Testing for the number of regimes in financial time series GARCH Volatility

Abdellah Tahiri[a,*], Brahim Benaid[a], Hassane Bouzahir[a], Naushad Ali Mamode Khan[b]

[a] *ISTI Lab, ENSA PO Box 1136, Ibn Zohr University-Agadir, Morocco*
[b] *Department of Economics and Statistics, University of Mauritius, Mauritius*

Abstract

This paper investigates the optimal number of regimes that can better describe the corresponding conditional variance to different stock market indices. We compared several GARCH models using the Deviance Information Criterion (DIC), provided by the Bayesian approach Markov chain Monte Carlo (MCMC), considering many stylized facts such as asymmetry (i.e., leverage effect), fat-tailed distributions, and volatility clustering. The results show clearly that the four selected models exhibit a leverage effect and have at least two regimes, whatever the GARCH specifications are. In addition, the optimal number of regimes in the conditional variance process may change from a series to another depending on their structure. A predictive test using the Value-at-Risk confirms that the selected processes provide accurate volatility forecasts.

Keywords: Volatility, Markov switching GARCH, Model selection, Bayesian Approach, DIC, Stock Market Indices.

JEL Classification: C1, C52, C58.

---

1. Introduction

Volatility has been a critical element in the risk management world and decision-making. Many mathematicians consider Louis Bachelier as the origin of modern finance (i.e., mathematical finance), particularly the "theory of speculation" thesis (1900). Since that time, several studies have been oriented toward forecasting and modeling volatility. The most popular used class of Auto-regression Conditional Heteroskedastic (ARCH) was introduced by Engel (1982)[15] and generalized by Bollerslev (1986)[11], gives more flexibility for the conditional variance process structure and allows more capturing to high persistence and clustering of conditional volatility. From a business perspective, negative shocks affect the conditional variance process more than positive ones (i.e., the asymmetry).

Consequently, Glosten-Jagannathan-Runkle (1993)[19] suggested a new specification (GJR-GARCH) in order to capture the leverage effect (asymmetry) exhibited in the conditional variance process. Different studies have investigated the forecasting difficulties of GARCH model class, such as Hamilton and Susmel (1994)[24] and Lamoureux and Lastrapes (1993)[28], which mentioned that the GARCH models class might lead to a worse multi-period volatility forecasting than even the constant variance models do. Friedman and Laibson (1989)[18] highlighted that the weakness of the GARCH class models in forecasting volatility comes from not

*Corresponding author
Email address: abdellahtahiri01@gmail.com (Abdellah Tahiri)

allowing the conditional variance process to be flexible with ”small” or ”large” shocks and structural breaks, which can lead to high estimated persistence.

Many researchers have focused on some features of stock market indices. The appropriate model should be able to capture stylized facts observed in financial markets; nonlinearity and asymmetry, for instance. Hamilton (1989) [23]introduced a Markov-switching model to capture structural breaks by allowing conditional variance parameters to switch across different regimes. A way to introduce the ”switch” to the return process is provided by the Markov-switching regimes GARCH (MS GARCH) models. These structures can adapt quickly to many variations in the volatility level (see Ardia, 2008[1] and Marcucci, 2005[29]).

Using MS GARCH models lead to path-dependence issue. Gray (1996)[20], Dueker (1997)[14], and Klaassen (2002)[26] handled this problem through approximation by collapsing the past regime conditional variance. Bauwens et al (2014)[5] recommended using the MCMC method to estimate complex models such as MS GARCH, avoiding the path-dependence issue.

In the present paper, we analyze four stock market indices from different financial markets. The primary purpose is to test the optimal number of regimes in volatility presented in each series. The idea was, for different series, the most appropriate model should take into consideration various specifications such as the number of regimes presented in the conditional variance process.

The remainder of this paper is organized as follows: presenting the Markov-switching GARCH (MS GARCH) model in Section 2, while in Section 3, we analyze and investigate about data and the methodology of this work. The empirical results are discussed in Section 4. Finally, Section 5 summarizes the results of this study.

2. Markov-Switching GARCH Model

One of Hamilton and Susmel’s (1994)[24] main recommendations is that ARCH models refer to much persistence in stock return volatility that leads to a poor forecast due to the leverage effect, i.e., the large shocks affect more the volatility than small breaks. The idea to avoid this phenomenon is to allow the ARCH model parameters to change and come from different regimes governed by Hidden Markov Chains (HMC).

By introducing Markov chains on ARCH specifications, we benefit from the primary Markov propriety:

$$P(S_{t+1} = j | S_t = i, S_{t-1} = i_{t-1}, ..., S_0 = i_0) = P(S_{t+1} = j | S_t = i), \tag{1}$$

with $(i_0, i_1, ..., i_{t-1}, i, j) \in \mathbb{N}^{n+2}$ and $(S_t)_{t\in\mathbb{N}}$ is a stochastic process.

Unfortunately, the moderate coefficient or ARCH term cannot capture the conditional heteroskedasticity within regimes. Then we need the GARCH term to explain better the volatility clustering in financial markets. This section introduces the Markov-Switching GARCH (MS GARCH) model following Bollerslev (1986)[11]. Let us consider a stock market index price P_t at time t, the log-return r_t is defined as the continuous rate of returns as:

$$r_t = \log(\mathrm{P}_t/\mathrm{P}_{t-1}),$$

we assume that $\mathbb{E}(r_t) = 0$ and (r_t) is serially uncorrelated.

The MS GARCH model following Ardia et al. (2018) can be defined as:

$$r_t | (S_t = k, \mathscr{I}_{t-1}) \sim f(0, \mathsf{V}_t^{(k)}, \Phi^{(k)}), \tag{2}$$

where $f(.)$ is a continuous distribution with zero mean and switching conditional variance $\mathsf{V}_t^{(k)}$ within the regime (k). $\Phi^{(k)}$ denotes all additional parameters and $\mathscr{I}_{t-1}$ is the accumulated information set at time $t-1$ generated by $\{r_{t-1}, r_{t-2}, ...\}$. In all models, the conditional variance $\mathsf{V}_t^{(k)}$ is allowed to switch across a Markov process $S_t \in \{1, 2, ...K\} \subset \mathbb{N}$. We define the matrix of transition for S_t as follows:

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1K} \\ p_{21} & p_{22} & \cdots & p_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{K1} & p_{K2} & \cdots & p_{KK} \end{bmatrix},$$

where $p_{ij} = \Pr[S_t = j | S_{t-1} = i]$ is the probability of switching from regime $(S_{t-1} = i)$ to regime $(S_t = j)$ and $\sum_{j=1}^{K} p_{ij} = 1$ for i fixed. When $(i = j)$, p_{ij} is called the persistence probability within the regime (i). In this study, we will consider up to three regimes (i.e., $S_t \in \{1, 2, 3\}$).

So far, the idea of MS GARCH specifications is to combine GARCH structure with parameters that change over time in order to capture structural breaks presented in the conditional variance. However, this operation leads to a path-dependence problem; as the conditional variance at time t depends on all regimes path $S_t(t = 1, ..., K)$. To avoid this issue, we refer to Haas et al. (2004)[22] and Ardia et al. (2017)[4].

Many studies have already shown that GARCH(1,1) provides better descriptions of market volatility behavior than even high order of ARCH specifications (see Haas, 2004)[22]. Following Bollerslev (1986)[11], the GARCH(1,1) conditional variance process expression can be written as follows:

$$r_t = \epsilon_t \mathsf{V}_t^{1/2} \qquad ; \quad \epsilon_t \overset{iid}{\sim} f(0, 1), \tag{3}$$

$$\mathsf{V}_t = \omega + \alpha r_{t-1}^2 + \beta \mathsf{V}_{t-1} = \psi(r_{t-1}, \mathsf{V}_{t-1}), \tag{4}$$

with $\omega > 0$, $\alpha \geqslant 0$ and $\beta \geqslant 0$ to ensure a positive variance, where (ϵ_t) is a sequence of independently and identically distributed (i.i.d) random variables with zero mean and unit variance. $f(.)$ a conditional distribution needs to be specified.

In all that will come, V_t is assumed to be a Markov-switching GARCH (1,1) process, governed by $S_t(t = 1, ..., K)$:

$$\mathsf{V}_t = \psi(r_{t-1}, \mathsf{V}_{t-1}, S_t), \tag{5}$$

We apply S_t on equation (4) and the conditional variance in switching regimes will be presented as:

$$\mathsf{V}_t = \omega^{(k)} + \alpha^{(k)} r_{t-1}^2 + \beta^{(k)} \mathsf{V}_{t-1}, \tag{6}$$

Gray (1996) [20] uses the information set at time $t-1$ for integrating out hidden regimes to avoid path-dependence issue as follows :

$$\mathsf{V}_t = \sum_{k=1}^{K} p_t^{(k)} \mathsf{V}_t^{(k)}, \tag{7}$$

where $\mathsf{V}_t^{(k)}$ is the conditional variance of r_t within the regime (k) and $p_t^{(k)} = \Pr(S_t = k | \mathscr{I}_{t-1})$ is the probability to be in the regime (k) according to the information up to $t-1$.

Let us assume that $S_t \in \{1, 2\}$, the conditional distribution of r_t is a switching of distributions $f(\mathsf{V}_t^{(k)})$ within regimes (k):

$$r_t|\mathscr{T}_{t-1} \sim \begin{cases} f(\mathsf{V}_t^{(1)}) & \text{With probability } p_t^{(1)} \\ f(\mathsf{V}_t^{(2)}) & \text{With probability } p_t^{(2)} \end{cases}, \tag{8}$$

$f(.)$ denotes a conditional distribution that should be specified (i.e., Normal, Student's-t, GED, ...).
We use equation (7) to replace equation (6) by specifying GARCH expression for each regime (k),

$$\mathsf{V}_t^{(k)} = \omega^{(k)} + \alpha^{(k)} r_{t-1}^2 + \beta^{(k)} \mathsf{V}_{t-1}, \tag{9}$$

V_{t-1} represents a regime-independent past variance. The matrix expression is used to simplify the expressions as follows:

$$\mathsf{V}_t = \Pi \times (\Omega \times \Upsilon), \tag{10}$$

with $\Pi = [p_t^{(1)}, p_t^{(2)}, ..., p_t^{(K)}]$, $\Omega = \begin{bmatrix} \omega^{(1)} & \alpha^{(1)} & \beta^{(1)} \\ \omega^{(2)} & \alpha^{(2)} & \beta^{(2)} \\ \vdots & \vdots & \vdots \\ \omega^{(K)} & \alpha^{(K)} & \beta^{(K)} \end{bmatrix}$ and $\Upsilon = [1, r_{t-1}^2, \mathsf{V}_{t-1}]^t$.
So far, one of the most interesting phenomena in financial markets is what we call the "leverage effect", caused by the fact that negative returns have an important impact on conditional volatility. However, the standard GARCH(1,1) structure gives the same considerations to negative as positive returns in terms of influencing the future volatility.
To avoid this issue, we will refer to the GJR-GARCH model proposed by Glosten et al. (1993)[19] to capture the asymmetry effect in our time series. The conditional variance process can be expressed as:

$$\mathsf{V}_t^{(k)} = \omega^{(k)} + (\alpha^{(k)} + \gamma^{(k)} \mathbb{1}(r_{t-1} < 0)) r_{t-1}^2 + \beta^{(k)} \mathsf{V}_{t-1}, \tag{11}$$

where $\mathbb{1}(r_{t-1} < 0) = 1$ if $r_{t-1} < 0$ and 0 otherwise. For $k \in \{1, ..., K\}$, $\gamma^{(k)} \geqslant 0$ is the asymmetry measure parameter in the conditional variance process.
One other dimension which influences the performance of our conditional volatility modeling is the distribution of standardized innovations (ϵ_t) that needs to be assumed and specified. In this study, we focus on three distributions, namely, normal, Student's-t, and generalized error distribution (GED). Regarding the presence of asymmetry, skewed distributions are required. We refer to Fernández & Steel (1998) [17] to define the skewed density as follows:

$$f_\xi(z) = \frac{2\sigma_\xi}{\xi + \xi^{(-1)}} f^*(z_\xi), \tag{12}$$

where

$$z_\xi = \begin{cases} \xi^{-1}(\sigma_\xi z + \mu_\xi) & \text{if } z \geqslant -\dfrac{\mu_\xi}{\sigma_\xi} \\ \xi(\sigma_\xi z + \mu_\xi) & \text{otherwise} \end{cases},$$

with $\mu_\xi = m(\xi - \xi^{-1})$, $\sigma_\xi^2 = (1 - m^2)(\xi^2 - \xi^{-2}) + 2m^2 - 1$ and $m = \int_0^\infty 2t f^*(t)dt$.
$0 < \xi < \infty$ is the degree of asymmetry parameter and $f^*(.)$ is a symmetric density function with zero mean and unit variance.

3. Data and Methodology

In this section, we present the data set used within this study to look at some features that describe the stock market indices series. We selected four major stock market indices, namely, S&P500, FTSE 100, CAC40, and Nikkei 225, respectively, reflecting the USA, UK, France, and Japan markets. All historical data have been obtained from the Yahoo Finance platform. We consider daily log-returns during the period range from January 02, 1990, until October 10, 2019, with around 7,500 observations to ensure that we have captured the maximum of the market behavior. The calculations used the formula; $r_t = \log(\mathrm{P}_t/\mathrm{P}_{t-1})$ in percentage, where P_t is the adjusted-close price at time t.

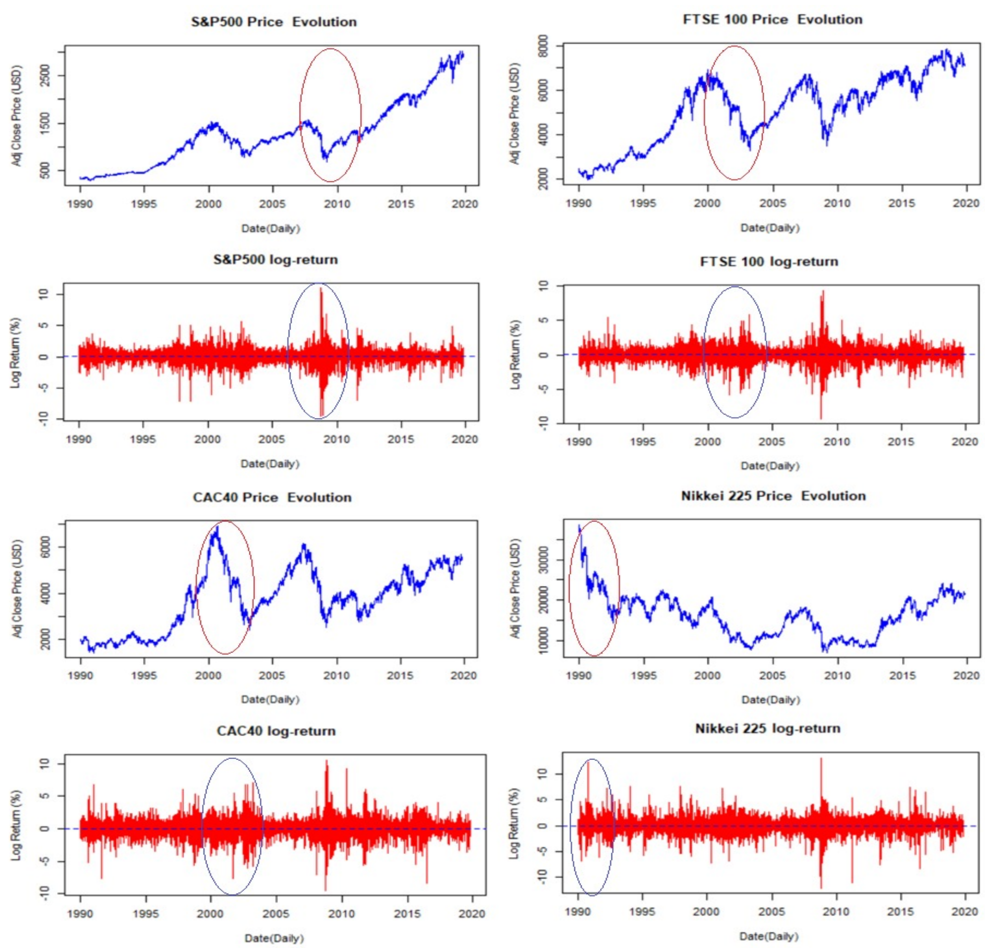

Figure 1: Presentation of stock market indices log-return (%) and price evolution

In Figure 1, we provide a graphic representation for our series to check the volatility clustering. Clearly, we see that there are ”low (high)” fluctuations followed by ”low (high)” fluctuations. In addition, the level of changes varies from one series to the other, while the level of changes is more significant for CAC40 and Nikkei 225 series compared to S&P500 and FTSE 100 series. The correlation between the level of fluctuations in log-returns and their stock market index evolution is presented as well.

Table 1 reports daily log-returns summary statistics for S&P500, FTSE 100, CAC40 and Nikkei 225:

Table 1: Descriptive statistics of the return data.

Statistic	**S&P500**	**FTSE 100**	**CAC40**	**Nikkei 225**
Mean (%)	0.028	0.014	0.013	-0.008
Median (%)	0.053	0.0316	0.035	0.0155
Maximum (%)	10.957	9.384	10.594	13.234
Minimum (%)	-9.469	-9.264	-9.471	-12.111
Std.Dev (%)	1.102	1.082	1.354	1.450
Skewness	-0.264	-0.131	-0.073	-0.149
Kurtosis	8.789	6.170	4.744	5.474
JB-Statistic	24250	12033	7095.5	9169
JB p-value	< 0.01	< 0.01	< 0.01	< 0.01
LM (12)	1972.7	1686.9	1222.6	1282.9
LM p-value	< 0.01	< 0.01	< 0.01	< 0.01

The mean is quite small and close to zero for S&P500, FTSE 100, CAC40, and Nikkei 225 with values of 0.028%, 0.014%, 0.013%, and -0.008%, respectively which defend the assumption of keeping the mean zero. The standard deviation (Std.Dev) is around unity. However, the Std.Dev for CAC40 and Nikkei 225 are more important with values of 1.35% and 1.45% respectively. A high standard deviation indicates high level of volatility (hypothesis to be verified).

The skewness (asymmetry coefficient) is significant and negative showing that the distribution tail spread to the left. The normalized kurtosis (excess kurtosis) is significantly higher than the normal distribution value of 0, indicating that the distributions have fatter tails.

Lagrange multiplier (LM) test demonstrates the presence of ARCH effect in all series, under the null hypothesis of ”no ARCH effect”, where $q = 12$ is the number of lags. The results show also that the distributions are far to be normal distributions according to Jarque-Bera test under the null hypothesis of ”normal distribution”.

The main purpose of this paper is to identify the appropriate number of regimes in each series, which better describe the behavior of each stock market index. The idea is to test various models with several specifications, where the number of regimes is one of the parameters that constitute model specifications. Therefore, the optimal number of regimes is that which corresponds to the best specification. We present the optimization problem as follows:

$$\hat{\Phi} \equiv \arg \min_{\Phi \equiv (G,f,K)} \text{DIC}(\Phi), \tag{13}$$

DIC is the Deviance Information Criterion (see Spiegelhalter et al. 2002[30]).
The output of this optimization refers to the optimal model specification $\hat{\Phi}$, where G is the GARCH model class[1], $f(.)$ denotes the distribution of the standardized innovations, and K is the number of regimes.
As we mentioned before, the introduction of MS-GARCH to model the dynamic volatility leads to the path-dependence issue. The solution is to orient toward the Bayesian estimation MCMC (Markov Chain Monte Carlo) technique, which requires evaluation of the likelihood function:

$$\mathcal{L}(\Theta, \mathscr{I}_t) = \prod_{t=1}^{T} f(r_t|\Theta, \mathscr{I}_{t-1}), \tag{14}$$

where Θ is the vector of the model parameters and $f(r_t|\Theta, \mathscr{I}_{t-1})$ is the conditional density function of r_t according to the available information in $t-1$; $\mathscr{I}_{t-1}$. In our case of the MS-GARCH model, the conditional density $f(.)$ can be written as:

$$f(r_t|\Theta, \mathscr{I}_{t-1}) = \sum_{i=1}^{K}\sum_{j=1}^{K} p_{ij} f(r_t|S_t = j, \Theta, \mathscr{I}_{t-1}) p_{t-1}^{(i)}, \tag{15}$$

where $p_{ij} = \Pr[S_t = j|S_{t-1} = i]$ is the transition probability from regime $(S_{t-1} = i)$ at $t-1$ to regime $(S_t = j)$ at t and $p_{t-1}^{(i)} = \Pr[S_{t-1} = i|\Theta, \mathscr{I}_{t-1}]$ is the filtered probability of regime $(S_{t-1} = i)$ at time $t-1$.
- For $K = 2$:

$$\begin{aligned} f(r_t/\Theta, \mathscr{I}_{t-1}) = {} & p_{t-1}^{(1)}[p_{11} f(r_t/S_t = 1, \Theta, \mathscr{I}_{t-1}) + (1-p_{11}) f(r_t/S_t = 2, \Theta, \mathscr{I}_{t-1})] \\ & + p_{t-1}^{(2)}[p_{22} f(r_t/S_t = 2, \Theta, \mathscr{I}_{t-1}) + (1-p_{22}) f(r_t/S_t = 1, \Theta, \mathscr{I}_{t-1})]. \end{aligned}$$

The estimation of $\hat{\Theta}$ used the Bayesian approach MCMC based on the adaptive random-Walk-Metropolis sampler (see Vihola, 2012[33]) to generate draws from the posterior distribution, where the model parameters are treated as random variables.
The power of MCMC comes from the fact that all inferences are based on the joint posterior distribution of the model parameters. Therefore, normality is not an essential requirement, and the convergence of the estimation algorithm to be judged by ourselves. The posterior estimation accuracy can be evaluated by the Monte Carlo Standard Error (MCSE) for each parameter. The more number of iterations increases, the more posterior estimation accuracy increases (i.e., MCSE $\to 0$).
In this study, we generated 20,000 draws, then we discarded the first 5,000 iterations as the burn-in sample, and we considered a lag of 5 to reduce the autocorrelation in the Markov chain. By the end, we built a sample of size 3,000. To ensure the convergence of our model-fitting algorithm, we check the output results with different seeds. [2].
In the present paper, we share the same logic with Box (1978); " all models are wrong. However, some of them are useful ". To compare the fitting-quality of our models, we are using DIC for the Bayesian model selection problems where the posterior distributions of the model parameters obtained by MCMC that can

[1]We consider two classes of GARCH model, the symmetric standard GARCH(1,1) against asymmetric GJR-GARCH(1,1).
[2]The R package provided by Ardia et al., (2017)[4] is used to estimate model parameters implementing the MCMC method.

be viewed as the AIC (Akaike Information Criterion) or the BIC (Bayesian Information Criterion):

$$\text{DIC} = \bar{D} + P_D, \tag{16}$$

where $\bar{D}$ is the posterior mean deviance that measures the adequacy of fit, while P_D is the effective number of parameters describes the model complexity (for further details, see Spiegelhalter et al., 2002[30]).
The DIC is not a criterion to identify the correct model, but it can be used to compare different specifications to determine the most appropriate one since the helpful model is the one with a smaller DIC, which means the goodness of fit is penalized by the model complexity.

4. Empirical results and discussion

So far, we have presented the most efficient auto-regression models (i.e., Standard GARCH and GJR-GARCH) with different specifications to get more flexibility in capturing the persistence of the stock market indices conditional volatility. We have also shown the best approach to fit complex models such as MS-GARCH models using MCMC/ Bayesian approach. The suitability of our models is judged with DIC criterion to select the most appropriate models.
In this section, we provide log-returns results analysis of S&P500, FTSE 100, CAC40, and Nikkei 225. In our in-sample analysis, we fit all models (i.e., $4 \times 18 = 72$ models) using the whole data set from January 01, 1990, until October 10, 2019.

Table 2: Deviance Information Criterion

		S&P500			**FTSE 100**		
		k=1	k=2	k=3	k=1	k=2	k=3
S-GARCH	**Sk-N**	19579.78	19418.08	19450.18	20060.04	19985.10	19981.48
	Sk-STD	19278.53	19272.02	19434.17	19936.01	19936.25	19967.61
	Sk-GED	19250.42	19234.04	19352.87	19947.17	19958.45	19994.52
GJR-GARCH	**Sk-N**	19320.12	19059.54	19174.34	19882.81	19833.83	19813.35
	Sk-STD	19067.71	19070.24	19028.94	19790.42	**19738.61**	19945.14
	Sk-GED	19054.80	**19014.57**	19112.39	19799.58	19798.73	20604.14
		CAC40			**Nikkei 225**		
		k=1	k=2	k=3	k=1	k=2	k=3
S-GARCH	**Sk-N**	23927.91	23814.4	23777.05	24998.95	24804.86	24789.35
	Sk-STD	23738.48	23738.33	23759.03	24748.38	24733.34	24710.19
	Sk-GED	23764.90	23736.23	23742.65	24760.17	24727.94	24716.21
GJR-GARCH	**Sk-N**	23734.96	23572.45	23540.55	24827.91	24644.45	24627.53
	Sk-STD	23547.75	23527.04	**23512.13**	24614.89	24597.53	24585.36
	Sk-GED	23593.04	23552.82	23685.67	24626.19	24599.99	**24543.25**

Before we start fitting models, we filtered our series with AR(1) model recommended by AIC (results not reported) to ensure the uncorrelated between log-returns (r_t) observations.

For the four stock market indices, the DIC of different specifications is reported in Table 2. Berg et al. (2004)[10] and Ardia (2008)[1] have shown many DIC advantages in selecting the adequate model to describe stochastic volatility better.

For the S&P500 index, almost all two-regimes models provide a better trade-off in terms of fitting adequacy and model complexity, except the GJR-GARCH skewed student's-t distribution, which can be preferred with three regimes. The comparison leads to select the GJR-GARCH skewed GED distribution model within two regimes that outperform all other specifications.

Regarding the FTSE 100 Index, the skewed student's-t and GED distributions with single-regime provide more adequacy for the standard symmetric GARCH (i.e., S-GARCH). However, two-regimes GJR-GARCH skewed student's-t distribution offer once again better adequacy compared to the complexity of the model.

For the CAC40 index, Markov switching GARCH specifications outperform signal-regime. The most appropriate model to describe CAC40 log-returns behavior is the three-regimes GJR-GARCH skewed student's-t distribution. In addition, using standard symmetric GARCH requires two-regime specifications.

Finally, for the Japanese stock market index Nikkei 225, as a result of the high market volatility, the comparison has shown that three-regimes are required, whatever the GARCH specifications and the conditional distribution are. In fact, there is an additional advantage for the GJR-GARCH skewed GED distribution model within three-regimes.

Table 3 summarizes the results of our comparison. The results show clearly that the MS GJR-GARCH is strongly favored to describe market stock indices volatility, a justification for the presence of leverage effect.

Table 3: Selected specifications

	Specification ($\hat{\Phi}$)			**Unconditional Probabilities (π_k)**			
	Class	**Dist**	**Regimes**	**Regime 1**	**Regime 2**	**Regime 3**	**UnVol (%)**
S&P500	GJR-GARCH	Sk-GED	K=2	0.7575	0.2425	-	1.031
FTSE 100	GJR-GARCH	Sk-STD	K=2	0.5515	0.4485	-	1.120
CAC40	GJR-GARCH	Sk-STD	K=3	0.3113	0.6766	0.0121	1.955
Nikkei 225	GJR-GARCH	Sk-GED	K=3	0.3494	0.6457	0.0048	1.445

We also present some additional proprieties such as the unconditional volatility (UnVol); the values are coherent with the standard deviation mentioned before. We observe high unconditional volatility in both CAC40 and Nikkei 225, with values of ($\simeq 1.95\%$) and ($\simeq 1.44\%$), respectively.

The estimation results show that the French index CAC40 and the one of Japan Nikkei 225 are more volatile. This logic explains the necessity of considering three regimes to capture three levels of change (i.e., low, medium, and high). The More volatility increases, the more we need a considerable number of regimes.

From the posterior Means of the MS GJR-GARCH models, the regime's (k)unconditional probabilities [3] are respectively, ($\simeq 76\%$) for the first regime and ($\simeq 24\%$) for the second one in case of S&P500, where we see more stability in the first regime. For FTSE 100, there is a significant balance between regimes. In the case

[3]For $K = 2 : \pi_1 = \frac{1-p_{22}}{1-p_{11}-p_{22}}$ and $\pi_2 = \frac{1-p_{11}}{1-p_{11}-p_{22}}$ where $\pi_1 + \pi_2 = 1$

of CAC40 and Nikkei 225, the results show more stability in the second regime with values of ($\simeq 68\%$) and ($\simeq 65\%$) respectively, and with significant stability in the first regime comparing to the third one as well. These remarks can be expressed with the help of the smoothed probabilities $\Pr[S_t = k|\mathscr{I}_T]$ presented in Figure 2, which are easier to interpret. From Figure 2, we can observe certain stability in the time within regimes for different series.

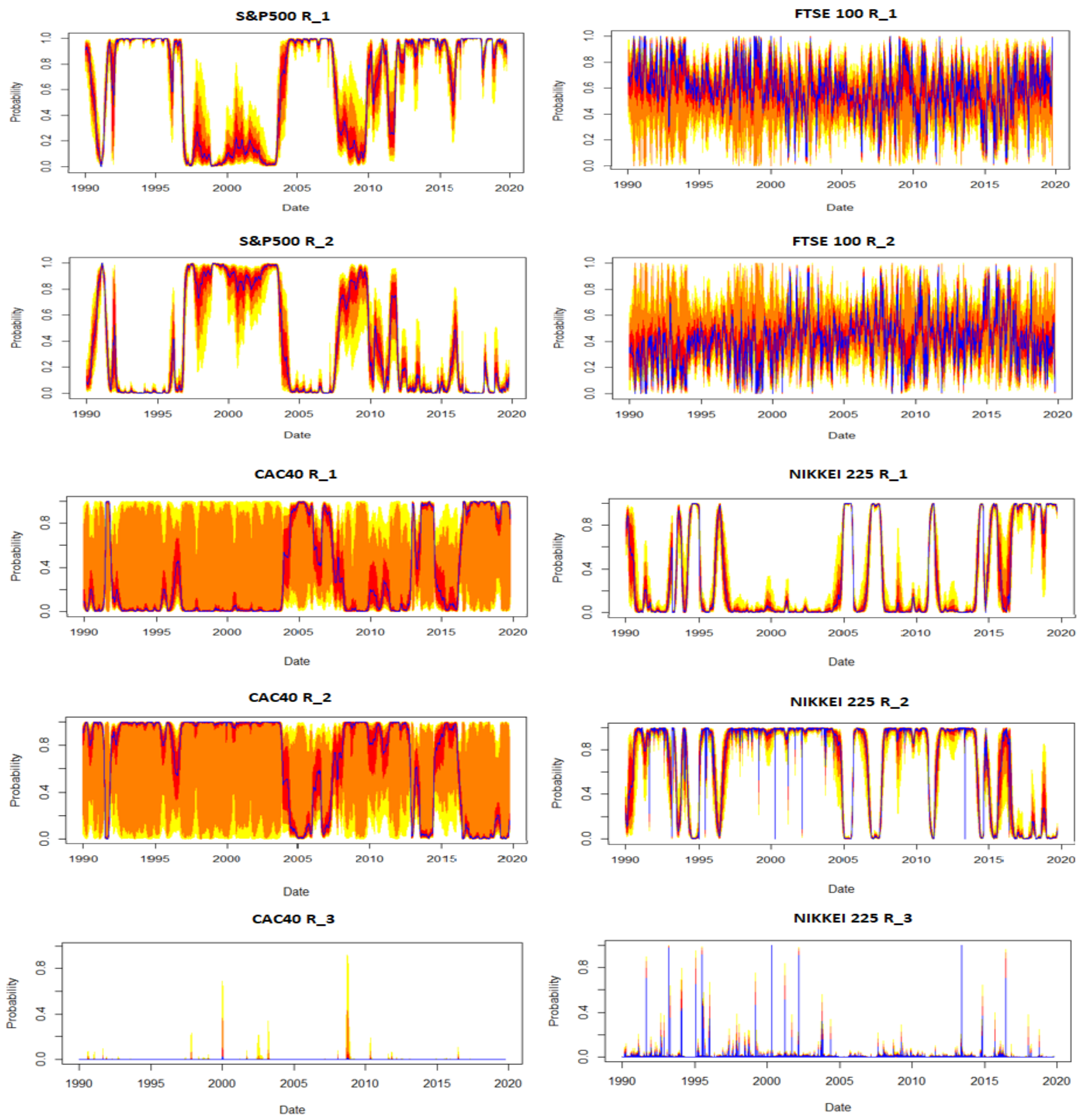

Figure 2: Smoothed probability

Table 4 presents the parameter estimates of the most appropriate MS GJR-GARCH specifications associated with S&P500, FTSE 100, CAC40, and Nikkei 225; two-regimes skewed GED, two-regimes skewed student's-t, three-regimes skewed student's-t, and three-regimes skewed GED, respectively. We report in Table 4, for each stock market indices and a given regime (k), the parameters of GJR-GARCH(1,1) ; $(\omega^{(k)}, \alpha^{(k)}, \gamma^{(k)}, \beta^{(k)})$, where the leverage effect is controlled by $\gamma^{(k)}$.Further, $\Phi^{(k)} \equiv (\eta^{(k)}, \xi^{(k)})$ are the parameters of the skewed distribution with $\eta^{(k)}$ and $\xi^{(k)}$ measure the tail and the asymmetry, respectively.

We also provided the elements of the transition matrix $p_{ij} = \Pr[S_t = j|S_{t-1} = i]$ where p_{kk} represents the regime's (k) persistence probability.

Table 4: Selected model parameters

	SP500			FTSE 100			CAC40			Nikkei 225		
Parameter	Mean	Std.Dev	MCSE	Mean	Std.Dev	MCSE	Mean	Std.Dev	MCSE	Mean	Std.Dev	MCSE
$\omega^{(1)}$	0.0196	0.0036	0.0001	0.0103	0.0096	0.0003	0.0667	0.0229	0.0007	0.1204	0.0415	0.0013
$\alpha^{(1)}$	0.0001	0.0000	0.0000	0.0140	0.0081	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\gamma^{(1)}$	0.1672	0.0223	0.0007	0.0929	0.0871	0.0028	0.2416	0.0716	0.0023	0.3035	0.0746	0.0024
$\beta^{(1)}$	0.8829	0.0129	0.0004	0.9265	0.0485	0.0015	0.8076	0.0606	0.0019	0.7339	0.0765	0.0024
$\eta^{(1)}$	1.3046	0.0526	0.0017	29.9671	33.5484	1.0609	9.3118	8.8656	0.2804	1.3459	0.0843	0.0027
$\xi^{(1)}$	0.9020	0.0170	0.0005	0.9327	0.0518	0.0016	0.8902	0.0364	0.0012	0.8743	0.0313	0.0010
$\omega^{(2)}$	0.0512	0.0085	0.0003	0.0254	0.0122	0.0004	0.0431	0.0110	0.0003	0.0864	0.0125	0.0004
$\alpha^{(2)}$	0.0002	0.0001	0.0000	0.0045	0.0083	0.0003	0.0000	0.0000	0.0000	0.0002	0.0001	0.0000
$\gamma^{(2)}$	0.1597	0.0175	0.0006	0.2356	0.1044	0.0033	0.1414	0.0621	0.0020	0.1339	0.0287	0.0009
$\beta^{(2)}$	0.8905	0.0054	0.0002	0.8531	0.0548	0.0017	0.9053	0.0371	0.0012	0.8928	0.0176	0.0006
$\eta^{(2)}$	1.6208	0.0823	0.0026	74.5030	34.7333	1.0984	24.1334	9.0266	0.2854	1.7187	0.0829	0.0026
$\xi^{(2)}$	0.8913	0.0292	0.0009	0.8540	0.0625	0.0020	0.9280	0.0327	0.0010	0.9675	0.0245	0.0008
$\omega^{(3)}$	-	-	-	-	-	-	11.2702	10.5682	0.3342	3.5627	1.1113	0.0351
$\alpha^{(3)}$	-	-	-	-	-	-	0.0000	0.0000	0.0000	0.0150	0.0120	0.0004
$\gamma^{(3)}$	-	-	-	-	-	-	0.1976	0.0577	0.0018	0.0302	0.0128	0.0004
$\beta^{(3)}$	-	-	-	-	-	-	0.8724	0.0312	0.0010	0.8188	0.0410	0.0013
$\eta^{(3)}$	-	-	-	-	-	-	76.9845	11.8973	0.3762	9.3191	3.3499	0.1059
$\xi^{(3)}$	-	-	-	-	-	-	4.7675	2.1488	0.0680	1.0110	0.3378	0.0107
p_{11}	0.9968	0.0014	0.0000	0.9438	0.0440	0.0014	0.9921	0.0023	0.0001	0.9923	0.0018	0.0001
p_{12}	-	-	-	-	-	-	0.0073	0.0021	0.0001	0.0049	0.0016	0.0001
p_{21}	0.0099	0.0026	0.0001	0.0691	0.0483	0.0015	0.0036	0.0025	0.0001	0.0041	0.0013	0.0000
p_{22}	-	-	-	-	-	-	0.9961	0.0027	0.0001	0.9903	0.0024	0.0001
p_{31}	-	-	-	-	-	-	0.0036	0.0095	0.0003	0.0147	0.1190	0.0038
p_{32}	-	-	-	-	-	-	0.0295	0.0157	0.0005	0.9461	0.1296	0.0041
$V_{per}^{(1)}$	0.967	-	-	0.987	-	-	0.928	-	-	0.886	-	-
$V_{per}^{(2)}$	0.971	-	-	0.975	-	-	0.976	-	-	0.960	-	-
$V_{per}^{(3)}$	-	-	-	-	-	-	0.971	-	-	0.849	-	-

In addition, for a given series, the volatility reaction to negative past returns is different from one regime to another, especially for FTSE 100 with values of ($\gamma^{(1)} \simeq 0.09$) and ($\gamma^{(2)} \simeq 0.24$) for the first and the second regime, respectively.

The estimation results for Nikkei 225 also imply that in comparison to the second regime ($\gamma^{(2)} \simeq 0.13$) and the third one ($\gamma^{(3)} \simeq 0.03$), the volatility reaction to past negative returns is stronger ($\gamma^{(1)} \simeq 0.30$) within the first regime.

A significant parameter also provided in Table 4 is the volatility persistence ($V_{per}^{(k)}$).We observe a balance between the first and second regimes for S&P500 and FTSE 100 where the corresponding values are close to one, indicating highly persistent in the volatility process. Regarding the CAC40 index, the persistence is higher within the second regime ($V_{per}^{(2)} \simeq 0.98$) and the third one ($V_{per}^{(3)} \simeq 0.97$).

The estimation shown that the volatility persistence is significantly important ($V_{per}^{(2)} \simeq 0.96$) within the second regime compared with the first ($V_{per}^{(1)} \simeq 0.89$) and the second one ($V_{per}^{(2)} \simeq 0.85$) concerning the Japanese index

Nikkei 225.

We have developed our regime-switching models that shown high flexibility regarding the persistence of volatility for different stock market indices. Moreover, we would evaluate the performance of the selected models from the out-of-sample analysis point of view.

The out-of-sample period started from January 3, 2014, to October 10, 2019, with approximately 1,470 log-returns for each stock market index. A valid model should be able to correctly forecast the Value-at-Risk (VaR) for a specified coverage level. Since we use a family of models that capture time-varying parameters, we consider a wide rolling window to better forecast the one-ahead Value-at-Risk of 5% based on the models selected previously. To test the accuracy of VaR coverage, we refer to the Unconditional Coverage (UC) test of Kupiec (1995)[21] based on the number of VaR violations (or hit) [4], the Conditional Coverage (CC) test of Christoffersen (1998)[13], and the Dynamic Quantile (DQ) test of Engle and Manganelli (2004)[16] founded on the number of violations in addition to the fact that the violation variable ($I_t(\alpha)$) should be distributed independently. Meanwhile, we refer to the Basel Committee on Banking Supervision (BCBS)[6] [7] requirements regarding the internal VaR model validation from a regulatory perspective.

The results presented in Table 5 clearly demonstrate the selected model's ability to predict the VaR accurately at the 5% risk level.

Table 5: VaR Back-Test Results for 95% confidence level

	Unconditional Coverage		Conditional Coverage		Dynamic Quantile	
	uc LRstat	uc LRp	cc LRstat	cc LRp	DQstat	DQp
S&P500	0.0213	0.884	0.3056	0.858	2.3090	0.941
FTSE 100	1.9537	0.162	3.9809	0.137	4.7031	0.696
CAC40	0.1088	0.742	0.8063	0.668	6.3093	0.504
Nikkei 225	0.0736	0.786	1.7790	0.411	10.3516	0.169

According to the UC, CC, and DQ tests, the results are in accepting the null hypothesis (p-value > 0.05) of the correct forecasting of the one-ahead VaR at a 5% coverage level. We plot the results of our back-test in Figure 3. The ability of the selected models to capture the high breaks in log-returns is provided as well.

[4] $I_t(\alpha) = 1 \quad \text{if} \quad r_t < \text{VaR}_t(\alpha)$ and zero otherwise

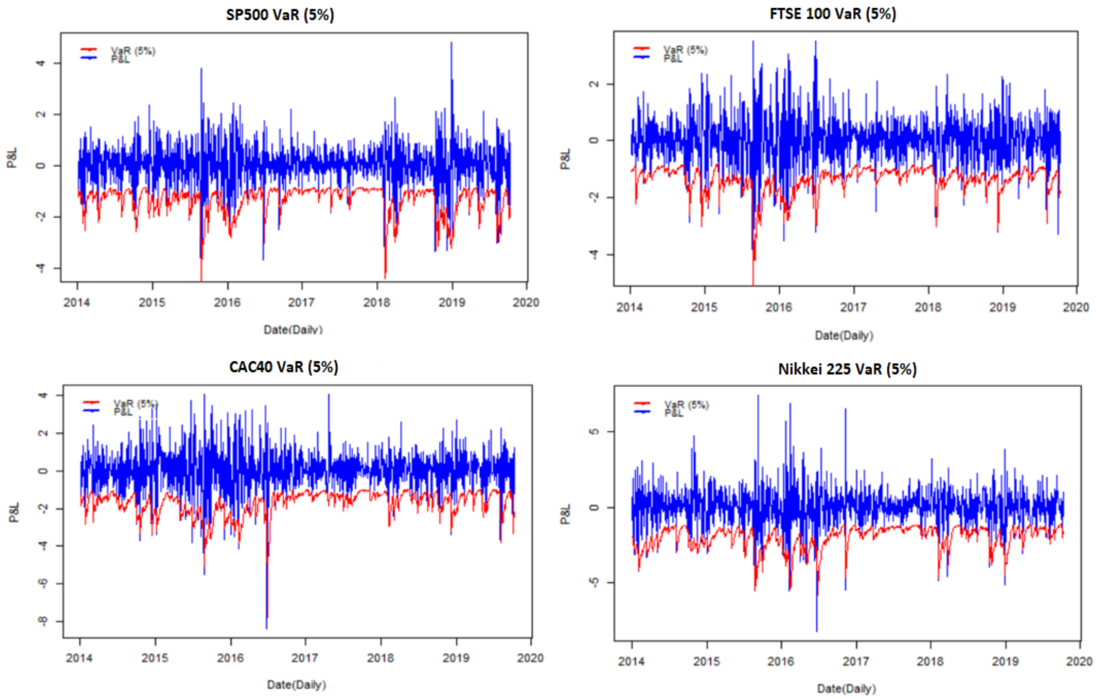

Figure 3: Stock Market Indices Value-at-Risk analysis based on the selected models.

On the other hand, for S&P500 (i.e., the widely popular stock market index), we test the selected model's forecasting accuracy to predict the VIX index. VIX is the volatility index that represents as closely as possible the implied volatility option based on the S&P500 index. Figure 4 clearly shows the flexibility of the MS GJR-GARCH (1,1) skewed GED distribution model with two regimes in capturing the VIX volatility.

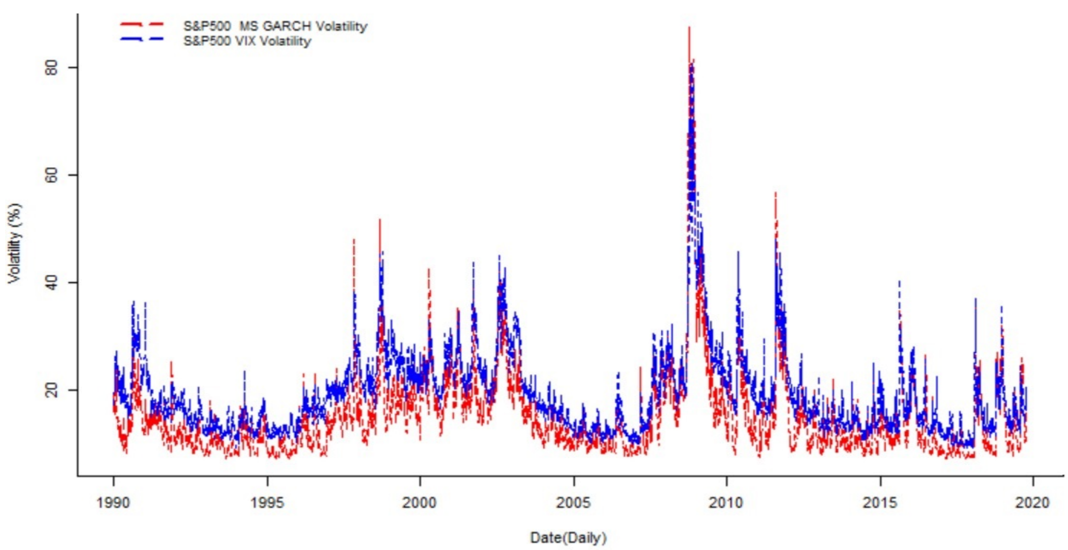

Figure 4: S&P500 MS-GARCH versus VIX volatility evolution

5. Conclusion

In this paper, we analyzed four stock market indices, namely, S&P500 (USA), FTSE 100 (UK), CAC40 (France), and Nikkei 225 (Japan), based on their daily log-returns from January 1990 until October 2019, about 30 years of data, in order to identify the optimal number of regimes using two GARCH-type specifications; the standard symmetric GARCH (1,1) against the asymmetry GJR-GARCH (1,1), with different skewed conditional distributions (Normal, student's-t and GED), where all parameters switch across a given number of regimes.

In the empirical part, we estimated models (about 72 models) using the Bayesian approach MCMC avoiding the path-dependence issue. Thus, we provided a simple comparison between these models from the Deviance Information Criterion (DIC) point of view, which measures the model's adequacy by comparing the fitting quality to model complexity. The stability of estimation is ensured by trying different seeds, where we judge the model's convergence by ourselves.

Overall, the results revealed that the asymmetry GJR-GARCH outperforms the standard symmetric GARCH(1,1) due to the presence of the leverage effect in the conditional variance process. Moreover, the regime-switching models provide more accuracy than the signal-regime, whatever the specifications are.

Furthermore, the skewed fat-tailed (Student's-t and GED) distributions are strongly preferred. In addition, the stock market indices, which are characterized by a high level of volatility (CAC40 and Nikkei 225), require three-regimes specifications; low, medium volatility regime, and the third one monitoring the high level of conditional variance with high persistence of volatility.

Finally, the validity of the selected models was verified from the out-of-sample analysis point of view, according to specific statistical tests (i.e., UC, CC, and DQ) in parallel with the Basel committee requirements and the ability to predict the S&P500 implied volatility (VIX).

Based on the empirical results, the main recommendation of the present paper is that the number of regimes in a conditional variance process should be founded on a statistical test and not chosen randomly.

An interesting topic for future research is to apply the results using the marginal likelihood since the DIC is sensitive to the prior distributions.

References

[1] Ardia, D., (2008). Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications. Springer.

[2] Ardia, D., Bluteau, K., Rüede, M., (2018). Regime changes in Bitcoin GARCH volatility dynamics, Finance Research Letters.

[3] Ardia, D., Bluteau, K., Boudt, K., Catania, L., (2018). Forecasting risk with Markov–switching GARCH models: A large–scale performance study. International Journal of Forecasting.

[4] Ardia, D., Bluteau, K., Boudt, K., Catania, L., Trottier, D. A., (2017). Markov–switching GARCH models in R: The MSGARCH Package. Journal of Statistical Software.

[5] Bauwens, Luc ; Dufays, Arnaud ; Rombouts, Jeroen., (2014). Marginal likelihood for Markov-switching and change-point GARCH models. Journal of Econometrics, Vol. 178, Part 3, p. 508-522.

[6] Basel Committee on Banking Supervision (2013). Fundamental review of the trading book: A revised market risk framework. Technical Report 265. Bank of International Settlements.

[7] Basel Committee on Banking Supervision (2016). Minimum capital requirements for market risk. techreport 352. Bank of International Settlements.

[8] B. Benaid and H. Bouzahir., (2018). A Statistical Test of Volatility Persistence in GARCH Models and their Application to Stock Exchange, Advances And Applications in Statistics, 52(6), pp 363-374.

[9] Benaid, B, (2011). Asymptotic distribution of least squares estimators for purely unstable ARMAX(p,s,q) process. International Journal of Applied Mathematics and Statistics, number J11, 1-14.

[10] Berg, A., Meyer, R., Yu, J., (2004). Deviance information criterion for comparing stochastic volatility models. Journal of Business & Economic Statistics 22 (1), 107–120.

[11] Bollerslev, Y.. (1986). Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 307–327.

[12] Box, G.E.P. and J.M. Jenkins, (1976), Time series analysis: Forecasting and control (Holden-Day, San Francisco, CA).

[13] Christoffersen PF (1998). "Evaluating Interval Rorecasts." International Economic Review, 39(4), 841-862.

[14] Dueker, M. J. (1997). Markov Switching in GARCH Processes and Mean-Reverting Stock-Market Volatility. Journal of Business & Economic Statistics, 15(1), 26–34.

[15] Engle, R.F., (1982), Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica 50, 987-1008.

[16] Engle RF and Manganelli S. (2004). "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles." Journal of Business & Economic Statistics, 22(4), 367-381.

[17] Fernández, C., Steel, M. F. J., Mar. (1998). On Bayesian modeling of fat tails and skewness. Journal of the American Statistical Association 93 (441), 359–371.

[18] Friedman, B.M. and D.1. Laibson, (1989), Economic implications of extraordinary movements in stock prices, Brookings Papers on Economic Activity 2, 137 189.

[19] Glosten, L. R., Jagannathan, R., Runkle, D. E., (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48 (5), 1779–1801.

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

[21] Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk measurement models. The journal of Derivatives, 3(2), 73-84.

[22] Markus Haas, (2004), A New Approach to Markov-Switching GARCH Models, Journal of Financial Econometrics, 2, (4), 493-530 .

[23] Hamilton, J. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica, 57(2), 357-384.

[24] James Hamilton and Raul Susmel, (1994), Autoregressive conditional heteroskedasticity and changes in regime, Journal of Econometrics, 64, (1-2), 307-333.

[25] Harvey, A., & Sucarrat, G. (2013). EGARCH models with fat tails, skewness and leverage. Computational Statistics & Data Analysis, 76 320-338.

[26] Klaassen, F., (2002). Improving GARCH volatility forecasts with regime-switching GARCH. Empirical Economics 27, 363–394 .

[27] Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change, and the GARCH model. Journal of Business and Economic Statistics, 8(2), 225-234.

[28] Lamoureux, C., & Lastrapes, W. (1993). Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities. The Review of Financial Studies, 6(2), 293-326.

[29] Marcucci, J. (2005). Forecasting stock market volatility with regimes witching GARCH models. Studies in Nonlinear Dynamics & Econometrics. 9(4), 1558–3708.

[30] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., van der Linde, A., (2002). Bayesian measures of model complexity and fit. Journal of Royal Statistical Society B 64, 585–616.

[31] Susmel, R. (2000), Switching volatility in private international equity markets. Int. J. Fin. Econ., 5: 265-283

[32] Trottier, D.-A., Ardia, D., Aug. (2016). Moments of standardized Fernández-Steel skewed distributions: Applications to the estimation of GARCH-type models. Finance Research Letters 18, 311–316.

[33] Vihola, M. Robust., (2012). Adaptive Metropolis algorithm with coerced acceptance rate. Stat Comput 22, 997–1008 .

[34] Zhu, Dongming & Galbraith, John W., (2010). "A generalized asymmetric Student-t distribution with application to financial econometrics," Journal of Econometrics, Elsevier, vol. 157(2), pages 297-305.