A Comparison of GARCH Option Pricing Models Using Bayesian Inference and Implied Calibration

Arvid T.L. Voormanns[†]

Under the supervision of
dr. Bart Diris [‡] (supervisor)
prof. dr. Dick van Dijk [§] (coreader)

In partial fulfillment for the degree Master of Science

Econometrics and Management Science - Quantitative Finance

Erasmus University Rotterdam

8 April 2016

Abstract

This paper presents a comparison between the pricing performance of option implied GARCH models and the Bayesian GARCH option pricing model and a generalization thereof that employs a regime switching feature. An algorithm is presented that shows how option prices can be computed using a Markov switching GARCH model from a Bayesian perspective and detailed information is presented about the inference techniques. The estimated models are tested empirically by forecasting out-of-sample S&P500 index option prices over the year 2014 and the results are compared with 2 competitive benchmark models. The results show that the option implied GARCH and the Markov switching GARCH models substantially outperform the simple GARCH model. The Markov switching GARCH can compete quite well with the option implied models and generally outperforms these models for medium and long maturities.

Keywords: *Bayesian inference, implied calibration, GARCH option pricing, regime switching.*

[†] Graduate student Econometrics and Management Science at the Erasmus School of Economics
Student number: 402084
Contact: atl.voormanns@gmail.com
[‡] Assistant professor of financial econometrics at the Econometric Institute of Erasmus University Rotterdam
[§] Professor of financial econometrics at the Econometric Institute of Erasmus University Rotterdam

Acknowledgements

A word of gratitude is in place for the supervisor of this thesis, dr. Bart Diris, and the co-reader, prof. dr. Dick van Dijk. Both provided me with useful insights that improved the quality of this thesis. I am also grateful to my dear friend Ben for guiding me through the rough patches in life and keeping me focused when needed. Finally, I am forever grateful to my parents for their endless motivation, love and support. There is no doubt in my mind that I could not have come this far without them.

"Je gaat het pas zien als je het door hebt."

\- Johan Cruijff

Contents

1 Introduction **4**

2 Theoretical Framework **8**

2.1 Risk Neutralization of a GARCH process . . . 9

2.2 GJR-GARCH . . . 10

2.3 Markov Switching GJR-GARCH . . . 12

3 Methodology **14**

3.1 Bayesian Estimation of the GJR-GARCH Model . . . 14

3.1.1 Full Conditional Posterior of β . . . 16

3.1.2 Full Conditional Posterior of θ . . . 17

3.2 Bayesian Estimation of the MS GJR-GARCH Model . . . 17

3.2.1 Full Conditional Posterior of $\mathbf{P}'_i$ . . . 18

3.2.2 Full Conditional Posterior of $\mathbf{S}$ . . . 19

3.2.3 Posterior Distributions of $\theta_{G,j}$ and β_j . . . 20

3.2.4 Complete Posterior Simulation Scheme MS GJR-GARCH . . . 21

3.3 Prior Specifications and the Label Switching Problem . . . 21

3.4 Pricing the Option . . . 24

3.5 Model Calibration . . . 26

3.5.1 Implied GJR . . . 26

3.5.2 FHS GJR . . . 27

4 Benchmark Models **29**

4.1 Ad-hoc Black-Scholes . . . 29

4.2 Heston Nandi GARCH . . . 29

5 Data and Empirical Results **31**

5.1 Data Description . . . 31

5.2 Bayesian Inference Results . . . 33

5.2.1 Inference Results GJR . . . 34

5.2.2 Inference Results MS GJR . . . 35

5.2.3 Prior Sensitivity . . . 38

5.3 Calibration Results Option Implied Models . . . 39

5.4 Predictive Option Pricing Results 41

5.4.1 Analysis of Pricing Errors 41

5.4.2 Aggregate Results 44

5.4.3 Performance over Time 48

6 Conclusion and Limitations **60**

7 Appendix I: Inference Results Benchmark Models **65**

8 Appendix II: Pricing Algorithm Markov Switching GJR **67**

9 Appendix III: Laplace Approximation of Bayes Factors **68**

10 Appendix IV: Bayesian Estimation of the Ad-hoc BS **70**

1 Introduction

Financial derivatives are a set of financial instruments that are, as the name suggests, derived from financial products such as equities, bonds or loans. These derivatives are used in various ways and combinations to reduce risk and to speculate on market behaviour. The market of exchange-traded derivatives is enormous and had an estimated daily turnover of more than $4.2 trillion in 2014[1], which signifies the importance of a proper valuation of these instruments. One of the most popular types of derivatives are options, which give the holder the right to buy or sell an underlying asset for a fixed price at a certain point in the future[2]. In 1973 Black and Scholes ("BS") introduced their famous formula for pricing European options. The BS framework assumes that the volatility of the underlying and the risk-free rate are constant and that the stock price follows a geometric Brownian motion. The power of the BS model lies in its simplicity, but it also leaves room for improvement. Mainly the modeling of the volatility parameter can improve pricing performance, as the assumption of constant volatility implies that heteroskedasticity is ignored.

In this paper the heteroskedasticity is captured by modeling the volatility of the underlying asset using the model of Glosten, Jaganathan and Runkle (1993) ("GJR"), which is an extension of the GARCH model introduced by Bollerslev (1986). In the GJR model the conditional variance is able to react asymmetrically to negative and positive shocks, enabling it to capture a stylized fact known as the leverage effect[3]. Other asymmetric GARCH models could also be considered, such as the Exponential GARCH (Nelson et al., 1991), the Quadratic GARCH (Sentana, 1991) or the Treshold GARCH (Zakoian, 1991) for example. However, Engle and Ng (1993) present evidence that the GJR is the best parametric GARCH model and is able to outperform other asymmetric models, such as the EGARCH for example. Following up on this argument in an option pricing context are the papers of Bauwens and Lubrano (2002), Duan et al. (2006) and Barone-Adesi et al. (2008) who document a satisfying pricing performance using the GJR model. Based on these arguments, the GJR is favored and selected as baseline model for option pricing. The specific models under consideration are (1) the simple GJR, (2) the Markov Switching GJR ("MS GJR"), (3) the option implied GJR ("Implied GJR") and (4) the option implied GJR using the Filtered Historical Simulation approach from Barone-Adesi et al. (2008) ("FHS GJR"). The choice for these models is motivated by the fact that they can substantially capture the stylized

[1] Source: BIS Triennial Surve

[2] In this paper European style options are considered, which can be exercised only on one specific date.

[3] The fact that negative returns have a stronger effect on volatility than positive returns of the same magnitude.

facts of financial returns. Besides the leverage effect, the GJR is also able to capture volatility clustering, while the MS GJR also captures leptokurtosis and volatility jumps. The option implied GJR models are calibrated on observed market option prices and hence capture return dynamics that are reflected by the market and are hard to capture otherwise[4]. Furthermore, the FHS GJR is free of any distributional assumptions and relies only on past historical innovations. This enables the model to fully capture the structure of the empirical return distribution.

The pricing methodology employed in this paper builds on the work of Duan (1995), who introduced an option pricing model for the case when returns follow a GARCH(1,1) process with normal errors. The Duan (1995) framework can easily be extended to accommodate the GJR and the MS GJR models, provided that the normality assumption is maintained. This approach is followed by many authors (see for example Hafner and Herwartz, 1999; Heston and Nandi, 2000; Bauwens and Lubrano, 2002; Christofferson and Jacobs, 2004 and more). A drawback of this method is that the normality assumption implies that the frequently observed fat tails in return series cannot be captured when using GARCH models. However, Ang and Timmermann (2011) show that the individual regimes in the Markov switching model can be described as a mixture of normals, which enables the model to capture fat tails. Hence, in spite of the normality assumption the MS GJR is still able to capture high kurtosis to some extent. Alternatively, a leptorkurtic distribution could be imposed, such as the (skewed) students t-distribution or the generalized error distribution. However, this complicates the option valuation method of Duan (1995) and therefore the normality assumption is maintained. The only exception to this assumption is the FHS GJR, which relies entirely on past innovations, meaning no structure is imposed on the return distributions. This also means the FHS GJR does not depend on the Duan (1995) valuation framework.

For model estimation of the GJR and MS GJR a Bayesian Markov Chain Monte Carlo ("MCMC") procedure, introduced by Henneke et al. (2006), is used. The MCMC method allows the unobserved state variables in the MS GJR model to be simulated along the model parameters in the Gibbs sampler by using a technique known as data augmentation. The Markov switching GARCH model used in this paper most resembles the specification used by Henneke et al. (2006) and Bauwens et al. (2010) who also study the model from a Bayesian perspective. For the option implied models a Monte Carlo calibration procedure is used in which an error criterion function is minimized in an attempt to move the model price as close as possible to the market price

[4] Throughout this paper the term "calibrated" refers to option implied.

by letting the GARCH parameters vary. The models are re-estimated every 4 weeks and are used to predict out-of-sample S&P 500 index option prices on weekly basis throughout the year 2014. This means that the models are kept fixed for a period of 4 weeks and the forecast horizon ranges up to 4 weeks. The results are compared with two benchmark models, the ad-hoc Black Scholes ("AHBS") from Dumas et al. (1998), and the Heston and Nandi (2000) ("HN") closed form GARCH option pricing model. Both benchmark models are also estimated using Bayesian inference to enhance the comparability of the models.

Many papers in option pricing literature are concerned with only a single method of model estimation, either option implied or inferred from historical asset returns. Since option implied models have nowadays become the standard in financial industry it is interesting to investigate whether this is justified. Directly comparing the pricing results of both methods can provide insights on which method yields better results. This particular topic is not intensively studied in financial literature, mainly because it is well known that option implied models yield accurate pricing performance, while GARCH models estimated with historical returns usually do not perform well in this context (see for example Weber and Prokupczuk (2011)). This study however, adds two interesting features that makes such GARCH models more competitive; the Bayesian estimation methodology and a regime switching property. Thus, the main goal of this paper is to compare and discuss the pricing performance of the option implied GARCH models relative to the Bayesian GARCH models, and conclude whether one of these methods is superior. The literature is somewhat limited on this particular topic, although some examples that slightly overlap are Hsieh and Ritchen (2000) and Christofferson and Jacobs (2004) who both present a comparison of the pricing performance of different kind of calibrated and non-calibrated GARCH models. However, these papers do not explicitly address differences due to estimation methods but focus more on parsimony and updating rules. More closely related is the work of Weber and Prokopczuk (2011), who employ an implied calibration scheme to infer GARCH parameters from observed market option prices and compare its pricing performance with a GARCH model estimated with historical return data. They record an improved pricing performance of the calibrated model for all moneyness categories and maturities. However, the approach of Weber and Prokopczuk (2011) differs in their use of American options, which are early exercisable and therefore more complex to value. Furthermore, Weber and Prokopczuk (2011) use only frequentist estimation methods.

The application of Markov switching models for options pricing starts with the work of Duan et al. (2002), who develop a backward recursion procedure which yields the regime switching GARCH as limiting case. More recent articles include the work of Chen and Hung (2009), who develop a lattice algorithm, and Fuh et al. (2012), who use an approximation to arrive at a closed form solution. The MS GJR model proposed in this paper deviates from most models used in the literature due to its Bayesian estimation process and its forth flowing option pricing algorithm. The latter is based on the work of Bauwens and Lubrano (2002) and the MS GJR can therefore be seen as an extension of their single regime GJR option pricing model. To current knowledge, the Markov switching GARCH option pricing model has not yet been studied from a Bayesian perspective and therefore is new in the option pricing literature. The contribution of this paper can now be stated in twofold as (1) presenting a comparison of the pricing performance of option implied GARCH models and GARCH models estimated with historical asset returns and (2) showing how options can be priced using a regime switching GARCH specification from a Bayesian perspective.

The foundations of the Bayesian GARCH option pricing model were laid by Bauwens and Lubrano (2002), who priced options using Bayesian inference in combination with asymmetric GARCH models. Their findings suggest that modeling asymmetry leads to increased pricing performance for in-the-money options. Moreover, they show that Bayesian inference leads to a better estimation of the underlying volatility which subsequently improves the pricing performance. Duan et al. (2002) and Yoa et al. (2006) use a Markov switching GARCH option pricing model which is able to switch between different GARCH specifications, providing a high degree of flexibility. Both papers find that the regime switching feature more accurately describes volatility which subsequently improves pricing performance. Combining the regime switching property with the asymmetric GJR and the Bayesian estimation method could potentially yield a model that, from a theoretical perspective, captures all stylized facts and completely accounts for estimation uncertainty. The option implied models on the other hand, are based more on empirical results, as indicated by their popularity among practitioners. This study therefore represents a comparison of a practical method versus a theoretically well founded method.

The pricing results show that the option implied models perform well, both models convincingly outperform the single regime GARCH model, while their mutual performance is quite similar. The Markov switching GARCH model also convincingly outperforms the GJR and competes quite well with the option implied models. For most samples the MS GJR is actually able to outperform both of the option implied models. Especially in-the-money options and options with long maturities are priced more accurately by the Markov switching GARCH. However, the ad-hoc Black Scholes appears to be a tough benchmark and clearly dominates the test models.

2 Theoretical Framework

To price derivatives a measure is required that is free of investors' risk preferences. This is referred to as a risk-neutral measure and is necessary because the estimated price would otherwise reflect the investors' level of risk aversion, which yields a subjective price that is inconsistent with the market price[5]. To price derivatives using GARCH models, the expected payoff has to be discounted under a martingale measure [6]. This martingale measure however, is not unique because the GARCH framework implies that markets are incomplete, which means there exist contingent claims that cannot exactly be replicated by a replicating portfolio. This results in an infinite number of risk-neutral specifications under which the payoff function can be evaluated. Furthermore, there exists no risk neutral measure that does not change the unconditional variance or the conditional variance for more than one period ahead for GARCH processes. The choice for a risk-neutral measure has to be based on economic justifications. Duan (1995) introduced the *Locally Risk Neutral Valuation Relationship* ("LRNVR"), which is based on the economic foundation that agents are utility maximizers and that utility is additive and time-separable. The LRNVR does not alter the one period ahead conditional variance and ensures the future expected return equals the riskfree rate. The assumption of normality implies equal conditional variance under both measures and is of paramount importance in this framework. More recent studies indicate that the assumption of normality is not realistic as this cannot describe the frequently observed fat tails in return distributions. Typically a leptokurtic distribution, such as the student or skewed t-distribution (Fergusson and Platen, 2006) or the Generalized Error Distribution (Gao et al., 2012), is considered more suitable for modeling financial time series. For the application at hand however, imposing a distribution other than the normal would imply that the requirement of equal one period ahead conditional variance under $\mathbb{P}$ and $\mathbb{Q}$ no longer holds. Hence, since the Duan (1995) framework is not compatible with leptokuric distributions, the normality assumption is maintained in this paper. Using the LRNVR, Heston and Nandi (2000) derived a closed-form GARCH option pricing model that is able to produce good out-of-sample results when the GARCH parameters are updated every period. Allowing for asymmetry showed

[5]See Gisiger (2010) for a thorough explanation of risk-neutral pricing

[6]i.e. a fair measure

to further improve the pricing performance. Other examples on asymmetric GARCH option pricing models are Bauwens and Lubrano (2002), Hafner and Herwartz (1999) and Christofferson (2006), who confirmed these findings as they record a decrease in pricing errors when accounting for asymmetry. The latter is explained by the fact the asymmetric GARCH models are able to capture the leverage effect.

For a better understanding of the risk-neutralization principles of a GARCH model, the next section briefly outlines the seminal work of Duan (1995). Next, the GJR model is discussed and the LRNVR is used to derive the corresponding risk-neutral specification and subsequently the MS GJR model and its risk-neutral version are defined.

2.1 Risk Neutralization of a GARCH process

In the Duan (1995) framework the one-period rate of return is assumed to be lognormally distributed under the real world measure (denoted $\mathbb{P}$) , i.e.

$$log(\frac{P_t}{P_{t-1}}) = r_f + \lambda\sigma_t - \frac{1}{2}\sigma_t^2 + z_t, \qquad z_t|\mathbb{F}_{t-1} \sim N(0, \sigma_t^2),$$

$$\sigma_t^2 = \gamma + \alpha z_{t-1}^2 + \phi\sigma_{t-1}^2 \tag{1}$$

in which P_t denotes the stockprice at time t, r_f is the continuously compounded risk free rate, λ can be interpreted as the unit risk premium and $\mathbb{F}_{t-1}$ the filtration (or information set) up to time *t-1*. Furthermore, the risk free rate is assumed to be constant and the stock is assumed to pay no dividends. Using the LRNVR the return process can be rewritten such that it follows a martingale and its expected return equals the risk free rate. To achieve this the conditional mean is replaced by the risk free rate (i.e. $r_f + \lambda\sigma_t$ is replaced by r_f) and a new risk neutral innovation term ξ_t is introduced. The GARCH equation is then altered by replacing z_t with $\xi_t - \lambda\sigma_t$ such that the model under the risk neutral measure (denoted $\mathbb{Q}$) becomes

$$log(\frac{P_t}{P_{t-1}}) = r_f - \frac{1}{2}\sigma_t^2 + \xi_t, \qquad \xi_t|\mathbb{F}_{t-1} \sim N(0, \sigma_t^2), \tag{2}$$

$$\sigma_t^2 = \gamma + \alpha(\xi_{t-1} - \lambda\sigma_{t-1})^2 + \phi\sigma_{t-1}^2 \tag{3}$$

As discussed by Duan (1995), this specification is a *local risk neutral valuation relationship*, and it implies two important properties: (1) the one period ahead conditional variance remains unchanged under both measures and (2) the expected one period ahead future return equals the riskfree rate r_f. More formally,

$$E_{\mathbb{Q}}[P_t/P_{t-1}|\mathbb{F}_{t-1}] = exp(r_f), \tag{4}$$

and

$$Var_{\mathbb{P}}(log(\frac{P_t}{P_{t-1}})|\mathbb{F}_{t-1}) = Var_{\mathbb{Q}}(log(\frac{P_t}{P_{t-1}})|\mathbb{F}_{t-1}). \tag{5}$$

A necessity for these properties to hold is that the innovations in (1) and (3) are assumed to be normally distributed. These properties are desirable because they imply that the conditional variance can be observed and estimated under the $\mathbb{P}$ measure and the conditional mean can be replaced by the risk free rate. The resulting model under $\mathbb{Q}$ is then a well specified model that is locally free of risk preferences. This measure is a generalization of the conventional risk neutral measure and accommodates heteroskedasticity.

2.2 GJR-GARCH

The popular class of GARCH models, introduced by Bollerslev (1986), were extended by Glosten Jagannathan and Runkle (1993) to accommodate asymmetry. Empirical studies have shown that volatility is usually higher after an unexpected negative return than after an unexpected positive return of the same magnitude. This is one of the stylized facts of return series and is frequently referred to as the leverage effect. It is explained by observing that when firms use both debt and equity to finance their business, the debt-equity ratio will increase as the stock price decreases, which then increases equity return volatility (Lee and Liu, 2014). The GJR model is more suitable of capturing the leverage effect because it allows for asymmetric responses to the shocks. In this paper the approach of Bauwens and Lubrano (2002) is followed, who define the returns in discrete time as $r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$. This simplifies the risk-neutral measure as no lognormal distribution has to be manipulated when introducing the risk neutral innovation term while the desirable properties of Duan (1995) still hold.

The basic GJR model under $\mathbb{P}$ is given by

$$r_t = \mu_t + z_t, \qquad z_t = \sigma_t \epsilon_t, \qquad \epsilon \sim N(0,1) \tag{6}$$

$$\sigma_t^2 = \gamma + \alpha z_{t-1}^2 + \phi \sigma_{t-1}^2 + \delta z_{t-1}^2 \mathbb{I}_{t-1}[z_{t-1} < 0] \tag{7}$$

in which μ_t is the conditional expectation of the returns and $\mathbb{I}_t$ is an indicator function which takes the value 1 if z_t <0, and 0 otherwise. To guarantee a positive variance estimate, the parameters γ, α, ϕ and δ are restricted to be strictly positive. Furthermore, the *unconditional* variance is given by

$$E[\sigma_t^2] = \frac{\gamma}{1 - \alpha - \phi - \frac{1}{2}\delta} \tag{8}$$

provided that $\alpha + \phi + \frac{1}{2}\delta < 1$. Using the LRNVR, this model can be re-written in a risk-neutral form by replacing the conditional mean with r_f and introducing a new innovation term. The Duan (1995) framework implies that $\mu_t = \mu + \lambda\sigma_t$ which implies that the risk neutral innovation term is given by ξ_t = z_t + μ + $\lambda\sigma_t$ - r_f. To obtain the risk neutral GJR specification z_t can then be replaced by ξ_t - μ - $\lambda\sigma_t$ + r_f in the GJR equation. However, Bauwens and Lubrano (2002) argue that the Duan (1995) framework generally does not have a good fit and leads to poor estimates of λ. Instead, they incorporate the lagged return in the return specification. Hafner and Herwartz (2001) showed that incorporating the lagged return indeed leads to better likelihood values than incorporating $\lambda\sigma_t$. Consequently, this approach is followed by setting $\mu_t = \mu + \rho r_{t-1}$ in equation (6) such that the selected model under $\mathbb{P}$ is

$$r_t = \mu + \rho r_{t-1} + u_t, \qquad u_t|\mathbb{F}_{t-1} \sim \mathbb{N}(0, \sigma_t) \tag{9}$$

$$\sigma_t^2 = \gamma + \alpha u_{t-1}^2 + \phi\sigma_{t-1}^2 + \delta u_{t-1}^2 \mathbb{I}_{t-1}[u_{t-1} < 0] \tag{10}$$

Under the $\mathbb{Q}$ measure a new innovation term based on the LRNVR is introduced, given by v_t = u_t + μ_t - r_f such that the expected return equals the risk free rate. The GJR specification is then altered by replacing u_t by v_t - μ_t + r_f such that the risk neutral specification is then given by

$$r_t = r_f + v_t, \qquad v_t|\mathbb{F}_{t-1} \sim \mathbb{N}(0, \sigma_t) \tag{11}$$

$$\sigma_t^2 = \gamma + \alpha(v_{t-1} - \mu_{t-1} + r_f)^2 + \phi\sigma_{t-1}^2 + \delta(v_{t-1} - \mu_{t-1} + r_f)^2 \mathbb{I}_{t-1}[(v_{t-1} - \mu_{t-1} + r_f) < 0]. \tag{12}$$

Bauwens and Lubrano show that the stationarity condition is given by $(\alpha + \frac{1}{2}\delta)(1 + \rho^2) + \phi < 1$ for the risk-neutral GJR.. The model under $\mathbb{P}$ can be used for inference and subsequently the model under $\mathbb{Q}$ can be used to simulate the terminal time stock price evolution required to price the option.

2.3 Markov Switching GJR-GARCH

The MS GARCH model proposed is similar to the model introduced by Franc and Zakoian (2001) and is an extension of Hamilton's (1989) original Markov Switching model. Franc and Zakoian (2001) provide an estimation method based on the generalized method of moments, while Henneke et al. (2006) and Bauwens et al. (2010) study the same model in a Bayesian context. The setting in this paper is slightly different as the GJR model requires an additional asymmetry parameter. The Markov switching property enables the conditional variance to switch between J different set of parameters at each point in time. This yields a more flexible specification of the conditional variance as this allows to capture more return dynamics as in the single regime case. In the MS GJR model the conditional variance is governed by a latent (i.e. unobservable) regime switching variable, denoted S_t for t = 1,...,T. Using frequentist methods the estimation of these latent variables is challenging. The reason is that the regime dependence makes the likelihood intractable analytically when the GARCH components are present. This is due to the fact that the conditional variance depends on all preceding conditional variances and therefore on the whole unobserved sequence of regimes. One way to deal with this is by conditioning on the regime at time t and apply an Expectation - Maximization step to estimate the model parameters and the regime probabilities [7]. Under Bayesian inference a more natural way of dealing with the latent variables can be applied. Instead of conditioning on the regime, the parameter space is augmented with the unobservable variable S_t such that the latent variable can be simulated together with the model parameters. Following the approach of Rachev et al. (2008, ch. 11), three regimes are used which can be given the economic interpretation of low volatility state, normal volatility state and high volatility state[8]. Although this is an ad-hoc way of a-priori selecting the number of states, a three state model is sufficient to capture most dynamics found in financial time series. This argument is supported by Fuh et al. (2012), who note that empirical work rarely indicates that more than three states are required. It might be possible two states may also be sufficient, however the three state model encompasses the two state model and should therefore be able to describe the market dynamics equally well in case two states are more likely. To establish the number of regimes a simple diagnostic measure is to plot the return data and check whether the magnitude of volatility differs between different periods.

[7] See Kole (2010) for more details

[8] The terms 'state' and 'regime' are used interchangeably throughout this paper

In a Bayesian context a more formal way to determine the appropriate number of states is to use Bayes factors, which indicate how likely one model is relative to another. In the empirical part both the return plot and the Bayes factors are used to substantiate the choice for the 3 state model.

The basic set up is similar as in the single-regime GJR process stated in equation (10), only now each parameter can take three values at each point in time. The model under $\mathbb{P}$ becomes

$$r_t = \mu(S_t) + \rho(S_t)r_{t-1} + u_t, \qquad u_t \sim \mathbb{N}(0, \sigma_t^2)$$

$$\sigma_t^2 = \gamma(S_t) + \alpha(S_t)u_{t-1}^2 + \phi(S_t)\sigma_{t-1}^2 + \delta(S_t)u_{t-1}^2 I[u_{t-1} < 0] \tag{13}$$

in which for each period t

$$(\mu(S_t), \rho(S_t), \gamma(S_t), \alpha(S_t), \phi(S_t), \delta(S_t)) = \begin{cases} (\mu_1, \rho_1, \gamma_1, \alpha_1, \phi_1, \delta_1), \text{ if } \mathrm{S}_t = 1 \\ (\mu_2, \rho_2, \gamma_2, \alpha_2, \phi_2, \delta_2), \text{ if } \mathrm{S}_t = 2 \\ (\mu_3, \rho_3, \gamma_3, \alpha_3, \phi_3, \delta_3), \text{ if } \mathrm{S}_t = 3 \end{cases}$$

To arrive at the complete risk neutral specification of (13), the Duan (1995) framework is insufficient as the Markov switching property implies the existence of regime shift risk. Duan (1999) accounts for this by letting the regime probabilities be time varying functions of the return distribution. This allows for a shift in the innovation term to obtain the risk neutral probability. In the Bayesian context this is not as straightforward. In contrast to Duan (1999), the probabilities are constant over time and are drawn from their posterior distribution such that it is not possible to model them as functions of the return distribution. For simplicity it is therefore assumed that regime shift risk is not priced such that the risk neutralization can proceed as described for the single regime model. This means the transition probabilities under $\mathbb{P}$ are maintained and the risk neutralization is identical as described earlier in equations (11) and (12) only applied to the MS GJR specification. This assumption is somewhat restrictive, but is imposed by many authors. For example, Bollen et al. (1999), Yao et al. (2003) and Satoyoshi and Mitsui (2010) show that pricing performance can still be satisfactory under this assumption.

The regime switching model uses a homogeneous first-order Markov chain with transition matrix $\mathbf{P}$, whose elements are given by $p_{ij} = Pr(S_{t+1} = j|S_t = i, \mathbf{P})$ in which the probabilities in each row sum up to 1. Due to the Markov property the volatility state that is visited at time

t only depends on time t-1, i.e. $Pr(S_t|S_{t-1}, S_{t-2}, \ldots, S_1) = Pr(S_t|S_{t-1})$. The Markov chain is assumed to be irreducible and aperiodic, which means that all states communicate and that each state can be entered directly from any other state. This allows the volatility to jump from the low state directly to the high state and vice-versa.

3 Methodology

Work that combines Bayesian inference with option pricing theory shows that Bayesian inference has some advantages over the traditional frequentist approach (see for example Rombouts and Stentoft 2014; Avellaneda 1999; Darsinos and Satchwell 2001; Vargas Mendoza 2011; Jones 2003; and more). One advantage is that the parameters are integrated over their entire parameter space in the Bayesian estimation process, such that all uncertainty is accounted for. An illustrative example of the latter is given by Rombouts and Stentoft (2014), who find that Bayesian inference leads to smaller pricing errors in small samples compared to frequentist methods. Furthermore, Bayesian inference can also be applied when limited data is available because economic theory and experts' opinions can be included in the analysis. Alternatively, information from earlier periods or closely related variables can also be included into the prior-distributions of the model parameters, allowing one to create an informative prior.

This section describes how Bayesian inference can be performed to estimate the GJR and the MS GJR models. For a basic introduction on Bayesian econometrics the interested reader is referred to Greenberg (2013). For applications of Bayesian methods in finance, Rachev, Hsu, Bagasheva and Fabozzi (2008) provide a good overview of popular techniques and for more specific literature on the Bayesian estimation of Markov Switching GARCH models see Henneke et al. (2006) and Bauwens et al. (2010).

3.1 Bayesian Estimation of the GJR-GARCH Model

To estimate the model using Bayesian inference the GARCH parameters are collected in the vector $\theta = (\gamma, \alpha, \phi, \delta)'$, and to make the relation with the GARCH process explicit the volatility is written as $\sigma_t^2 = \sigma_t^2(\theta)$. Using the relationships implied by the LRNVR, the model can be estimated under $\mathbb{P}$. The parameters μ and ρ are collected in the vector β such that the return specification can conveniently be written as

$$r_t = \begin{pmatrix} 1 \\ r_{t-1} \end{pmatrix}' \begin{pmatrix} \mu \\ \rho \end{pmatrix} + \mathrm{u}_t \quad = x'_{t-1}\beta + u_t. \tag{14}$$

The likelihood function is then obtained by using that $p(\mathbf{r}|\theta,\beta) = \prod_{t=1}^{T} p(r_t|r_{t-1},\beta,\theta)$ which, together with independent standard normally distributed innovations, can be written as

$$p(\mathbf{r}|\theta) = \prod_{t=1}^{T} \sigma_t^{-1}(\theta) exp\Bigg(-\frac{1}{2}\sum_{t=1}^{T} \frac{(r_t - x'_t\beta)^2}{\sigma_t^2(\theta)} \Bigg) \tag{15}$$

which in matrix form can be rewritten as

$$p(\mathbf{r}|\beta,\theta) \propto |\mathbf{D}|^{-\frac{1}{2}} exp\Bigg(-(\mathbf{r}-\mathbf{X}\beta)'\mathbf{D}^{-1}(\mathbf{r}-\mathbf{X}\beta)/2 \Bigg) \tag{16}$$

in which $\mathbf{X}$ is the $T \times k$ matrix with the first column a vector of ones and k-1 vectors of predictor variables and $\mathbf{D} = \mathrm{diag}(\sigma_1^2(\theta), \ldots, \sigma_T^2(\theta))$. The priors for β and θ are a-priori chosen to be independent such that $\mathrm{p}(\beta,\theta) = \mathrm{p}(\beta)\mathrm{p}(\theta)$. For some prior views on β the normal prior is chosen. For θ, a uniform prior on the space S_θ is chosen to ensure a stationary model is constructed. Formally this is stated as

$$p(\beta) \sim \mathbb{N}(\mathbf{b},\mathbf{B}),$$

$$p(\theta) \propto \mathbb{I}[\theta \in S_\theta],$$

with $S\{\theta\} = \{\theta \in \mathbb{R}^4 : \gamma,\alpha,\phi,\delta > 0,\ \alpha+\phi+\frac{1}{2}\delta < 1\}$. The joint posterior distribution is obtained as product of the likelihood and the prior distributions, i.e.

$$\begin{aligned} p(\beta,\theta|\ \mathbf{r}) \propto \quad & |\mathbf{D}|^{-\frac{1}{2}} exp\Bigg(-(\mathbf{r}\ \text{-}\ \mathbf{X}\beta)'\mathbf{D}^{-1}(\mathbf{r}\ \text{-}\ \mathbf{X}\beta)/2 \Bigg) \\ \times \quad & exp\Bigg(-(\beta-\mathbf{b})'\mathbf{B}^{-1}(\beta-\mathbf{b})/2 \Bigg)\mathbb{I}[\theta \in S(\theta)]. \end{aligned} \tag{17}$$

To retrieve the distributional properties the joint posterior can be split up into two conditional posteriors which can then be used to simulate the parameters separately with a Markov Chain Monte Carlo posterior simulation method, also known as the Gibbs sampler.

3.1.1 Full Conditional Posterior of β

Greenberg (2013, ch. 4.3) shows that using some algebraic manipulation[9], the kernel of the conditional posterior of β appears to be multivariate normal;

$$p(\beta|\mathbf{r},\theta) \propto |\mathbf{D}|^{-\frac{1}{2}} exp\Bigg(-(\beta-\hat{\beta})'\hat{\Sigma}_{\beta}(\beta-\hat{\beta})/2\Bigg) \tag{18}$$

with mean and variance given by

$$\widehat{\beta} = (\mathbf{X}'\mathbf{D^{-1}}\mathbf{X}+\mathbf{B^{-1}})^{-1}(\mathbf{X}\mathbf{D^{-1}}\mathbf{r}+\mathbf{B^{-1}}\mathbf{b}), \tag{19}$$

$$\widehat{\Sigma}_{\beta} = (\mathbf{X}'\mathbf{D^{-1}}\mathbf{X}+\mathbf{B^{-1}})^{-1} \tag{20}$$

The conditional posterior of β is not exactly multivariate normal as $\mathbf{D}$ (indirectly) depends on β. However, for a proposal distribution this will suffice, such that a Metropolis-Hastings step can be used to simulate β. The first step is to find the candidate generating distribution with pdf denoted by $g(\beta|\beta^{(m)})$, in which $\beta^{(m)}$ is the current draw of β. This distribution can then be used to draw from the target distribution f(β) = $p(\beta|\mathbf{r}$, $\theta^{(m)})$. Since the full conditional posterior of β is multivariate normal, this distribution is the logical choice for the proposal distribution. The simulation scheme is presented in table 1.

Sampling Scheme for β

1. Simulate proposal value β^* from $\mathbb{N}(\widehat{\beta}, \widehat{\Sigma}_{\beta})$
2. Compute the Metropolis-Hastings acceptance probability as

$$\text{q}(\beta^*, \beta^{(m)}) = \min((\text{f}(\beta^*)/\text{g}(\beta^*|\beta^{(m)}))/ (\text{f}(\beta^{(m)})/\text{g}(\theta^{(m)}|\beta^*)),1)$$

3. set $\beta^{(m+1)}$ =β^* with probability q

   set $\beta^{(m+1)}$ =$\beta^{(m)}$ with probability 1 - q

Table 1: Metropolis-Hastings sampling algorithm for β. A random draw from the uniform (0,1) distribution is used to decide whether β^* is accepted or rejected.

[9] This is also referred to as the decomposition rule

3.1.2 Full Conditional Posterior of θ

The GARCH parameters θ yield a posterior distribution that is non-standard. To overcome this issue, the solution suggested by Rachev et al. (2008, ch.5) is used, and simulation is performed from a distribution centered at the mode $\widehat{\theta}$ of the log posterior $\log(p(\theta\,|\mathbf{r},\beta))$, with a scale matrix proportional to the negative inverse Hessian matrix of $\log(p(\theta\,|\mathbf{r},\,\beta))$ evaluated in $\widehat{\theta}$ and denoted by $-H(\widehat{\theta})^{-1}$. The posterior mode can be obtained by maximizing the log posterior, i.e.

$$\max_{\theta} \qquad -\frac{1}{2}log(|\mathbf{D}|) - \frac{1}{2}(\mathbf{r}\text{ - }\mathbf{X}\beta)'\mathbf{D}^{-1}(\mathbf{r}\text{ - }\mathbf{X}\beta) \tag{21}$$

and the Hessian is obtained by evaluating the matrix of second order derivatives of $\log(p(\theta|\mathbf{r},\beta))$ at $\widehat{\theta}$ [10]. A convenient choice for the proposal distribution of θ is the multivariate normal, which is denoted by pdf g and is used to draw from the target distribution $\mathrm{f}(\theta) = p(\theta\,|\mathbf{r},\,\beta^{(m)})$. The simulation sequence is outlined in table 2. An accept-reject step is used to deal with the truncation, which lets the algorithm repeatedly draw a proposal value θ^* until the parameter restrictions are satisfied. Furthermore, the scalar c is used to inflate the variance and is set equal to 1.2^2, following Geweke (1994).

Sampling Scheme for θ

1. Simulate proposal value θ^* from $\mathrm{N}(\widehat{\theta}, \text{- c } \times H(\widehat{\theta})^{-1}) \times \mathbb{I}[\theta \in S(\theta)]$
2. Compute the Metropolis-Hastings acceptance probability as

$$\mathrm{q}(\theta^*,\,\theta^{(m)}) = \min(\ (\mathrm{f}(\theta^*)/\mathrm{g}(\theta^*))/\ (\mathrm{f}(\theta^{(m)})/\mathrm{g}(\theta^{(m)})),1)$$

3. set $\theta^{(m+1)} = \theta^*$ with probability q

   set $\theta^{(m+1)} = \theta^{(m)}$ with probability 1 - q

Table 2: Metropolis-Hastings sampling algorithm for the GARCH parameters θ. A random draw from the uniform (0,1) distribution is used to decide whether θ^* is accepted or rejected. The scalar c can be used to tune the acceptance rate. In this setting the approach of Geweke (1994) is followed by setting $c = 1.2^2$.

3.2 Bayesian Estimation of the MS GJR-GARCH Model

To estimate the Markov switching model data augmentation is applied. This technique uses the complete data likelihood to form the posterior distribution and samples the latent variables S_t alongside the model parameters. To apply data augmentation the latent variables are included

[10] In this paper Matlab's Fmincon function is used, which produces the Hessian matrix as byproduct.

in the parameter vector, i.e. θ ={$\theta_{G,j}$', β'_j,$\mathbf{P}'_i$, $\mathbf{S}$} with $\theta_{G,j} = \{\gamma_j, \alpha_j, \phi_j, \delta_j\}$, $\mathbf{P}'_i$ the i-th row of $\mathbf{P}$ for i=1,. . .,J, $\mathbf{S} = (S_1, \ldots, S_T)$' the regime path for all periods and β_j the regime dependent model parameters. The priors for β_j and $\theta_{G,j}$ are again a priori assumed to be independent. Furthermore, the transition probabilities in each row of $\mathbf{P}$ need to sum to 1 and are elements of (0,1). A suitable prior for each row $\mathbf{P}_i$ therefore is the Dirichlet distribution, while assuming independence between different rows. To reflect prior beliefs about the effect the regimes have on the GJR parameters, proper normal priors are imposed on $\theta_{G,j}$ and for β_j. Finally, the parameter constraints can be imposed through an indicator function as before. More formally,

$$p(\theta_{G,j}) \sim \mathbb{N}(\mu_{G,j}, \Sigma_{G,j}) I[\theta_{G,j} \in S\{\theta_{G,j}\}]$$

$$p(\mathbf{P}_i) \overset{ind}{\sim} \mathbb{D}ir(a_{i1}, a_{i2}, \ldots, a_{ij})$$

$$p(\beta_j) \sim \mathbb{N}(\mathbf{b}_j, \mathbf{B}_j).$$

where $S\{\theta_{G,j}\} = \{\theta_{G,j} \in \mathbb{R}^4 : \gamma_j, \alpha_j, \phi_j, \delta_j > 0,\ \alpha_j + \phi_j + \frac{1}{2}\delta_j < 1\}_{j=1}^{J}$. The joint posterior is then given by the product of the complete data likelihood and the prior distributions,

$$\prod_{j=1}^{J} \Bigg[|\mathbf{D}|^{-\frac{1}{2}} exp\bigg(-\frac{1}{2}(\mathbf{r} - \mathbf{X}\beta_j)'\mathbf{D}^{-1}(\mathbf{r} - \mathbf{X}\beta_j)\mathbb{I}[S_t = j]\bigg)$$

$$\times exp\bigg(-\frac{1}{2}(\beta_j - \hat{\beta}_j)\hat{\Sigma}_{\beta_j}^{-1}(\beta_j - \hat{\beta}_j)\mathbb{I}[S_t = j]\bigg)$$

$$\times exp\bigg(-\frac{1}{2}(\theta_{G,j} - \mu_{G,j})'\Sigma_{G,j}(\theta_{G,j} - \mu_{G,j})\mathbb{I}[S_t = j]\bigg)\Bigg]$$

$$\times \prod_{i=1}^{J}\prod_{j=1}^{J} p_{ij}^{a_{ij}+n_{ij}-1} \tag{22}$$

where the number of transitions from state i to state j in the sample of $\mathbf{S}$ is denoted by n_{ij}.

3.2.1 Full Conditional Posterior of $\mathbf{P}'_i$

Collecting the terms involving p_{ij} in (22), the full conditional posterior of the transition probabilities per row $\mathbf{P}_i$ is given by

$$p(\mathbf{P}'_i|\mathbf{r}, \theta_{-\mathbf{P}'_i}) \propto \prod_{j=1}^{J} p_{ij}^{a_{ij}+n_{ij}-1} \tag{23}$$

for i the row number and with $\theta_{-\mathbf{P}_i'}$ the vector of all parameters except $\mathbf{P}_i$. Equation (23) can be recognized as the kernel of a Dirichlet distribution with parameters $(a_{i1}+n_{i1},\ldots,a_{iJ}+n_{iJ})$. The parameters a_{ij} are set a-priori while the parameters n_{ij} are the number of transitions from state i to j. Sampling from the Dirichlet distribution is then done by sampling three independents observations from the Chi-square distribution with degree of freedom equal to $2(a_{ij}+n_{ij})$, and normalizing the draws. Table (3) describes the sampling algorithm when assuming three regimes.

Sampling Scheme for $\boldsymbol{P}_i$

1. For each row i sample three independent observations

$$y_{i1} \sim \chi^2_{2(a_{i1}+n_{i1})} \qquad y_{i2} \sim \chi^2_{2(a_{i2}+n_{i2})} \qquad y_{i1} \sim \chi^2_{2(a_{i3}+n_{i3})}$$

2. set $\quad p_{i1} = \frac{y_{i1}}{\sum_{k=1}^{J} y_{ik}} \qquad p_{i2} = \frac{y_{i2}}{\sum_{k=1}^{J} y_{ik}} \qquad p_{i3} = \frac{y_{i3}}{\sum_{k=1}^{J} y_{ik}}$

Table 3: Sampling algorithm for the transition probabilities. A draw from the $Dir(a_{i1}+n_{i1},\ldots,a_{iJ}+n_{iJ})$ distribution is obtained by drawing J random observations from the $\chi^2_{2(a_{ij}+n_{ij})}$ distribution and dividing each χ^2_j draw by the sum of the χ^2 draws.

3.2.2 Full Conditional Posterior of S

The number of regime paths that could generate the terminal regime S_T increases with the amount of regimes imposed. In this application there are 3^T possibilities that lead to the terminal state, which makes it practically impossible to draw from the $1 \times T$ vector $\mathbf{S}$ at once. Instead a T-step procedure is applied in which the elements are drawn one at a time. Using the rules of conditional probability, Rachev et al. (2008, ch. 11) show that the full conditional posterior of S_t is given by

$$p(S_t = j, |\mathbf{r}, \theta_{-S}, \mathbf{S}_{-t}) = \frac{p(\mathbf{r}|\theta_{-S}, \mathbf{S}_{-t}, S_t = j)p_{ij}p_{jk}}{\sum_{s=1}^{J} p(\mathbf{r}|\theta_{-S}, \mathbf{S}_{-t}, S_t = s)p_{is}p_{sk}} \tag{24}$$

for $j = 1,\ldots,J$. This probability is computed by conditioning on all states other than time t and is used to draw a regime path as described in table (4).

*Simulation Scheme for **S***

1. At $t = 1$, compute the probability in (24) for $j = 1,\ldots,J$.
2. Divide the interval (0,1) into three sub-intervals with lengths proportional to the computed probabilities.
3. Draw an observation u from the uniform distribution $U(0,1)$
4. Depending on which interval u falls into, set $S_1^{(m)} = j$.
5. Update $\mathbf{S}^{(m)}$ with $S_1^{(m)}$.
6. Repeat for $t = 2,\ldots, T$.

Table 4: Sampling algorithm for the latent regime path variables. The procedure is stepwise and becomes computational more intensive as the sample size T increases.

3.2.3 Posterior Distributions of $\theta_{G,j}$ and β_j

The posterior distribution for $\theta_{G,j}$ is not available analytically due to the regime dependence of the conditional variance. However, using the informative prior the kernel of the posterior distribution is given by

$$p(\theta_{G,j}|\theta_{-\theta_{G,j}},\mathbf{r}) \propto \prod_{j=1}^{J}\Bigg[|\mathbf{D}|^{-1/2}exp\Bigg(-\frac{1}{2}(\mathbf{r}-\mathbf{X}\beta_j)'\mathbf{D}^{-1}(\mathbf{r}-\mathbf{X}\beta_j)\mathbb{I}[S_t=j]\Bigg)$$

$$\times exp\Bigg(-\frac{1}{2}(\theta_{G,j}-\mu_{G,j})'\Sigma_{\theta,j}^{-1}(\theta_{G,j}-\mu_{G,j})\mathbb{I}[S_t=j]\Bigg)\Bigg]. \tag{25}$$

Conditional on the regime path **S** the only difference with the single regime GJR model is the informative prior of $\theta_{G,j}$. For the GARCH parameters, the proposal distribution is again the multivariate normal and the simulation procedure is similar to the one described for the single-regime GJR, only now the log of equation (25) is maximized over the regime parameters $\theta_{G,j}$. This optimization procedure is not as straightforward as for the single regime GJR model. First the regime path **S** is initialized by selecting a state for each time period with uniform probabilities, which corresponds to uniform prior beliefs about the transition probabilities. Furthermore, when maximizing all GARCH-regime parameters at the same time, the algorithm tends to get stuck such that the solution does not move much from its initial values. To overcome this issue, a grid of starting values is generated and conditional optimization is applied in which the GJR parameters are optimized by one regime at a time. The procedure starts by optimizing the regime 1 parameters while fixing the GJR parameters for regimes 2 and 3 to the current grid values. Next the GJR parameters for regime 2 are optimized conditional on the solution for the GJR parameters of regime 1 and fixing the regime 3 GJR parameters to the current grid value. Then the GJR parameters for regime 3 are optimized conditional on the solutions for regime 1 and 2. This process is repeated for all simulated grids. Finally, the set of solutions that yield the highest log likelihood value are selected to simulate from. After running some pre-runs of the Bayesian algorithm, a representative estimate of $\mathbf{S}$ is obtained which can be used to re-do the optimization procedure. The estimates that this produces will then be used for Bayesian inference of $\theta_{G,j}$. The simulation procedure for β_j remains the same, only now the parameters are drawn for each regime.

3.2.4 Complete Posterior Simulation Scheme MS GJR-GARCH

Combining the above yields the complete simulation scheme, given in table (5). These steps are repeated a large amount of times until the MCMC algorithm is converged to the stationary distribution.

Simulation Scheme for the MS GJR

1. For $j = 1,\ldots,J$, draw p_{ij} from (23)
2. Draw $\mathbf{S}^{(m)}$ from (24)
3. Draw $\beta_j^{(m)}$ as in table (1)
4. Draw $\theta^*_{G,j}$ from the proposal distribution for $j = 1,\ldots,J$.
5. Repeat the previous step until the parameter restrictions are satisfied
6. Compute the acceptance probability and accept or reject $\theta^*_{G,j}$ for $j = 1,\ldots,J$.

Table 5: Bayesian Markov Chain Monte Carlo algorithm. This procedure loops through the sampling schemes of the full conditional posteriors of the model parameters to draw from the posterior distribution of the Markov Switching GJR model.

3.3 Prior Specifications and the Label Switching Problem

For the GJR model the informative prior for β is an ordinary least squares estimate from an earlier period starting in December 2001 and ranging up until the first sample period in January 2006. The prior-sample then rolls on as the estimation window rolls on, such that the prior is estimated with approximately 1250 observations for all samples. For the parameters β_j and $\theta_{G,j}$ in the MS GJR model, the informative priors are inferred from sub-samples corresponding to times with low, normal and high volatility by estimating the single regime GJR model for these periods. The posterior means are then taken as hyperparameters for the prior means in the MS GJR model. The posterior standard deviations of β are used to construct $\mathbf{B}_j$, while the prior variances $\Sigma_{G,j}$ are taken as scaled identity matrices due to the fact that the posterior variances were rather small. The identity matrices are multiplied by 0.01, 0.1 and 1 for the low, medium and high regimes respectively. This setting reflects a somewhat strong prior intuition, but clearly defines the different volatility states. For the transition probabilities in $\mathbf{P}$, the Dirichlet prior is imposed with parameters $a_{ij} = 1$ for $i, j = 1, 2, 3$ which reflects uniform beliefs about the transition probabilities p_{ij}. These priors are kept fixed for all estimation windows. The subsample periods that correspond to different levels of volatility are presented in figure (1) and an overview of the obtained priors for the MS GJR model is given in Table (6).

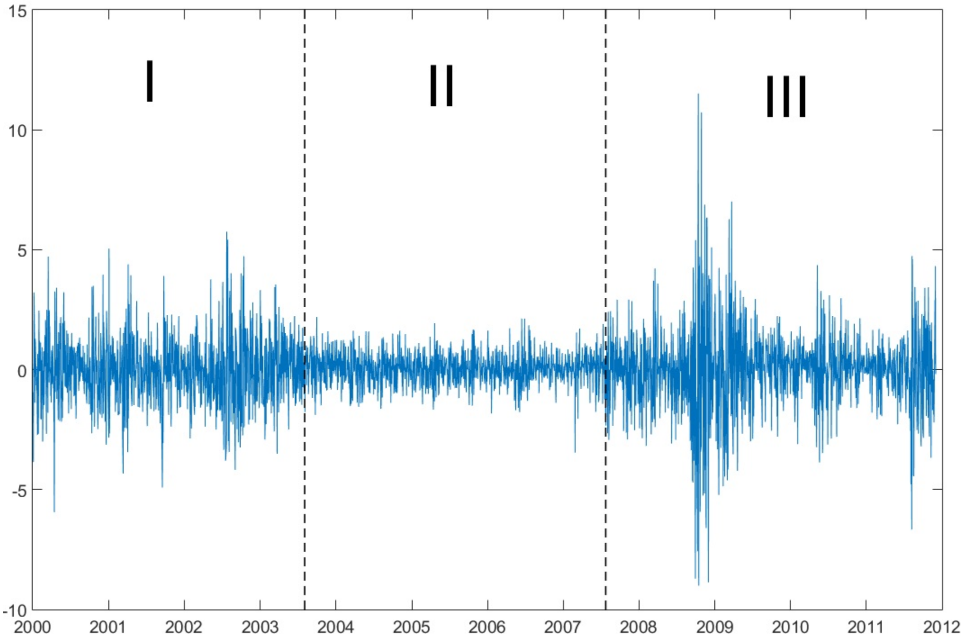


Figure 1: Plot of percentage returns of the pre-sample of the S&P 500 index. The plot shows different periods of volatility throughout time. The priors used for the MS GJR are determined by estimating the single regime GJR model for the low volatility period (II), medium volatility period (I) and high volatility period (III). All subsamples consists of approximately 1000 observations, which corresponds to roughly $3\frac{1}{2}$ years.

In Markov switching models a possible complication may arise from the label switching problem, which is caused by the likelihood being invariant to permutations of the components' indices. This means the model is not identified because the parameter estimates and the regime labels can simply be switched. Multiple solutions have been suggested to resolve the label switching issue, with the three most popular methods being identification constraints (Stephens, 1997; McLachlan and Peel, 2000; Geweke 2007), relabelling algorithms (Stephens, 2000; Celeux et al., 2000; Hurn et al., 2003;), and probabilistic approaches (Jasra, 2005; Sperrin, 2009). The simplest of these approaches are the identification constraints, which can be easily imposed, for example as $\mu_2 > \mu_1$. Without such identification constraints the sampler can generate bimodal distributions for the means and variances, which makes the posterior moments meaningless. In this paper the

Prior Overview for the MS GJR

Low Regime

	μ	ρ		γ	α	ϕ	δ
$\mathbf{b}_1$	0.035	-0.0285	$\mu_{G.1}$	0.0325	0.0784	0.8233	0.0773
				0.01	0.00	0.00	0.00
$\mathbf{B}_1$	0.0042	-0.0001	$\Sigma_{G.1}$	0.00	0.01	0.00	0.00
	-0.0001	0.0035		0.00	0.00	0.01	0.00
				0.00	0.00	0.00	0.01

Medium Regime

	μ	ρ		γ	α	ϕ	δ
$\mathbf{b}_2$	0.0538	-0.0038	$\mu_{G.2}$	0.0408	0.002	0.8842	0.1885
				0.1	0.0	0.0	0.0
$\mathbf{B}_2$	0.0064	0.0002	$\Sigma_{G.2}$	0.0	0.1	0.0	0.0
	0.0002	0.0035		0.0	0.0	0.1	0.0
				0.0	0.0	0.0	0.1

High Regime

	μ	ρ		γ	α	ϕ	δ
$\mathbf{b}_3$	0.071	-0.0628	$\mu_{G.3}$	0.1330	0.0164	0.8721	0.1058
				1	0	0	0
$\mathbf{B}_3$	0.0234	0.0003	$\Sigma_{G.3}$	0	1	0	0
	0.0003	0.0035		0	0	1	0
				0	0	0	1

Table 6: Priors used for the MS GJR model. The hyperparameters of the priors of β and θ_j are Bayesian estimates from sub-periods which correspond to low, medium and high volatility periods respectively. The prior-variances of θ_j are set ad-hoc and are based on prior beliefs about the effect the regimes will have.

approach of Bauwens et al. (2010) is adopted, who argue that such strict inequalities correspond to a maximum likelihood set-up and that in a Bayesian context such constraints need not hold strictly for the different regimes to be sufficiently separated. Instead, the parameter support can be allowed to overlap, provided that the overlap is not too large, and the identification can be imposed through uniform priors similar to those used for the stationarity condition. The example used by Bauwens et al. (2010) states that the support of μ_1 can be restricted to (-0.09, +0.03) while μ_2 could then be restricted to (-0.03, +0.09). The regimes must be sufficiently separated for identification, which means some parameters need to be different amongst regimes. In this setting it is required that the level of the conditional variance, γ_j, increases for the ascending regimes. The remaining parameters can be restricted somewhat looser as there may be a trade-off between the magnitude of these values between different regimes, as suggested by Rachev et al. (2008, ch. 11). This approach requires some caution concerning the wideness of the prior supports. If the interval is too wide the posterior draws will be sampled inefficiently, while if the interval is too narrow the posterior density may get truncated. The prior supports are therefore initiated sufficiently wide and are narrowed when possible or widened when needed.

3.4 Pricing the Option

The theoretical call option price at time t with strike price K and maturity T is given by

$$C(P_t, K, T) = e^{-r(t-T)} \int_0^\infty max(P_T - K, 0) f^{\mathbb{Q}}(P_T) dP_T \tag{26}$$

in which $f^{\mathbb{Q}}(P_T)$ denotes the risk neutral density of the underlying asset price at time T. As the volatility of the process is assumed to follow a GARCH process there exists no closed-form solution to (26) such that it has to be approximated through simulation. First the asset returns have to be sampled from the risk neutral predictive density. Next the sampled returns have to be aggregated to obtain the predicted price density. Then the expected call price given in (26) can be approximated by the sample average. For expiration at the next time-step, the predictive density can be expressed as

$$f^{\mathbb{Q}}(\mathrm{r}_{t+1}|\mathbf{r}) = \int f^{\mathbb{Q}}(\mathrm{r}_{t+1}|\mathbf{r}, \theta) p(\theta|\mathbf{r}) d\theta. \tag{27}$$

An analytical solution to (27) is not available but using the algorithm of Geweke (1989) it can be approximated by

$$\widehat{f^{\mathbb{Q}}}(r_{t+1}|\mathbf{r}) = \frac{1}{N} \sum_{n=1}^{N} \mathbb{N}(r_{t+1}|r_f, \sigma_{t+1}^{2(n)}) \tag{28}$$

where $\mathbb{N}(.)$ denotes the normal distribution, (n) indicates the n-th draw of the governing GARCH parameters and N is the total number of draws. This means that when simulating $\theta^{(n)} \sim p(\theta|\mathbf{r}, \mathrm{r}_{t+1}^{(n)})$, simultaneously $\mathrm{r}_{t+1}^{(n)} \sim f^{\mathbb{Q}}(r_{t+1}|\theta^{(n)}, \mathbf{r})$ is simulated. This simulation procedure was generalized for maturities s periods ahead by Bauwens and Lubrano (2002), who state that the predictive density of r_{t+s} under $\mathbb{Q}$ is given by

$$f^{\mathbb{Q}}(r_{T+s}|\mathbf{r}) = \int_\theta \Bigg(\int_{R^{\tau-1}} f^{\mathbb{Q}}(r_T|\theta, r_{T-1}) f^{\mathbb{Q}}(r_{T-1}|\theta, r_{T-2}) \times \dots$$

$$\dots \times f^{\mathbb{Q}}(r_{t+1}|\theta, r_t) p(\theta|\mathbf{r}) dr_{T-1} \dots dr_{t+1} \Bigg) d\theta. \tag{29}$$

The unobserved $r_{t+1}, \dots, r_{T-1}$ have to be simulated sequentially for the integral to be evaluated. The individual elements in (29) are normal with mean r_f and variance σ_{t+s}^2 but the resulting density of the inner integral is not a known distribution. Bauwens and Lubrano (2002) propose a simulation scheme that draws from the predictive distribution τ periods ahead which, for the application at hand, requires minor alterations to suit the regime switching model.

The algorithm starts by initiating a loop running from $mc = 1$ to MC in which MC denotes the number of Monte Carlo simulations. This loop represents the integral over θ and as such the transition probabilities, regime paths and the model parameters $\beta^{(mc)}$ and $\theta_{G,j}^{(mc)}$ from the different regimes are drawn at every iteration. Next, a loop running from $n = 1$ to N is initiated which represents the inner integral over the risk neutral return densities. Within the n loop a vector ε with length τ consisting of standard normally distributed realizations is drawn. Next, a loop running from $t = 1$ to τ is initiated, which represents the time to maturity of the option under consideration. Within the t loop, first the probabilities given by equation (24) are computed for $i = 1, 2, 3$. A draw from the uniform (0,1) distribution is then obtained to decide the governing regime. Then the conditional variance can be constructed and subsequently the future risk neutral return can be simulated as following

$$\sigma_{s,t}^2(S_{s,t}) = \gamma_{mc}(S_{s,t}) + \alpha_{mc}(S_{s,t}){v_{s,t-1}^*}^2 + \phi_{mc}(S_{s,t})\sigma_{s,t-1}^2 + \delta_{mc}(S_{s,t}){v_{s,t-1}^*}^2\mathbb{I}[v_{s,t-1}^* < 0]$$

$$r_{s,t} = r_f + \varepsilon_{n,t}\sigma_{s,t}(S_{s,t}) \tag{30}$$

in which $v_{s,t-1}^* = r_{s,t-1} - \mu_{mc} - \rho_{mc}r_{s,t-2}$, $S_{s,t}$ is the governing regime and the subscript s denotes a simulated value. The algorithm needs to be initiated for $t = 1, 2$ using the last two returns of the sample and the estimated GARCH variance on the last day of the sample. Then the return path until time $t + \tau$ can be simulated and the terminal asset price is computed as

$$P_T = P_t\prod_{t=1}^{\tau}(1 + r_{s,t}/100). \tag{31}$$

The option price is then computed as the expected discounted payoff;

$$C_t^{(mc)}(P_t, T, K) = (1 + r_f)^{-\tau}max(P_T - K, 0). \tag{32}$$

This process is repeated N times, at which the n loop is exited and the average discounted payoff corresponding to a single draw of θ is obtained by averaging over N. Then the algorithm draws new observations of the transition probabilities and the model parameters and repeats the whole process MC times. The final option price is then computed as the average discounted payoffs of all the simulations,

$$\widehat{C}_t(P_t, T, K) \approx \frac{1}{MC} \sum_{mc=1}^{MC} C_t^{(mc)}(P_t, T, K) \tag{33}$$

In case of the single-regime GJR the posterior reduces to $p(\theta|\mathbf{r}, \beta)$ and the state variables $S_{s,t}$ reduce to 1 such that only parameters from the posterior of the single regime GJR model are drawn. The algorithm then reduces to the Bauwens and Lubrano (2002) case[11]. To speed up the algorithm a variance reduction method known as the method of antithetic variates is applied by using the standard normal draws twice but with an opposite sign the second time. Other variance reduction methods may also be implemented [12]. Option prices corresponding to a single draw of θ are then averaged to obtain the price corresponding to draw mc. Geweke (1989) shows that when $MC \to \infty$ the value of N can be set equal to 1 while the results will remain consistent. Following this argument, the settings used in this application are $MC = 30000$ and $N = 1$. Table (19) in the appendix provides a more detailed pseudo code of the described algorithm.

3.5 Model Calibration

3.5.1 Implied GJR

To estimate the Implied GJR model, the GJR is calibrated on observed market option prices through minimizing an error criterion function, for which several specifications exist. Popular choices for the objective function are the root mean squared error, the mean absolute error, and the mean squared or absolute percentage errors. Although Weber and Prokopczuk (2011) argue that a percentage error criterion is easier to interpret, a disadvantage of this approach is that the percentage errors can become substantial when the option prices are close to 0. For this reason the mean squared error is preferred and the considered objective function is given by

$$\min_{\theta} \qquad MSE = \frac{1}{L} \sum_{l=1}^{L} (\widehat{C}_l(t, T, K) - C_l(t, T, K))^2 \tag{34}$$

in which C_l is the observed market price, $\widehat{C}_l$ is the price as computed in equation (33), l is the option under consideration and L is the total number of options in the calibration sample. For a proper Monte Carlo calibration the options prices have to converge to their limiting value, meaning a large number of Monte Carlo iterations are required which makes the calibration a

[11] See table 3 in Bauwens and Lubrano (2002)

[12] Some examples are the method of control variates and Duan's (1998) empirical martingale measure

high dimensional problem. For out-of-sample option pricing, as described in table (19), 30000 Monte Carlo runs are used. This amount is quite burdensome for the calibration process however, and therefore the settings here are slightly different. The main difference lays in the fact that the price is computed as a classical maximum likelihood estimate as there is only one model estimate, such that $MC = 1$. The value of N can then be set to 30000 to achieve the same number of iterations, but to relieve the computational burden N is set to 10000 instead. Some trial runs showed that the differences are negligible while the calibration is up to 5 times faster. To initiate the calibration, the parameters in θ are set equal to their posterior mean, which is obtained as inference result for the GJR model. For the calibration process one day of option data on the last Wednesday of the sample is used. In a proper implied calibration scheme the calibration sample should consist out of a sample of options with varying time to maturities, strike prices and moneyness categories. This ensures the calibrated model captures the market dynamics sufficiently well. The amount of options to be used is a an empirical issue which has no clear guidelines, although to speed up the calibration it is a good idea to not use an excessive amount of options. Therefore, the sample size is kept moderate, the minimum amount of options used for the calibration is 30, while the maximum is 50, depending on the available options on a particular Wednesday.

3.5.2 FHS GJR

The calibration for the FHS GJR is slightly different than that of the implied GJR as it is obtained by using filtered historical simulations. This method was introduced by Barone-Adesi et al. (2008), who state the model as

$$log(\frac{P_t}{P_{t-1}}) = r_t = \mu_t^* + \epsilon_t \tag{35}$$

$$\sigma_t^2 = \gamma^* + \alpha^* \epsilon_{t-1}^2 + \phi^* \sigma_{t-1}^2 + \delta^* \epsilon_t^2 \mathbb{I}[\epsilon_t < 0] \tag{36}$$

in which μ_t^* is such that the expected return equals the risk free rate and $\epsilon_t = \sigma_t z_t$, where z_t is the historical innovation term, i.e. $z_t \sim f_t(0, 1)$. The historical innovation terms are drawn from the empirical distribution which is obtained by dividing the empirical return innovation by the corresponding estimated conditional volatility, i.e. $z_t = \widehat{\varepsilon}_t / \widehat{\sigma}_t$. The risk neutral GARCH parameters in θ^* are possibly different than those under $\mathbb{P}$ and are obtained by calibration on observed market option prices. In this risk neutral specification the historical $\mathbb{P}$ measure innovations are maintained as the calibrated parameters in θ^* represent the approximate change of measure from $\mathbb{P}$ to $\mathbb{Q}$. The strength of this benchmark model is that it does not assume a parametric distribution and is able to simulate scenarios that did not occur in the sample even though it uses only past innovations. For this reason the strength of the filtered historical simulation is widely recognized in quantitative finance, for example in Value at Risk calculations.

For option pricing the algorithm consists of two steps; (1) a Monte Carlo simulation step and (2) a calibration step. The algorithm starts by initiating the GJR parameters in θ^*. In this approach the posterior means of the GJR parameters are used as initial values, unlike Barone-Adesi et al. (2008), who use maximum likelihood estimates to initiate the model. The historical innovations are then divided by the corresponding GJR volatilities to obtain a standardized series of empirical innovations, $z_t = \hat{\varepsilon}_t/\hat{\sigma}_t$. Then, for every Monte Carlo iteration n, a new vector $\mathbf{z}^*$ is created by randomly drawing from z_t with uniform probabilities and with replacement. Now the GJR variances for time t to time $T = t + \tau$ can be simulated as following

$$\sigma_{n,t}^2 = \gamma^* + \alpha^* {z_{n,t}^*}^2 + \phi^* \sigma_{n,t}^2 + \delta^* z_t^{*2} \mathbb{I}[z_{n,t}^* < 0]. \tag{37}$$

At the terminal time T the τ-period simulated return is given by

$$P_T/P_t = exp(\tau\mu^* + \sum_{i=1}^{\tau} \sigma_{n,t+i} z_{n,t+i}^*) \tag{38}$$

which can be used to compute the simulated payoff of the option at time T. After repeating this process N times the price can be computed as the discounted average payoff as

$$\widehat{C}(t,T,K) = exp(-r_f\tau)\frac{1}{N}\sum_{n=1}^{N} max(P_T^{(n)} - K, 0) \tag{39}$$

with $P_T^{(n)}$ the simulated asset price and N the total number of simulated sampled paths. The procedure continues by calibrating the GJR parameters in θ^* on observed market prices. This is done by computing the prices as described in step (1) and letting the GJR parameters in θ^* vary in order to minimize a mean squared error function identical to equation (34)[13]. Once the calibration is finished the calibrated FHS GJR estimates can be used to price options as in step (1). Just as for the implied GJR, the number of Monte Carlo runs used for the calibration exercise is 10000, as also for the FHS GJR taking N larger (e.g. 30000) does not change the

[13] The minimization algorithm applied for both option implied models is the Nelder-Mead algorithm, which was also used in Barone-Adesi et al. (2008)

results much but is more time consuming. Unlike Barone-Adesi et al. (2008), who calibrate the model using only out-of-the-money options, the same set of in-sample-options used for the Implied GJR is used to estimate the model. This allows for a better in-between comparison of the two calibrated models.

4 Benchmark Models

4.1 Ad-hoc Black-Scholes

The first benchmark model is the ad-hoc Black Scholes model from Dumas et al. (1998). It works by smoothing the implied Black Scholes volatilities of the cross section of the selected options across the strikes and maturities by fitting the following equation

$$\sigma_{BS} = a_0 + a_1 K + a_2 K^2 + a_3 \tau + a_4 \tau^2 + a_5 K \tau + \epsilon_i \tag{40}$$

in which σ_{BS} is the Black-Scholes implied volatility for an option with strike price K and time to maturity τ. The parameters are estimated using a set of in-sample options with varying strike prices, times to maturity and moneyness categories one week prior to the test sample. Out-of-sample option prices can then be computed by fitting equation (40) and plugging the fitted implied volatilities into the Black-Scholes formula. Although this method is theoretically not consistent, ad-hoc Black-Scholes models are widely used in the industry because they incorporate information from different implied volatilities in the option pricing process and are easy to estimate. This makes the ad-hoc Black-Scholes a more challenging benchmark model than the regular Black Scholes model, which uses only a single volatility estimate. The ad-hoc Black Scholes is estimated using Bayesian methods to enhance the comparability of the pricing results. A detailed explanation is provided in the appendix on the Bayesian estimation of the ad-hoc Black Scholes.

4.2 Heston Nandi GARCH

The second benchmark is the closed form GARCH model from Heston and Nandi (2000). This model allows for asymmetry by introducing an asymmetry parameter in the GARCH specification. Heston and Nandi (2000) show that the model is able to outperform the simple Black Scholes model and other GARCH models in forecasting option prices out-of-sample. They describe that its dominance is due to the ability of the GARCH model to capture correlation of volatility and spot returns as well as the path dependence in volatility. The model specifications under the risk neutral measure $\mathbb{Q}$ are given by

$$log(\frac{P_t}{P_{t-1}}) = r_t = r_f - \frac{1}{2}\sigma_t^2 + \sigma_t z_t^*, \qquad z_t^* \sim \mathbb{N}(0,1), \tag{41}$$

$$\sigma_t^2 = \omega + \alpha_{HN}(z_{t-1}^* - \gamma_{HN}^* \sigma_{t-1})^2 + \beta_{HN}\sigma_{t-1}^2, \tag{42}$$

where

$$z_t^* = z_t + (\lambda + \frac{1}{2})\sigma_t, \qquad \text{and} \qquad \gamma_{HN}^* = \gamma_{HN} + \lambda + \tfrac{1}{2}.$$

The risk premium is given by λ, the asymmetry parameter is denoted by γ_{HN} and γ_{HN}^* is its risk neutral version. Note that the *HN* subscript is used to differentiate between parameters used in the GJR specification. Like the GJR model, the HN model also allows for asymmetric reaction to shocks. Instead of using an indicator variable, an additional asymmetry parameter is incorporated in the ARCH component which raises the variance more after a negative shock than after a positive shock. Decomposing the ARCH term in (41) gives

$$\alpha_{HN}(z_{t-1}^* - \gamma_{HN}^* \sigma_t)^2 \ = \ \alpha_{HN}(z_{t-1}^{2^*} + \gamma_{HN}^{2^*}\sigma_{t-1}^2 - 2\gamma_{HN}^*\sigma_{t-1}z_{t-1}^*) \tag{43}$$

which shows that a negative shock z_{t-1}^* turns the last term in the parentheses positive and thus increases the conditional variance estimate. Heston and Nandi (2000) show that the conditional variance and the stock price are correlated as following

$$Cov_{t-i}(\sigma_{t+i}^2, log(P_t)) \ = \ -2\alpha\gamma\sigma_t^2 \tag{44}$$

which, given positive estimates of α and γ, postulates a negative correlation between the conditional variance and the log spot price. This implies that the γ parameter describes the asymmetry in the distribution of the log returns, which means the model accounts for the leverage effect.

Heston and Nandi (2000) derive the call price as

$$C_t = \frac{1}{2}S_t + \frac{exp(-r\tau)}{\pi}\left(\int_0^\infty \mathbb{R}\Big[\frac{K^{-i\phi}\xi^*(i\phi+1)}{i\phi}\Big]d\phi\right)$$

$$-e^{-r\tau}K\bigg(\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathbb{R}\Big[\frac{K^{-i\phi}\xi^*(i\phi)}{i\phi}\Big]d\phi\bigg) \tag{45}$$

where $\xi^*(\phi)$ denotes the conditional moment generating function at time t of $log(P_T)$ given by

$$\xi^*(\phi) = E_{\mathbb{Q}}[e^{\phi log(P_T)}|\mathbb{F}_t] = P_t^{\phi}exp(A_t + B_t\sigma_{t+1}^2), \tag{46}$$

and $\mathbb{R}[.]$ is the real part of a complex number. The coefficients A_t and B_t are computed using a recursion, starting at the terminal time T condition $A_T = B_T = 0$. At time t the coefficients are then given by

$$A_t = A_{t+1} + \phi r_f + B_{t+1}\omega - \frac{1}{2}log(1-2\alpha B_{t+1}) \tag{47}$$

$$B_t = \phi(\lambda+\gamma) - \frac{1}{2}\gamma^2 + \beta B_{t+1} + \frac{(1/2)(\phi-\gamma)^2}{1-2\alpha B_{t+1}} \tag{48}$$

The HN GARCH parameters are estimated using Bayesian inference in the same way as explained in section (3.1.2), with the only difference being the GARCH specification. The posterior means are used to compute the recursions such that the integrals in (46) can be evaluated. In this paper only the main result is used in order to price options and compare results. For a thorough discussion of the model the interested reader is referred to Heston and Nandi (2000).

5 Data and Empirical Results

5.1 Data Description

For the empirical analysis the S&P500 index ("SPX") is selected. The options on the SPX are one of the most actively traded and are frequently used in empirical research. Index prices are obtained from CRSP and option data is provided by Optionmetrics. The interest rate is assumed to be fixed at an annual rate of 0.1%, considering the fact that the models are compared based solely on heteroskedasticity. The returns are daily and are multiplied by 100 to obtain percentage returns. Option prices are taken as the average of the closing bid-ask spread over all exchanges corresponding to a particular trading day. To ensure a representative sample is selected the following data is discarded; option prices smaller than $ 0.05, options with implied volatility larger than 70%, options with a trading volume lower than 100, and options with times to maturity smaller than 5 days or larger than 360 days. Furthermore, to reduce possible weekend effects only options traded on Wednesdays are selected. As the applied framework assumes no dividends are paid over the life of the option, the spot prices should be corrected for any dividend payments. A standard approach, introduced by Harvey and Whaley (1992) is used, in which cash dividends paid during the lifetime of the option are used as proxy for the expected dividend payments. The present value can then be calculated and subtracted from the index level to obtain the dividend corrected index level.

For model estimation a rolling window is used over the period January 1 2006 to December 24 2014 in which the models are re-estimated every third Wednesday of the month in 2014, yielding 12 series of estimated models. The estimation window consist of 2000 observations (approximately 8 years of return data) and is chosen sufficiently large such that the communication between the different states in the MS GJR model is properly captured. The model parameters are not expected to change much over a period of 4 weeks (corresponding to a shift of approximately 20 observations in the rolling window) such that the model parameters are kept fixed over a period of 4 weeks. This means each estimated model forecast option prices for horizons of 1, 2, 3 and 4 weeks ahead [14]. This approach yields 52 sets of predictive option prices with the minimum amount of options on a particular Wednesday being 466, the maximum amount being 2368, while the total amount of options after filtering is 36635. The option sets used for the calibration processes of the Implied GJR and FHS GJR are identical and have a price date equal to the last day in the sample and consists of options with varying strike prices, maturities and moneyness categories. These calibration sets are kept moderately in size to accommodate the calibration procedure, and consists of a minimum of 30 options and a maximum of 50 options. The same set of in-sample options is also used to estimate the ad hoc Black Scholes model.

The selected options do not possess any wild card features and are categorized based on short maturity ($\leq$ 60 days), medium maturity ((60,160] days) and long maturity ((160,360] days) such that performance amongst different times to expiration can be compared. Moneyness is defined as the strikeprice divided by the spotprice, $M_{\$} = \frac{K}{S_t}$, and is used to define the following categories: out-of-the-money if $M_{\$} \geq 1.015$, at-the-money if $0.985 < M_{\$} < 1.015$ and in-the-money if $M_{\$} \leq 0.985$. Deep in- and out-of-the money options are typically thinly traded and may suffer from liquidity issues and are excluded for that reason.

[14]The parameter estimates indeed hardly changed on a week-to-week basis. This is as expected because the rolling window only shifts 5 observations each week, relative to an estimation window of 2000 observations. Imposing fixed parameters for 4 weeks therefore only reduces accuracy to a minor degree while it greatly relieves the computational burden ((52-12) $\times$ 6 = 240 models less to estimate). For the option implied models the parameter changes are slightly larger but still small enough to keep the models fixed over a period of 4 weeks.

Descriptive Statistics SPX			
Mean	0.031	**Max**	4.734
Std. Dev.	1.338	**Skewness**	-0.082
Mode	-0.0429	**Kurtosis**	13.37
Median	0.077	**J-B pval.**	0.0001
Min	-9.01	**Interq. range**	1.042

Table 7: Descriptive statistics for the S&P 500 index percentage returns.

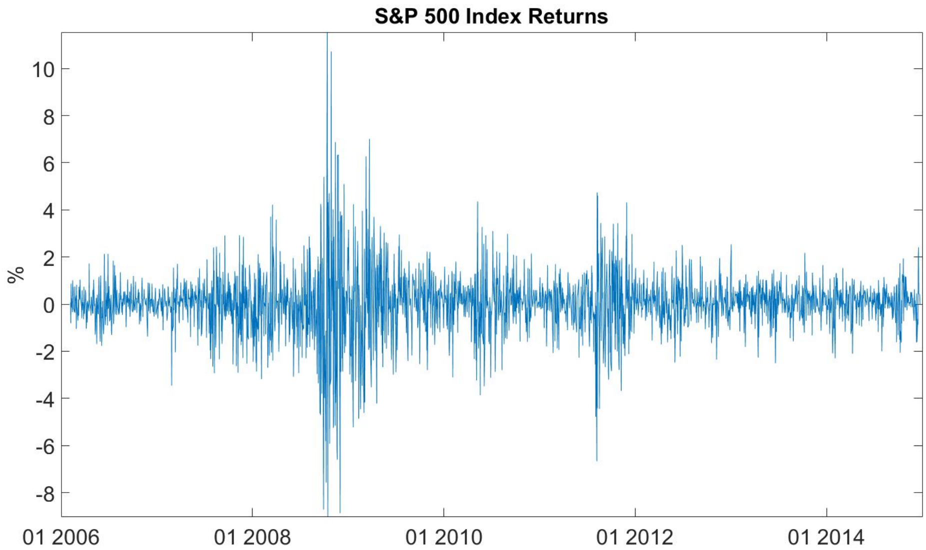


Figure 2: Plot of percentage returns on the S&P 500 index. The total sample used ranges from January 2006 to December 2014 and consists of 2238 observations.

5.2 Bayesian Inference Results

In the Gibbs sampler 35000 iterations are used and, even though the chain appears to converge rather quickly for both models, the first 5000 iterations are discarded as burn-in values out of precaution. The algorithms generally appear to converge to the stationary distribution well within the first 5000 iterations. The acceptance rates are high for the conditional mean parameters in β, with values around 0.95, which is an indication the proposal distribution is close to the posterior. For the GARCH parameters the acceptance rates vary between 0.6 and 0.7 for both models. To substantiate whether the number of states used in the MS GJR model is sufficient the Bayes factors are computed, which are ratios of marginal likelihoods that can be used to compare two models[15]. Only in rare cases can Bayes factors be computed analytically, however, using the Laplace method a decent approximation can be computed relatively easy. When testing model A versus model B, a-posteriori model A is more likely when $BF_{A|B} > 1$, assuming a-priori

[15] Bayes factors can loosely be interpreted as a 'Bayesian likelihood ratio test'.

model A is equally likely as model B. Higher values of the computed Bayes factor correspond to stronger evidence in favor of model A. The computed Bayes factors for the MS GJR models with different number of regimes are as following (with the subscript denoting the number of states); $\text{BF}_{3|1} = 61.58$, $\text{BF}_{2|1} = 23.66$ and $\text{BF}_{3|2} = 2.60$. Thus, the Bayes factors suggest that both the 3 and 2 state models are more likely than the 1 state model, while the 3 state model also seems more likely than the 2 state model. The improvement of $\text{BF}_{3|2}$ over $\text{BF}_{3|1}$ is only marginal, which suggests that adding more states will not improve performance much more. Therefore this can be interpreted as a rough indication that three states are indeed sufficient to model the conditional volatility. More details on the Bayes factors and their approximation are given in the appendix.

5.2.1 Inference Results GJR

The inference results for the GJR model presented in table (8) show that the return level represented by μ slightly increases over 2014 while the first order autocorrelation ρ behaves quite steady around a value of -0.07. The γ parameter can be interpreted as the level of the conditional variance and has relatively low values, suggesting a low volatility period during the estimation sample. The effect of the unexpected shocks, measured by α, is quite low relative to the reaction towards negative unexpected shocks, measured by δ. The values of α are quite low but seem to decline even further from a value of 0.09 in March to a value of 0.007 in the last sample in December. In contrast to this, the values of δ vary between 0.15 and 0.23 and appear to contribute quite well to describing the conditional variance. Thus, it appears the leverage effect is quite pronounced for the given data set. Meanwhile, the dependence on last times' conditional variance slightly increases as the estimation window rolls on with estimated values of 0.84 in January upwards to 0.90 in December. The estimated models appear to be highly persistent, as is usually the case for GARCH models. Overall, the parameter estimates do not appear to change too much over time.

Inference Results GJR

	μ	ρ	γ	α	ϕ	δ	PM
Jan	**0.0136**	**-0.0689**	**0.0543**	**0.0083**	**0.8376**	**0.2343**	**0.9630**
	0.0135	*0.0127*	*0.0045*	*0.006*	*0.006*	*0.0073*	*0.0052*
Feb	**0.0135**	**-0.0682**	**0.0366**	**0.0451**	**0.8428**	**0.1780**	**0.9769**
	0.0125	*0.0127*	*0.0109*	*0.0241*	*0.0124*	*0.0032*	*0.0105*
Mar	**0.0138**	**-0.0727**	**0.0457**	**0.0899**	**0.8117**	**0.1648**	**0.9840**
	0.0127	*0.0128*	*0.0198*	*0.0732*	*0.0667*	*0.0133*	*0.0038*
Apr	**0.0139**	**-0.0713**	**0.0147**	**0.0219**	**0.8932**	**0.1530**	**0.9916**
	0.0122	*0.0126*	*0.0052*	*0.014*	*0.0099*	*0.0219*	*0.0053*
May	**0.0137**	**-0.0732**	**0.0170**	**0.0146**	**0.8987**	**0.1510**	**0.9887**
	0.0118	*0.0123*	*0.0028*	*0.0109*	*0.0057*	*0.0167*	*0.0047*
Jun	**0.0157**	**-0.0711**	**0.0294**	**0.0134**	**0.8753**	**0.1741**	**0.9758**
	0.0123	*0.0126*	*0.0015*	*0.0045*	*0.0022*	*0.0022*	*0.0061*
Jul	**0.0156**	**-0.0719**	**0.0176**	**0.0077**	**0.9004**	**0.1550**	**0.9855**
	0.0120	*0.0124*	*0.0026*	*0.0065*	*0.0085*	*0.0080*	*0.0034*
Aug	**0.0172**	**-0.0723**	**0.0373**	**0.0071**	**0.8785**	**0.1702**	**0.9707**
	0.0126	*0.0124*	*0.0104*	*0.0056*	*0.0044*	*0.0161*	*0.0085*
Sep	**0.0175**	**-0.0733**	**0.0250**	**0.0072**	**0.8838**	**0.1838**	**0.9830**
	0.0120	*0.0125*	*0.0021*	*0.0057*	*0.0038*	*0.0098*	*0.0048*
Oct	**0.0175**	**-0.0740**	**0.0168**	**0.0096**	**0.9015**	**0.1517**	**0.9869**
	0.0120	*0.0125*	*0.0026*	*0.0075*	*0.0057*	*0.0113*	*0.0043*
Nov	**0.0196**	**-0.0739**	**0.0411**	**0.0054**	**0.8632**	**0.2125**	**0.9748**
	0.0125	*0.0129*	*0.0197*	*0.0044*	*0.0307*	*0.0392*	*0.0085*
Dec	**0.0194**	**-0.0731**	**0.0223**	**0.0075**	**0.8972**	**0.1684**	**0.9889**
	0.0122	*0.0127*	*0.0063*	*0.0060*	*0.0069*	*0.0093*	*0.0043*

Table 8: Inference results for the GJR model. Given are the posterior means (in bold) and posterior standard deviations (in italic) of the posterior parameter distributions. The models are re-estimated every the third Wednesday of the month during the year 2014. In the Gibbs sampler 35000 iterations are used of which 5000 are used as burn-in sample. "PM" Denotes the persistence measure and is defined as $\alpha + \phi + \delta/2$.

5.2.2 Inference Results MS GJR

The posterior inference results of the regime parameters are given in table (9). The different regimes are properly identified in both the conditional mean and the GARCH parameters. Moreover, no bimodalities are observed in the posterior histograms (not reported here), indicating the model does not suffer from the label switching problem. Considering the conditional mean, it holds that $\mu_3 > \mu_2 > \mu_1$, indicating a slightly higher return level during the different regimes. The differences however, are only of small magnitude. The auto-correlation parameter is roughly the same for all regimes, although it is slightly less negative in the medium volatility regime. The level of the conditional variance is clearly identified among the different regimes as $\gamma_3 > \gamma_2 > \gamma_1$. Noteworthy is the fact that γ_1 in the MS GJR in general is higher than the γ values observed in the GJR model, suggesting a possible downward bias in the description of the level of the conditional variance in the GJR model. For the ascending regimes the effects of the unexpected shocks, α_j, and the asymmetry effect, δ_j, increases. For instance, the effect of the unexpected shocks in the medium regime is up to 10 times as high as for the low regime, while the high regime has α values up to twice as large as the medium regime. The δ values are quite substantial in all regimes, although they increase for more volatile regimes. The dependence on last periods' conditional variance diminishes for more volatile regimes, which coincides with larger unexpected shocks in more volatile periods which are being captured by higher values of α_j and δ_j. Thus, the results roughly comply with the prior intuition about the interpretation of the regimes describing low, medium and high volatility periods. In comparison to the single regime GJR model, the GARCH estimates in the MS GJR are less persistent, which mainly stems from their lower dependence on the previous conditional variances. Bauwens et al. (2010) point out that standard GARCH models generally suffer from an upward bias in the persistence parameter, which may be why the estimated GJR models are close to being integrated.

The posterior transition probability matrices are given in table (10) and indicate a high probability to remain in a given state once it is entered. From an economic perspective this can be explained as different periods of volatility lasting for a certain amount of time, which is typically the case in, for example, expansion or recession. Indeed, the return plot in figure (2) suggests transitions between states do not occur continuously but typically persist for quite some time. A plot of the estimated probabilities of being in state i over time is given in figure (3).

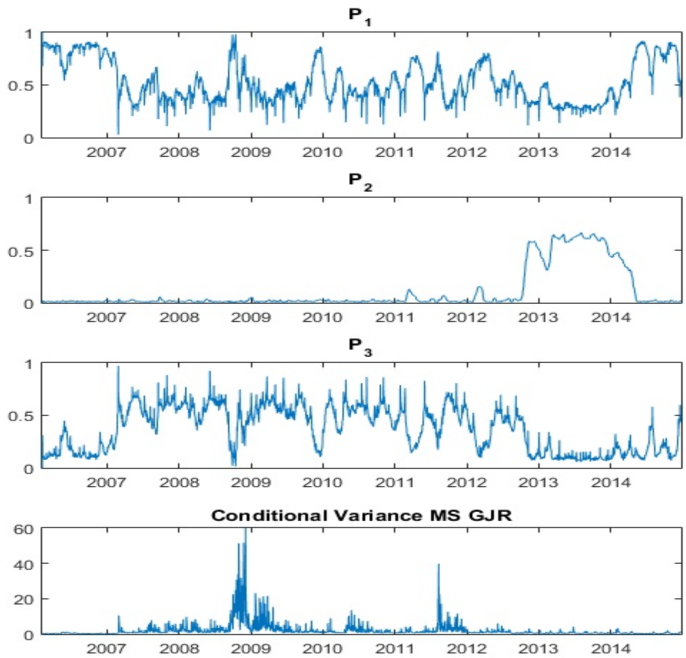


Figure 3: Posterior regime probabilities for the total estimation window. The top figure (P_1) shows the posterior probability of being in state 1, the second figure (P_2) shows the posterior probability of being in state 2, the third figure (P_3) shows the posterior probability of being in state 3 and the bottom figure displays the fitted MS GJR conditional variance.

Inference Results Markov Switching GJR

	Regime 1							**Regime 2**							**Regime 3**						
	μ	ρ	γ	α	ϕ	δ	**PM**	μ	ρ	γ	α	ϕ	δ	**PM**	μ	ρ	γ	α	ϕ	δ	**PM**
Jan	**0.037**	**-0.068**	**0.061**	**0.027**	**0.691**	**0.253**	**0.845**	**0.045**	**-0.057**	**0.164**	**0.200**	**0.487**	**0.291**	**0.833**	**0.049**	**-0.072**	**0.465**	**0.235**	**0.413**	**0.369**	**0.833**
	0.016	*0.022*	*0.017*	*0.020*	*0.046*	*0.182*	*0.053*	*0.014*	*0.022*	*0.026*	*0.041*	*0.044*	*0.055*	*0.039*	*0.014*	*0.025*	*0.084*	*0.120*	*0.036*	*0.038*	*0.104*
Feb	**0.037**	**-0.065**	**0.099**	**0.132**	**0.609**	**0.154**	**0.818**	**0.043**	**-0.051**	**0.164**	**0.199**	**0.489**	**0.291**	**0.834**	**0.048**	**-0.068**	**0.502**	**0.225**	**0.272**	**0.307**	**0.650**
	0.017	*0.024*	*0.026*	*0.086*	*0.063*	*0.074*	*0.015*	*0.014*	*0.025*	*0.032*	*0.028*	*0.037*	*0.070*	*0.048*	*0.013*	*0.029*	*0.100*	*0.121*	*0.059*	*0.069*	*0.207*
Mar	**0.027**	**-0.068**	**0.071**	**0.035**	**0.671**	**0.248**	**0.830**	**0.034**	**-0.059**	**0.175**	**0.193**	**0.484**	**0.265**	**0.810**	**0.038**	**-0.070**	**0.403**	**0.224**	**0.315**	**0.234**	**0.655**
	0.016	*0.022*	*0.016*	*0.025*	*0.045*	*0.044*	*0.044*	*0.013*	*0.022*	*0.046*	*0.046*	*0.042*	*0.149*	*0.109*	*0.012*	*0.025*	*0.046*	*0.120*	*0.036*	*0.108*	*0.203*
Apr	**0.041**	**-0.063**	**0.069**	**0.043**	**0.678**	**0.255**	**0.848**	**0.047**	**-0.051**	**0.165**	**0.201**	**0.485**	**0.292**	**0.832**	**0.051**	**-0.070**	**0.500**	**0.302**	**0.246**	**0.387**	**0.741**
	0.016	*0.025*	*0.014*	*0.028*	*0.030*	*0.035*	*0.038*	*0.013*	*0.025*	*0.035*	*0.056*	*0.071*	*0.155*	*0.084*	*0.013*	*0.029*	*0.114*	*0.092*	*0.088*	*0.025*	*0.181*
May	**0.043**	**-0.066**	**0.067**	**0.033**	**0.676**	**0.260**	**0.839**	**0.049**	**-0.051**	**0.164**	**0.199**	**0.489**	**0.291**	**0.833**	**0.054**	**-0.072**	**0.461**	**0.398**	**0.294**	**0.274**	**0.829**
	0.017	*0.024*	*0.008*	*0.023*	*0.025*	*0.036*	*0.019*	*0.014*	*0.025*	*0.018*	*0.016*	*0.007*	*0.034*	*0.032*	*0.013*	*0.028*	*0.088*	*0.030*	*0.055*	*0.104*	*0.121*
Jun	**0.036**	**-0.054**	**0.069**	**0.021**	**0.664**	**0.251**	**0.810**	**0.044**	**-0.042**	**0.164**	**0.200**	**0.489**	**0.291**	**0.834**	**0.048**	**-0.058**	**0.460**	**0.235**	**0.349**	**0.237**	**0.702**
	0.017	*0.024*	*0.012*	*0.013*	*0.023*	*0.019*	*0.020*	*0.014*	*0.024*	*0.020*	*0.040*	*0.029*	*0.030*	*0.027*	*0.013*	*0.028*	*0.078*	*0.121*	*0.023*	*0.111*	*0.185*
Jul	**0.047**	**-0.072**	**0.078**	**0.039**	**0.623**	**0.231**	**0.777**	**0.053**	**-0.058**	**0.164**	**0.200**	**0.488**	**0.290**	**0.833**	**0.058**	**-0.078**	**0.348**	**0.339**	**0.374**	**0.304**	**0.865**
	0.017	*0.026*	*0.016*	*0.026*	*0.043*	*0.030*	*0.044*	*0.015*	*0.025*	*0.020*	*0.020*	*0.013*	*0.026*	*0.028*	*0.014*	*0.029*	*0.027*	*0.070*	*0.021*	*0.086*	*0.102*
Aug	**0.033**	**-0.038**	**0.074**	**0.059**	**0.513**	**0.301**	**0.723**	**0.039**	**-0.029**	**0.166**	**0.197**	**0.496**	**0.282**	**0.834**	**0.042**	**-0.043**	**0.253**	**0.357**	**0.418**	**0.309**	**0.930**
	0.015	*0.022*	*0.051*	*0.040*	*0.105*	*0.057*	*0.039*	*0.012*	*0.022*	*0.010*	*0.032*	*0.083*	*0.128*	*0.026*	*0.011*	*0.024*	*0.089*	*0.054*	*0.061*	*0.083*	*0.049*
Sep	**0.045**	**-0.064**	**0.076**	**0.028**	**0.666**	**0.208**	**0.798**	**0.051**	**-0.052**	**0.164**	**0.200**	**0.489**	**0.291**	**0.834**	**0.055**	**-0.071**	**0.254**	**0.223**	**0.448**	**0.258**	**0.799**
	0.016	*0.024*	*0.012*	*0.017*	*0.122*	*0.119*	*0.078*	*0.014*	*0.023*	*0.039*	*0.019*	*0.018*	*0.068*	*0.038*	*0.013*	*0.027*	*0.073*	*0.123*	*0.065*	*0.098*	*0.121*
Oct	**0.018**	**-0.070**	**0.069**	**0.057**	**0.657**	**0.232**	**0.830**	**0.027**	**-0.058**	**0.166**	**0.202**	**0.486**	**0.288**	**0.832**	**0.032**	**-0.070**	**0.392**	**0.279**	**0.465**	**0.209**	**0.848**
	0.017	*0.024*	*0.007*	*0.019*	*0.014*	*0.009*	*0.009*	*0.013*	*0.024*	*0.037*	*0.058*	*0.055*	*0.068*	*0.041*	*0.012*	*0.028*	*0.032*	*0.092*	*0.065*	*0.123*	*0.090*
Nov	**0.051**	**-0.088**	**0.069**	**0.067**	**0.652**	**0.237**	**0.837**	**0.057**	**-0.073**	**0.164**	**0.200**	**0.490**	**0.290**	**0.835**	**0.062**	**-0.095**	**0.408**	**0.385**	**0.182**	**0.367**	**0.751**
	0.018	*0.025*	*0.019*	*0.011*	*0.019*	*0.003*	*0.030*	*0.015*	*0.025*	*0.030*	*0.018*	*0.064*	*0.044*	*0.058*	*0.015*	*0.030*	*0.040*	*0.033*	*0.105*	*0.036*	*0.141*
Dec	**0.032**	**-0.073**	**0.068**	**0.063**	**0.660**	**0.252**	**0.849**	**0.039**	**-0.063**	**0.165**	**0.200**	**0.488**	**0.291**	**0.834**	**0.043**	**-0.076**	**0.322**	**0.279**	**0.401**	**0.212**	**0.786**
	0.017	*0.023*	*0.028*	*0.025*	*0.029*	*0.029*	*0.018*	*0.013*	*0.023*	*0.027*	*0.024*	*0.040*	*0.038*	*0.046*	*0.012*	*0.025*	*0.050*	*0.098*	*0.035*	*0.102*	*0.172*

Table 9: Inference results for the Markov Switching GJR model. Given are the posterior means (in bold) and the posterior standard deviations (in italic) of the parameters. The models are estimated monthly on the third Wednesday of the month during the year 2014 using 8 years of return data, corresponding to approximately 2000 observations. The persistence measure is denote by "PM" and is defined as $\alpha + \phi + \delta/2$.

Posterior Transition Probabilities MS GJR model

Jan	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.99**	**0.00**	**0.01**
	4E-04	*9E-07*	*4E-04*
$p_{2.k}$	**0.00**	**0.96**	**0.03**
	0.003	*0.021*	*0.021*
$p_{3.k}$	**0.02**	**0.00**	**0.98**
	0.006	*5E-05*	*0.006*

Feb	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.92**	**0.01**	**0.08**
	0.007	*0.001*	*0.006*
$p_{2.k}$	**0.17**	**0.81**	**0.02**
	0.016	*0.019*	*0.002*
$p_{3.k}$	**0.02**	**0.00**	**0.98**
	0.004	*1E-04*	*0.005*

Mar	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.97**	**0.00**	**0.03**
	0.002	*7E-04*	*0.001*
$p_{2.k}$	**0.19**	**0.77**	**0.05**
	0.016	*0.02*	*0.004*
$p_{3.k}$	**0.02**	**0.00**	**0.98**
	0.005	*7E-05*	*0.005*

Apr	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.96**	**0.00**	**0.04**
	0.003	*9E-06*	*0.003*
$p_{2.k}$	**0.02**	**0.94**	**0.03**
	0.009	*0.023*	*0.025*
$p_{3.k}$	**0.03**	**0.00**	**0.97**
	0.006	*3E-06*	*0.006*

May	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.91**	**0.04**	**0.05**
	0.005	*0.004*	*0.001*
$p_{2.k}$	**0.14**	**0.85**	**0.01**
	0.006	*0.008*	*6E-04*
$p_{3.k}$	**0.02**	**0.00**	**0.97**
	0.005	*7E-04*	*0.006*

Jun	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.98**	**0.00**	**0.02**
	7E-04	*8E-05*	*6E-04*
$p_{2.k}$	**0.02**	**0.93**	**0.05**
	0.001	*0.006*	*0.006*
$p_{3.k}$	**0.02**	**0.01**	**0.97**
	0.006	*0.001*	*0.007*

Jul	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.94**	**0.00**	**0.06**
	9E-04	*1E-06*	*0.001*
$p_{2.k}$	**0.00**	**0.98**	**0.01**
	0.003	*0.010*	*0.01*
$p_{3.k}$	**0.04**	**0.00**	**0.96**
	0.008	*5E-07*	*0.008*

Aug	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.96**	**0.00**	**0.04**
	0.001	*1E-04*	*0.002*
$p_{2.k}$	**0.13**	**0.80**	**0.08**
	0.010	*0.010*	*0.003*
$p_{3.k}$	**0.02**	**0.01**	**0.97**
	0.005	*0.001*	*0.007*

Sep	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.85**	**0.02**	**0.13**
	0.004	*0.001*	*0.003*
$p_{2.k}$	**0.12**	**0.80**	**0.08**
	0.002	*0.006*	*0.010*
$p_{3.k}$	**0.03**	**0.00**	**0.97**
	0.004	*1E-04*	*4E-03*

Oct	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.98**	**0.00**	**0.02**
	8E-04	*3E-06*	*0.001*
$p_{2.k}$	**0.02**	**0.87**	**0.11**
	0.01	*0.03*	*0.02*
$p_{3.k}$	**0.02**	**0.00**	**0.98**
	0.006	*3E-05*	*0.006*

Nov	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.90**	**0.00**	**0.10**
	0.003	*6E-04*	*0.003*
$p_{2.k}$	**0.09**	**0.82**	**0.09**
	0.02	*0.02*	*0.01*
$p_{3.k}$	**0.01**	**0.00**	**0.99**
	0.002	*7E-05*	*0.001*

Dec	$p_{j.1}$	$p_{j.2}$	$p_{j.3}$
$p_{1.k}$	**0.98**	**0.00**	**0.02**
	0.003	*1E-06*	*0.003*
$p_{2.k}$	**0.01**	**0.98**	**0.01**
	0.01	*0.01*	*0.01*
$p_{3.k}$	**0.02**	**0.00**	**0.98**
	0.001	*1E-06*	*0.001*

Table 10: Posterior transition probability matrices for the estimated Markov Switching GJR models. Given are the posterior means (in bold) and the posterior standard deviations (in italic) of the estimated transition probabilities. The matrices correspond to the estimated regime parameters displayed in table (9).

5.2.3 Prior Sensitivity

As a robustness check the prior sensitivity of the MS GJR model is considered. Since the model employs informative priors some caution has to be taken concerning the prior influence. If the priors dominate the posterior distribution the inference results may become too subjective and may not be useful. To assess the influence of the informative priors on the GARCH parameters, the MS GJR model of the first sample in January is re-estimated using flat priors. The inference results are compared with the results of the MS GJR model for January that uses the informative prior. For the large estimation window used the prior influence is expected to be negligible, meaning the parameters should not deviate much from those observed in table (13). Large differences between the estimated parameters may indicate a possible misspecification of the model. The resulting estimates of the MS GJR with flat priors are presented in table (11). Some minor differences in the parameters of the non-informative model are observed. For example, the parameters of the conditional mean are slightly larger in magnitude, whereas the γ parameters are slightly lower for regimes 1 and 2 and slightly higher for regime 3. The differences however, are very marginal and not larger than 0.025. The other parameters in regimes 1 and 2 also do not show any notable differences. In the high volatility regime the differences are more pronounced; α is approximately 0.1 higher and δ is approximately 0.14 lower in the non-informative model. Hence, there appears to be a difference in the trade-off between the reaction towards shocks in general and towards negative shocks, which may be related to the informative prior. Overall, the informative prior seems to have only a small influence on the posterior distribution, which is mainly limited to the parameters in the high volatility regime. The posterior transition probabilities are also very similar, and also indicate the states are very persistent.

Estimated MS GJR Model January using Non-Informative Priors

μ	ρ	γ	α	ϕ	δ	**PM**
Regime 1						
0.0359	**-0.0694**	**0.0347**	**0.0215**	**0.8508**	**0.0798**	**0.9122**
0.0175	*0.0227*	*0.0082*	*0.0166*	*0.0186*	*0.0349*	*0.0200*
Regime 2						
0.0439	**-0.0575**	**0.1390**	**0.0484**	**0.6618**	**0.2882**	**0.8542**
0.0148	*0.0233*	*0.0375*	*0.0342*	*0.0415*	*0.1757*	*0.0752*
Regime 3						
0.0482	**-0.0713**	**0.3794**	**0.2286**	**0.5659**	**0.2285**	**0.9087**
0.0136	*0.0269*	*0.1843*	*0.0790*	*0.0105*	*0.1265*	*0.0697*

Posterior Transition Probabilities

	$\mathbf{p}_{j.1}$	$\mathbf{p}_{j.2}$	$\mathbf{p}_{j.3}$
$\mathbf{p}_{1.k}$	**0.926**	**0.032**	**0.041**
	0.004	*0.001*	*0.004*
$\mathbf{p}_{2.k}$	**0.038**	**0.958**	**0.004**
	0.005	*0.006*	*0.001*
$\mathbf{p}_{3.k}$	**0.033**	**0.005**	**0.962**
	0.007	*0.001*	*0.008*

Table 11: Inference results for the MS GJR model in January using flat priors. Given are the posterior means (in bold) and the posterior standard deviations (in italic). In the Gibbs sampler 35000 iterations are used of which 5000 serve as burn-in values.

5.3 Calibration Results Option Implied Models

Table (13) presents the calibration results of the Implied GJR and the FHS GJR. The models are re-calibrated every month on the same set of options one week prior to the options under consideration. Giving a clear interpretation to the calibrated parameters is difficult as these are obtained as possibly sub-optimal solutions to a highly dimensional problem. Furthermore, since the parameters are estimated from option data they reflect the risk-neutral structure, which is different from the real world measure. This makes it even harder to give an interpretation to the parameter estimates. Nonetheless, a cautious attempt can be done to describe any differences and similarities. The γ parameter is quite low for both models and is similar to values observed in the GJR model. Both models have a relatively high ϕ parameter, although in the FHS model this parameter is slightly higher and more steady over time. The α parameter is quite low, especially in the FHS model where values close to 0 are observed. The values of δ show some large fluctuations in between months in both the FHS and Implied GJR models, which seem to be part of a trade-off between ϕ and δ. High values of δ are compensated by lower values of ϕ and vice-versa. Hence, the models mainly capture the return dynamics through ϕ and δ, while the α and γ parameters are typically very low.

Estimation Results Option Implied Models

	Implied GJR					**FHS GJR**				
	γ	α	ϕ	δ	**PM**	γ	α	ϕ	δ	**PM**
Jan	**0.072**	**0.011**	**0.631**	**0.371**	**0.828**	**0.028**	**1E-04**	**0.902**	**4E-04**	**0.902**
	0.006	*0.001*	*0.017*	*0.033*	*0.024*	*0.003*	*0.002*	*0.003*	*0.005*	*0.004*
Feb	**0.019**	**0.000**	**0.909**	**0.013**	**0.916**	**0.064**	**0.057**	**0.779**	**0.328**	**0.999**
	0.006	*0.001*	*0.021*	*0.002*	*0.021*	*0.007*	*0.012*	*0.032*	*0.029*	*0.037*
Mar	**0.000**	**0.027**	**0.805**	**0.085**	**0.874**	**0.028**	**1E-04**	**0.902**	**4E-04**	**0.902**
	0.000	*0.002*	*0.002*	*0.001*	*0.003*	*3E-04*	*0.007*	*0.003*	*5E-04*	*0.008*
Apr	**0.020**	**0.018**	**0.770**	**0.269**	**0.922**	**0.030**	**1E-04**	**0.944**	**0.113**	**0.999**
	0.002	*0.002*	*0.009*	*0.002*	*0.010*	*0.007*	*0.006*	*0.092*	*0.006*	*0.093*
May	**0.000**	**0.103**	**0.850**	**0.020**	**0.963**	**0.031**	**0.020**	**0.760**	**0.302**	**0.931**
	0.003	*0.001*	*0.005*	*0.006*	*0.006*	*4E-04*	*0.017*	*0.022*	*0.006*	*0.023*
Jun	**0.000**	**0.015**	**0.625**	**0.324**	**0.803**	**0.032**	**1E-04**	**0.874**	**3E-04**	**0.873**
	0.000	*0.003*	*0.028*	*0.010*	*0.029*	*0.0135*	*0.0026*	*0.003*	*0.01*	*0.006*
Jul	**0.001**	**0.000**	**0.902**	**0.053**	**0.928**	**0.014**	**1E-05**	**0.983**	**0.007**	**0.986**
	0.000	*0.000*	*0.005*	*0.002*	*0.005*	*0.0011*	*4E-04*	*0.002*	*0.002*	*0.002*
Aug	**0.065**	**0.004**	**0.538**	**0.324**	**0.704**	**0.020**	**1E-04**	**0.729**	**0.094**	**0.776**
	0.002	*0.010*	*0.001*	*0.030*	*0.018*	*0.0011*	*0.007*	*0.001*	*0.005*	*0.007*
Sep	**0.021**	**0.002**	**0.905**	**0.000**	**0.908**	**0.031**	**0.008**	**0.811**	**0.231**	**0.935**
	0.002	*0.017*	*0.003*	*0.001*	*0.017*	*0.0023*	*0.001*	*0.008*	*0.005*	*0.008*
Oct	**0.027**	**0.002**	**0.919**	**0.000**	**0.921**	**0.017**	**1E-05**	**0.967**	**8E-04**	**0.966**
	0.047	*0.014*	*0.032*	*0.010*	*0.036*	*8E-04*	*2E-04*	*6E-04*	*1E-05*	*7E-04*
Nov	**0.026**	**0.004**	**0.685**	**0.530**	**0.954**	**0.005**	**1E-05**	**0.973**	**1E-04**	**0.973**
	0.026	*0.004*	*0.0001*	*0.001*	*0.004*	*0.001*	*3E-04*	*9E-04*	*1E-04*	*0.001*
Dec	**0.037**	**0.009**	**0.739**	**0.417**	**0.957**	**1E-04**	**1E-05**	**0.924**	**0.068**	**0.955**
	0.002	*0.001*	*0.006*	*0.018*	*0.011*	*1E-04*	*1E-04*	*0.004*	*0.003*	*0.004*

Table 13: Estimates of the calibrated GJR and FHS-GJR models. Given are the calibration results of weekly re-calibrating both models using 10000 Monte Carlo runs and one day of option data. Note that PM is the persistence measure, which is defined as $\alpha + \phi + \delta/2$.

5.4 Predictive Option Pricing Results

Option prices for the GJR and MS GJR are computed using the algorithm described in section 3.4. The outer integral in equation (29) over θ is approximated by $MC = 30000$ Monte Carlo iterations. This number appears large enough to obtain consistent results such that the inner integral over the risk neutral densities can be approximated by setting $N = 1$. For the Implied GJR and the FHS GJR the settings are $MC = 1$ and $N = 30000$ which coincides with a classical maximum likelihood scheme. The computed option prices are then compared based on the Mean Absolute Percentage Error ("MAPE"),

$$MAPE = \frac{1}{N}\sum_{n=1}^{N}\frac{|C_n(t,T,K) - \widehat{C}_n(t,T,K)|}{C_n(t,T,K)}, \tag{49}$$

the Mean Mispricing Error ("MME"),

$$MME = \frac{1}{N}\sum_{n=1}^{N}\frac{C_n(t,T,K) - \widehat{C}_n(t,T,K)}{C_n(t,T,K)} \tag{50}$$

and the Root Mean Squared Error ("RMSE"),

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(C_n(t,T,K) - \widehat{C}_n(t,T,K))^2}. \tag{51}$$

These evaluation metrics are chosen to compare the results on different criteria. The MAPE measures the magnitude of the pricing errors and represent a measure of accuracy, the MME indicates the direction of the pricing error (i.e. over- or underpricing) and describes the bias, while the RMSE describes a combination of bias and volatility of the pricing errors.

5.4.1 Analysis of Pricing Errors

As initial analysis the box and whisker plots of the pricing errors, given in figure (4), are inspected for any systematic bias patterns. If the models are correctly specified the pricing errors should be centered approximately around 0 across moneyness categories and time to expiration. For the GJR model, the plot reveals that this is not the case. Only for short and medium maturity in-the-money options the pricing errors are centered somewhat around 0. However, even for these categories the GJR produces substantial negative pricing errors. When options move away

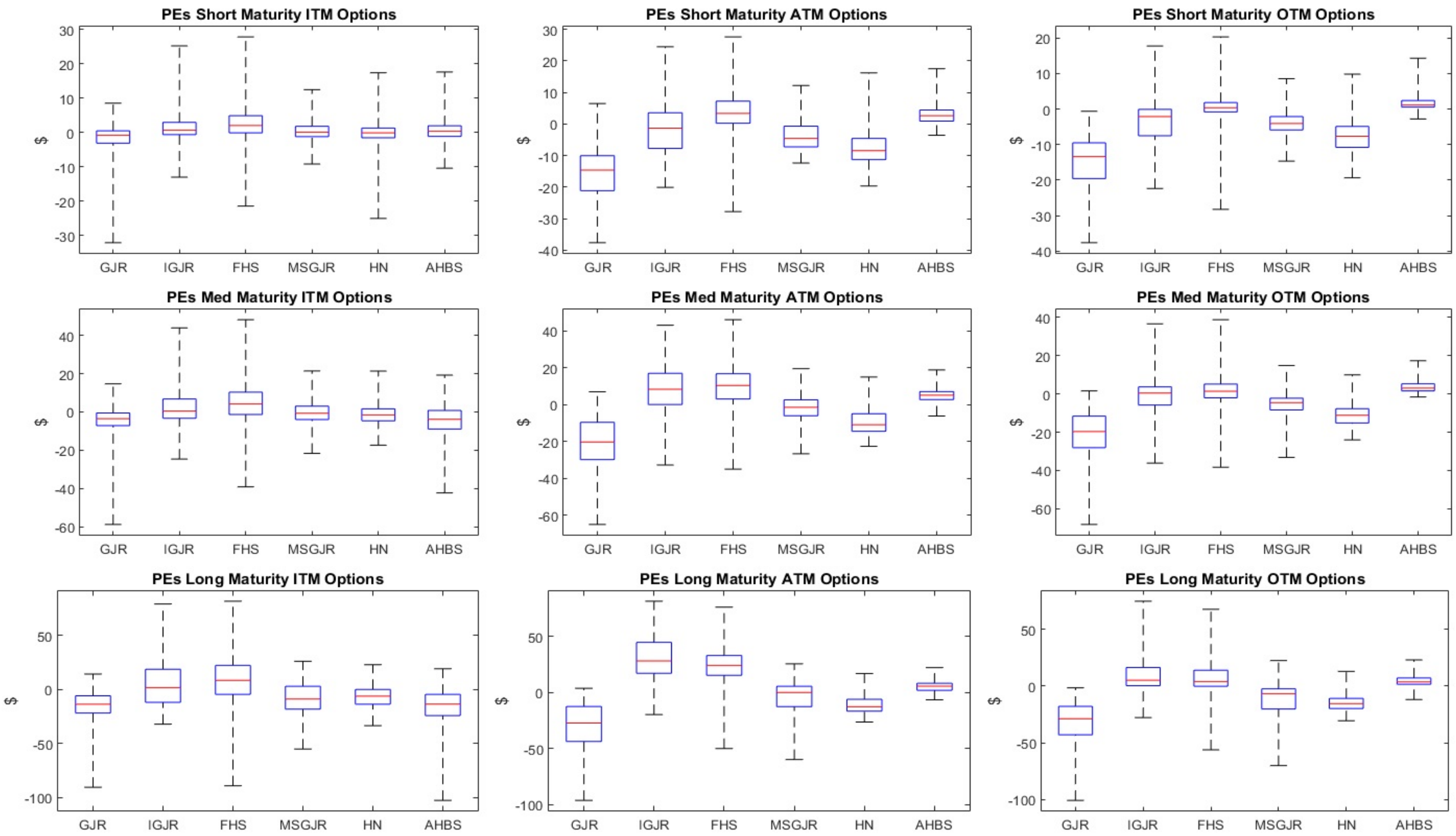


Figure 4: This figure shows the box and whiskers plots of aggregate pricing errors ("PEs"). The box and whiskers plot provide a visual interpration of the spread of the data and consists of the minimum, the first quartile, the median (red line), the third quartile and the maximum of the data set. The x-axis display the models; GJR, Implied GJR (IGJR), FHS GJR (FHS), MS GJR, HN and Ad-Hoc BS (AHBS), while the y-axis displays the pricing error in U.S. $ currency.

from in-the-money pricing performance of the GJR deteriorates in a rapid fashion. For at-the-money and out-of-the-money options the pricing errors are centered far below 0, suggesting the GJR is possibly misspecified. In addition to this, negative pricing errors are realised more often and are of much larger magnitude than positive errors, indicating a negative skewness in the pricing errors. This indicates that the GJR has a systematic upward bias, meaning it has a general tendency to overprice options. The plot also reveals a pattern for the pricing errors of both the option implied models. In- and out-of-the-money residuals are centered approximately around zero, while at-the-money options are centered above zero for both the Implied GJR and FHS GJR. This means at-the-money options are more subject to underpricing by the option implied models, which becomes more pronounced as maturity increases. The pricing errors of the FHS GJR reveal a symmetric pattern, suggesting overpricing happens in equal proportion to underpricing for the categories of which the residuals are centered around 0. In contrary, the pricing errors of the Implied GJR appear to be more asymmetric and indicate an underpricing behaviour. The pricing errors corresponding to MS GJR are centered relatively closer around 0 across moneyness categories in comparison with the GJR, Implied GJR and FHS GJR, which suggests a more consistent pricing performance. However, the median value appears to shift slightly below 0 as options move towards out-of-the-money and some asymmetries are observed for the long maturity pricing errors. Similar to the GJR, the MS GJR tends to overprice options as maturity increases, although the overpricing appears to be less severe.

The plot suggests that the Implied GJR, FHS GJR and MS GJR remove a large portion of the systematic bias observed in the GJR model, which is especially pronounced for medium and long maturities and at- and out-of-the-money options. However, unlike the MS GJR, the option implied models have a downwards bias for at-the-money options, meaning they have a tendency to underprice options in this category. This pattern amplifies for longer maturities, where the pricing errors of both option implied models are centered substantially above 0. Furthermore, the minimum and maximum pricing errors of the MS GJR are substantially lower than those of the GJR, Implied GJR, and FHS GJR for most categories. A closer look into the data revealed that relatively large pricing also errors appear to occur less frequent for the MS GJR.

When considering the benchmark models, the HN and ad-hoc BS, it appears these are performing quite well and prove hard to beat. The HN pricing errors have a relatively small spread and are centered closely around 0 for in-the-money options and medium and long maturity at-the-money options. The remaining categories suffer from a slight upward pricing bias as the pricing errors are concentrated slightly below 0. The pricing errors of the ad-hoc BS are also centered closely around 0 and show a very tight spread in case of at-the-money and out-of-the-money options. In-the-money options are priced less accurate by the ad-hoc BS as the spread of the pricing errors is relatively large, while also some relatively large negative pricing errors are observed in this category. The magnitude of the pricing errors of the ad-hoc BS seems to decrease as options move more towards out-of-the-money.

Overall the box and whiskers plots suggest the following; (1) the simple GJR model produces heavily biased pricing forecasts, (2) incorporating past option data in the model removes a large portion of this bias and (3) incorporating a regime switching property reduces this bias even further. These observations can be attributed to the fact that the option implied models capture more dynamics from the risk-neutral distribution than the simple GJR (recall that in a GARCH setting there are an infinite amount of risk-neutral measures) while the performance of the MS GJR can be ascribed to its the ability to capture the well discussed stylized facts such as leptokurtosis and volatility jumps. The improvement is most profound for out-of and at-the-money options, where the GJR appears to be completely misspecified in comparison to the MS GJR, Implied GJR and FHS GJR. Furthermore, the MS GJR appears to perform slightly better than the option implied models in terms of bias and accuracy as these pricing errors are centered closer around 0 and show a slightly smaller spread.

5.4.2 Aggregate Results

The aggregate results of the MME, MAPE and RMSE over 2014, displayed in table (14), show a clear pattern; a relative good pricing performance of in-the-money options which decreases as time to maturity increases and a relatively bad pricing performance of at- and out-of-the-money options, which improves as time to maturity increases. Furthermore the RMSEs of the pricing errors increase as time to maturity increases, which is as expected as the volatility of the simulated asset price increases along the maturity horizon. This pattern seems to hold for all models, although there are some differences in performance between different categories among the models. Overpricing is quite substantial for the GJR model given the negative MMEs observed across all maturity and moneyness categories. For out-of-the-money options the GJR performs worst of all, it overprices short maturity options in this category by 600% on average. The performance of the GJR in this moneyness bucket seems to improve slightly as time to maturity increases. When comparing the RMSEs, the GJR is again outperformed by the other models, which indicates a relatively high bias and/or volatility of the pricing errors. The severe overpricing, as observed for the GJR model, is less pronounced for the remaining models. The MS GJR mostly overprices out-of-the-money options, although the extent of overpricing reduces as time to maturity increases. Like the GJR, the MS GJR is able to price in-the-money options fairly accurate while at- and out-of-the-money options are priced less good in terms of accuracy and volatility. The performance in these categories however, is still strikingly better than that of the GJR model.

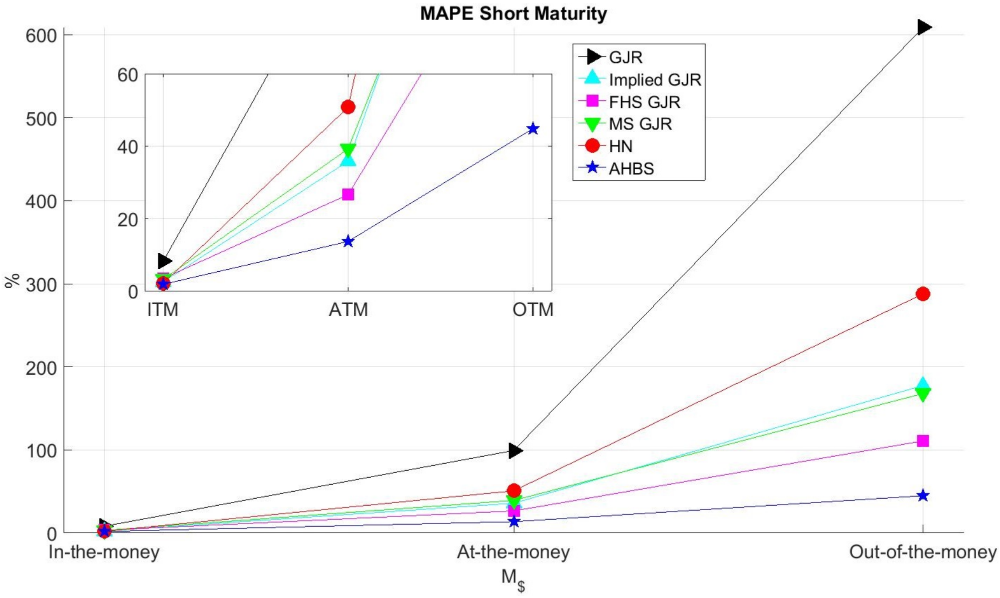
Aggregate Option Pricing Results

		Days to Expiration								
		≤ 60			(60,160]			> 160		
Model	$\boldsymbol{M_\$}$	**MME**	**MAPE**	**RMSE**	**MME**	**MAPE**	**RMSE**	**MME**	**MAPE**	**RMSE**
	ITM	-0.073	0.081	9.199	-0.119	0.128	18.393	-0.199	0.211	39.511
GJR	ATM	-0.991	0.993	19.222	-0.566	0.566	29.311	-0.469	0.482	47.979
	OTM	-6.086	6.086	18.414	-4.554	4.554	29.577	-1.557	1.558	51.930
	ITM	0.013	0.024	4.170	0.027	0.039	8.768	0.054	0.070	20.847
Implied GJR	ATM	-0.216	0.358	7.929	0.140	0.285	16.254	0.356	0.391	37.489
	OTM	-1.629	1.777	7.513	-0.719	1.216	11.774	0.185	0.757	19.199
	ITM	0.021	0.034	6.143	0.032	0.052	12.310	0.049	0.079	27.773
FHS GJR	ATM	0.057	0.266	8.992	0.139	0.295	16.579	0.233	0.320	31.216
	OTM	-0.648	1.108	6.875	-0.570	1.176	11.134	-0.194	0.911	19.130
	ITM	0.007	0.031	4.247	0.020	0.046	7.215	-0.042	0.092	20.300
MS GJR	ATM	-0.343	0.392	6.761	-0.069	0.141	8.594	-0.078	0.166	21.501
	OTM	-1.625	1.679	5.553	-1.075	1.128	9.389	-0.431	0.513	24.374
	ITM	-0.005	0.020	3.336	-0.004	0.020	5.339	-0.013	0.024	11.965
HN	ATM	-0.489	0.508	9.380	-0.231	0.244	11.575	-0.143	0.151	13.342
	OTM	-2.866	2.879	8.880	-2.323	2.327	12.391	-1.770	1.772	16.336
	ITM	0.011	0.018	3.506	-0.003	0.025	9.163	-0.031	0.041	25.775
AHBS	ATM	0.115	0.136	4.648	0.113	0.119	6.826	0.066	0.073	7.279
	OTM	0.394	0.448	2.660	0.497	0.511	4.746	0.294	0.362	6.417

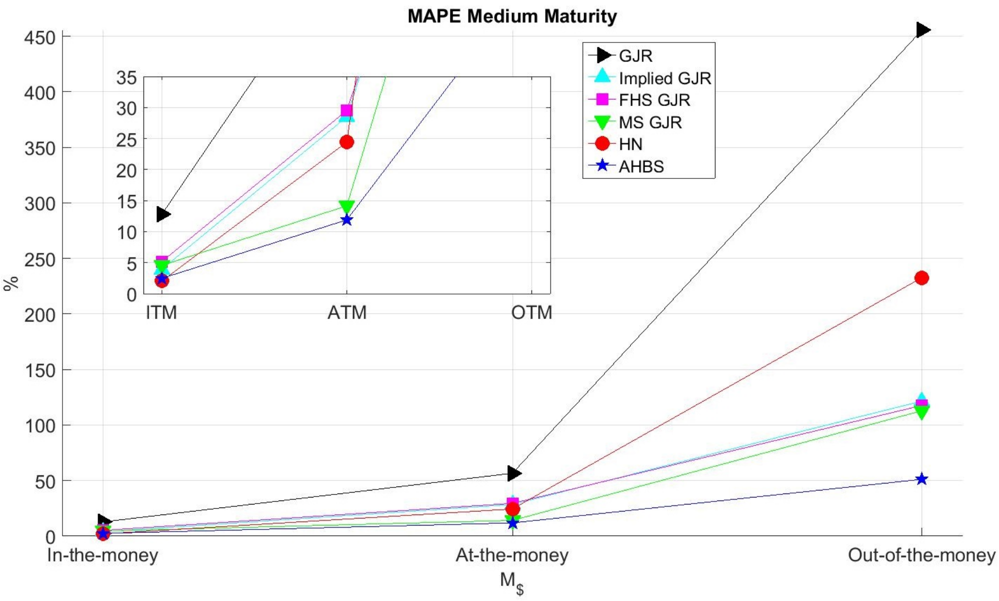
Table 14: Predictive option pricing results for the aggregate categorized sample over 2014. This table displays the Mean Mispricing Errors (in %), the Mean Absolute Percentage Errors (in %) and the Root Mean Squared Errors (in U.S. $ currency). The results are categorized based on moneyness (in-the-money (ITM), at-the-money (ATM), out-of-the-money (OTM)) and days to expiration ($\leq$ 60 days, (60,160] days, $>$ 160 days).

Overall, the MS GJR shows a substantial improvement over the single regime GJR. It outperforms the GJR across all maturities and moneyness categories and thus seems to remove a large portion of the bias of the GJR model, which was also suggested by the box and whiskers plot. The option implied models show a similar improvement over the GJR model. Both the FHS and Implied GJR outperform the GJR for all categories, especially at- and out-of-the-money options are priced more accurately by the FHS and Implied GJR. This performance tends to increase as time to expiration increases, which is also the case for the MS GJR model. When considering only the test models the GJR clearly performs worst while the MS GJR model competes quite well with the calibrated models. Based on accuracy the FHS and Implied GJR perform similar to the MS GJR for short maturities but are caught up by the MS GJR as time to maturity increases, while in terms of volatility and bias the MS GJR scores better in almost all categories. The plots in figure (5) and (6) provide a visual interpretation of the absolute mispricing across moneyness buckets, categorized by time to expiration.

None of the test models is able to systematically outperform the benchmark models. Considering the pricing accuracy, the MS GJR outperforms the HN model for at- and out-of-the-money options, and performs similar for in-the-money options compared to both the HN and ad-hoc BS. The HN shows a low pricing accuracy for out-of-the-money options and is outperformed by all models in this category, except for the GJR. The best performing model is the ad-hoc BS. This may seem surprising as this model is theoretically inconsistent, whereas the GARCH models are well founded from a theoretical perspective. However, as noted by Heston and Nandi (2000), the ad-hoc BS is designed such that it fits the term structure of implied volatilities as well as the volatility smile over strike prices. This gives the ad-hoc BS more flexibility compared to the GARCH models. Hence, the ad-hoc BS is a challenging benchmark which is hard to beat, as was also postulated by Dumas et al. (1998), who note that its performance is as good as that of sophisticated models that allow the volatility to be a deterministic function of asset price and time. Indeed, the ad-hoc BS outperforms the theoretical consistent GARCH models, even when these models are inferred from option data or include a regime switching property. These findings are consistent with the work of Duan et al. (2002), who find that the ad-hoc BS performs equally well, or even better, as their Markov switching GARCH model. Nonetheless, the MS GJR still performs quite well even though the normality assumption is maintained and regime shift risk is assumed to be absent. Its outperformance of the GJR model is readily observed, but the performance of the MS GJR is generally also better than that of the option implied models. This suggests that calibrated models are relatively less capable of inferring market dynamics from option data as opposed to a regime switching model estimated with historical returns. A big difference between regime switching models and calibrated models is that the latter cannot account for volatility jumps. Also, the option implied models incorporate mainly information from the previous time-step, which means they mostly reflect the market momentum. This possibly leads to a lower accuracy when simulating the volatility evolution for medium and long maturities. The regime switching model does not suffer from this momentum bias as it incorporates information from a long sample period. This could potentially contribute to a more accurate simulation of the underlying asset price over longer maturity horizons, which leads to an increased pricing performance of medium and long maturity options.



(a)



(b)

Figure 5: Performance over moneyness categories. This figure shows the Mean Absolute Percentage Errors across the moneyness spectrum for short maturity options (figure a) and medium maturity options (figure b).

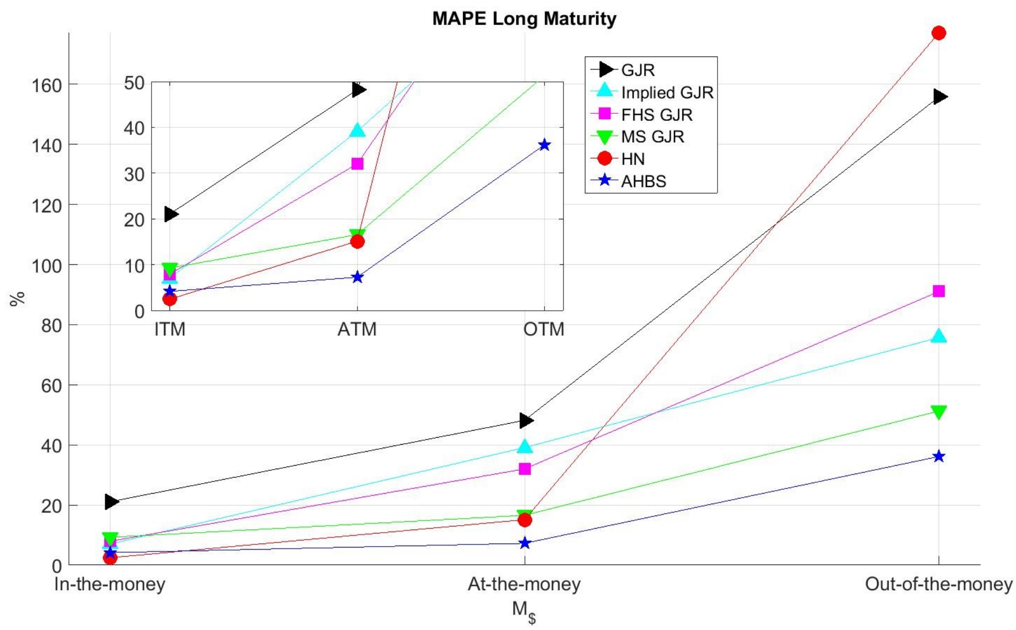
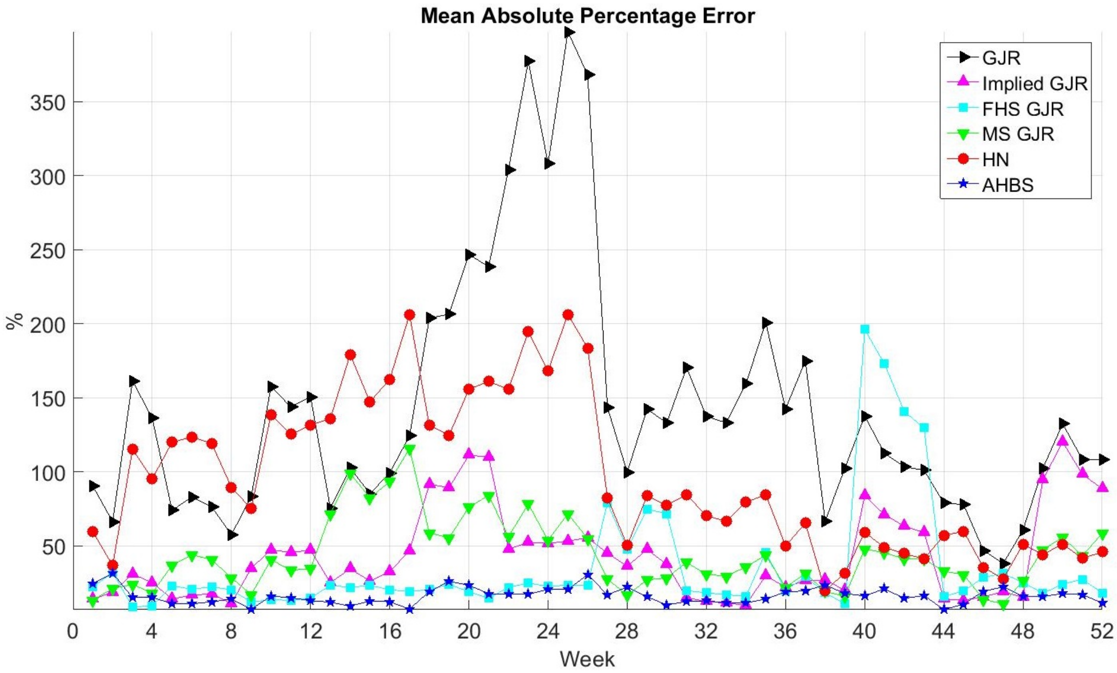


Figure 6: Performance over moneyness categories. This figure shows the Mean Absolute Percentage Errors across the moneyness spectrum for long maturity options.

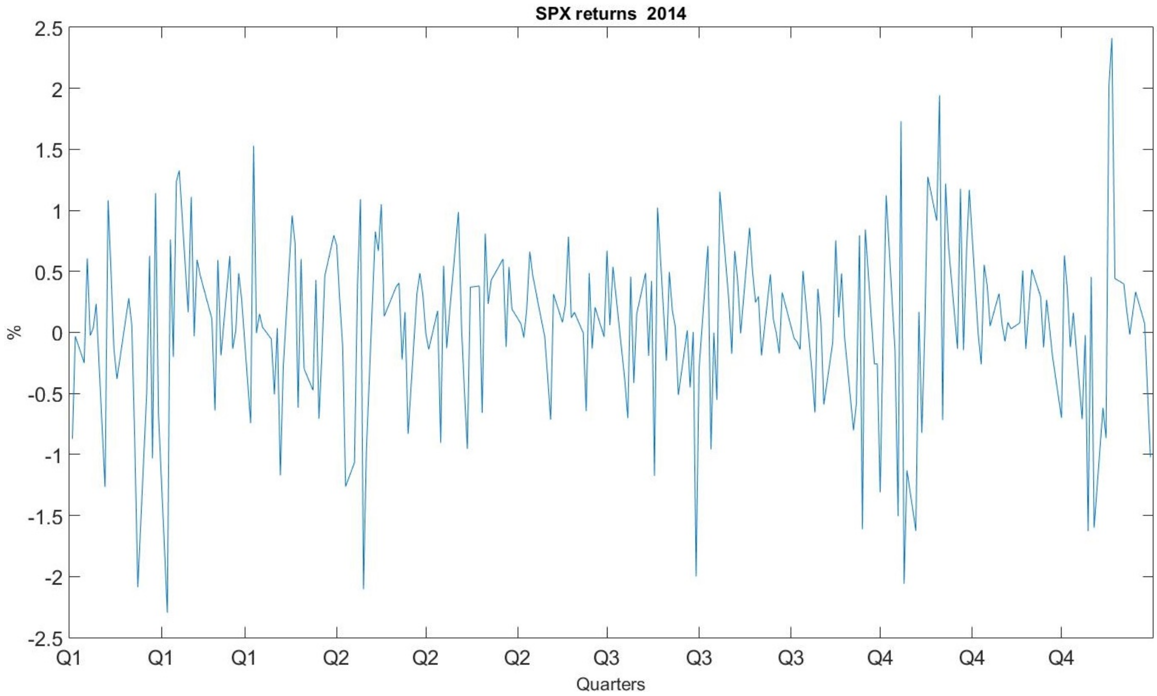
5.4.3 Performance over Time

The pricing performance over time is considered by inspecting the MAPEs of the aggregate weekly option samples over time, plotted in figure (7a). The figure does not differentiate between different maturity and moneyness categories and therefore some caution needs to be taken when interpreting the plot. Nonetheless, the plot can indicate how pricing performance generally behaves over time. Furthermore, the plot can indicate whether there are any notable differences between the different forecast horizons within each 4 week subsample that corresponds to a specific estimated model.

It is desirable that pricing performance remains consistent over time because this implies the models sufficiently capture the changing market conditions. The plot shows that only the ad-hoc BS and the FHS GJR perform consistent throughout the given time-span. The MS GJR and the Implied GJR perform less stable and appear to have a sudden jump in the percentage errors in the second quarter of 2014, which then recovers halfway through 2014. The performance then remains more consistent for the remainder of 2014. The HN and GJR appear to perform very inconsistent, especially in the first half of 2014 large jumps over 4 week periods are observed. In the figure a sharktooth-pattern can be recognized for the HN and GJR models, which indicates



(a)



(b)

Figure 7: Figure (a) displays the aggregate performance over time. The plot shows the aggregate MAPE values across moneyness categories and maturities over the year 2014 and indicates the overall consistency of the pricing performance of the models throughout time. Figure (b) shows a plot of the S&P500 index percentage returns over 2014, with the x-axis indicating the corresponding quarter

relative large jumps from one sample to another. For the HN model this pattern diminishes in the second half of 2014, while the GJR is subject to this bias throughout the entire year. This inconsistent performance is likely caused by the severe mispricing of out-of-the-money options by both the HN and GJR.

When considering the performance of the models within each 4 week sub-sample that corresponds to a specific estimated model, there appears not to be any systematic patterns for the Implied GJR, FHS GJR, MS GJR and ad-hoc BS. Although there is some discrepancy in the results of the MS GJR and the Implied GJR in the second quarter of 2014, this does not appear to be a repetitive trend. Thus, for the aforementioned models the MAPE values do not appear to change much in between the 4 week periods. This is an indication that there is no systematic bias related to the increasing forecast horizon within the 4 week subsamples for these models. Contrary to this is the performance of the HN and GJR models, which display the readily observed sharktooth-pattern. This indicates the models perform inconsistent within some of the 4 week subsamples and possibly suffer from a bias related to the expanding forecast horizon.

The plot in figure (7a) suggests there are some differences in the performance over the year. For example, the absolute mispricing in the second quarter is of larger magnitude than in the remainder of the year. Especially after week 28 the HN, GJR, Implied GJR and MS GJR models perform substantially better. Therefore, as a final analysis the quarterly results are considered. Tables (15) and (16) present the quarterly results, while figures (8) to (13) show the mean absolute percentage errors across moneyness buckets for the different maturities for each quarter. The quarterly results indeed reveal some differences. From quarter 1 to quarter 2 an increase in the bias and the volatility of the pricing errors is observed for the models that are estimated using historical asset returns (GJR, MS GJR, HN). Surprisingly, the option implied GARCH models and the ad-hoc BS appear to perform relatively more constant from quarter 1 to quarter 2. This is unexpected because the parameters in these models were less stable over time compared to the GJR, MS GJR and HN models. Possibly the models estimated with the historical returns do not sufficiently capture market events around this point in time. The plot in figure (7b) indeed shows some large negative returns at the beginning of the second quarter. In the second quarter the magnitude of the pricing errors seems to be substantially larger for at- and out-of-the-money options than for in-the-money-options. However, as maturity increases the differences between the first and second quarter appear to diminish. From quarter 2 to quarter 3 the results appear to improve. For short maturity options the performance increases in quarter

3 in both terms of bias and accuracy, for the medium maturity options this improvement is less pronounced and for long maturity options pricing performance is slightly less good as in the second quarter. The option implied models again perform relatively more constant compared to the other models. In quarter 4 the FHS GJR shows a large decrease in pricing accuracy which is caused by the severe mispricing of out-of-the-money options between weeks 40 and 44. The other models also appear to misprice out-of-the-money options during this period, although to a lesser extent, while short and medium maturity in- and at-the-money options are priced similarly in quarter 3 and quarter 4. The mispricing of out-of-the-money options is possibly caused by the increased volatile behaviour of the S&P500 at the start of the fourth quartile. This is also reflected by the fact that out-of-the-money options are mispriced to a larger extent for short maturities than for long maturities in the fourth quarter. This pattern makes sense because long maturity options have more time to end up in-the-money and therefore are more traded, which reduces inconsistencies in the market price.

The models show some inconsistencies over time. For example, the option implied models outperform the MS GJR for short maturities in the second and third quarter, while in the first and fourth quarter the MS GJR outperforms the option implied model for this category. Nevertheless, the quarterly results show some support the main findings implied by the aggregate results. In all quarters it holds that the GJR performs bad and is substantially outperformed by the option implied models and the MS GJR. Throughout the different quarters the mutual performance of the option implied models is similar, while the performance relative to the MS GJR is also comparable. Although there is no clear winner among the test models, the MS GJR appears to perform slightly better across moneyness categories and maturities. Finally, the ad-hoc BS dominates all samples and is only occasionally outperformed.

Results Aggregated per quarter: Q1 - Q2

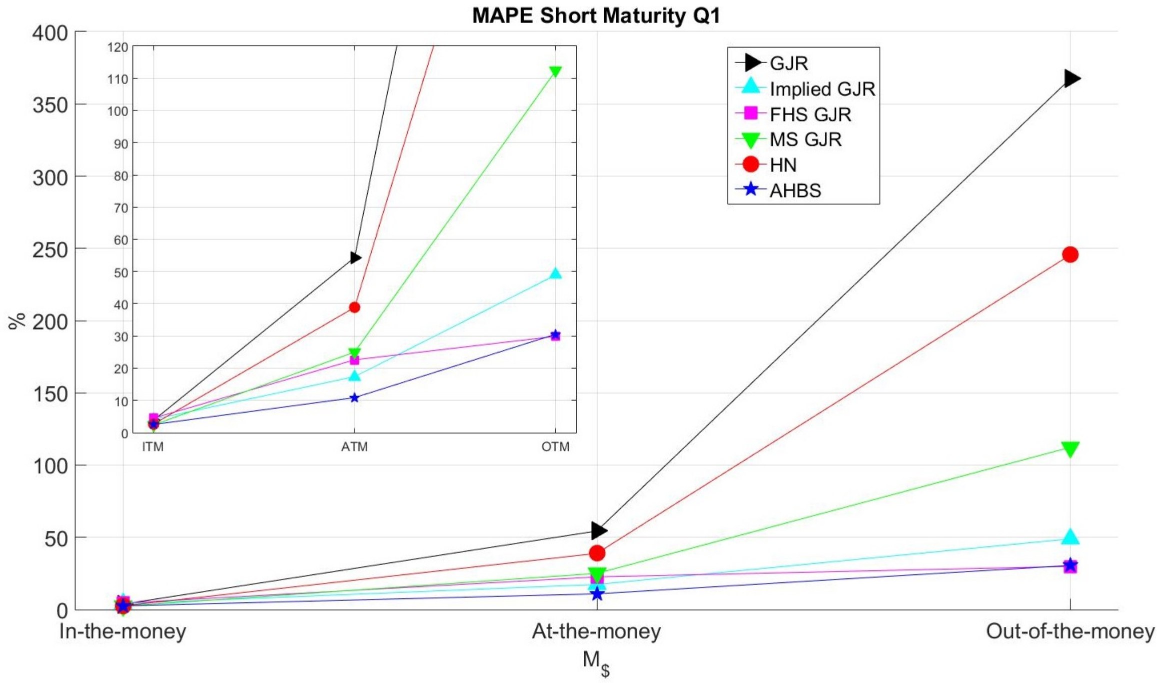
		Days to Expiration								
		$\leq$ **60**			**(60,160]**			$>$ **160**		
Model	$M_{\$}$	**MME**	**MAPE**	**RMSE**	**MME**	**MAPE**	**RMSE**	**MME**	**MAPE**	**RMSE**
Quarter 1										
	ITM	-0.031	0.038	3.790	-0.029	0.046	5.847	-0.081	0.124	20.306
GJR	ATM	-0.544	0.544	10.518	-0.272	0.272	12.764	-0.254	0.278	27.973
	OTM	-3.676	3.676	10.522	-2.471	2.471	14.624	-0.874	0.876	33.043
	ITM	0.040	0.041	4.597	0.096	0.096	12.180	0.139	0.139	26.985
Implied GJR	ATM	0.011	0.174	4.664	0.292	0.293	15.517	0.424	0.424	41.128
	OTM	-0.314	0.490	2.053	0.302	0.438	6.375	0.447	0.492	18.797
	ITM	0.045	0.045	4.907	0.085	0.085	10.916	0.120	0.120	22.698
FHS GJR	ATM	0.220	0.226	6.332	0.216	0.216	11.093	0.269	0.269	24.072
	OTM	0.119	0.299	2.336	0.105	0.255	4.372	0.227	0.284	10.934
	ITM	0.007	0.024	2.811	0.046	0.046	6.117	0.100	0.100	15.733
MS GJR	ATM	-0.199	0.250	4.632	0.037	0.073	4.811	0.137	0.137	13.376
	OTM	-1.025	1.124	3.556	-0.385	0.490	3.584	0.082	0.141	8.296
	ITM	-0.014	0.026	2.683	-0.017	0.030	4.030	-0.012	0.025	5.193
HN	ATM	-0.381	0.389	7.874	-0.251	0.254	11.202	-0.139	0.145	12.646
	OTM	-2.456	2.456	7.666	-2.361	2.361	12.250	-1.831	1.832	16.401
	ITM	0.024	0.025	2.826	0.015	0.025	3.855	-0.058	0.068	23.223
AHBS	ATM	0.065	0.109	3.416	0.083	0.086	5.165	0.057	0.060	5.949
	OTM	0.120	0.306	1.604	0.181	0.280	3.131	0.175	0.312	6.507
Quarter 2										
	ITM	-0.134	0.134	12.215	-0.131	0.144	19.044	-0.210	0.249	42.227
GJR	ATM	-1.451	1.451	21.099	-0.674	0.675	30.312	-0.560	0.585	51.932
	OTM	-7.792	7.792	19.269	-5.325	5.325	27.840	-2.088	2.092	52.205
	ITM	0.013	0.046	4.084	0.095	0.095	11.580	0.156	0.156	27.558
Implied GJR	ATM	-0.356	0.515	7.601	0.339	0.347	17.686	0.534	0.534	43.976
	OTM	-1.688	1.935	5.507	0.169	0.655	7.073	0.645	0.689	20.939
	ITM	0.042	0.043	4.141	0.095	0.095	11.259	0.139	0.139	24.041
FHS GJR	ATM	0.086	0.141	3.791	0.234	0.234	10.972	0.323	0.323	25.620
	OTM	-0.052	0.297	1.313	0.166	0.253	4.022	0.457	0.457	12.431
	ITM	-0.059	0.059	4.438	-0.018	0.032	4.045	0.016	0.036	5.943
MS GJR	ATM	-0.718	0.718	8.987	-0.165	0.170	7.287	-0.011	0.065	6.320
	OTM	-3.239	3.239	7.788	-1.580	1.580	8.387	-0.276	0.285	9.279
	ITM	-0.066	0.066	5.220	-0.064	0.064	7.105	-0.048	0.049	8.467
HN	ATM	-0.856	0.856	12.060	-0.429	0.429	16.544	-0.243	0.243	18.112
	OTM	-4.872	4.872	11.810	-3.675	3.675	16.916	-2.425	2.425	20.617
	ITM	0.023	0.023	2.048	0.021	0.025	3.193	-0.007	0.037	9.546
AHBS	ATM	0.106	0.122	2.740	0.116	0.116	5.098	0.077	0.079	6.684
	OTM	0.328	0.362	1.313	0.417	0.417	3.067	0.271	0.278	4.355

Table 15: Predictive option pricing results for the aggregate categorized samples of the first and second quarter in 2014. This table displays the Mean Mispricing Errors (in %), the Mean Absolute Percentage Errors (in %) and the Root Mean Squared Errors (in U.S. $ currency). The results are categorized based on moneyness (in-the-money (ITM), at-the-money (ATM), out-of-the-money (OTM)) and days to expiration ($\leq$ 60 days, (60,160] days, $>$ 160 days).

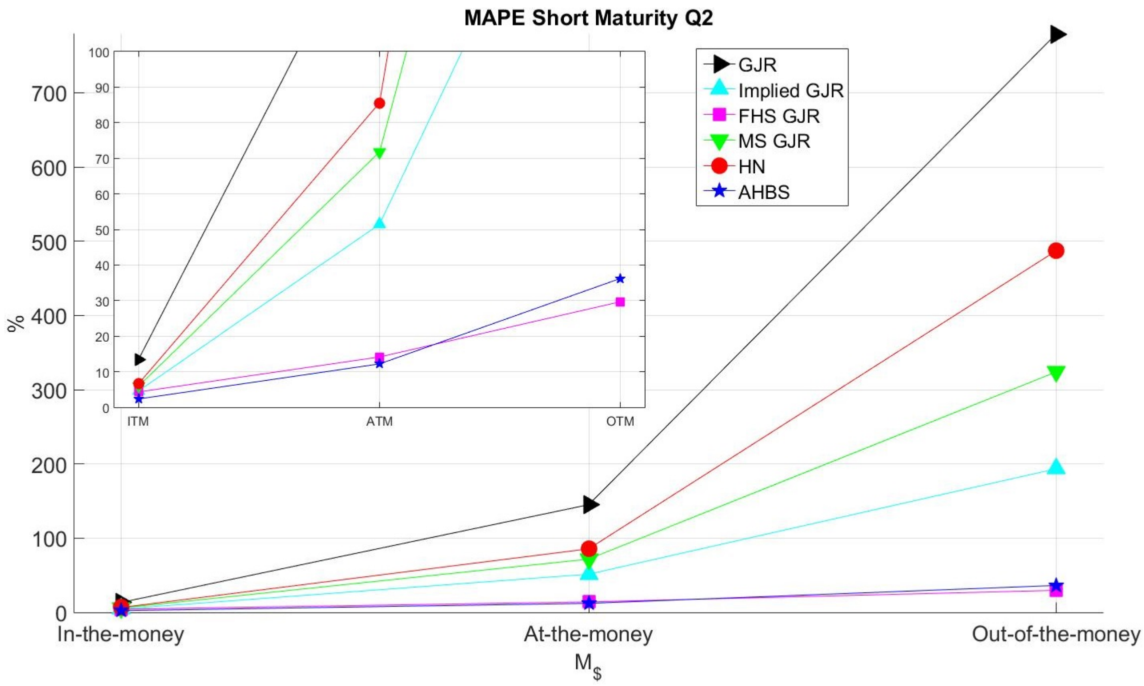
Results Aggregated per quarter: Q3 - Q4

		Days to Expiration								
		≤ 60			**(60,160]**			**> 160**		
Model	$M_{\$}$	**MME**	**MAPE**	**RMSE**	**MME**	**MAPE**	**RMSE**	**MME**	**MAPE**	**RMSE**
Quarter 3										
	ITM	-0.088	0.096	10.559	-0.187	0.188	24.952	-0.282	0.282	49.858
GJR	ATM	-1.134	1.143	22.526	-0.789	0.789	39.090	-0.653	0.653	62.352
	OTM	-7.721	7.721	21.852	-6.609	6.609	39.383	-2.083	2.083	67.486
	ITM	0.014	0.019	4.090	0.029	0.037	9.101	0.036	0.058	20.735
Implied GJR	ATM	-0.175	0.307	7.011	0.238	0.243	15.289	0.395	0.395	38.800
	OTM	-1.355	1.467	5.004	0.131	0.453	6.803	0.620	0.622	19.922
	ITM	0.014	0.017	3.618	0.020	0.024	6.429	0.023	0.032	12.368
FHS GJR	ATM	0.020	0.192	6.402	0.073	0.196	10.972	0.124	0.195	20.666
	OTM	-0.608	0.956	4.496	-0.712	1.107	7.432	-0.789	1.183	12.810
	ITM	0.014	0.031	4.910	0.052	0.054	9.038	-0.022	0.056	10.731
MS GJR	ATM	-0.253	0.339	6.779	0.027	0.123	8.250	-0.109	0.148	14.800
	OTM	-1.479	1.577	5.072	-0.667	0.789	5.879	-0.502	0.533	18.692
	ITM	0.000	0.015	3.295	-0.002	0.018	5.681	-0.011	0.022	13.629
HN	ATM	-0.411	0.457	9.032	-0.197	0.231	11.100	-0.093	0.119	10.767
	OTM	-2.839	2.890	8.662	-2.385	2.396	11.784	-1.668	1.674	13.456
	ITM	0.007	0.014	3.087	-0.011	0.024	11.434	-0.043	0.050	36.588
AHBS	ATM	0.158	0.159	4.943	0.140	0.140	6.959	0.095	0.095	9.227
	OTM	0.511	0.511	3.160	0.576	0.576	5.344	0.320	0.443	7.449
Quarter 4										
	ITM	-0.059	0.070	8.261	-0.071	0.087	11.751	-0.149	0.155	30.555
GJR	ATM	-0.778	0.778	19.282	-0.439	0.439	23.207	-0.377	0.378	41.043
	OTM	-5.695	5.695	19.564	-3.414	3.414	24.957	-1.194	1.194	48.403
	ITM	0.005	0.020	4.136	0.008	0.027	7.462	0.012	0.033	15.561
Implied GJR	ATM	-0.299	0.403	10.573	-0.092	0.287	16.570	0.056	0.192	21.390
	OTM	-2.735	2.836	11.427	-2.092	2.309	16.735	-0.903	1.199	17.074
	ITM	0.015	0.043	8.106	0.027	0.065	15.803	0.023	0.094	38.956
FHS GJR	ATM	-0.062	0.479	14.499	0.121	0.435	23.029	0.195	0.471	46.541
	OTM	-1.541	2.227	11.004	-0.981	1.911	16.016	-0.593	1.639	31.398
	ITM	0.013	0.029	3.914	-0.005	0.040	5.836	-0.125	0.144	29.616
MS GJR	ATM	-0.191	0.251	5.641	-0.141	0.168	10.347	-0.310	0.323	37.569
	OTM	-1.303	1.322	5.632	-1.458	1.459	12.822	-0.982	1.047	42.254
	ITM	0.005	0.015	3.049	0.004	0.016	4.905	-0.003	0.019	12.671
HN	ATM	-0.306	0.324	7.801	-0.158	0.163	8.914	-0.073	0.077	8.514
	OTM	-2.116	2.116	7.990	-1.756	1.756	10.809	-1.219	1.219	14.327
	ITM	0.008	0.018	4.155	-0.003	0.025	7.968	-0.018	0.027	15.277
AHBS	ATM	0.123	0.149	6.296	0.102	0.115	7.886	0.034	0.055	6.972
	OTM	0.537	0.550	3.314	0.580	0.580	5.229	0.401	0.404	6.800

Table 16: Predictive option pricing results for the aggregate categorized samples of the third and fourth quarter in 2014. This table displays the Mean Mispricing Errors (in %), the Mean Absolute Percentage Errors (in %) and the Root Mean Squared Errors (in U.S. $ currency). The results are categorized based on moneyness (in-the-money (ITM), at-the-money (ATM), out-of-the-money (OTM)) and days to expiration (≤ 60 days, (60,160] days, > 160 days).

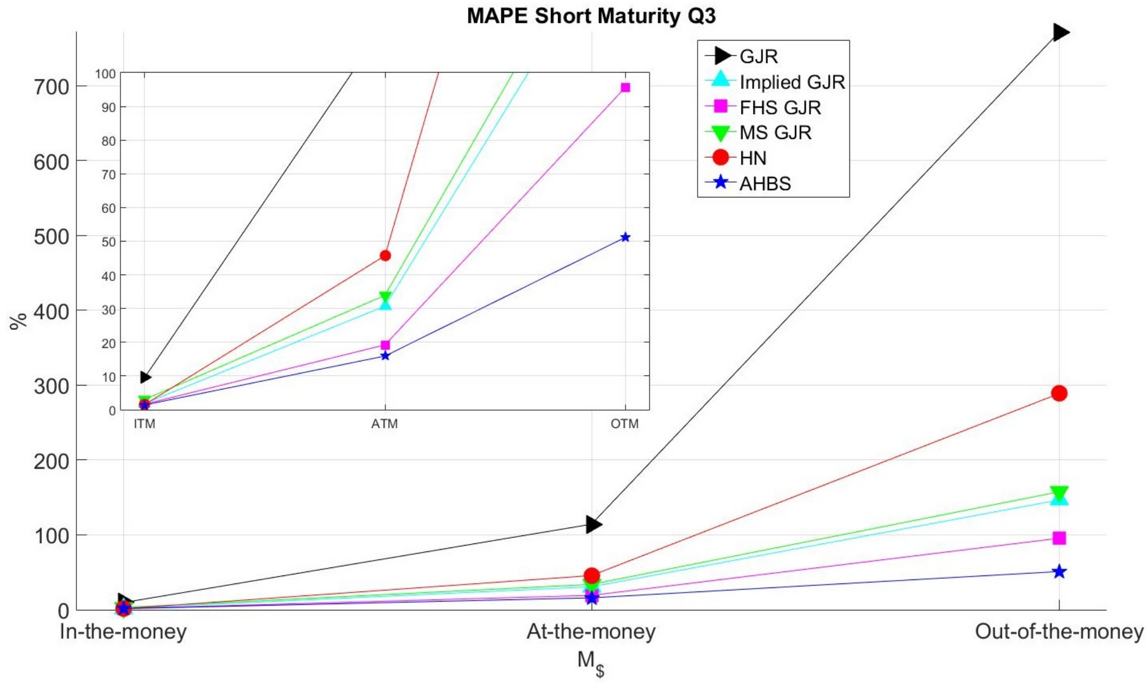


(a)

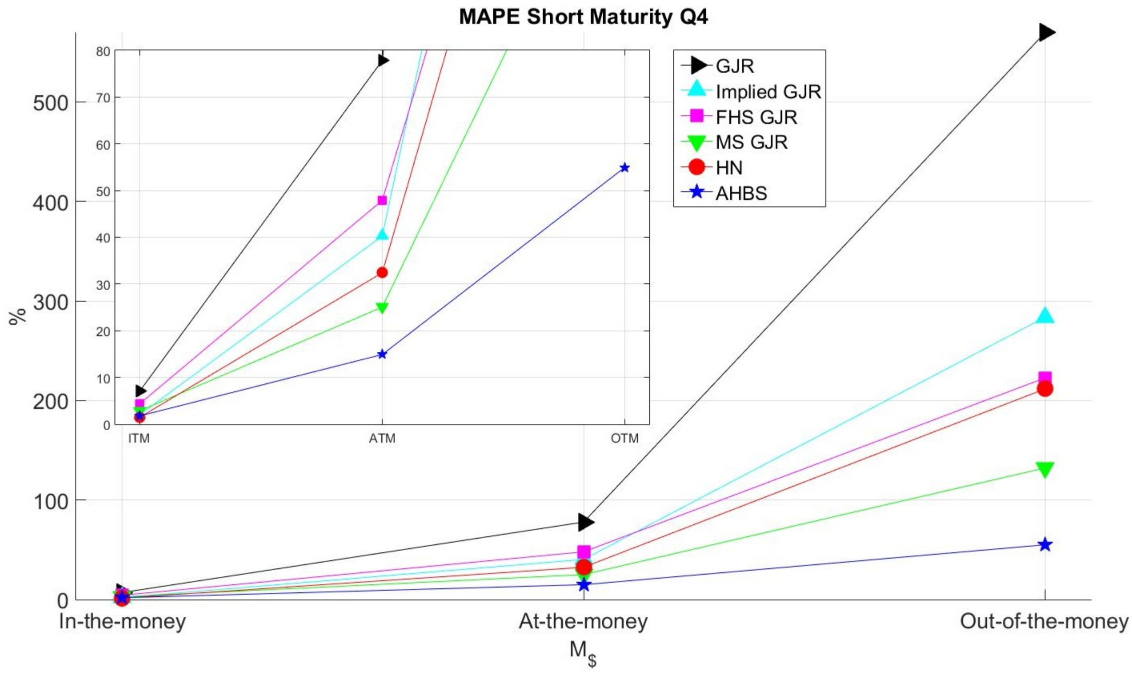


(b)

Figure 8: Mean Absolute Percentage Errors across moneyness categories for short maturities in the first quarter (figure a) and second quarter (figure b) of 2014.

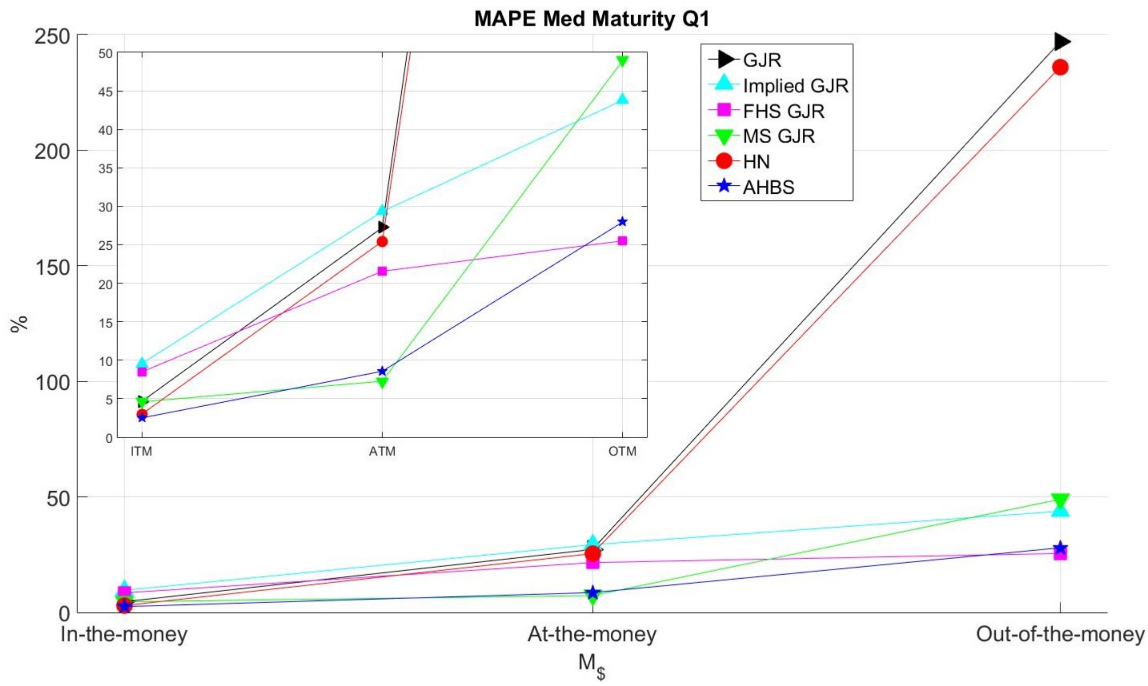


(a)

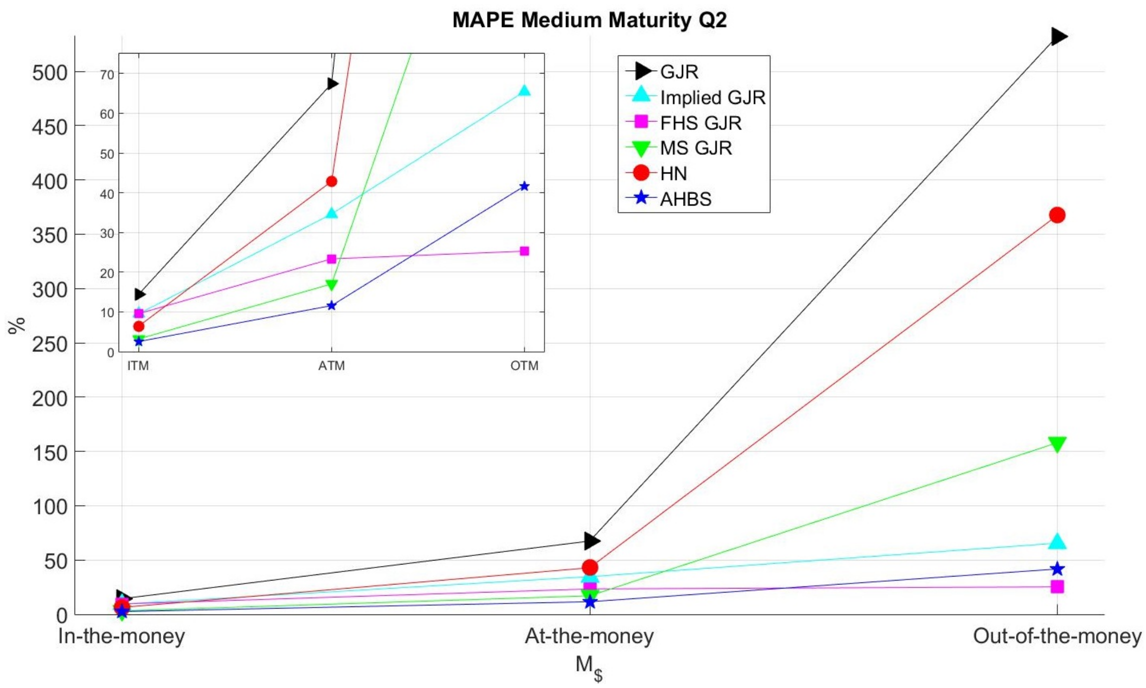


(b)

Figure 9: Mean Absolute Percentage Errors across moneyness categories for short maturities in the third quarter (figure a) and fourth quarter (figure b) of 2014.

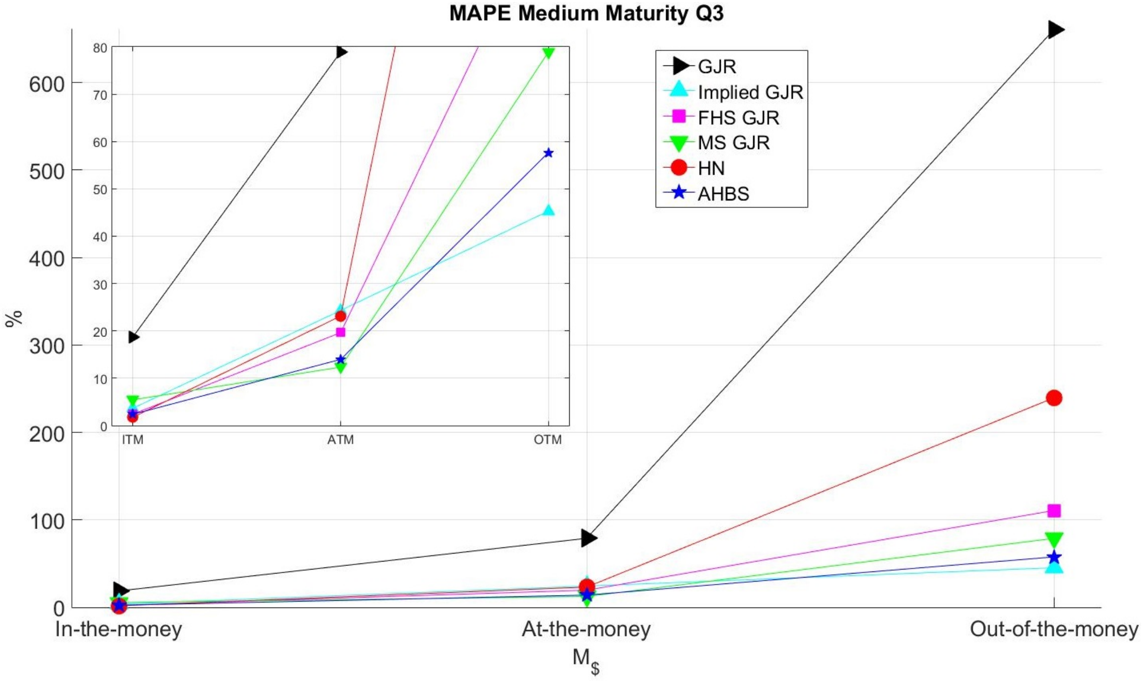


(a)

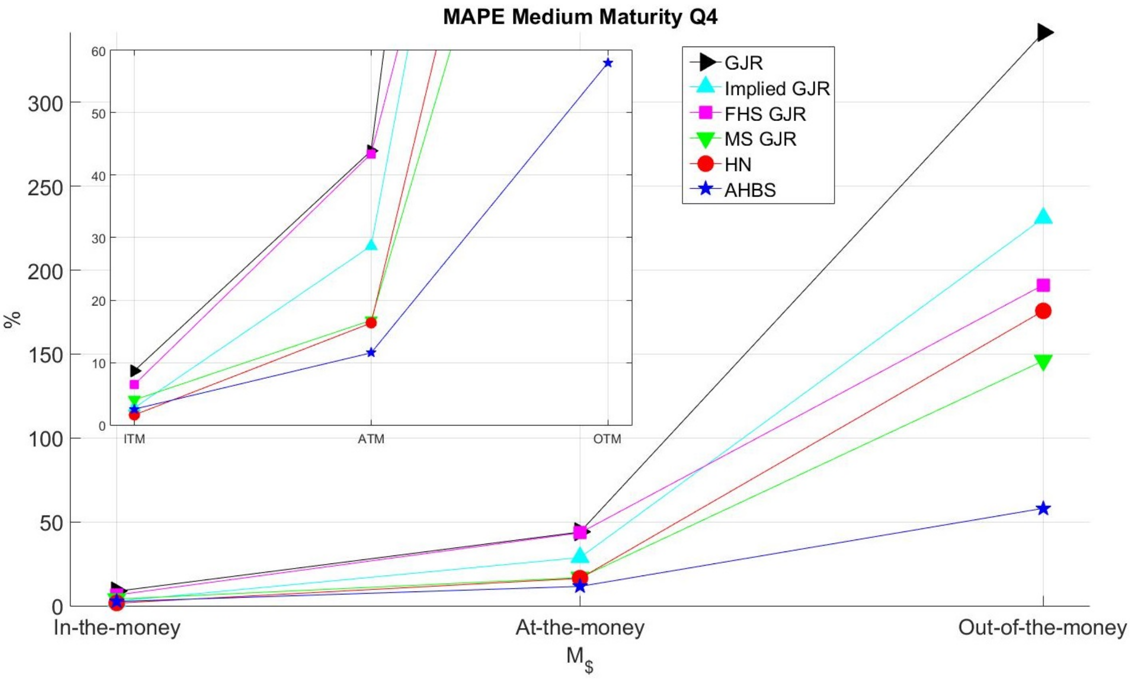


(b)

Figure 10: Mean Absolute Percentage Errors across moneyness categories for medium maturities in the first quarter (figure a) and second quarter (figure b) of 2014.

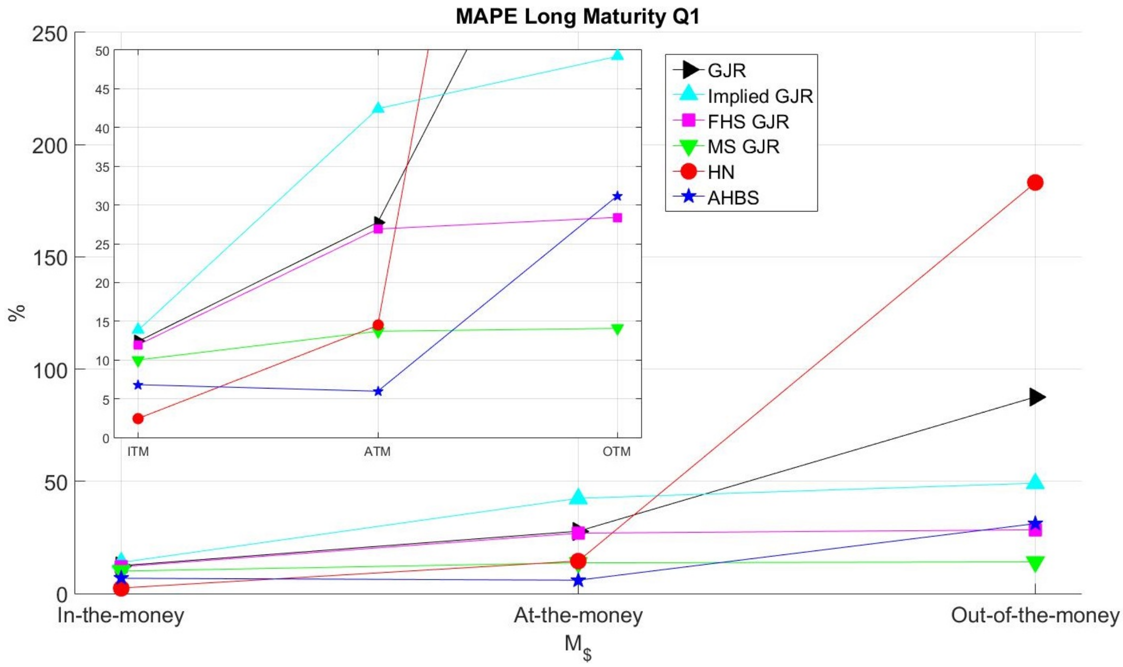


(a)



(b)

Figure 11: Mean Absolute Percentage Errors across moneyness categories for medium maturities in the third quarter (figure a) and fourth quarter (figure b) of 2014.



(a)

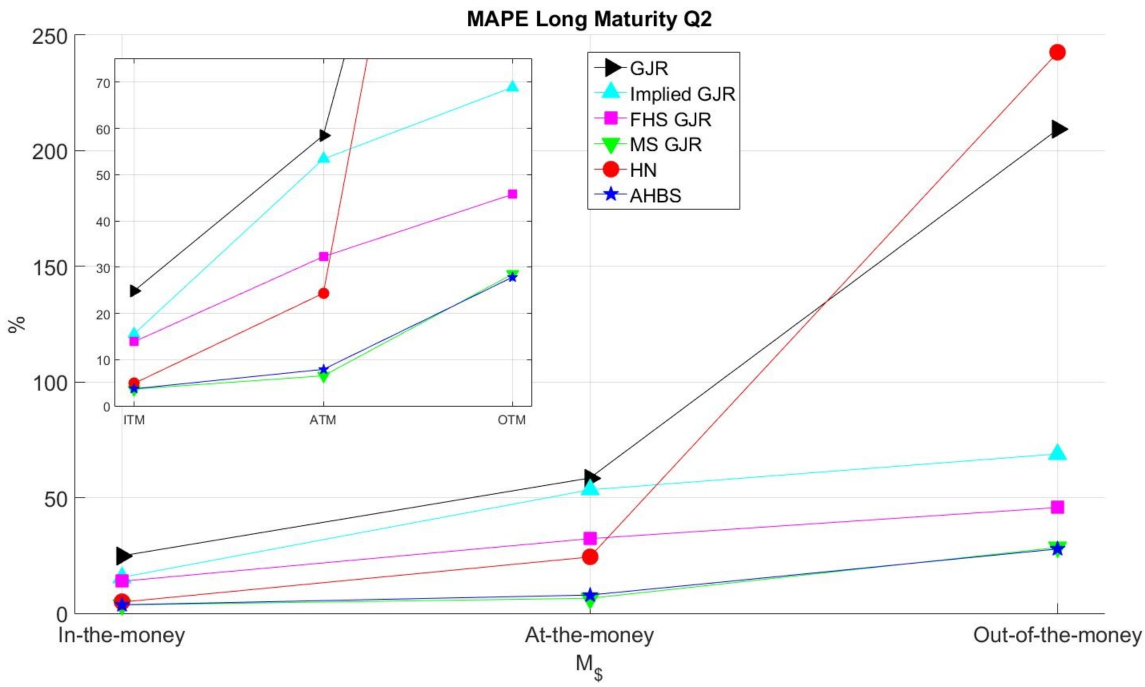
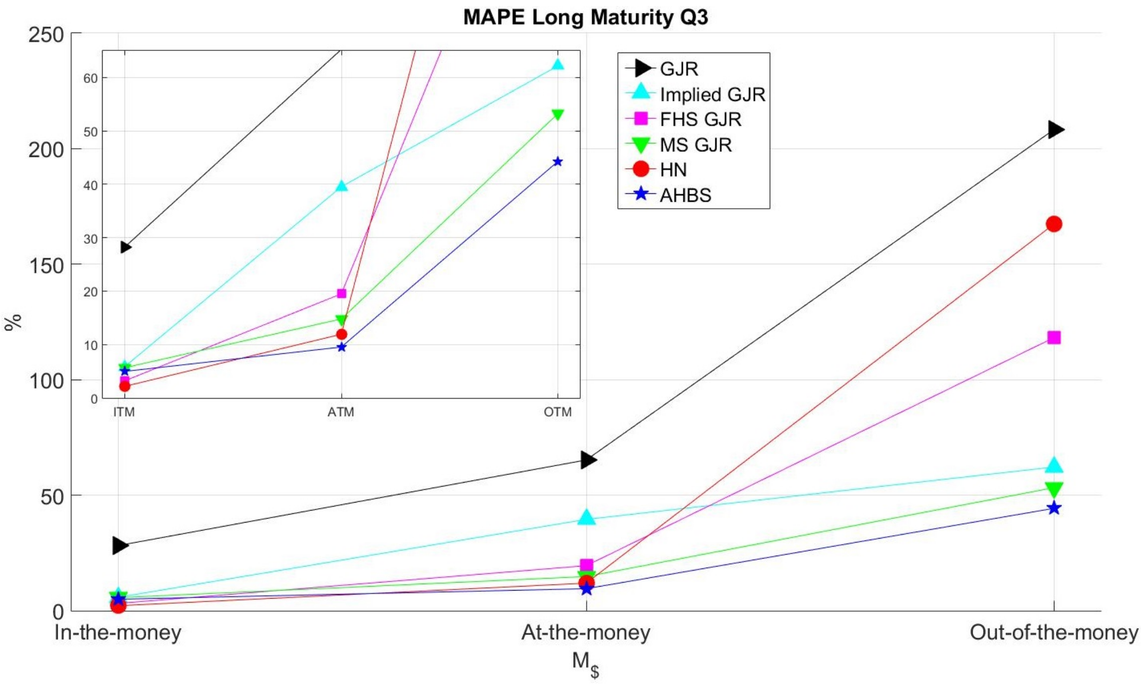
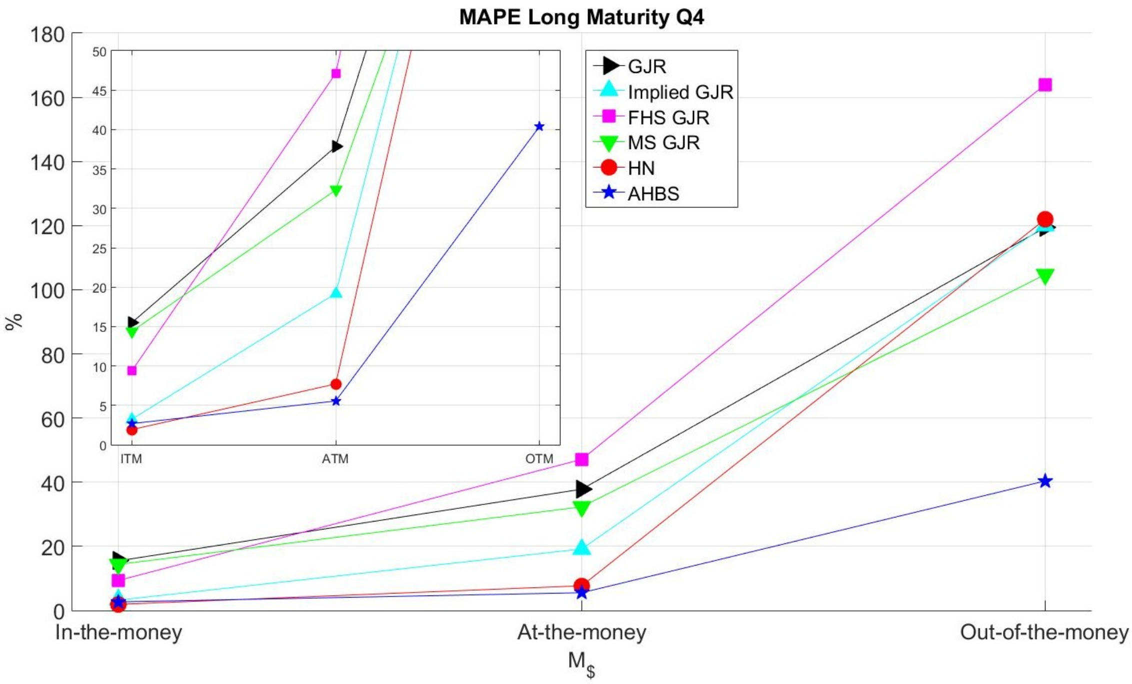


Figure 12: Mean Absolute Percentage Errors across moneyness categories for long maturities in the first quarter (figure a) and second quarter (figure b) of 2014.



(a)



(b)

Figure 13: Mean Absolute Percentage Errors across moneyness categories for long maturities in the third quarter (figure a) and fourth quarter (figure b) of 2014.

6 Conclusion and Limitations

This paper compares the option pricing performance of calibrated and non-calibrated GARCH models. The main objective is to investigate whether the Bayesian GARCH option pricing model, and in particular the Bayesian Markov switching GARCH, is able to outperform the option implied GARCH models in terms of forecasting option prices. This comparison results in a horse-race between a class of practical models and a class of theoretically founded models. Especially the MS GJR seems promising as it accounts for most of the stylized facts of asset returns. Furthermore, its application to derivative pricing is fairly new and has not been studied yet from a Bayesian perspective. To this end the Bauwens and Lubrano (2002) pricing algorithm is extended to the case of a regime switching GARCH. This results in a Bayesian algorithm that is generally compatible with all types of regime switching GARCH specifications with only minor alterations. It was argued that the Markov switching property allows the model to capture more market dynamics, such as volatility jumps and leptokurtosis. This statement seems to have some justification, as the inference results show the MS GJR parameters for the different regimes clearly resemble different levels of volatility. The pricing results show a satisfactory performance of the MS GJR and indicate it is superior over the single regime GJR, while it is also able to compete quite well with the option implied models. The MS GJR outperforms the Implied GJR and FHS GJR for most categories, although its performance seems not to be overwhelmingly better. The pricing performance of the MS GJR is most profound for in-the-money options, although at-the-money and out-of-the-money options are also priced accurately for longer maturities, compared to the benchmark models. The option implied models show a similar performance, although for long maturities their performance is less pronounced than that of the MS GJR. The test models are unable to beat the benchmark models in a systematic manner. Only for out-of-the-money options the HN is outperformed by the option implied models and the MS GJR, while for at-the-money options only the MS GJR performs slightly better than the HN model. None of the models is able to consistently beat the ad-hoc BS model. The MS GJR is able to outperform the ad-hoc BS across moneyness categories occasionally but generally performs less accurate. For long maturity options the performance becomes more similar. Overall it can be concluded that the regime switching GARCH model yields satisfactory pricing performance and competes quite well with the popular class of the option implied GARCH models. It therefore seems that GARCH models estimated with historical returns do not necessarily always perform less good compared to option implied GARCH models, as was argued by Weber and Prokupczuk (2011).

It should be noted however, that such comparison is slightly unfair when one of the models has a regime switching feauture. Nevertheless, this result is promising because the Bayesian Markov switching GARCH option pricing model introduced in this paper has room for improvement.

Some final remarks can be made considering limitations and suggestions for future research. To accommodate the Duan (1995) risk-neutralization principle, the returns are assumed to be normally distributed. However, from empirical research it is known that returns exhibit fat tails and therefore the student's t-distribution would be more appropriate. This would require a different risk neutral measure, of which two examples are the extended Girsanov principle, introduced by Elliot and Madan (1998), and the Escher transform method, pioneered by Bühlmann et al. (1996). For the risk neutralization of the MS GJR the regime shift risk was ignored for simplicity, which means a possible bias in the pricing results of the regime switching model may be present. This issue can be overcome by deriving the risk neutral transition probabilities, which can be done by, for example, inferring them from observed option data as suggested by Duan et al. (2002). However, applying this in a Bayesian context is not so straightforward and requires more in depth research. Furthermore, it would be interesting to inspect whether Bayesian methods can be employed in the calibration process. This could decrease parameter uncertainty and would enhance the comparability with the other models. Martin et al. (2003) introduce a method to perform implicit Bayesian inference from option prices which possibly could be used in the calibration process. Finally, it would be interesting to see whether it is possible to estimate the MS GARCH model from option data. This would yield a model that has the flexibility of the regime switching model and at the same time captures the risk-neutral structure directly from option prices.

References

1. Ang, A., Timmermann, A. (2011). Regime changes and financial markets (No. w17182). National Bureau of Economic Research.

2. Avellaneda, Marco, and Peter Laurence. Quantitative modeling of derivative securities: from theory to practice. CRC Press, 1999.

3. Bakshi, Gurdip, Charles Cao, and Zhiwu Chen. "Empirical performance of alternative option pricing models." The Journal of Finance 52.5 (1997): 2003-2049.

4. Bates, David S. "Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options." Review of financial studies 9.1 (1996): 69-107.

5. Barone-Adesi, Giovanni, Robert F. Engle, and Loriano Mancini. "A GARCH option pricing model with filtered historical simulation." Review of Financial Studies (2008).

6. Bauwens, Luc, Arie Preminger, and Jeroen VK Rombouts. "Theory and inference for a Markov switching GARCH model." The Econometrics Journal 13.2 (2010): 218-244.

7. Bauwens, Luc, and Michel Lubrano. "Bayesian option pricing using asymmetric GARCH models." Journal of Empirical Finance 9.3 (2002): 321-342.

8. Bollen, N. P. (1998). Valuing options in regime-switching models. The Journal of Derivatives, 6(1), 38-49.

9. Bollerslev, Tim. "Generalized autoregressive conditional heteroskedasticity." Journal of econometrics 31.3 (1986): 307-327.

10. Byun, Suk Joon, and Byungsun Min. "Conditional Volatility and the GARCH Option Pricing Model with NonNormal Innovations." Journal of Futures Markets 33.1 (2013): 1-28.

11. Casarin, Roberto. Bayesian Inference for Markov Switching Stochastic Volatility Models. working paper, CEREMADE, forthcoming, 2003.

12. Celeux, G., Hurn, M., Robert, C. P. (2000). Computational and inferential difficulties with mixture posterior distributions. Journal of the American Statistical Association, 95(451), 957-970.

13. Chen, ChaoChun, and MingYang Hung. "Option pricing under Markovswitching GARCH processes." Journal of Futures Markets 30.5 (2010): 444-464.

14. Christoffersen, Peter, Steve Heston, and Kris Jacobs. "Option valuation with conditional skewness." Journal of Econometrics 131.1 (2006): 253-284.

15. Darsinos, T. Satchell, SE, 2001. "Bayesian Analysis of the Black-Scholes Option Price," Cambridge Working Papers in Economics 0102, Faculty of Economics, University of Cambridge.

16. Duan, Jin-Chuan. "The GARCH option pricing model." Mathematical finance 5.1 (1995): 13-32.

17. Duan, Jin-Chuan, and Hua Zhang. "Pricing Hang Seng Index options around the Asian financial crisis–A GARCH approach." Journal of Banking Finance 25.11 (2001): 1989-2014.

18. Duan, Jin-Chuan, and Jean-Guy Simonato. "Empirical martingale simulation for asset prices." Management Science 44.9 (1998): 1218-1233.

19. Duan, J., Gauthier, G., Simonato, J., Sasseville, C. (2006). Approximating the GJR-GARCH and EGARCH option pricing models analytically. Journal of Computational Finance, 9(3), 41.

20. Dueker, Michael J. "Markov switching in GARCH processes and mean-reverting stock-market volatility." Journal of Business Economic Statistics 15.1 (1997): 26-34.

21. Dumas, Bernard, Jeff Fleming, and Robert E. Whaley. "Implied volatility functions: Empirical tests." The Journal of Finance 53.6 (1998): 2059-2106.

22. Elliott, Robert J., and Carlton-James U. Osakwe. "Option pricing for pure jump processes with Markov switching compensators." Finance and Stochastics 10.2 (2006): 250-275.

23. Engle, R. F., Ng, V. K. (1993). Measuring and testing the impact of news on volatility. The journal of finance, 48(5), 1749-1778.

24. Fergusson, K., Platen, E. (2006). On the Distributional Characterization of Daily LogReturns of a World Stock Index. Applied Mathematical Finance, 13(01), 19-38.

25. Francq, C., Zakoıan, J. M. (2001). Stationarity of multivariate Markov–switching ARMA models. Journal of Econometrics, 102(2), 339-364. Fuh, C. D., Ho, K. W. R., Hu, I., Wang, R. H. (2012). Option pricing with Markov switching. Journal of Data Science, 10, 483-509.

26. Gao, Y., Zhang, C., Zhang, L. (2012). Comparison of GARCH Models based on Different Distributions. Journal of Computers, 7(8), 1967-1973.

27. Geweke, J. (2007). Interpretation and inference in mixture models: Simple MCMC works. Computational Statistics Data Analysis, 51(7), 3529-3550.

28. Geweke, John. "Bayesian comparison of econometric models." Federal Reserve bank of Minneapolis working paper 532 (1994).

29. Geweke, John. "Bayesian inference in econometric models using Monte Carlo integration." Econometrica: Journal of the Econometric Society (1989): 1317-1339.

30. Gisiger, Nicolas. "Risk-neutral probabilities explained." (2010).

31. Glosten, Lawrence R., Ravi Jagannathan, and David E. Runkle. "On the relation between the expected value and the volatility of the nominal excess return on stocks." The journal of finance 48.5 (1993): 1779-1801.

32. Greenberg, Edward. Introduction to Bayesian econometrics. Cambridge University Press, 2013.

33. Hafner, Christian M., and Helmut Herwartz. "Time-varying market price of risk in the CAPM." (1999).

34. Hafner, Christian M., and Helmut Herwartz. "Option pricing under linear autoregressive dynamics, heteroskedasticity, and conditional leptokurtosis." Journal of Empirical Finance 8.1 (2001): 1-34.

35. Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica: Journal of the Econometric Society, 357-384.

36. Harvey, C. R., Whaley, R. E. (1992). Dividends and SP 100 index option valuation. Journal of Futures Markets, 12(2), 123-137.

37. Henneke, J., S. Rachev, and F. Fabozzi. MCMC based estimation of MS-ARMAGARCH models. Technical Reports, Department of Probability and Statistics, UCSB, 2006.

38. Henneke, J. S., Rachev, S. T., Fabozzi, F. J., Nikolov, M. (2011). MCMC-based estimation of Markov Switching ARMA–GARCH models. Applied Economics, 43(3), 259-271.

39. Heston, Steven L., and Saikat Nandi. "A closed-form GARCH option valuation model." Review of Financial Studies 13.3 (2000): 585-625. Hurn, M., Justel, A., Robert, C. P. (2003). Estimating mixtures of regressions. Journal of Computational and Graphical Statistics, 12(1), 55-79.

40. Jasra, A., Holmes, C. C., Stephens, D. A. (2005). Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling. Statistical Science, 50-67.

41. Jones, Christopher S. "The dynamics of stochastic volatility: evidence from underlying and options markets." Journal of econometrics 116.1 (2003): 181-224.

42. Kaehler, Jürgen, and Volker Marnet. Markov-switching models for exchange-rate dynamics and the pricing of foreign-currency options. Physica-Verlag HD, 1994.

43. Kole, E. Regime Switching Models: An Example for a Stock Market Index. Unpublished manuscript, Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam April 2010

44. Lee, Dong Wook, and Mark H. Liu. "Does more information in stock price lead to greater or smaller idiosyncratic return volatility?." Journal of Banking Finance 35.6 (2011): 1563-1580.

45. Martin, Gael M., Catherine S. Forbes, and Vance L. Martin. "Implicit Bayesian inference using option prices." Journal of Time Series Analysis 26.3 (2005): 437-462.

46. McLachlan, G., Peel, D. (2000). Finite Mixture Models: Wiley Series in Probability and Mathematical Statistics.

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

48. Rachev, Svetlozar T., et al. Bayesian methods in finance. Vol. 153. John Wiley Sons, 2008.

49. Rombouts, Jeroen VK, and Lars Stentoft. "Bayesian option pricing using mixed normal heteroskedasticity models." Computational Statistics Data Analysis 76 (2014): 588-605.

50. Satoyoshi, K., Mitsui, H. (2011). Empirical study of nikkei 225 options with the Markov switching Garch model. Asia-Pacific Financial Markets, 18(1), 55-68.

51. E. Sentana, Quadratic ARCH models. Review of Economic Studies 62, 639–661 (1995)

52. Sperrin, M., Jaki, T., Wit, E. (2010). Probabilistic relabelling strategies for the label switching problem in Bayesian mixture models. Statistics and Computing, 20(3), 357-366.

53. Stephens, M. (2000). Dealing with label switching in mixture models. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 62(4), 795-809.

54. Vargas Mendoza, A., 2011. "Bayesian Estimation and Option Mispricing". Massachusets Institute of Technology.

55. Weber, Michael, and Marcel Prokopczuk. "American option valuation: Implied calibration of GARCH pricing models." Journal of Futures Markets 31.10 (2011): 971-994.

56. Yao, David D., Qing Zhang, and Xun Yu Zhou. "A regime-switching model for European options." Stochastic processes, optimization, and control theory: applications in financial engineering, queueing networks, and manufacturing systems. Springer US, 2006. 281-300.

57. Zakoian, J. M. (1994). Threshold heteroskedastic models. Journal of Economic Dynamics and control, 18(5), 931-955.

7 Appendix I: Inference Results Benchmark Models

Inference results ad-hoc Black Scholes

	C	**K**	$\mathbf{K}^2$	τ	τ^2	$\mathbf{K} \times \tau$
Jan	**2.286**	**-2.28E-03**	**5.92E-07**	**1.21E-03**	**5.71E-07**	**-7E-07**
	0.348	*3.45E-04*	*8.51E-08*	*2.96E-04*	*3.08E-07*	*1.7E-07*
Feb	**1.309**	**-1.11E-03**	**2.44E-07**	**-2.25E-04**	**-3.4E-07**	**2.25E-07**
	0.140	*1.52E-04*	*4.12E-08*	*9.46E-05*	*4.57E-08*	*5.23E-08*
Mar	**2.467**	**-2.32E-03**	**5.64E-07**	**3.00E-04**	**-1.2E-07**	**-1E-07**
	0.676	*7.16E-04*	*1.89E-07*	*3.43E-04*	*1.2E-07*	*1.82E-07*
Apr	**1.395**	**-1.14E-03**	**2.39E-07**	**-4.11E-04**	**-4.4E-07**	**3.49E-07**
	0.372	*3.87E-04*	*1.01E-07*	*1.94E-04*	*5.74E-08*	*1.06E-07*
May	**2.103**	**-1.89E-03**	**4.33E-07**	**-5.39E-05**	**-7.7E-07**	**2.11E-07**
	0.204	*2.10E-04*	*5.39E-08*	*1.62E-04*	*6.96E-08*	*8.38E-08*
Jun	**0.565**	**-2.32E-04**	**-1.1E-08**	**-1.48E-03**	**-8.2E-07**	**9.39E-07**
	0.146	*1.47E-04*	*3.7E-08*	*1.39E-04*	*5.48E-08*	*7.27E-08*
Jul	**0.820**	**-5.12E-04**	**6.8E-08**	**-6.59E-04**	**-6.5E-07**	**5.06E-07**
	0.249	*2.41E-04*	*5.85E-08*	*1.37E-04*	*6.07E-08*	*7.03E-08*
Aug	**2.174**	**-1.87E-03**	**4.1E-07**	**-7.62E-05**	**-6.1E-07**	**2.08E-07**
	0.436	*4.30E-04*	*1.06E-07*	*2.09E-04*	*5.22E-08*	*1.04E-07*
Sep	**2.337**	**-1.76E-03**	**3.16E-07**	**-3.04E-03**	**-2.1E-07**	**1.62E-06**
	0.271	*2.47E-04*	*5.62E-08*	*2.67E-04*	*2.76E-08*	*1.33E-07*
Oct	**0.086**	**4.39E-04**	**-2.1E-07**	**-1.91E-03**	**-2.8E-07**	**1.06E-06**
	0.333	*3.42E-04*	*8.83E-08*	*1.67E-04*	*7.98E-08*	*8.26E-08*
Nov	**0.649**	**-3.50E-04**	**3.5E-08**	**-4.46E-04**	**-1.1E-06**	**4.59E-07**
	0.781	*7.43E-04*	*1.78E-07*	*6.78E-04*	*3.31E-07*	*3.54E-07*
Dec	**1.091**	**-3.60E-04**	**-5.2E-08**	**-3.40E-03**	**-3.4E-07**	**1.73E-06**
	0.385	*3.82E-04*	*9.62E-08*	*4.49E-04*	*2.15E-07*	*2.26E-07*

Table 17: Bayesian inference results of the ad-hoc Black Scholes model. Given are the posterior means (in bold) and the posterior standard deviations (in italic). In the Gibbs sampler 35000 iterations are run, of which the first 5000 are used as burn-in sample.

Inference results Heston - Nandi GARCH

	ω	β_{HN}	α_{HN}	λ_{HN}	γ_{HN}	**PM**
Jan	**6.3E-08**	**0.699**	**2.70E-06**	**3.06E-04**	**315.427**	**0.968**
	6.1E-08	*0.005*	*5.13E-08*	*2.41E-04*	*8.16E-05*	*0.001*
Feb	**4.8E-07**	**0.719**	**3.69E-06**	**2.10E-03**	**253.749**	**0.956**
	3.7E-07	*0.021*	*2.41E-07*	*9.14E-04*	*3.33E-05*	*0.005*
Mar	**1.2E-07**	**0.741**	**3.11E-06**	**6.40E-04**	**267.851**	**0.964**
	1.1E-07	*0.002*	*2.52E-08*	*9.12E-06*	*9.25E-06*	*0.001*
Apr	**1.8E-07**	**0.718**	**3.58E-06**	**4.37E-04**	**259.983**	**0.960**
	1.7E-07	*0.018*	*2.21E-07*	*1.91E-04*	*1.32E-04*	*0.003*
May	**1.0E-07**	**0.700**	**2.16E-06**	**2.40E-04**	**354.554**	**0.971**
	8.9E-08	*0.010*	*7.62E-08*	*6.50E-05*	*1.03E-04*	*0.001*
Jun	**3.3E-07**	**0.672**	**2.08E-06**	**4.30E-02**	**379.033**	**0.971**
	2.7E-07	*0.007*	*4.33E-08*	*7.36E-05*	*8.76E-05*	*0.003*
Jul	**5.3E-07**	**0.698**	**2.34E-06**	**2.41E-02**	**336.117**	**0.963**
	4.8E-07	*0.039*	*2.57E-07*	*8.34E-04*	*0.001*	*0.010*
Aug	**5.3E-07**	**0.698**	**2.34E-06**	**2.41E-02**	**336.117**	**0.963**
	4.8E-07	*0.039*	*2.57E-07*	*8.34E-04*	*0.001*	*0.010*
Sep	**1.4E-09**	**0.593**	**2.38E-06**	**1.21E-02**	**399.474**	**0.973**
	1.6E-22	*0.000*	*9.95E-20*	*1.60E-16*	*3.79E-04*	*0.000*
Oct	**2.2E-07**	**0.671**	**3.52E-06**	**2.05E-03**	**287.316**	**0.962**
	1.9E-07	*0.004*	*4.16E-08*	*1.88E-05*	*2.22E-05*	*0.002*
Nov	**2.2E-07**	**0.601**	**3.35E-06**	**2.23E-03**	**328.710**	**0.963**
	1.8E-07	*0.014*	*1.05E-07*	*1.12E-04*	*1.42E-04*	*0.003*
Dec	**8.8E-08**	**0.660**	**2.08E-06**	**4.29E-04**	**390.390**	**0.978**
	9.5E-08	*0.013*	*8.43E-08*	*3.69E-04*	*5.37E-04*	*0.001*

Table 18: Bayesian inference results of the Heston Nandi GARCH model. Given are the posterior means (in bold) and the posterior standard deviations (in italic). "PM" denotes the persistence measure and is defined as $\beta_{HN} + \alpha_{HN}\gamma_{HN}^2$. In the Gibbs sampler 35000 iterations are run, of which the first 5000 are used as burn-in sample.

8 Appendix II: Pricing Algorithm Markov Switching GJR

Pricing Algorithm MS GJR

let mc = 1 . . . MC

Draw β_{mc} from p(β|$\mathbf{r}$,$\theta_{-\beta}$)

Draw $\mathbf{P}_{mc}$ from p($\mathbf{P}$|$\mathbf{r}$,$\theta_{-\mathbf{P}}$)

Draw $\theta_{mc,G,i,j}$ from p($\theta_{G,j}|\mathbf{r}, \theta_{-\theta_{G,j}}$) for j = 1,. . .,J

let n = 1 . . . N

Draw $\varepsilon \sim N\tau(0, \text{I}\tau)$

let t = 1

Compute the probabilities in (24) for i = 1,2,3

Draw a uniform random variable from (0,1) and determine the governing regime $S_{s,1}$

Compute $v_0^* = r_{Tn}$ - μ_{mc} - $\rho_{mc} r_{Tn-1}$

Construct $\sigma_{s,1}^2(S_{s,1}) = \gamma_{mc}(S_{s,1}) + \alpha_{mc}(S_{s,1}) v_0^{*2} + \phi_{mc}(S_{s,1})\sigma_{s,Tn}^2 + \delta_{mc}(S_{s,1}) v_0^{*2}\mathbb{I}[v_0^* < 0]$

Construct $\text{r}_{s,1} = r_f + \varepsilon_{n,1}\sigma_{s,1}(S_{s,1})$

let t = 2

Compute the probabilities in (24) for i = 1,2,3

Draw a uniform random variable from (0,1) and determine the governing regime $S_{s,2}$

Compute $v_1^* = r_{s,1}$ - μ_{mc} - $\rho_{mc} r_{Tn}$

Construct $\sigma_{s,2}^2(S_{s,2}) = \gamma_{mc}(S_{s,2}) + \alpha_{mc}(S_{s,2}) {v_{s,1}^*}^2 + \phi_{mc}(S_{s,2})\sigma_{s,1}^2 + \delta_{mc}(S_{s,2}) {v_{s,1}^*}^2\mathbb{I}[v_{s,1}^* < 0]$

Construct $\text{r}_{s,2} = r_f + \varepsilon_{n,2}\sigma_{s,2}(S_{s,2})$

let t = 3 . . .τ

Compute the probabilities in (24) for i = 1,2,3

Draw a uniform random variable from (0,1) and determine the governing regime $S_{s,t}$

Compute $v_{t-1}^* = r_{s,t-1}$ - μ_{mc} - $\rho_{mc} r_{s,t-2}$

Construct $\sigma_{s,t}^2(S_{s,t}) = \gamma_{mc}(S_{s,t}) + \alpha_{mc}(S_{s,t}) {v_{s,t-1}^*}^2 + \phi_{mc}(S_{s,t})\sigma_{s,t-1}^2 + \delta_{mc}(S_{s,t}) {v_{s,t-1}^*}^2\mathbb{I}[v_{s,t-1}^* < 0]$

Construct $\text{r}_{s,t} = r_f + \varepsilon_{n,1}\sigma_{s,t}(S_{s,t})$

set t = t +1

Construct $\text{P}_T = P_t \prod_{t=1}^{\tau}(1 + r_{s,t}/100)$

Compute $C_t^{(mc)}(P_t, T, K) = (1 + r_f)^{-\tau} \max(\text{P}_T - K, 0)$

set n = n + 1

set mc = mc + 1

Compute the final price as

$$\widehat{C}_t(P_t, T, K) \approx \frac{1}{MC} \sum_{mc=1}^{MC} C_t^{(mc)}(P_t, T, K)$$

.

Table 19: Monte carlo simulation procedure for pricing European call options using a Markov switching GJR model. For the single regime GJR model the drawing of the transition probabilities is omitted such that the state variables reduce to 1 and only the posterior parameters from the single regime model are drawn. The remainder of the algorithm is the same. Note that Tn denotes the last date of the return sample.

9 Appendix III: Laplace Approximation of Bayes Factors

Bayes factors can be used to compare two or more nested or non-nested models. When comparing model A with model B the ratio of the marginal likelihoods is considered,

$$BF_{A|B} = \frac{p_A(r)}{p_B(r)} = \frac{\int_{\theta_A} p_A(\theta_A)p_A(r|\theta_A)d\theta_A}{\int_{\theta_B} p_B(\theta_B)p_B(r|\theta_B)d\theta_B}. \tag{52}$$

A-posteriori model A is more likely than model B when $BF_{A|B}$ is larger than 1 when assuming a-priori both models are equally likely. To compute the marginal likelihoods the integrals in (52) have to be evaluated, which typically cannot be done analytically. The marginal likelihoods can be estimated using Monte Carlo techniques, however a more convenient method is to approximate the integrals using Laplace's method. The following is a short description of Laplace's method.

Consider an integral of the following form

$$I = \int_a^b e^{-\lambda g(\theta)} dy.$$

with $g(\theta)$ a smooth vector function with a local minimum at $\widehat{\theta}$ in the interval (a,b) and λ a large scalar. Expanding a Taylor series of the function g around $\widehat{\theta}$ gives

$$g(\theta) \approx g(\widehat{\theta}) + g'(\widehat{\theta})(\theta - \widehat{\theta}) + (\theta - \widehat{\theta})' g''(\widehat{\theta})(\theta - \widehat{\theta})/2.$$

Note that $\widehat{\theta}$ is a local minimum which means that $g'(\widehat{\theta}) = \mathbf{0}$ (the null vector) and that $g''(\widehat{\theta})$ (the inverse Hessian matrix) is semi-positive definite such that

$$g(\theta) \approx g(\widehat{\theta}) + (\theta - \widehat{\theta})' g''(\widehat{\theta})(\theta - \widehat{\theta})/2.$$

Substituting this in the integral this becomes

$$I \approx \int_a^b e^{-\lambda(g(\widehat{\theta}) + (\theta-\widehat{\theta})' g''(\widehat{\theta})(\theta-\widehat{\theta})/2)} dy$$

$$= e^{-\lambda g(\widehat{\theta})} \int_a^b e^{-\lambda(\theta-\widehat{\theta})' g''(\widehat{\theta})(\theta-\widehat{\theta})/2} dy$$

which can be recognized as the integral over the kernel of a multivariate normal *pdf*, with mean $\widehat{\theta}$ and variance matrix $g''(\widehat{\theta})^{-1}$, multiplied by a constant. Scaling the integral properly and taking a = -∞ and b = ∞ the integral integrates nicely over the normal *pdf*,

$$= e^{-\lambda g(\widehat{\theta})}(2\pi)^{k/2}\frac{1}{\sqrt{|\lambda g''(\widehat{\theta})|}}\int_{-\infty}^{\infty}(2\pi)^{-k/2}\sqrt{|\lambda g''(\widehat{\theta})|}e^{-\lambda(\theta-\widehat{\theta})'g''(\widehat{\theta})(\theta-\widehat{\theta})/2}dy$$

$$= e^{-\lambda g(\widehat{\theta})}(2\pi\lambda)^{k/2}\frac{1}{\sqrt{|g''(\widehat{\theta})|}}\lambda^{-k/2}\times 1 \tag{53}$$

in which k denotes the number of model parameters. An approximation can be obtained by taking $\widehat{\theta}$ as the posterior mode and letting the function g denote the negative log posterior, i.e. $g(\widehat{\theta}) = \frac{-1}{N}\ ln(p(r|\widehat{\theta}))$ - $\frac{1}{N}\ ln(p(\widehat{\theta}))$ and $\lambda = N$. The variance-covariance matrix is then given by the inverse Hessian of the negative log posterior, $\Sigma(\widehat{\theta}) = g''(\widehat{\theta})^{-1}$. Substituting this in (53), the marginal likelihood is approximated by

$$ML_A \approx p(r|\widehat{\theta}_A)p(\widehat{\theta}_A)(2\pi)^{k_A/2}|\Sigma(\widehat{\theta}_A)|^{1/2}N^{-k_A/2} \tag{54}$$

This approximation can then be used to compute the Bayes factors of the MS GJR model with 1, 2 and 3 regimes. Additionally, the Bayesian Information Criterion ("BIC"), given by -$2ln(p(r|\widehat{\theta}_A)p(\widehat{\theta}_A))$ +k $ln(N)$, is also computed . Table (20) shows the computed Bayes factors for the MS GJR model.

	Log Like.	**Marg. Like.**	**BIC**	**Bayes**	**Factors**
1 State	-1184.00	-8.82E-26	2414.27	$\mathbf{BF}_{3\|1}$	61.58
2 States	-1111.25	-1.33E-25	2299.64	$\mathbf{BF}_{3\|2}$	2.60
3 States	-1040.72	-4.89E-25	2189.43	$\mathbf{BF}_{2\|1}$	23.66

Table 20: The table displays the log likelihood value, the approximated marginal likelihood value, the Bayesian information criterion and the Bayes factors.

Following the scale of Jeffreys (1961), given in table (21), the computed Bayes factors indicate that 3 regimes are more likely than 1 and 2 regimes and that 2 regimes are more likely than 1 regime. It also appears that the strength of the evidence in favor of 3 regimes versus 2 regimes is relatively weaker as opposed to 3 regimes versus 1 regime. Hence, the Bayes factors suggest that out of the considered number of regimes, 3 regimes are most likely.

Scale for interpreting the Bayes factor

$\mathbf{BF}_{A\|B}$	**Ln** $\mathbf{BF}_{A\|B}$	**Evidence for** $\mathbf{M}_A$
<1	<0	Negative
1-3	0-2.2	Not worth more than a bare mention
3-20	2.2-6	Positive
20-150	6-10	Strong
>150	>10	Very strong

Table 21: Scale for interpreting the Bayes factors.

10 Appendix IV: Bayesian Estimation of the Ad-hoc BS

Equation (40) is a simple linear regression model and, assuming standard normal residuals ϵ_i, the model can be estimated using Bayesian inference. Collecting the covariates in the $n \times k$ matrix X, in which the first column is a vector of ones, and the parameters in β_{AH}, this model can be written as

$$\sigma_{BS} = X\beta_{AH} + \epsilon \tag{55}$$

where $\epsilon \sim \mathbb{N}(0, I_N\sigma^2)$ such that the likelihood function is given by

$$p(\sigma_{BS}|\beta_{AH}, \sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N exp\left(-\frac{1}{2\sigma^2}(\sigma_{BS} - X\beta_{AH})'(\sigma_{BS} - X\beta_{AH})\right). \tag{56}$$

Since no prior beliefs about the parameters exists the priors imposed for β_{AH} and σ are uninformative and given by

$$p(\beta_{AH}) \propto 1 \text{ and } p(\sigma^2) \propto \sigma^{-2}.$$

Multiplying the priors with the likelihood function yields the posterior distribution, which is proportional to

$$\left(\frac{1}{\sigma}\right)^{N+2} exp\left(-\frac{1}{2\sigma^2}(\sigma_{BS} - X\beta_{AH})'(\sigma_{BS} - X\beta_{AH}\right). \tag{57}$$

To sample the model parameters, the full conditional posteriors of β_{AH} and σ need to be determined. Using the decomposition rule, as explained in Greenberg (2013, ch. 4.3), the full conditional posterior of β_{AH} can be written as

$$p(\beta_{AH},\sigma|\sigma_{BS}) \propto \left(\frac{1}{\sigma}\right)^{N+2} exp\left(-\frac{1}{2\sigma^2}((\sigma_{BS}-X\widehat{\beta}_{AH})'(\sigma_{BS}-X\widehat{\beta}_{AH})+(\beta_{AH}-\widehat{\beta}_{AH})'X'X(\beta_{AH}-\widehat{\beta}_{AH}))\right)$$

$$\propto exp\left(-\frac{1}{2\sigma^2}(\beta_{AH}-\widehat{\beta}_{AH})'X'X(\beta_{AH}-\widehat{\beta}_{AH})\right) \tag{58}$$

which is the kernel of a multivariate normal distribution with mean $\widehat{\beta}$ (the ordinary least squares estimator) and covariance matrix $\sigma^2(X'X)^{-1}$. Collecting the terms with σ in equation (57), the full conditional posterior of σ^2 is given by

$$p(\sigma^2|\beta_{AH},\sigma_{BS}) \propto \left(\frac{1}{\sigma}\right)^{N+2} exp\left(-\frac{1}{2\sigma^2}(\sigma_{BS}-X\beta_{AH})'(\sigma_{BS}-X\beta_{AH})\right) \tag{59}$$

which is the kernel of an inverted Gamma-2 distribution with parameter $(\sigma_{BS}-X\beta_{AH})'(\sigma_{BS}-X\beta_{AH})$ and N degrees of freedom. The full conditional posteriors can then be used to construct the Gibbs sampler as given in table (22).

Sampling Scheme for β_{AH} and σ

1. Set m = 0 and initialize the Gibbs sampler by setting $\beta_{AH}^{(0)} = \widehat{\beta}_{AH}$.
2. Draw a random value from the Chi-square distribution with N degrees of freedom,

   $$\text{y} \sim \chi^2(N).$$

3. Divide the inverted Gamma-2 parameter by the simulated Chi-square value to obtain a draw of ${\sigma^2}^{(m+1)}$ from the inverted Gamma-2 distribution,

   $${\sigma^2}^{(m+1)} \sim (\sigma_{BS}-X\beta_{AH}^{(m)})'(\sigma_{BS}-X\beta_{AH}^{(m)})/\text{y}.$$

4. Simulate $\beta_{AH}^{(m+1)}$ from the multivariate normal distribution,

   $${\beta_{AH}}^{(m+1)} \sim \mathbb{N}(\widehat{\beta}_{AH},{\sigma^2}^{(m+1)}(X'X)^{-1}).$$

5. set m = m + 1 and go back to step 2.

Table 22: Gibbs sampling routine for the Ad Hoc Black Scholes model. A draw from the inverted Gamma-2 distribution is obtained by using that $(\sigma_{BS}-X\beta_{AH})'(\sigma_{BS}-X\beta_{AH})/\sigma^2 \sim \chi^2(N)$.