effort has been directed towards modelling volatility for financial time series. This has to be done taking in consideration the stylized fact, especially high persistence in the autocorrelation of squared observations and leptokurtosis (Taylor, 2008). The two main classes in which volatility can be modelled are the generalized autoregressive conditional heteroscedasticity (GARCH) and stochastic volatility method. The stochastic volatilty model can be assumed to be more flexible than the GARCH model since it allows for two error processes unlike the GARCH model that uses a single error process. However, Hafner & Preminger (2010) preferred the GARCH model when they compared the ability of the two models in fitting the characteristic features observed in high frequency financial data. The likelihood function of the stochastic volatility is usually intractable. This prohibits direct evaluation of the model.

Hamilton (2010) studied Markov regime in markets resulting from dramatic breaks in the time series. Hillebrand (2005) indicate that a distinct error occurs in the GARCH parameters estimates if regime change is not taken into consideration. Markov regime switching GARCH (MRS-GARCH) models belong to a class of models proposing that volatility is characterized by structural changes driven by a Markov chain. GARCH model has been widely used to model volatility. MRS-GARCH model allows the parameters of a GARCH model to change over time in order to allow for structural changes in a data series.

Little has been done on application of the MRS-GARCH in agricultural commodities since most research have focused on the traditional investments, stocks, currencies and interest rates, and energy commodity. Agirculture plays a key role in most under-developed and developing economies. Structural models used to analyse commodity markets are based on microeconomics and econometrics theories that require model specification, estimation and simulation. A challenges with the structural models is dealing with uncertainty in the markets. Greater concerned being macroeconomic influences. Levels of production of agricultural commodities play a significant role in determination of their prices and volatility. Agricultural commodity prices respond quickly to expected changes in supply and demand conditions. Schnepf (2005) highlight three characteristics that set them apart from most volatile prices of non-farm goods and services. The characteristics are seasonality of production, the derived nature of their demand and price-inelastic demand and supply functions. Schnepf (2005) further indicate that the speed and efficiency with which the various price adjustments occur depend largely on the market structure within which a commodity is being traded.

Tea and coffee are the only products that are traded in organized exchanges in Kenya. Tea is traded at the Mombasa Tea Auction and has been the country's leading foreign exchange earner. Using the prices of tea, one can be able to review the application of the MRS-GARCH model in a commodity market. Kenya is the third leading producer of black tea in the world and the largest exporter of tea in the world (Chang, 2015). The East African Tea Traders Association (EATTA) brings together players in the tea value chain and is the umbrella under which the Mombasa Tea Auction is conducted. Auctions are held weekly with the main grades auction being held on Tuesdays and secondary grades auction being held on Mondays. Mombasa Tea Auction is second largest black tea auction center in the world and it is located in a region where production is throughout the year.

Volatility in commodities is not directly observable, however it exhibits features in the returns. Mandelbrot (1963) noted the existance of volatility clustering, large changes tend to be followed by large changes, and small changes tend to be followed by small changes. In addition, volatility evolves over time in a continuous manner implying that jumps are rare. This means that volatility is often stationary, it does not diverge to infinity. Finally, volatility seems to react differently to a big price increase or a big price drop. This is referred to as the leverage effect. These properties indicate that evolution of the variance can be determined using an exact function.

Engle (1982) developed the first model that provides a systematic framework for volatility modelling, known as the ARCH model. The basic idea of ARCH models is that the innovations of an asset return is serially uncorrelated, but dependent, and the dependence can be described by a simple quadratic function of its lagged values. Tsay (2005) highligted several weakness with the ARCH model: the model assumes that positive and negative shocks have the same effects on volatility; the ARCH model is restrictive which limits the ability of ARCH models with Gaussian innovations to capture excess kurtosis; the ARCH model does not provide any new insight for understanding the source of variations of a financial time series; and, ARCH models are likely to overpredict the volatility because they respond slowly to large isolated shocks in the retrun series. Another limitation is that the ARCH models are linear in the squares of innovations. Some sort of nonlinearity may need to be accommodated (Friedman et al., 1989).

Bollerslev (1986) proposed the generalized ARCH (GARCH) model to reduce the number of parameters that adequately describe the volatility process in an ARCH model. Apart from reducing the number of parameters, the GARCH models encounters the same weaknesses as the ARCH model. To overcome some weaknesses of the GARCH model, in particular to allow for asymmetric effects between positive and negative returns, in handling financial time series, Nelson (1991) proposes the exponential GARCH (EGARCH) model. Another volatility model commonly used to handle leverage effects is the GJR-GARCH model (Glosten et al., 1993).

Elliott et al. (1998) assume that the volatility of the assets is driven by some common states in the economy. The states were unobservable and can be represented by a hidden Markov chain. Following the arguments of Schwert (1989), this study seek to come up with a MRS-GARCH model that can be used to model volatility for the price of tea traded at the Mombasa tea auction. Forecasted volatility when applied to the mean equation of a time-series can be used to obtain future prices of tea traded at the auction. The performance of the MRS-GARCH model is then compared with the performance of other single-regime GARCH models. The single-regime models considered are the GARCH, EGARCH and GJR-GARCH models. The comparison is in terms of both the in-sample and out-of-sample fit.

This study's addition to existing literature, on the application of the MRS-GARCH model, will be in three fold. The focus on the Mombasa tea auction will provide an insight to the specific tea market. Time of investigation is until June 2017 with an out-of-sample test that covers upto December 2017. This was a recent period and therefore unique. Furthermore, by limiting the data used to be based on a single commodity type, in this case tea, it increased the chance of distinguishing the superiority of the model since there might be different models that are best at forecasting the volatility of the different commodities or asset types. The results from the study could also be beneficial in the market. Out-of-sample perfomance is critical for real world perfomance. By comparing various volatility models, a better model for risk management can be recommended. This will be useful in determining appropriate level inventories to hold, and the best time to come into the market ofr traders. Volatility modelling is also instrumental for traders of volatility related products, such as options, once such an exchange becomes operational.

The thesis is arranged as follows: Chapter 2 provides a theoritical framework for the single-regime GARCH and the MRS-GARCH models, estimation of MRS-GARCH models and a basis for comparing models. Chapter 3 highlights the methodology applied to the study. Chapter 4 describes the data and the corresponding return series is decomposed to obtain the trend, the cyclic and residual components. The parameters of the single-regime GARCH and MRS-GARCH models are estimated in Chapter 5. Chapter 6 compares the perfomance of the MRS-GARCH with other single-regime GARCH models using both the in-sample and out-of-sample forecasts. The final section offers some concluding remarks and recommends areas for future work.

Chapter 2

Literature Review

There are various models built on regime changes. Schwert (1989) highlight the fluctuation of aggregates stock returns with either a high or low variance, with the switches between the states is determined by a two-state Markov process. Hamilton & Susmel (1994) and Cai (1994) explored the possibility of changing volatility and allowed the parameters of an autoregressive conditional heteroskedasticity (ARCH) process to come from one of several different regimes, with transitions between regimes governed by an unobserved Markov chain in order to take into account sudden changes in the level of the conditional variance. Regime switching in the volatility of returns have been found by Hamilton & Susmel (1994), Hamilton & Lin (1996), Edwards & Susmel (2001), and Kanas (2005).

Gray (1996) proposes a tractable regime-switching GARCH models for short-term interest rates with time-varying probability, but estimates an approximation to the model. Modifications to this model have been suggested by Haas et al. (2004), Dueker (1997), Klaassen (2002) and Bollen et al. (2000). Abramson & Cohen (2007) gives stationarity conditions for some of the tractable models.

The MRS-GARCH model has been applied in several studies. Bauwens et al. (2006) used the NASDAQ daily return series to develop a regime-switching univariate GARCH model with a time-varing probability of switching between a non-explosive regime and an explosive one. Zhang et al. (2015) evaluated the forecast performance of single-regime GARCH models and the two-regime Markov Regime Switching GARCH model for crude oil price volatility. The results indicate that the two-regime MRS-GARCH model beats the single-regime GARCH type models and nonlinear GARCH models exhibit greater accuracy than the linear GARCH model. Reher et al. (2011) combines Gray (1996) and Klaassen (2002) Markov-switching framework with Hentschel (1995) approach of nesting alternative single-regime GARCH models to establish a two-regime Markov-switching GARCH model which enables estimation of functional GARCH specifications within each regime.

This section gives a theoretical background for the MRS-GARCH model. This includes definition of the common conditional distribution of the standardized innovations in each regime. The section starts with the single-regime GARCH models as the foundation for MRS-GARCH model. Stochastic volatility model, as an alternative in volatility modelling, is also discussed. This section further indicates the estimation options available and how models can be compared.

2.1 Single-Regime GARCH Models

The ARCH model, developed by Engle (1982), was the first framework to model volatility. Volatility is modelled as a deterministic function. Consider a return series, $r_t = \mu_t + a_t$, where μ_t is conditional mean of the series and a_t are the innovations at time t. An ARCH models has a_t serially uncorrelated, but dependent. The dependence of a_t is indicated in Equation (2.1). An ARCH(n) model has the form

$$a_t = \sigma_t \epsilon_t, \text{ and } \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + ... + \alpha_n a_{t-n}^2, \tag{2.1}$$

where σ_t^2 is the conditional variance of the series, ϵ_t are independent and identically distributed random variables with zero mean zero and unit variance, $\alpha_0 > 0$, and $\alpha_i \geq 0$ for $i > 0$.

Bollerslev (1986) proposed the generalized ARCH (GARCH) model to reduce the number of parameters that adequately describe the volatility process in an ARCH model. A GARCH(n, s) model has the innovations, a_t, in the form

$$a_t = \sigma_t \epsilon_t, \text{ and } \sigma_t^2 = \alpha_0 + \sum_{i=1}^{n} \alpha_i a_{t-i}^2 + \sum_{j=1}^{s} \beta_j \sigma_{t-j}^2, \tag{2.2}$$

where ϵ_t are indepedent and identically distributed random variables with zero mean and unit variance, $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, and $\sum_{i=1}^{\max(\mathrm{m,s})} (\alpha_i + \beta_j) < 1$. The constraint on $\alpha_i + \beta_j$ implies that the unconditional variance of a_t is finite and the conditional variance evolves over time. As before, ϵ_t is often assumed to have a standard normal or standardized Student-t or generalised error distribution. The GARCH(n,s) model will reduce to an ARCH(n) model when $s = 0$. α_i is referred to as the ARCH parameter and β_j is referred to as the GARCH parameter.

Nelson (1991) developed the exponential GARCH (EGARCH) model to allow for asymmetric effects between positive and negative asset returns. This was a limitation of the GARCH model. The EGARCH model has a weighted innovation in the form

$$g(\epsilon_t) = \theta \epsilon_t + \gamma [|\epsilon_t| - \mathbb{E}(|\epsilon_t|)], \tag{2.3}$$

where θ and γ are real constants. Both ϵ_t and $|\epsilon_t| - \mathbb{E}(|\epsilon_t|)$ are zero-mean identically and independent distribution sequences with continuous distributions. Therefore, $\mathbb{E}[g(t)] = 0$. The asymmetry of $g(t)$ can easily be seen by rewriting it as

$$g(\epsilon_t) = \begin{cases} (\theta + \gamma) - \gamma \mathbb{E}(|\epsilon_t|) & \text{, if } \epsilon_t \geq 0 \\ (\theta - \gamma) - \gamma \mathbb{E}(|\epsilon_t|) & \text{, if } \epsilon_t < 0. \end{cases} \tag{2.4}$$

An EGARCH(n, s) model can be written as

$$a_t = \sigma_t \epsilon_t, \ \ln(\sigma_t^2) = \alpha_0 + \frac{1 + \beta_1 B + ... + \beta_{s-1} B^{s-1}}{1 - \alpha_1 B - ... - \alpha_n B^n} g(\epsilon_{t-1}) \tag{2.5}$$

where α_0 is a constant, B is the back-shift (or lag) operator such that $Bg(\epsilon_t) = g(\epsilon_{t-1})$, and $1 + \beta_1 B + ... + \beta_{s1} B^{s-1}$ and $1 - \alpha_1 B - ... - \alpha_n B^n$ are polynomials with zeros outside the unit circle and have no common factors.

Another volatility model commonly used to handle leverage effects is the GJR-GARCH model Glosten et al. (1993). A GJR-GARCH(n, s) model assumes the form

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{s} (\alpha_i + \gamma_i N_{t-i}) a_{t-i}^2 + \sum_{j=1}^{n} \beta_j \sigma_{t-j}^2 \tag{2.6}$$

where N_{t-i} is an indicator for negative a_{t-i}, that is,

$$N_{t-i} = \begin{cases} 1 & , \text{if } a_{t-1} < 0 \\ 0 & , \text{if } a_{t-1} \geq 0 \end{cases} \tag{2.7}$$

and α_i, γ_j, and β_j are non-negative parameters satisfying conditions similar to those of GARCH models. From the model, it is seen that a positive a_{t-i} contributes $\alpha_i a_{t-i}^2$ to σ_t^2, whereas a negative a_{t-i} has a larger impact $(\alpha_i + \gamma_i) a_{t-i}^2$ with $\gamma_i > 0$. The model uses zero as its threshold to separate the impacts of past shocks.

$$\sigma_t = \alpha_0 + \sum_{i=1}^{s} (\alpha_i + \gamma_i N_{t-i}) a_{t-i}^2 + \sum_{j=1}^{m} \beta_j \sigma_{t-j}. \tag{2.8}$$

Negative returns have a greater influence on future volatility than do positive returns. This leverage effect is reflected in the EGARCH and GJR-GARCH models. From equations (2.6) and (2.7), the value of N_{t-i} captures the leverage effect in the GJR-GARCH model. A positive error will have a weight of 0 and a negative one will be assigned a weight of 1.

2.2 Markov Regime Switching GARCH Model

An MRS-GARCH process, $\{y_t\}$, for $t = 1, ..., T$, has the form:

$$y_t = \varepsilon_t$$

with

$$\varepsilon_t = \eta_t \sqrt{(h_t(\Delta_t))}.$$

η_t is an identically and independently distributed random variable with zero mean and unit variance and there exist $\alpha_0(\Delta_t), \alpha_i(\Delta_t), i = 1, ..., q$ and $\gamma_l(\Delta_t), l = 1, ..., p$ such that

$$h_t(\Delta_t) = \alpha_0(\Delta_t) + \sum_{i=1}^{q} \alpha_i(\Delta_t)\varepsilon_{t-i}^2 + \sum_{l=1}^{p} \gamma_l(\Delta_t) h_{t-l}. \tag{2.9}$$

Δ_t is a variable indicating the state at time t and follows a Markov chain with finite state space $S = 1, ..., k$, and a transition matrix P. The probability of switching from between regimes depend on the transition matrix, P, indicated below

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1k} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{k1} & p_{k2} & p_{k3} & \cdots & p_{kk} \end{bmatrix}$$

with $p_{ij} = p(\Delta_t = j \mid \Delta_{t-1} = i)$ the probability of being in state j at time t given state i at time $t-1$.

Calculation of the likelihood function for such a sample is not feasible. The quasi maximum likelihood method can then not be used to estimate the model since it requires the integration of k^T possible regime paths where k is the number of regimes (Hamilton & Susmel, 1994) (Cai, 1994). To circumvent the path dependence problem, Gray (1996) substitutes h_{t-1} with the conditional variance of the error term ε_{t-1} given the information up to time $t-2$:

$$h_t(\Delta_t) = \alpha_0(\Delta_t) + \alpha(\Delta_t)\varepsilon_{t-1}^2 + \gamma(\Delta_t)\sum_{k}^{i=1} p(\Delta_{t-1} = i \mid \Omega_{t-2})h_{i,t-1}. \tag{2.10}$$

The model of Haas et al. (2004) contrasts with this approach because each specific conditional variance depends only on its own lag,

$$h_t(\Delta_t) = \alpha_0(\Delta_t) + \alpha(\Delta_t)\varepsilon_{t-1}^2 + \gamma(\Delta_t)h_{t-1}(\Delta_t). \tag{2.11}$$

This model can be rewritten in matrix form:

$$\mathbf{h}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1\varepsilon_{t1}^2 + \gamma\mathbf{h}_{t1},$$

where $\boldsymbol{\alpha}_0 = [\alpha_{01}, \alpha_{02}, ..., \alpha_{0k}]'$, $\boldsymbol{\alpha}_1 = [\alpha_{11}, \alpha_{12}, ..., \alpha_{1k}]'$ and $\boldsymbol{\gamma} = diag(\gamma_1, \gamma_2, ..., \gamma_k)$. $\mathbf{h}_t$ is thereby a vector of $k \times 1$ components. These MS-GARCH models can be easily estimated by Maximum Likelihood (ML) estimation following the work of Hamilton & Susmel (1994).

In a GARCH(1,1) model, persistence of a shock to the conditional variance is indicated by the sum of α and γ. An estimated that is close to one indicates a highly persistent volatility process. Mikosch & Stărică (2004) indicate that high persistence in the volatility process may be due to structural changes on the parameters of the model over a period of time due to different regimes. The MRS-GARCH allows for regime changing of the parameters.

The MRS-GARCH model can be interpreted as a Markov chain with transition kernel that is a mixture of distributions. Some of the assumptions are.
A1 η_t is identically and independently distributes and has a continous positive

density on $\mathbb{R}$ with $\mathbb{E}(\eta_t) = 0$ and $Var(\eta_t) = 1$.

A2 $\alpha_j > 0$ and $\gamma_j > 0$ for $j = 1, 2, ..., n$.

A3 $\beta_1 + \gamma_1 < 1$ i.e. the first regime is stable.

Assumtion A1 is standard and is satisfied in commonly used distributions for GARCH models. Assumption A2 is slightly stronger than the usual non-negative conditions ($\gamma_{s_t} \geq 0$, $\beta_{s_t} \geq 0$).

2.2.1 Conditional distribution

Model specification is completed by the definition of the conditional distribution of the standardize innovations $\eta_{t,k}$ in each regime of the Markov chain. The most common distributions employed to model financial logreturns are the normal distribution, the Student-t distribution and GED distribution. Each distribution is standardized to have a zero mean and a unit variance. The probability density function (PDF) of the standard Normal distribution is given by:

$$f_N(\eta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\eta^2}, \eta \in \mathbb{R}. \tag{2.12}$$

The normal distribution does not take into consideration the heavy tails of financial time series. This limits its application.

The PDF of the standardized Student-t distribution is given by:

$$f_S(\eta; \nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{(\nu - 2)\pi}\Gamma(\frac{\nu}{2})} \Big(1 + \frac{\eta^2}{(\nu - 1)}\Big)^{\frac{\nu+1}{2}}, \eta \in \mathbb{R} \tag{2.13}$$

where $\Gamma()$ is the Gamma function. The constraint $\nu > 2$ is imposed to ensure that the second order moment exists. The kurtosis of this distribution is higher for lower ν. The probability density function is symmetric and the degrees of freedom will determine distribution at the tails.

The PDF of the standardized generalized error distribution (GED) is given by:

$$f_{GED}(\eta;\nu) = \frac{\nu e^{-\frac{1}{2}|\eta/\lambda|^{\eta}}}{\gamma 2^{(1+1/\eta)}\Gamma(1/\eta)}, \lambda = \left(\frac{\Gamma(1/\nu)}{4^{1/\nu}\Gamma(3/\nu)}\right)^{\frac{1}{2}}, \eta \in \mathbb{R} \tag{2.14}$$

where $\nu > 0$ is the shape parameter. GED is a symmetrical distribution defined by three parameters indicating the mode of the distribution, dispersion of the distribution and the shape parameter that controls the skewness.

2.3 An Alternative of the MRS-GARCH Model

An alternative of the GARCH-type models is the family of stochastic volatility models where volatility is assumed to follow a stochastic process. An example is a model proposed by Heston (1993) where the underlying asset behavior is characterized by the following risk-neutral dynamics

$$\begin{aligned} &\frac{dS_t}{S_t} = rdt + \sqrt{V_t}dW_t^1 \\ &dV_t = a(\bar{V} - V_t)dt + \eta\sqrt{V_t}dW_t^2 \\ &dW_t^1 dW_t^2 = \rho dt \end{aligned} \tag{2.15}$$

where

S is the price of the underlying asset at time t

r is the risk free rate

V_t is the variance at time t

$\bar{V}$ is the long-term variance

a is the variance mean-reversion speed

η is the volatility of the variance process

dW_t^1, dW_t^2 are two correlated Weiner processes, with correlation coefficient ρ

This model exhibit some desirable financial time series properties. It models volatility as a mean-reverting process which is consistent with financial markets behaviour. It introduces correlation on the shocks between asset returns and volatility. This allows modelling the statistical dependence between the underlying asset and its volatility, which is a prominent feature of financial markets.

Due to intractability of the likelihood function in stochastic volatility models, other methods other than maximum likelihood should be used. Harvey et al. (1994) and Ruiz (1994) propose the quasi maximum likelihood. Monfardini (1998) propose for the usage of indirect inference and Andersen et al. (1999) applies efficient method of moments.

Elliott, Siu, et al. (2007) and Elliott, Kuen Siu, & Chan (2007) extend the Heston model by incorporating regime switching in the volatilty process. They priced volatility derivatives by using a mean reverting level of volatility governed by a Marlov chain. Goutte et al. (2017) considers that a hidden Markov chain governs volality's speed of mean reversion, the mean reversion level, the volatility of volatility, and the correlation with the stock index in pricing S&P 500 and VIX options. Regime switching stochastic volatility models have also been studied by Biswas & Goswami (2017), So et al. (1998), and Exterkate et al. (2017). The stochastic approach is applied in the risk-nuetral framework and that is why it is usually used to price derivatives.

2.4 Estimation

Estimation of MRS-GARCH models can be done either by maximum likelihood or by Markoc chain Monte Carlo (MCMC) Bayesian techniques. Both approaches require the evaluation of the likelihood function.

Let $\boldsymbol{\Psi} = (\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1, \gamma, P)$ be the vector of model parameters. The likelihood function is:

$$\mathcal{L}(\Psi|\mathfrak{I}_T) = \prod_{t=1}^{T} f(y_t|\Psi, \mathfrak{I}_{t-1}) \tag{2.16}$$

where $f(y_t|\Psi, \mathfrak{I}_{t-1})$ denotes the density of y_t given past observations, $\mathfrak{I}_{t1}$, and model parameters Ψ. For MRS-GARCH, the conditional density of y_t is:

$$f(y_t|\Psi, \mathfrak{I}_{t-1}) = \sum_{i=1}^{K}\sum_{j=1}^{K} p_{i,j} z_{i,t-1} f_{\mathcal{D}}(y_t|s_t = j, \Psi, \mathfrak{I}_{t-1}), \tag{2.17}$$

where $z_{i,t1} = P[s_{t1} = i|\boldsymbol{\Psi}, \mathfrak{I}_{t1}]$ represents the filtered probability of state i at time $t-1$ obtained via Hamiltons filter (Hamilton & Susmel, 1994).

The ML estimator $\hat{\boldsymbol{\Psi}}$ is obtained by maximizing the logarithm of Equation (2.16). In the case of MCMC estimation, we follow Ardia et al. (2008), by combining the likelihood with a diffuse (truncated) prior $f(\boldsymbol{\Psi})$ to build the kernel of the posterior distribution $f(\boldsymbol{\Psi}|\mathfrak{I}_T)$. As the posterior is of an unknown form (the normalizing constant is numerically intractable), it must be approximated by simulation techniques.

2.5 Comparing model perfomance

A well-fitting model will result in the predicted values being close to the observed data values. Forecast evaluation is key in evaluating the perfomance of a model. Evaluation of competing volatility models can be difficult because, as remarked by Bollerslev et al. (1994) and Lopez et al. (2001), there does not exist a unique criterion capable of selecting the best model. According to Marcucci (2005) and Wei et al. (2010), the real loss function has been used by several researchers to evaluate volatility forecast. Instead of choosing a particular statistical loss function as the best and unique criterion, one can use different interepretations and can lead to a more complete forecast evaluation of the competing models. The statistical loss functions are:

$$MSE_1 = n^{-1}\sum_{t=1}^{n}(\hat{\sigma}_{t+1} - \hat{h}_{t+1|t}^{\frac{1}{2}})^2 \tag{2.18}$$

$$MSE_2 = n^{-1}\sum_{t=1}^{n}(\hat{\sigma}_{t+1}^2 - \hat{h}_{t+1|t})^2 \tag{2.19}$$

$$QLIKE = n^{-1}\sum_{t=1}^{n}(\log\hat{h}_{t+1} + \hat{\sigma}_{t+1}^2\hat{h}_{t+1|t}^{-1}) \tag{2.20}$$

$$R2LOG = n^{-1}\sum_{t=1}^{n}[\log\hat{\sigma}_{t+1}^2\hat{h}^{-1} - t + 1|t]^2 \tag{2.21}$$

$$MAD_1 = n^{-1}\sum_{t=1}^{n}|\hat{\sigma}_{t+1} - \hat{h}_{t+1|t}^{\frac{1}{2}}| \tag{2.22}$$

$$MAD_2 = n^{-1}\sum_{t=1}^{n}|\hat{\sigma}_{t+1}^2 - \hat{h}_{t+1|t}| \tag{2.23}$$

where $\hat{h}_{t+1|t}$ is the h-step volatility forecast and $\hat{\sigma}_{t+1}^2$ is volatility as time $t+h$. The criteria in equations (2.18) and (2.19) are typically mean squared error metrics. The criteria in Equations (2.19) and (2.21) are equivalent to using the R^2 metric in the Mincer-Zarnowitz regression of $\hat{\sigma}_{t+1}^2$ on a constant $\hat{h}_{t+1|t}$ and of log $(\hat{\sigma}_{t+1}^2)$ on a constant and log $\hat{h}_{t+1|t}$, respectively provided that the forecasts are unbiased. Moreover, the $R2LOG$ loss function has the particular feature of penalizing volatility forecast asymmetrically in low volatility and high high volatility periods as pointed out by (Pagan & Schwert, 1990) who put forward equation (2.20), calling it logarithm loss function. The loss function in equation (2.20) corresponds to the loss implied by a gaussian likelihood and is suggested by Bollerslev et al. (1994). The Mean Absolute Deviation (MAD) criteria in equations (2.10) and (2.11) are useful because they are generally more robust to the possible presence of outliers than the MSE criteria, but they impose the same penalty on over- and under-predictions and are not invariant to scale transformations.

Chapter 3

Methodology

3.1 Data description

This study is about forecasting volatility associated with the price of tea at the Mombasa Tea Auction. Weekly weighted average spot prices is used to obtain the results as trading of the main grade takes place on Teusdays. The spot price data range from early 2010 to end of 2017. The forecast power of the MRS-GARCH model is evaluated on both in-sample and out-of-sample data, which is then compared to the single regime GARCH-type models. Observations till end of June 2017 are used as in-sample data for estimating the models, while the remaining observations, to December 2017, are selected as out-of-sample data to evaluate the forecasting performance. The data was obtained from the EATTA.

3.2 Decomposition of the time series

If P_t is the price at time t, the natural logarithm of the returns, r_t, is calculated for the price data.

$$r_t = \ln\, P_t - \ln\, P_{t-1} = p_t - p_{t-1} \tag{3.1}$$

Working with logarithm of returns, the additive model was used to decompose the time series. From Equation (3.1), the resulting representation of the returns is:

$$r_t = T_t + S_t + \varepsilon_t \tag{3.2}$$

where T_t is the trend, S_t is the seasonality component and ε_t is the residual term.

The trend is estimated using either the autoregressive or moving average techniques. Estimation is done through maximum likelihood technique and examination of the goodness-of-fit determines the best fit.

When trend is remove form the series in Equation (3.2), the resulting series will be as shown below:

$$r_t - T_t = S_t + \varepsilon_t \tag{3.3}$$

A parametric seasonal pattern in Equation (3.3) is genarated based on a sinusoidal pattern. In this study, a sinusoidal function that has monthly and yearly components is chosen. Global supply of tea is affected by colder and warmers time of the year. Given that the Mombasa Tea Auction resides in a region where producion is all year round, it is expected that the demand will be affected by the time of the year, especially for periods when the other tea auctions in the would are not operating. In addition, given that the ultimate consumer demand is determined by the time of the month, tea demand will exhibit some monthly season pattern. Like Erlwein (2008) in her application of hidden markov model to model electricity prices, the seasonal component is be given by:

$$\begin{aligned} S_t = \sum_{h=1}^{3} \Bigg(& d1_h \sin(s_h \frac{2\pi}{50.5} t) + d2_h \cos(s_h \frac{2\pi}{50.5} t) \\ & + d3_h \sin(s_h \frac{2\pi}{4.21} t) + d4_h \cos(s_h \frac{2\pi}{4.21} t) \Bigg) \end{aligned} \tag{3.4}$$

for $s1 = 1$, $s2 = 2$ and $s3 = 4$ and the constants $d1_h$, $d2_h$, $d3_h$, and $d4_h$ are to be determined. The model in equation (3.4) assumes that a year has on average 50.4 weeks and a month has on average 4.21 weeks.

The residuals can now be obtained by removing seasonality component from Equation (3.3)

$$r_t - T_t - S_t = \varepsilon_t \tag{3.5}$$

3.3 Estimation Markov Regime Switching GARCH (MRS-GARCH) model

The resulting time series, $y_t = \varepsilon_t$, has no trend and seanality components. This is a zero mean process. The estimation of the MRS-GARCH model is tackled by maximum likelihood. The MRS-GARCH model is implemented according to Haas et al. (2004) specification. This implies that there were 2 separate single-regime conditional variance processes, possibly 2 separate conditional distributions, a Markov chain dictating the switches between regimes and a zero mean process.

The normal distribution, the Student-t distribution and GED distribution are considered for the conditional distribution of the standardize innovations $\eta_{t,k}$ in each regime of the Markov chain.

3.4 Evaluation criteria for forecast perfomance

From the dataset, GARCH$(1,1)$, EGARCH$(1,1)$, and GJR-GARCH(1,1) are estimated. Forecasts are generated from the models both in-sample and the out-of-sample data.

Goodness of fit is used to evaluate in-sample forecast perfomance and the loss functions as the evaluation criteria for the out-of-sample forecast performance.

The loss functions used are the Mean Squared Error (MSE), and Mean Absolute Error (MAE).

Chapter 4

Data

The weekly weighted-average spot prices, P_t, is obtained from the weekly turnover and volume as indicated in equation (4.1).

$$P_t = \frac{\text{week turnover}}{\text{week volume}} \tag{4.1}$$

4.1 Tea Prices

Empirical studies on financial time series involve returns rather than prices. Campbell et al. (1997) indicates why returns are preffered. In addition, the limited liability assumption implies that gross returns have a lognormal distribution. This makes their logarithm normally distributed. The weekly prices are transformed into continuously compounded returns. This is done by taking the log differences of the prices. Figure 1 highlights the price (in USD) and the weekly (log) return series for the prices in the study period.

Descriptive statistics of the log return series are represented in Table 8 in Appendix A. As table shows, the commodity has a weekly average return of -0.0177 % with a standard deviation of 3.7264%. The series also displays a positive skewness of 0.257 and a kurtosis of 7.147. These values indicate that the returns are not normally distributed, namely it has fatter tails. Also, p-value of the Shapiro-Wilk normality test statistic confirms the non-normality of price returns.

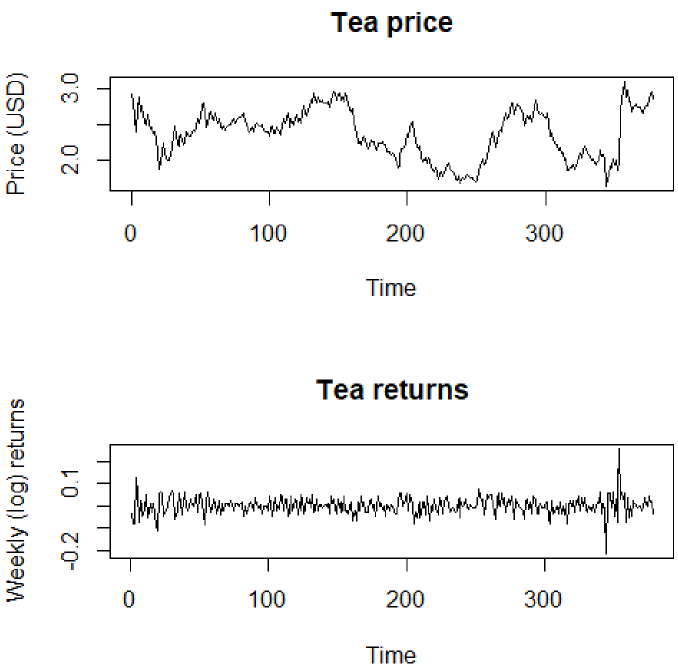

Figure 1: The graphs plots the price (in USD) and log returns for the main grade tea that was traded in the Mombasa Tea Action from 2010 to 2017

4.2 Return Series Decomposition

In order to analyse volatility of a commodity price, the time series is first decomposed to extract the trend and seasonality components.

The Akaike Information Criteria and log likelihood statistics are used in selecting a trend model. Table 9 in the Appendix presents perfomance of various models. $ARIMA(2,1,3)$ is the best performing model under both criteria. Table 1 indicate the parameters of the $ARIMA(2,1,3)$ model.

Table 1: ARIMA(2,1,3) model parameters

	ar1	ar2	ma1	ma2	ma3	intercept
estimate	0.3768	-0.9601	-0.286	0.9336	0.1517	-2.00e-04
s.e.	0.0175	0.0172	0.0532	0.0258	0.0534	2.00e-03
t-statistic	21.53143	-55.8198	-5.37594	36.18605	2.840824	-0.1

The sinusoidal function in Equation (3.1) is fit to capture seasonality on the residuals of the ARIMA(2,1,3) model

$$S_t = \sum_{h=1}^{3} \left(d1_h \sin(s_h \frac{2\pi}{50.5} t) + d2_h \cos(s_h \frac{2\pi}{50.5} t) + d3_h \sin(s_h \frac{2\pi}{4.21} t) + d4_h \cos(s_h \frac{2\pi}{4.21} t) \right) \tag{4.2}$$

where

$s_1 = 1; s_2 = 2; s_3 = 4;$

with $d1_h$, $d2_h$, $d3_h$ and $d4_h$ being the coefficients to be fitted.

The parameters for the sinusoidal function are indicated in table 10 in Appendix A. The coefficient that are not significantly different from zero have been dropped off.

Figure 2 shows the residuals series after extracting the trend and seasonality components in the return series.

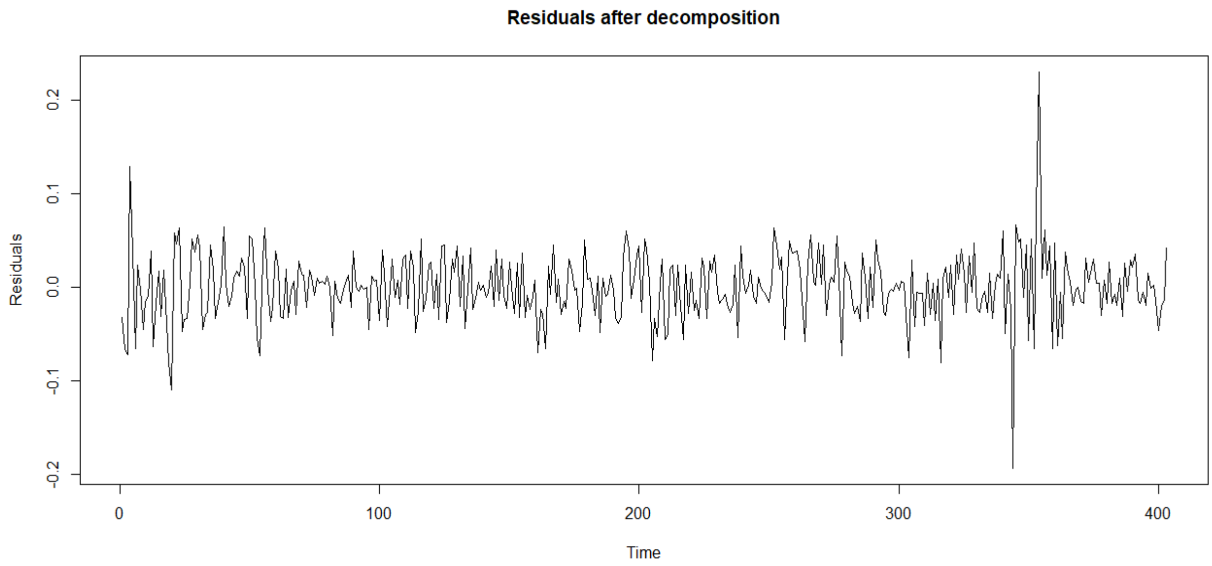

Figure 2: Residuals from the (log) return series after time series decomposition

4.3 Section summary

The logarithm of the return series has been obtained from the price series. The resulting series is then decomposed where the trend and seasonality components are extracted. $ARIMA(2,1,3)$ gives the best fit for trend and a sinusoidal function is used to capture seasonality. The residuals are applied to model volatility in the subsequent sections.

Chapter 5

Estimation of Model Parameters

This section uses maximum likelihood estimation to estimate the parameters for the MRS-GARCH and single-regime GARCH models.

5.1 Single-Regime GARCH Models

The procedures are computed numerically by using the R Package *rugarch*.

Table 11 in Appendix B presents estimated parameters for the uniregime GARCH models. From the table, and at 5% level of significance, the constant parameter is significant for the EGARCH models with normal and GED innovations, the ARCH parameter is significant for the standard GACRH models with GED innovations, the GARCH parameter is significant for all the models and the asymmetry effect term is significantly different from zero for the EGARCH model and not the GJR-GARCH model. Asymmetry indicates negative returns have higher conditional variance as compared to positive returns of similar size.

The degree of volatility persistence for GARCH models can be obtained by summing ARCH and GARCH parameters estimates $(\alpha_1 + \beta_1)$. For EGARCH (1, 1) and GJR-GARCH (1, 1), persistence is equal to β_1 and $(\alpha_1 + \gamma_1)/2 + \beta_1$ respectively. All models display strong persistence in volatility ranging from 0.89 to 0.98. This implies that once volatility has increase it tends to remain high.

If distribution assumptions for standardized errors are compared, it reveals that normality assumption is highly outperformed by other two fat-tailed distributions in terms of loglikelihood values apart from the EGARCH model. It is an anticipated result because of the fat tails property of log reurns. Overall, the EGARCH model with GED distribution has the largest log-likelihood among uniregime GARCH models.

5.2 Markov Regime Switching GARCH Models

The resulting time series, $y_t = \varepsilon_t$, has no trend and seasonality components. This is a zero mean process. Estimation of the MRS-GARCH model is tackled by maximum likelihood. The MRS-GARCH model is implemented according to Haas et al. (2004) specification and the procedures are computed numerically by using the R Package $MSGARCH$. This implies that there were two separate single-regime conditional variance processes, with two separate conditional distributions, a Markov chain dictating the switches between regimes and a zero mean process. The normal distribution, the Student-t distribution and GED distribution are considered for the conditional distribution of the standardize innovations $\eta_{t,k}$ in each regime of the Markov chain.

Estimation results and summary statistics of MRS-GARCH models are presented in Table 2, Table 3 and Table 4. Almost all parameter estimates are significantly different from zero atleast at 95% confidence level. The long term volatility level depends on the estimates of constant parameter α_0. Results are consistent with this argument and display that there are huge differences between α_0 estimates of each volatility regime. The parameter estimates α_0 in high volatility regimes are considerably greater than parameter estimates α_0 in low volatility regimes. Moreover, short run dynamics of volatility is determined by the ARCH parameter

α_1 and GARCH parameter β_1. Large estimates of α_1 suggest that effect of shocks to future volatility die out in a long time, so volatility is persistent. Large values of α_1 display reaction of volatility to the recent price changes.

Comparing the low and high volatility regimes in all MRS-GARCH models, the former volatility regimes have higher α_1 estimates and higher β_1 estimates than latter volatility regimes have, apart from some case where innovations in high volatility regime have GED conditional distribution. So, the GARCH processes in the low volatility regimes are more reactive and more persistent than that in the high volatility regime. In addition, it is interesting to notice that in most cases the degree of volatility persistence $(\alpha_1 + \beta_1)$ within low volatility regime is higher compared to the high volatility regime. Persistence within each regime is calculated as $\alpha_{i1} + \beta_{i1}$ where $i = 1, 2$.

5.3 Section summary

Table 11 in Appendix B presents parameters for the uniregime GARCH models. Parameters of MRS-GARCH models are presented in Table 2, Table 3 and Table 4. Important to note is that MRS-GARCH accurately describes the two regimes based on the different pattern of adjustment of the returns volatility. Estimation of the MRS-GARCH model indicates that the probability of stagnating within states is high. In addition, the probability of low volatility regime being be followed by a low volatility regime greater than the probability of a high volatility regime being followed by a high volatility regime.

Table 2: MRS-GARCH model with the low volatility regime having a normal conditional distribution

Volatility regime Conditional distribution		Low volatility Normal	High volatility Normal	Low volatility Normal	High volatility Student-t	Low volatility Normal	High volatility GED
α_0							
	estimate	0.0002	0.006	0.0002	0	0.0002	0.0013
	std. error	0	0.0065	0	0	0	0.0047
	t-statistic	38.8127	0.9259	24.1814	9.8672	40.0562	0.2827
	$(> \lvert t \rvert)$	<1e-16	1.77E-01	<1e-16	<1e-16	<1e-16	3.89E-01
α_1							
	estimate	0.143	0.0002	0.0488	0.3786	0.1523	0
	std. error	0.0066	0.0006	0.0039	0.3938	0.0065	0
	t-statistic	21.7091	0.3098	12.4229	0.9615	23.431	0.0655
	t-statistic	21.7091	0.3098	12.4229	0.9615	23.431	0.0655
	$\Pr(> \lvert t \rvert)$	<1e-16	3.78e-01	<1e-16	1.68e-01	<1e-16	4.74e-01
β_1							
	estimate	0.6127	0.2121	0.7751	0.6213	0.6014	0.8136
	std. error	0.0064	0.8508	0.0074	0	0.0063	0.6593
	t-statistic	95.4001	0.2493	105.0808	84434.192	94.8401	1.2341
	$\Pr(> \lvert t \rvert)$	<1e-16	4.02E-01	<1e-16	<1e-16	<1e-16	1.09E-01
df							
	estimate				5.6378		1.4473
	std. error				0.131		0.0303
	t-statistic				43.0442		47.8026
	$\Pr(> \lvert t \rvert)$				<1e-16		<1e-16
p_{11}	estimate						
	estimate	0.9948		0.8105		0.9951	
	std. error	0.0039		0.0128		0.0037	
	t-statistic	255.5282		63.173		265.5316	
	$\Pr(> \lvert t \rvert)$	<1e-16		<1e-16		<1e-16	
p_{21}							
	estimate	0.0796		0.38		0.0748	
	std. error	0.0003		0.0064		0.0003	
	t-statistic	282.1961		59.5095		284.5336	
	$\Pr(> \lvert t \rvert)$	<1e-16		<1e-16		<1e-16	
LogLikelihood		746.5646		748.6227		746.8806	
Persistence		0.7557	0.2123	0.8239	0.6213	0.7537	0.8136

Table 3: MRS-GARCH model with the low volatility regime having a student-t conditional distribution

Volatility regime Conditional distribution		Low volatility Student-t	High volatility Normal	Low volatility Student-t	High volatility Student-t	Low volatility Student-t	High volatility GED
α_0							
	estimate	0.0002	0.006	0.0002	0	0.0002	0.0013
	std. error	0	0.0065	0	0	0	0.0047
	t-statistic	38.8127	0.9259	24.1814	9.8672	40.0562	0.2827
	$\Pr(>\lvert t\rvert)$	<1e-16	1.77e-01	<1e-16	<1e-16	<1e-16	3.89e-01
α_1							
	estimate	0.143	0.0002	0.0488	0.3786	0.1523	0
	std. error	0.0066	0.0006	0.0039	0.3938	0.0065	0
	t-statistic	21.7091	0.3098	12.4229	0.9615	23.431	0.0655
	$\Pr(>\lvert t\rvert)$	<1e-16	3.78e-01	<1e-16	1.68e-01	<1e-16	4.74e-01
β_1							
	estimate	0.6127	0.2121	0.7751	0.6213	0.6014	0.8136
	std. error	0.0064	0.8508	0.0074	0	0.0063	0.6593
	t-statistic	95.4001	0.2493	105.0808	84434.192	94.8401	1.2341
	$\Pr(>\lvert t\rvert)$	<1e-16	4.02e-01	<1e-16	<1e-16	<1e-16	1.09e-01
df							
	estimate				5.6378		1.4473
	std. error				0.131		0.0303
	t-statistic				43.0442		47.8026
	$\Pr(>\lvert t\rvert)$				<1e-16		<1e-16
p_{11}							
	estimate	0.9948		0.8105		0.9951	
	std. error	0.0039		0.0128		0.0037	
	t-statistic	255.5282		63.173		265.5316	
	Pr(>—t—	<1e-16		<1e-16		<1e-16	
p_{21}							
	estimate	0.0796		0.38		0.0748	
	std. error	0.0003		0.0064		0.0003	
	t-statistic	282.1961		59.5095		284.5336	
	$\Pr(>\lvert t\rvert)$	<1e-16		<1e-16		<1e-16	
LogLikelihood		746.5646		748.6227		746.8806	
Persistence		0.7557	0.2123	0.8239	0.6213	0.7537	0.8136

Table 4: MRS-GARCH model with the low volatility regime having a GED conditional distribution

Volatility regime Conditional distribution		Low volatility GED	High volatility Normal	Low volatility GED	High volatility Student-t	Low volatility GED	High volatility GED
α_0							
	estimate	0.0002	0.0038	0.0002	0.0011	0.0002	0.0017
	std. error	0	0.0391	0	0.0029	0	0.0046
	t-statistic	34.1863	0.0976	34.7695	0.3813	37.7262	0.3682
	Pr($>\|t\|$)	<1e-16	4.61e-01	<1e-16	3.52e-01	<1e-16	3.56e-01
α_1							
	estimate	0.1187	0	0.1293	0	0.1369	0
	std. error	0.0077	0.0004	0.008	0	0.0073	0
	t-statistic	15.4579	0.056	16.1058	0.0797	18.8091	0.0739
	Pr($>\|t\|$)	<1e-16	4.78e-01	<1e-16	4.68e-01	<1e-16	4.71e-01
β_1							
	estimate	0.6333	0.4893	0.619	0.8371	0.6077	0.7612
	std. error	0.0072	5.2355	0.0071	0.4272	0.0068	0.6486
	t-statistic	88.1613	0.0935	86.6458	1.9593	89.3631	1.1735
	Pr($>\|t\|$)	<1e-16	4.63e-01	<1e-16	2.50e-02	<1e-16	1.20e-01
df							
	estimate	2.2459		2.2253	5.1385	2.2161	1.4895
	std. error	0.0164		0.0158	0.2644	0.0149	0.0315
	t-statistic	137.0798		140.7271	19.4327	148.3549	47.2446
	Pr($>\|t\|$)	<1e-16		<1e-16	<1e-16	<1e-16	<1e-16
p_{11}							
	estimate	0.9933		0.9943		0.9951	
	std. error	0.0047		0.0042		0.0037	
	t-statistic	212.0432		236.3434		265.5316	
	Pr($>\|t\|$)	<1e-16		<1e-16		<1e-16	
p_{21}							
	estimate	0.0876		0.0779		0.0748	
	std. error	0.0004		0.0003		0.0003	
	t-statistic	228.5858		245.5163		284.5336	
	Pr($>\|t\|$)	<1e-16		<1e-16		<1e-16	
LogLikelihood		746.9488		747.2084		746.8806	
Persistence		0.752	0.4893	0.7483	0.8371	0.7446	0.7612

Chapter 6

MRS-GARCH Model Perfomance

The data is divided into a six and a half year in-sample model estimation period (379 observations) and a subsequent half year out-of-sample forecasting period (25 observations). From the dataset, GARCH$(1,1)$, EGARCH$(1,1)$, and GJR-GARCH(1,1) have been estimated. Forecasts are generated from the models both in-sample and the out-of-sample data. Goodness of fit is used to evaluate in-sample forecast perfomance and the loss functions is the evaluation criteria for the out-of-sample forecast performance. The loss functions used are the Mean Squared Error (MSE), and Mean Absolute Error (MAE).

6.1 In samples

Table 5 provides a summary of the goodness-of-fit statistics that are considered in analysing the in-sample estimation performance of the volatility models. MRS-GARCH model with normal and student-t conditional distribution for the low and high volatility regimes respectively gives the best fit. All the MRS-GARCH models rank above the single regime GARCH mdoels. Thus, evaluating in sample estimation results according the goodness-of-fit statistics, the MRS-GARCH models perform better than single regimes GARCH models in describing the tea price volatility. In addition, comparing persistence of single regime GARCH models and MRS-GARCH models, it is observed that the high persistence in the former specification is reduced by latter models. This result indicates that high persistence in volatility of GARCH models is caused by regime shifts in the volatility process.

Table 5: In-sample evaluation

model	Conditional distribution of the low volatility regime	Conditional distribution of the high volatility regime	N_Par	Log(L)	Rank
GARCH	Normal		4	649.3951	18
GARCH	Student-t		5	676.7214	16
GARCH	GED		5	658.2845	17
EGARCH	Normal		5	727.6112	11
EGARCH	Student-t		6	709.4254	13
EGARCH	GED		6	734.5286	10
GJR	Normal		5	707.5468	15
GJR	Student-t		6	711.9217	12
GJR	GED		6	708.6446	14
MRS-GARCH	Normal	Normal	8	746.5646	8
	Normal	Student-t	9	748.6227	1
	Normal	GED	9	746.8806	6
	Student-t	Normal	9	746.4203	9
	Student-t	Student-t	10	746.9982	4
	Student-t	GED	10	746.7436	7
	GED	Normal	9	746.9488	5
	GED	Student-t	10	747.5236	2
	GED	GED	10	747.2084	3

6.2 Out-of-Sample Evaluation

This section investigates the ability of MRS-GARCH models and the single regime GARCH models to forecast tea price volatility at different future time horizons using the daily squared forecast error as actual volatility. The forecast horizons of 1, 2, 3, 5, 10, 15, 20 and 25 weeks were considered. Table 6 and Table 7 present the forecast errors in the volatility statistics.

Table 6: Mean absolute error in the forecast for various time horizon

	Low vol. regime	High vol. regime	1-week ahead		2-weeks ahead		3-weeks ahead		5-weeks ahead		10-weeks ahead		15-weeks ahead		20-weeks ahead		25-weeks ahead	
model			MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank
GARCH	Normal		0.021433	7	0.021433	1	0.014704	1	0.015676	1	0.014593	1	0.010879	1	0.010805	1	0.009887	1
	Student-t		0.021653	12	0.021653	3	0.014774	3	0.01581	3	0.014783	3	0.011057	3	0.011005	3	0.010058	3
	GED		0.021453	8	0.021453	2	0.014711	2	0.015686	2	0.014607	2	0.010892	2	0.010819	2	0.009898	2
EGARCH	Normal		0.025043	14	0.025043	7	0.018343	10	0.021272	11	0.024008	18	0.023259	18	0.025713	18	0.026212	18
	Student-t		0.026393	18	0.026393	18	0.018556	11	0.019442	4	0.018112	4	0.014399	4	0.014651	4	0.01358	4
	GED		0.023968	13	0.023968	4	0.017043	4	0.019548	5	0.021365	9	0.019887	9	0.021772	10	0.021828	10
GJRGARCH	Normal		0.025748	16	0.025748	16	0.018316	8	0.01987	8	0.019596	6	0.016371	6	0.016767	6	0.015534	6
	Student-t		0.025743	15	0.025743	15	0.018313	7	0.019864	7	0.019588	5	0.016361	5	0.016758	5	0.015524	5
	GED		0.025748	16	0.025748	16	0.018316	8	0.01987	8	0.019596	6	0.016371	6	0.016767	6	0.015534	6
MRS-GARCH	Normal	Normal	0.021334	4	0.025206	9	0.018685	14	0.021481	15	0.023251	14	0.021759	14	0.023684	14	0.023624	16
		Student-t	0.020973	1	0.024383	5	0.017577	5	0.019835	6	0.021149	8	0.019236	8	0.020707	8	0.020449	8
		GED	0.021378	6	0.025332	11	0.018754	15	0.021462	14	0.023131	12	0.021588	12	0.023461	12	0.023352	12
	Student-t	Normal	0.021312	3	0.025183	8	0.01865	12	0.02145	13	0.023235	13	0.021754	13	0.023672	13	0.023607	14
		Student-t	0.021009	2	0.024478	6	0.017777	6	0.020091	10	0.021709	10	0.020108	10	0.021751	9	0.021768	9
		GED	0.021351	5	0.025224	10	0.018655	13	0.021394	12	0.023099	11	0.021564	11	0.023442	11	0.023339	11
	GED	Normal	0.021464	9	0.025406	12	0.018811	16	0.021653	18	0.023496	17	0.02207	17	0.02402	17	0.02394	17
		Student-t	0.021491	11	0.025511	14	0.018929	18	0.0216	16	0.02334	15	0.021859	16	0.023747	16	0.023612	15
		GED	0.021485	10	0.025445	13	0.018868	17	0.02163	17	0.023348	16	0.021798	15	0.02369	15	0.023569	13

Table 7: Mean squared error in the forecast for various time horizon

	low vol. regime	high vol. regime	1-week ahead		2-weeks ahead		3-weeks ahead		5-weeks ahead		10-weeks ahead		15-weeks ahead		20-weeks ahead		25-weeks ahead	
model			MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank	MAE	rank
	Normal		0.000469	10	0.000469	1	0.000313	1	0.000327	1	0.000261	1	0.000183	1	0.000171	1	0.000149	1
GARCH	Student-t		0.000478	12	0.000478	3	0.000319	3	0.000334	3	0.000268	3	0.000188	3	0.000176	3	0.000153	3
	GED		0.000469	11	0.000469	2	0.000313	2	0.000327	2	0.000262	2	0.000183	2	0.000171	2	0.000149	2
	Normal		0.000643	14	0.000643	7	0.000437	7	0.000543	11	0.00064	18	0.000589	18	0.00072	18	0.000744	18
EGARCH	Student-t		0.000704	18	0.000704	18	0.000472	18	0.000478	5	0.000386	4	0.000279	4	0.000274	4	0.000244	4
	GED		0.000588	13	0.000588	4	0.000395	4	0.000473	4	0.000514	9	0.000443	9	0.000525	10	0.000526	10
	Normal		0.000673	16	0.000673	16	0.000453	14	0.00049	8	0.000438	6	0.00033	6	0.000334	6	0.0003	6
GJRGARCH	Student-t		0.000673	15	0.000673	15	0.000452	13	0.000489	7	0.000438	5	0.00033	5	0.000333	5	0.000299	5
	GED		0.000673	16	0.000673	16	0.000453	14	0.00049	8	0.000438	6	0.00033	6	0.000334	6	0.0003	6
		Normal	0.000455	4	0.00065	9	0.000444	9	0.000551	15	0.000598	14	0.00052	14	0.000612	14	0.000608	16
	Normal	Student-t	0.00044	1	0.000606	5	0.000409	5	0.000482	6	0.000502	8	0.000419	8	0.000477	8	0.000466	8
		GED	0.000457	6	0.000657	11	0.000449	11	0.000551	14	0.000592	12	0.000513	12	0.000601	12	0.000595	12
		Normal	0.000454	3	0.000649	8	0.000443	8	0.00055	13	0.000597	13	0.00052	13	0.000612	13	0.000607	15
MRS-GARCH	Student-t	Student-t	0.000441	2	0.000611	6	0.000414	6	0.00049	10	0.000526	10	0.000455	10	0.000523	9	0.000523	9
		GED	0.000456	5	0.000651	10	0.000444	10	0.000547	12	0.00059	11	0.000512	11	0.0006	11	0.000595	11
		Normal	0.000461	7	0.000661	12	0.000451	12	0.000561	18	0.000611	17	0.000534	17	0.000629	17	0.000624	17
	GED	Student-t	0.000462	9	0.000667	14	0.000456	17	0.000557	16	0.000602	15	0.000525	16	0.000615	16	0.000607	14
		GED	0.000462	8	0.000663	13	0.000453	16	0.000559	17	0.000602	16	0.000523	15	0.000612	15	0.000606	13

The MAE and MSE were used to evaluate the models. The error statistics are consistent in there ranking. The forecast error statistics suggest that MRS-GARCH models provide the most accurate volatility forecasts for 1-period ahead, and the GARCH models were better at longer horizon (for 2-weeks to 25-weeks ahead).

Chapter 7

Conclusion

The aim of this study is two fold. It seeks to develop a model to describe regime switching in volatility of return for tea traded at the Mombasa tea Auction using weekly data for the period between 2010 and 2017. This is achieved through a MRS-GARCH model. The second aim is to compare the perfomance of the MRS-GARCH model with other single-regime GARCH models. Several findings result from the presented analysis. There is evidence of a regime switching GARCH model in the volatility of tea prices. In addition, the estimation of the MRS-GARCH describes the two regimes based on the different parameters; and the estimated model captures all the events that are responsible for the presence of nonlinear features in the returns. Moreover, regime clustering is observed. A low volatility regime is more likely to be followed by a low volatility regime than for a high volatility regime to be followed by a high volatility regime. Lastly, consider several competing models to forecast returns volatility by obtaining the 1-, 2-, 3-, 5-,10-, 15-, 20- and 25- step ahead forecast and comparing the out-of-sample performance of the models on the basis of forecasting accuracy by applying statistical loss function, the results suggest that MRS-GARCH models has priority over single regime GARCH models for a period ahead, and the single regime GARCH processes bet the MRS-GARCH processes for longer time horizon.

For future work, Bayesian algorithm using a Gibbs sampling algorithm can be used to estimate the MRS-GARCH model, as an alternative to maximum likelihood estimation. Moreover, there are other techniques used to model volatility, such as those governed by a stochastic equation. Markov regime switching can be applied in such models.

References

Abramson, A., & Cohen, I. (2007). On the stationarity of markov-switching garch processes. *Econometric Theory*, *23*(3), 485–500.

Andersen, T. G., Chung, H.-J., & Sørensen, B. E. (1999). Efficient method of moments estimation of a stochastic volatility model: A monte carlo study. *Journal of econometrics*, *91*(1), 61–87.

Ardia, D., et al. (2008). Financial risk management with bayesian estimation of garch models. *Lecture notes in economics and mathematical systems*, *612*.

Bauwens, L., Preminger, A., & Rombouts, J. V. (2006). Regime switching garch models.

Biswas, A., & Goswami, A. (2017). Option pricing in a regime switching stochastic volatility model. *arXiv preprint arXiv:1707.01237*.

Bollen, N. P., Gray, S. F., & Whaley, R. E. (2000). Regime switching in foreign exchange rates:: Evidence from currency option prices. *Journal of Econometrics*, *94*(1-2), 239–276.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). Arch models. *Handbook of econometrics*, *4*, 2959–3038.

Cai, J. (1994). A markov model of switching-regime arch. *Journal of Business & Economic Statistics*, *12*(3), 309–316.

Campbell, J. Y., Lo, A. W.-C., & MacKinlay, A. C. (1997). *The econometrics of financial markets.* princeton University press.

Chang, K. (2015). World tea production and trade: Current and future development. *A publication by the Food and Agricultural Organization of the United Nations, Rome. Available online at www. fao. org (date accessed 29 May 2016).*

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

Edwards, S., & Susmel, R. (2001). Volatility dependence and contagion in emerging equity markets. *Journal of Development Economics*, *66*(2), 505–532.

Elliott, R. J., Hunter, W. C., & Jamieson, B. M. (1998). Drift and volatility estimation in discrete time. *Journal of Economic Dynamics and Control*, *22*(2), 209–218.

Elliott, R. J., Kuen Siu, T., & Chan, L. (2007). Pricing volatility swaps under heston's stochastic volatility model with regime switching. *Applied Mathematical Finance*, *14*(1), 41–62.

Elliott, R. J., Siu, T. K., Chan, L., & Lau, J. W. (2007). Pricing options under a generalized markov-modulated jump-diffusion model. *Stochastic Analysis and Applications*, *25*(4), 821–843.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Erlwein, C. (2008). *Applications of hidden markov models in financial modelling* (Unpublished doctoral dissertation). Brunel University, School of Information Systems, Computing and Mathematics.

Exterkate, P., Knapik, O., et al. (2017). *A regime-switching stochastic volatility model for forecasting electricity prices* (Tech. Rep.). Department of Economics and Business Economics, Aarhus University.

Friedman, B. M., Laibson, D. I., & Minsky, H. P. (1989). Economic implications of extraordinary movements in stock prices. *Brookings Papers on Economic Activity*, *1989*(2), 137–189.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The journal of finance*, *48*(5), 1779–1801.

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.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, *42*(1), 27–62.

Haas, M., Mittnik, S., & Paolella, M. S. (2004). A new approach to markov-switching garch models. *Journal of Financial Econometrics*, *2*(4), 493–530.

Hafner, C. M., & Preminger, A. (2010). Deciding between garch and stochastic volatility via strong decision rules. *Journal of Statistical Planning and Inference*, *140*(3), 791–805.

Hamilton, J. D. (2010). Regime switching models. In *Macroeconometrics and time series analysis* (pp. 202–209). Springer.

Hamilton, J. D., & Lin, G. (1996). Stock market volatility and the business cycle. *Journal of applied econometrics*, 573–593.

Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of econometrics*, *64*(1-2), 307–333.

Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *The Review of Economic Studies*, *61*(2), 247–264.

Hentschel, L. (1995). All in the family nesting symmetric and asymmetric garch models. *Journal of Financial Economics*, *39*(1), 71–104.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, *6*(2), 327–343.

Hillebrand, E. (2005). Neglecting parameter changes in garch models. *Journal of Econometrics*, *129*(1-2), 121–138.

Kanas, A. (2005). Regime linkages in the us/uk real exchange rate–real interest differential relation. *Journal of International Money and Finance*, *24*(2), 257–274.

Klaassen, F. (2002). Improving garch volatility forecasts with regime-switching garch. In *Advances in markov-switching models* (pp. 223–254). Springer.

Lopez, J. A., et al. (2001). Evaluating the predictive accuracy of volatility models. *Journal of Forecasting*, *20*(2), 87–109.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The journal of business*, *36*(4), 394–419.

Marcucci, J. (2005). Forecasting stock market volatility with regime-switching garch models. *Studies in Nonlinear Dynamics & Econometrics*, *9*(4).

Mikosch, T., & Stărică, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the igarch effects. *Review of Economics and Statistics*, *86*(1), 378–390.

Monfardini, C. (1998). Estimating stochastic volatility models through indirect inference. *The Econometrics Journal*, *1*(1), 113–128.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347–370.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of econometrics*, *45*(1-2), 267–290.

Reher, G., Wilfling, B., et al. (2011). Markov-switching garch models in finance: a unifying framework with an application to the german stock market. *Center for Quantitative Economics (CQE), University of Muenster*, 24–25.

Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of econometrics*, *63*(1), 289–306.

Schnepf, R. D. (2005). Price determination in agricultural commodity markets: a primer..

Schwert, G. W. (1989). Why does stock market volatility change over time? *The journal of finance*, *44*(5), 1115–1153.

So, M. E. P., Lam, K., & Li, W. K. (1998). A stochastic volatility model with markov switching. *Journal of Business & Economic Statistics*, *16*(2), 244–253.

Taylor, S. J. (2008). *Modelling financial time series.* world scientific.

Tsay, R. S. (2005). *Analysis of financial time series* (Vol. 543). John Wiley & Sons.

Wei, Y., Wang, Y., & Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using garch-class models. *Energy Economics*, *32*(6), 1477–1484.

Zhang, Y.-J., Yao, T., & He, L.-Y. (2015). Forecasting crude oil market volatility: can the regime switching garch model beat the single-regime garch models? *arXiv preprint arXiv:1512.01676*.

A Data

Table 8: Summary statistics for (log) returns

item name	(log) returns
number valid	403
mean	-0.000177
standard deviation	0.037264
median	0
trimmed mean	0.000203
mean absolute deviation	0.035024
minumum	-0.21334
maximum	0.257377
skewness	0.267529
kurtosis	7.417612
standard error	0.001856
Shapiro-Wilk normality test	W = 0.93642 p-value = 4.265e-12

Table 9: Evaluating model perfomance for the trend

Model	Log (Likelihood)	Rank	AIC	Rank
ARIMA(1,1,0) or AR(1)	755.82	15	-1505.64	10
ARIMA(2,0,0) or AR(2	758.03	10	-1508.07	2
ARIMA(3,0,0) or AR(3	758.4	7	-1506.8	4
ARIMA(0,1,1) or MA(1)	756.15	14	-1506.29	7
ARIMA(0,1,2) or MA(2)	757.58	12	-1507.15	3
ARIMA(0,1,3) or MA(3)	758.38	8	-1506.76	5
ARIMA(1,1,1) or AR(1) MA(1)	756.71	13	-1505.43	11
ARIMA(1,1,2) or AR(1) MA(2)	757.96	11	-1505.92	8
ARIMA(2,1,3) or AR(1) MA(3)	758.53	6	-1505.06	13
ARIMA(2,1,1) or AR(2) MA(1)	758.26	9	-1506.53	6
ARIMA(2,1,2) or AR(2) MA(2)	758.91	4	-1505.83	9
ARIMA(2,1,3) or AR(2) MA(3)	763.19	1	-1512.39	1
ARIMA(3,1,1) or AR(3) MA(1)	758.57	5	-1505.14	12
ARIMA(3,1,2) or AR(3) MA(2)	758.95	3	-1503.89	14
ARIMA(3,1,3) or AR(3) MA(3)	759.08	2	-1502.16	15

Table 10: Parameters for the sinusoidal function

	Estimate	Std. Error	t value	Pr($> \|t\|$)
$d0(= intercept)$	1.34e-05	1.81e-03	0.007	0.9941
$d1_1$	-4.26e-03	2.56e-03	-1.665	0.0966
$d2_1$	-8.31e-04	2.56e-03	-0.324	0.7459
$d4_1$	-5.61e-04	2.56e-03	-0.219	0.8266
$d1_2$	1.95e-03	2.56e-03	0.761	0.447
$d3_2$	1.53e-03	2.56e-03	0.599	0.5495
$d4_2$	1.20e-03	2.56e-03	0.469	0.6392
$d1_3$	-4.69e-03	2.55e-03	-1.837	0.067
$d2_3$	-5.62e-04	2.57e-03	-0.219	0.8268
$d3_3$	8.93e-04	2.56e-03	0.348	0.7278
$d4_3$	1.25e-03	2.56e-03	0.488	0.6259

B Single-Regime GARCH Model Parameters

Table 11: Parameters for the standard GARCH models

		GARCH			EGARCH			GJR-GARCH		
conditional distribution		normal	student-t	GED	normal	student-t	GED	normal	student-t	GED
μ										
	estimate	0	0	0	0	0	0	0	0	0
	std. error	0.004272	0.002245	0.003482	0.001955	0.001225	0.001793	0.03097	0.105047	0.067624
	t-statistic	0	0	0	0	0	0	0	0	0
	$\Pr(>\lvert t\rvert)$	1	1	1	1	1	1	1	1	1
α_0										
	estimate	0.000001	0.000001	0.000001	-0.62961	-0.77241	-0.66396	0.000001	0.000001	0.000001
	std. error	0.000006	0.000002	0.000002	0.237781	0.637307	0.249642	0.000006	0.000003	0.000007
	t-statistic	0.22299	0.79051	0.5991	-2.64786	-1.21199	-2.65966	0.23562	0.53675	0.207609
	$\Pr(>\lvert t\rvert)$	0.82354	0.42923	0.549109	0.0081	0.225518	0.007822	0.813729	0.59144	0.835534
α_1										
	estimate	0.051124	0.052016	0.05117	0.009298	0.069662	-0.00117	0.050014	0.050013	0.050014
	std. error	0.053796	0.010134	0.013237	0.056342	0.083204	0.048881	0.271332	0.94387	0.437512
	t-statistic	0.95034	5.13279	3.8656	0.16503	0.83725	-0.02399	0.18433	0.052987	0.114314
	$\Pr(>\lvert t\rvert)$	0.34194	0	0.000111	0.868917	0.402455	0.980861	0.853758	0.957742	0.908989
β_1										
	estimate	0.9002	0.90015	0.900212	0.899854	0.899687	0.899952	0.89981	0.899813	0.899811
	std. error	0.12517	0.023295	0.035616	0.040652	0.08227	0.037202	0.189021	0.246493	0.258295
	t-statistic	7.1918	38.64125	25.2755	22.1355	10.93583	24.19078	4.76038	3.650456	3.483653
	$\Pr(>\lvert t\rvert)$	0	0	0	0	0	0	0.000002	0.000262	0.000495
γ_1										
	estimate				0.558456	1.156824	0.337742	0.050543	0.050471	0.05054
	std. error				0.268302	0.435122	0.116069	0.434688	1.98074	0.960866
	t-statistic				2.08145	2.65862	2.909835	0.11627	0.025481	0.052599
	$\Pr(>\lvert t\rvert)$				0.037393	0.007846	0.003616	0.907435	0.979671	0.958052
df										
	estimate		4.048854	1.91059		3.998909	1.989976		4.005752	1.97621
	std. error		0.153306	0.262925		2.225083	0.41181		1.396173	1.027163
	t-statistic		26.41034	7.2667		1.7972	4.832266		2.869094	1.92395
	$\Pr(>\lvert t\rvert)$		0	0		0.072305	0.000001		0.004116	0.054361
LogLikelihood		649.3951	676.7214	658.2845	727.6112	709.4254	734.5286	707.5468	711.9217	708.6446
Information Criteria										
Akaike		-3.4148	-3.5541	-3.4565	-3.8233	-3.7218	-3.8546	-3.7172	-3.735	-3.7177
Bayes		-3.3731	-3.502	-3.4045	-3.7713	-3.6594	-3.7922	-3.6651	-3.6726	-3.6552
Persistence		0.951324	0.952165	0.951381	0.899854	0.899687	0.899952	0.975095	0.975061	0.975095