The frequency of regime switching in financial market volatility

Ahmed BenSaïda
LaREMfiQ Laboratory, IHEC, Sousse University, B.P. 40, Route de la ceinture, Sahloul 3, Sousse 4054, Tunisia

Article info
Article history:
Received 30 May 2014
Received in revised form 11 January 2015
Accepted 16 March 2015
Available online xxxx

JEL classification:
G17
C22
C58

Keywords:
Volatility
Risk response
Simulation
Skewed generalized t
Switching regime

Abstract

The mechanism of risk responses to market shocks is considered as stagnant in recent financial literature, whether during normal or stress periods. Since the returns are heteroskedastic, a little consideration was given to volatility structural breaks and diverse states. In this study, we conduct extensive simulations to prove that the switching regime GARCH model, under the highly flexible skewed generalized t (SGT) distribution, is remarkably efficient in detecting different volatility states. Next, we examine the switching regime in the S&P 500 volatility for weekly, daily, 10-minute and 1-minute returns. Results show that the volatility switches regimes frequently, and differences between the distributions of the high and low volatility states become more accentuated as the frequency increases. Moreover, the SGT is highly preferable to the usually employed skewed t distribution.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Financial market risk, usually measured by the volatility of the returns, is changing over time. Consequently, researchers have introduced a large variety of heteroskedastic GARCH-type models to investigate the ever fluctuating volatility. However, these models consider that the mechanism of risk responses to market’s multiple shocks remains stagnant by fixing the coefficients generating the conditional volatility. Andreou & Ghysels (2002), among others, have argued that financial returns are known to exhibit sudden jumps in their volatility, a phenomenon caused essentially by structural breaks, and cannot be captured by regime-invariant parameters such as the single-state GARCH-type models. Abdymomunov (2013) and Augustyniak (2014) have confirmed that the volatility is indeed subject to two regimes: high and low (or normal), where the high risk regime is considered as a financial stress and closely related to periods of crisis. Alternatively, Hillebrand (2005) has affirmed that the nearly integrated behavior, generally observed in classical GARCH models, is the consequence of structural changes. Besides, structural breaks in volatility dynamics can be the consequence of changes in risk perception. In fact, given the same information to an investor, the risk is perceived differently during periods of crisis with higher risk, and during normal periods with lower risk (Hoffmann et al., 2013).

The literature dealing with structural changes in volatility has emerged since the seminal paper of Hamilton & Susmel (1994). Mainly, two branches exist; some researchers consider that the risk is changing over different pre-identified periods, yet its structure is invariant during the same period by applying a time varying GARCH model (TV-GARCH), where breaks in the volatility are known, and the coefficients of the conditional volatility are held constant during the same period (Ichiu & Koyama, 2011; Karanasos et al., 2014; Liu et al., 2012). The second branch, however, considers that the volatility is subject to multiple unobservable or hidden
regimes – typically two regimes – and the transition between the different states is defined by a probability matrix, this approach applies a Markov-switching regime GARCH model, MS-GARCH henceforth (Bauwens et al., 2014; Geweke & Amisano, 2011; Marcucci, 2005). MS-GARCH models are more difficult to estimate due to the path-dependency or tractability problem explained by Cai (1994) and Hamilton & Susmel (1994). Indeed, since the regimes are unobservable, the conditional variance at time $t$ depends on all paths engendered by the different regimes, owing to the recursive property of GARCH process. Thus, the sample likelihood function is computed by integrating over all possible regime paths, which increase exponentially with time $t$ and the number of switching regimes, and its maximum becomes intractable. This problem is solved by Gray (1996) and improved by Klaassen (2002) by incorporating the conditional expectations of the lagged conditional variances into the GARCH formulation, therefore allowing the tractability of the MS-GARCH model. However, to our knowledge, all studies dealing with the switching regime volatility have merely considered weekly or at most daily returns under the normal or the Student’s $t$ distribution, a convenient assumption due to the simplicity and analytical tractability of these distributions (Ardia, 2009; Litimi & BenSaïda, 2014; Sun & Zhou, 2014; Wllfing, 2009). The asymmetry is usually captured by estimating the GJR model of Glosten et al. (1993) (Ardia, 2009; Daouk & Guo, 2004; Marcucci, 2005), although Alexander & Lazar (2009) have mentioned that the leverage effect has no influence when a skewed distribution is used. Moreover, low frequency data cannot adequately detect the rapidity of the volatility’s regime shifting. Hence, for the various aforementioned reasons, the main objective of this study is to investigate the frequency of regime switching in the financial market risk over different time scales, and under the highly flexible skewed generalized $t$ (SGT) distribution.

The contribution is threefold; first, we verify the efficiency of the MS-GARCH model under the highly flexible skewed generalized $t$ distribution in capturing both high-stress regime and normal-stress regime through extensive simulations; second, we estimate a GARCH switching regime model for weekly, daily, and intra-daily financial returns, and corroborate the likelihood tractability of the used approach; and third, we illustrate the effectiveness of the SGT over the skewed $t$ distribution. The remainder of this paper is as follow: Section 2 describes the methodology to perform a tractable MS-GARCH estimation; in Section 3 we simulate SGT pseudorandom numbers to generate high-volatility and normal-volatility returns into one single sample to confirm the effectiveness of the MS-GARCH model; Section 4 presents the data and summary statistics; Section 5 discusses the results; and finally we conclude in Section 6.

2. Concepts and methodology
2.1. Switching regime GARCH model

In general, the volatility evolves according to two different regimes: a high-stress regime, where the risk tends to be higher — generally during periods of financial crisis, and a low-stress regime, where the risk tends to be normal. As pointed out by Hoffmann et al. (2013), investors’ risk perception fluctuates significantly during financial crisis, more than it does during normal non-crisis periods. Consequently, and in alignment with Abdymomunov (2013) and Augustyniak (2014), we suspect that during these high-stress periods, the volatility tends to be higher than during normal periods; hence, we see no reason to extend our analysis for more than two regimes.

Our model is the MS-GARCH with orders $(1,1)$. There are many alternatives in choosing other GARCH-type models; however, Hansen & Lunde (2005) found no evidence that the GARCH model with orders $(1,1)$ is outperformed by other more sophisticated GARCH-type models.

Let $(s_t)$ be a state variable indicating a Markov chain, i.e., $s_t$ represents the diverse regimes for a time-dependent variable. The state variable is supposed to evolve according to a first-order Markov chain with a probability transition matrix $P$, which indicates the probability of being in state $j$ at time $t$ knowing that at time $t-1$ the state was $i$. For numerous regimes, each element of the transition matrix is defined in Eq. (1).

$$p_{ij} = \Pr(s_t = j | s_{t-1} = i)$$  \hspace{1cm} (1)

In the case of two regimes, $s_t = (1,2)$, the probability transition matrix is defined in Eq. (2), where each column sums up to one.

$$P = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \hspace{1cm} (2)$$

The ergodic probability, or the unconditional probability of being in state $s_t = 1$, is given by Eq. (3).

$$\pi_1 = \frac{1-q}{2-p-q} \hspace{1cm} (3)$$

The model to be estimated is a MS-GARCH$(1,1)$ defined in Eq. (4).

$$\begin{align*}
    r_t &= c_k + u_{t,s_t} \\
    u_{t,s_t} &= \varepsilon_{t,k} \sqrt{h_{t,s_t}} \\
    h_{t,s_t} &= \alpha_{0,k} + \alpha_{1,k} u_{t-1,s_t}^2 + \beta_{1,k} h_{t-1,s_t} \\
\end{align*} \hspace{1cm} (4)$$

Please cite this article as: BenSaïda, A., The frequency of regime switching in financial market volatility, J. Empir. Finance (2015), http://dx.doi.org/10.1016/j.jempfin.2015.03.005
where $r_t$ is the return of the financial index, and $c_k$ is a drift term representing different means of the returns depending on the $k$th regime ($k = 1, \ldots, K$). All coefficients are constant and time-invariant within the same unknown regime $k$, but are different from one regime to another. Klauaassen (2002) has developed a framework to obtain tractable maximum likelihood estimation by using a state-independent average of past conditional variance; so in the variance equation, past volatility $h_{t-1}$ is replaced by its expectation given the information available at time $t - 1$ in Eq. (5), where $i,j = 1, 2$.

$$E_{t-1}(h_{t-1,1} | s_t) = \bar{p}_{i,i-1} \left( \eta_i^2 + h_{t-1,1} \right) + \bar{p}_{j,i-1} \left( \eta_j^2 + h_{t-1,1} \right) - \left( \bar{p}_{i,i-1} \eta_i + \bar{p}_{j,i-1} \eta_j \right)^2$$

(5)

And the probabilities are computed as $p_{j,t} = \frac{p_j p_{jt}}{p_0}$, where $p_{jt}$ is the ex-ante probability such that $p_{jt} = \Pr(s_t = j | h_{t-1})$, which is the probability of being in regime $j$ at time $t$ given the information available at time $t - 1$.

$$p_{1,t} = \frac{p \left[ f(r_{t-1} | s_t = 1) p_{1,t-1} \right]}{f(r_{t-1} | s_t = 1) p_{1,t-1} + f(r_{t-1} | s_t = 2) \left( 1 - p_{1,t-1} \right)}$$

(6)

where $f(\cdot)$ stands for the likelihood function.

A more analytically tractable version of the MS-GARCH model is given by Haas et al. (2004). They assume that each state is characterized by its own volatility, and that each volatility process depends only on its lagged values and the squared residuals within every regime. We adopt this method since it overcomes the problems of severe estimation difficulty, and fuzzy dynamic properties usually found in Gray’s (1996) and Klauaassen’s (2002) frameworks. More precisely, suppose we have $K$ regimes, the conditional variances of a MS-GARCH($1,1$) model at time $t$ can be written as:

$$\begin{pmatrix} h_{1,t} \\ \vdots \\ h_{K,t} \end{pmatrix} = \left( \begin{array}{c} \alpha_{0,1} \\ \vdots \\ \alpha_{0,K} \end{array} \right) + \left( \begin{array}{c} \alpha_{1,1} \\ \vdots \\ \alpha_{1,K} \end{array} \right) \cdot \left( \begin{array}{c} r_{t-1} - \left( \begin{array}{c} \approx_1 \\ \vdots \\ \approx_K \end{array} \right) \right)^2 + \left( \begin{array}{c} \beta_{1,1} \\ \vdots \\ \beta_{1,K} \end{array} \right) \cdot \left( \begin{array}{c} h_{1,t-1} \\ \vdots \\ h_{K,t-1} \end{array} \right)$$

(7)

where $\circ$ denotes the Hadamard product or element-by-element multiplication. The coefficients $\alpha_{0,k}$, $\alpha_{1,k}$, $\approx_k$, and $\beta_{1,k}$ are relative to the $k$th regime ($k = 1, \ldots, K$). This formulation has conceptual benefits; in fact, the Markov-switching regime models are primarily designed to dissociate between diverse states and to allow for different GARCH parameterizations in each regime, thus, to detect the changes in volatility dynamics during high and low stress periods. Therefore, a clear association between the parameters $\alpha_{0,k}$, $\alpha_{1,k}$, $\approx_k$, and $\beta_{1,k}$ within regime $k$ and the corresponding $h_{k,t}$ allows for a straightforward interpretation of volatility dynamics.

2.2. The skewed generalized t distribution

The error term $e_{t, s_t}$ in Eq. (4) is independently and identically distributed (IID) with mean $E(e_{t, s_t}) = 0$, and variance $V(e_{t, s_t}) = 1$. The distribution used in this paper is the skewed generalized t (SGT) introduced by Theodossiou (1998) in Theorem 3, and presented in Eq. (8). As far as we know, the SGT has never been used in a switching regime GARCH model, although it is a highly flexible univariate distribution. Cheng & Hung (2011) used it to forecast the Value-at-Risk (VaR) of petroleum and metal asset using a GARCH model, and found a superior predictive power over the generalized error and normal distributions. Applications to financial data, however, are limited; Harris et al. (2004) used it to model oil international indexes and found a substantial improvement in the GARCH and EGARCH models over the generalized error, Student’s $t$, and normal distributions; Bali & Theodossiou (2007) used it to estimate various GARCH-type models applied on daily S&P 500 and compared their performances in computing accurate VaR. Allowing for a more flexible distribution within each regime enhances the stability of the states and permits to focus on the conditional variance’s behavior instead of capturing some irregularities.

$$f(x; \eta, \psi, \lambda, \mu, \sigma) = \frac{\eta}{2\pi \sigma \beta} \left( \frac{\psi}{\eta} \right)^{\frac{\lambda | x - \mu |^\psi}{1 + | x - \mu |^\psi}}$$

(8)

The SGT is defined for any $x \in \mathbb{R}$. The shape parameters $\eta > 0$ and $\psi > 2$, and the skewness parameter $|\lambda| < 1$. $B(\cdot)$ is the beta function, and $\text{sgn}(\cdot)$ is the sign function. The term $\mu$ is a location parameter, and $\theta$ is a normalizing constant to ensure that the variance of $x$ equals $\sigma^2$. The moment function about $\mu$ of the SGT, denoted $\mathbb{M}_r$, is defined in Eq. (9).

**Lemma 1.** Suppose that $x$ follows a SGT as described in Eq. (8), the $r$-th moment about $\mu$, provided that $r < \psi$, is given by:

$$\mathbb{M}_r = E(x - \mu)^r = \frac{\sigma^r \theta^r B \left( \frac{1 + r}{\eta} \frac{\psi - r}{\eta} \right)}{2\theta \left( \frac{1 + r}{\eta} \frac{\psi - r}{\eta} \right)} \left[ (1 + \lambda)^{1+r} + (-1)^r (1-\lambda)^{1+r} \right]$$

(9)

Please cite this article as: BenSaïda, A., The frequency of regime switching in financial market volatility, J. Empir. Finance (2015), http://dx.doi.org/10.1016/j.jempfin.2015.03.005
**Proof.** The moment function can be easily derived by applying Gradshteyn & Ryzhik (2007, p. 322), Section 3.241.4,\[
\int_0^\infty \frac{x^p}{(1 + ax)^q} dx = \frac{1}{q a^{p+1}} B \left( \frac{p+1}{q}, n-p+1 \right), \quad \text{under the conditions that} \quad -1 < p < n q - 1 \quad \text{and} \quad a \neq 0.
\]

The central moment function about the mean \(m\), denoted \(M_r\), is obtained by a simple application of the binomial theorem, and is presented in Eq. (10).

**Lemma 2.** The central moment function of the SGT is defined as:

\[
M_r = E(x-m)^r = \sum_{i=0}^r \binom{r}{i} (\mu-m)^{r-i} M_i.
\]

**Proof.**

\[
E(x-m)^r = E(x-\mu-m+\mu)^r = E \left[ \sum_{i=0}^r \binom{r}{i} (\mu-m)^{r-i} (x-\mu)^i \right] = \sum_{i=0}^r \binom{r}{i} (\mu-m)^{r-i} E(x-\mu)^i.
\]

We can next deduce \(\theta\) in Eq. (11) such that \(\text{Var}(x) = \sigma^2\).

\[
\theta = \frac{B \left( \frac{1}{\eta}, \frac{\psi}{\eta} \right)}{\sqrt{(1 + 3\lambda^2) B \left( \frac{1}{\eta}, \frac{\psi}{\eta} \right) B \left( \frac{2}{\eta}, \frac{\psi-2}{\eta} \right) - 4\lambda^2 B \left( \frac{2}{\eta}, \frac{\psi-1}{\eta} \right)^2}}
\]

The relationship between the location parameter \(\mu\) and the mean \(m\) is presented in Eq. (12). Furthermore, to be suitable for GARCH modeling, the mean and variance of the unconditional distribution must equal 0 and 1 respectively. So we shall set \(m = 0\) and \(\sigma = 1\) for the remaining of this paper.

\[
\mu = m - 2\sigma \lambda \frac{B \left( \frac{2}{\eta}, \frac{\psi-1}{\eta} \right)}{B \left( \frac{1}{\eta}, \frac{\psi}{\eta} \right)}
\]

The asymmetric generalized t (SGT) can considerably improve the efficiency gain in many conventional models (Hansen et al., 2010). Indeed, it can nest a large variety of other distributions including the generalized t distribution (GTD) (McDonald & Newey, 1988), Hansen’s (1994) skewed t, and the generalized error distribution (GED), among many others (Fig. 1). Adding the parameter \(\lambda\) allows for skewness in the distribution and increases the performance of risk measuring models (Haas, 2010). When \(\lambda > 0\), the SGT becomes skewed to the right, so large positive changes are more likely to occur; and when \(\lambda < 0\), the SGT becomes skewed to the left, so large negative changes are more likely to occur, see Theodossiou (1998) for further details.

2.3. Numerical approach

The distribution’s specific coefficients to be estimated are the shape parameters \(\eta\) and \(\psi\), and the skewness parameter \(\lambda\). For each regime we have different parameters, which means different distributions. To jump start the numerical optimization, the initial values of \(\eta, \psi, \lambda\) are the closest solution to the theoretical skewness and kurtosis of the SGT given in Eqs. (13) and (14), both are derived from the central moment function defined in Eq. (10), i.e., these theoretical equations are set to equal the sample skewness and kurtosis, respectively, and are solved for \(\eta, \psi, \lambda\).

\[
\text{Skewness} = 4\theta^3 \lambda \left( 1 + \lambda^2 \right) \left[ \frac{B \left( \frac{4}{\eta}, \frac{\psi-3}{\eta} \right)}{B \left( \frac{1}{\eta}, \frac{\psi}{\eta} \right)} - 2\theta \lambda \frac{B \left( \frac{2}{\eta}, \frac{\psi-1}{\eta} \right)}{B \left( \frac{1}{\eta}, \frac{\psi}{\eta} \right)} \right]^2
\]

1 The GED is also known as the Box–Tiao, or the Subbotin’s power exponential, or the generalized Gaussian distribution.

2 The parameter \(\theta\) satisfies Eq. (11).
Kurtosis \( \theta = \frac{\lambda}{\eta} + \frac{5}{\eta^2} + \frac{\lambda^2}{\eta^2} / C_{16}/C_{17} \eta; \psi - 4 \eta / C_{18}/C_{19} \eta / C_{16}/C_{17} \eta^2 \psi - \frac{4}{\eta} \eta / C_{18}/C_{19} \eta^2 \psi - \frac{3}{\eta} \eta / C_{2} \bigg)\end{equation}

Existing numerical procedures to estimate switching regime models are essentially based on the expectation–maximization (EM) algorithm developed by Hamilton (1990). Unfortunately, the EM algorithm cannot be implemented while estimating MS-GARCH models (Hamilton & Susmel, 1994), since it requires a constant and known volatility. Even when the volatility is constant and unknown, it is impossible to deploy the EM algorithm to estimate the volatility jointly with other parameters (Hahn et al., 2010). Some researchers used the Bayesian estimation based on the Markov chain Monte Carlo (MCMC) technique and the Gibbs sampler procedure to overcome path dependency (Augustyniak, 2014; Bauwens et al., 2014; Billio et al., in press; Hahn et al., 2010). However, to achieve convergence of the simulated distribution to the posterior one, we need to iteratively generate an arbitrary "sufficient" large number of draws, which can be time consuming for complex problems such as the MS-GARCH, and the regimes must be sufficiently separated to be identified. Therefore, we will use a direct approach to compute the maximum log-likelihood, which consists of directly maximizing the function given in Eq. (15).

\[
\text{LLF} = \sum_{t=1}^{T} \ln \left[ p_{1,t} f(r_t | s_t = 1) + p_{2,t} f(r_t | s_t = 2) \right]
\]

where \( f(r_t | s_t = i) \) is the conditional density function given that regime \( i \) occurs at time \( t \).

We develop the msgarch toolbox from scratch under MATLAB® programming language to conduct the estimations (The MathWorks, Inc., 2014); it can handle a large variety of parameterizations. The msgarch toolbox is composed of many functions that can estimate and investigate different univariate ARMAX/GARCH-type models under the SGT and its nested distributions, and for multiple hidden Markov-switching regimes. To enhance estimation speed, some iterative subroutines are written in C++ and compiled to be directly callable from MATLAB functions.

Fig. 1. Skewed generalized t distribution tree.
3. Simulations

To the best of our knowledge, the effectiveness of MS-GARCH models has never been verified. Existing literature merely compared a single-state model to a multi-state model and found a superior performance of the switching regime GARCH model. In this section, we simulate different returns under two regimes: a high-stress regime and a low-stress regime. Then, we combine both paths of each regime to form one single two-regime sample. Finally, we estimate the constructed sample using a MS-GARCH and compare the average fitted coefficients to the initial ones via the Wald test. The effectiveness of the MS-GARCH model is derived from its ability in correctly estimating the true coefficients and distributions of both regimes.

The procedure consists of generating 3 different sets of simulated returns as described in Eq. (4), each set contains 1000 samples. For the first set, every sample contains 3000 observations, for the second set 12,000 observations, and for the last set 30,000 observations. The high-volatility regime represents one-third of the total sample. The error term \( \varepsilon_t \) has a GARCH model with estimated persistence parameters \( \alpha_i \). The Wald test, at 5% significance level, is applied on the re-estimated coefficients for each sample to compare similarities between the original and the re-estimated model. The rate of acceptance of the null hypothesis, that all estimated coefficients equal the initial ones, is further reported.

3.1. Distribution and quantile functions of the SGT

Although the SGT has been introduced by Theodossiou (1998), and lately studied by Arslan & Genç (2009) and Nadarajah & Pogany (2012), none of these studies has correctly developed its cumulative distribution function (CDF), which is crucial in generating pseudorandom numbers. In this paper, we develop a closed form CDF of the SGT in Eq. (16) along with the quantile function in Eq. (17).

**Theorem 1.** Let \( x \) be a random variable following a SGT as described in Eq. (8), the corresponding distribution function or CDF has the closed form:

\[
F(x) = \frac{1}{2} - \frac{\lambda + \text{sgn}(x - \mu)}{2} \int_{\frac{\mu - \lambda}{\sqrt{1 - \lambda^2}}}^{\frac{1}{\sqrt{1 - \lambda^2}}} \left( \frac{1}{\eta^2} \right) \Psi \, \text{d}t
\]

where \( \text{sgn}(x) \) is the sign function, \( I_z(a, b) \) is the regularized incomplete beta function that satisfies: \( I_z(a, b) = \frac{1}{B(a, b)} B_z(a, b) \), with \( B_z(a, b) = \int_0^z t^{a-1}(1-t)^{b-1} \, \text{d}t \) is the incomplete beta function.

**Proof.** We can derive the above CDF by applying Gradshteyn & Ryzhik (2007, p. 322), Section 3.241.4.

**Theorem 2.** The quantile function of the SGT for any \( z \in [0, 1] \) is given by:

\[
F^{-1}(z) = \mu + \sigma \left[ \frac{\text{sgn} \left( z \frac{1-\lambda}{2} \right) + \lambda}{\left[ I_{-\frac{\lambda}{\sqrt{1-\lambda^2}}} \left( \frac{1}{\eta^2} \right) \right]^{-1} - 1} \right]
\]

where \( I_z^{-1}(a, b) \) is the inverse regularized incomplete beta function.

**Proof.** If we have \( F(x) = z \), then the inverse CDF is simply \( x = F^{-1}(z) \). After some manipulations, and noting that when \( (x - \mu) \geq 0 \) we have \( (z - \frac{1-\lambda}{2}) \geq 0 \), we derive the quantile function. Detailed proof is in Appendix B.

3.2. Simulation results

The results are summarized in Table 1, every average of 1000 coefficient estimates in each set is compared to the initial one used for the simulation. The Wald test, at 5% significance level, is applied on the re-estimated coefficients for each sample to compare similarities between the original and the re-estimated model. The rate of acceptance of the null hypothesis, that all estimated coefficients except the transition probabilities equal the initial ones, is further reported.

---

4 Actually, Arslan & Genç (2009) have introduced the CDF of a modified version of the SGT; their CDF, however, lacks detailed proof and is far from the distribution function’s usual S shape.

5 Uniform pseudorandom numbers are generated using the Mersenne-Twister algorithm with a super-astronomical period of \( 2^{19937} - 1 \), the initial seed is set to the computer clock at the time when the simulations are conducted to ensure maximum efficiency.
The MS-GARCH model is capable, at an acceptable rate, of detecting the regimes that govern the returns. For the first set \((N = 3000)\), the averages of the estimated coefficients are slightly far from the initial ones, and the Wald null hypothesis acceptance rate is rather low. For the reason that the high stress regime contains one third of the sample, which makes only 1000 observations, while the model contains a relatively large number of parameters, so the initial coefficients cannot be correctly discovered. As we increase the sample size, the goodness of fit, abridged by the Wald test acceptance rate, is improved. In fact, the results are better for \(N = 12,000\), and are best for \(N = 30,000\) observations. In this latter case, we have estimated all parameters at a remarkable precision, even the unconditional probability of the high-volatility regime, which is almost one-third of the simulated sample size.

4. Data and summary statistics

The data set is the Standard & Poor's 500 (S&P 500) from January 1, 2002 until June 30, 2014. We collected four different frequencies from DataStream: weekly, daily, 10-minute, and 1-minute. Analyzing different frequencies for a single index can be informative in understanding the dynamics of switching regime in the volatility. We couldn't extend the sample prior to 2002 due to the limited availability of 1-minute data. Index prices are known to be non-stationary, so we compute the returns. If \(P_t\) is the index, the continuous return is given by the natural logarithmic difference \(r_t = \ln(P_t) - \ln(P_{t-1})\).

Descriptive statistics of the returns are presented in Table 2. The average return of one week is greater than of one day, and so on. As expected, normality is overwhelmingly rejected for all frequencies by the Jarque–Bera test. The distributions of the returns are asymmetric and skewed to the left, except for intraday returns, where the distributions are skewed to the right. As argued by Alexander & Lazar (2009), conditional skewness and kurtosis of financial time series are time varying, and better captured by a MS-GARCH model.

Historical prices of the S&P 500 for all frequencies are plotted in Fig. 2. The returns are plotted in Fig. 3. As the frequency increases, the plot becomes more and more detailed, since it contains more information about the price. For example, one observation from weekly data (Fig. 2a) contains, on average, information about 5 trading days (Fig. 2b), and one daily observation contains thirty-six

Table 1
Simulation results.

<table>
<thead>
<tr>
<th></th>
<th>Coefficients</th>
<th>Initial</th>
<th>Average estimated for (N = 3000) obs.</th>
<th>Average estimated for (N = 12,000) obs.</th>
<th>Average estimated for (N = 30,000) obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Regime 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(High volatility)</td>
<td>(\alpha_0)</td>
<td>0.02</td>
<td>0.016752</td>
<td>0.019274</td>
<td>0.019981</td>
</tr>
<tr>
<td></td>
<td>(\alpha_1)</td>
<td>0.05</td>
<td>0.041805</td>
<td>0.049902</td>
<td>0.049947</td>
</tr>
<tr>
<td></td>
<td>(\beta_1)</td>
<td>0.85</td>
<td>0.88863</td>
<td>0.94743</td>
<td>0.95029</td>
</tr>
<tr>
<td></td>
<td>(\lambda)</td>
<td>4</td>
<td>4.822</td>
<td>3.9603</td>
<td>3.9982</td>
</tr>
<tr>
<td></td>
<td>(\phi)</td>
<td>8</td>
<td>12.1742</td>
<td>8.0252</td>
<td>8.0099</td>
</tr>
<tr>
<td></td>
<td>(\psi_4)</td>
<td>-0.2</td>
<td>-0.1753</td>
<td>-0.1987</td>
<td>-0.2007</td>
</tr>
<tr>
<td></td>
<td>(\psi_8)</td>
<td>7</td>
<td>10.2079</td>
<td>7.1032</td>
<td>7.0102</td>
</tr>
<tr>
<td></td>
<td>(\lambda)</td>
<td>-0.1</td>
<td>-0.0732</td>
<td>-0.0994</td>
<td>-0.1001</td>
</tr>
<tr>
<td><strong>Transition matrix</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(P_{1,1}) = p</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(P_{2,2}) = q</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Unconditional probability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\pi_1) of regime 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/3</td>
<td>0.4562</td>
<td>0.3437</td>
<td>0.3416</td>
</tr>
<tr>
<td><strong>Wald coefficient test</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(H_0) acceptance rate</td>
<td></td>
<td></td>
<td>63.3%</td>
<td>88.1%</td>
<td>94.5%</td>
</tr>
</tbody>
</table>

Table 2
Descriptive statistics.

<table>
<thead>
<tr>
<th>S&amp;P 500 return</th>
<th>One week</th>
<th>One day</th>
<th>10-minute</th>
<th>1-minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.00083486</td>
<td>0.00015686</td>
<td>4.3629e-06</td>
<td>4.6908e-07</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>0.025153</td>
<td>0.01222</td>
<td>0.0018956</td>
<td>0.00055164</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.85102</td>
<td>-0.21255</td>
<td>0.083904</td>
<td>0.93956</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>11.134</td>
<td>13.621</td>
<td>69.447</td>
<td>406.94</td>
</tr>
<tr>
<td>Minimum</td>
<td>-0.20084</td>
<td>-0.094695</td>
<td>-0.062162</td>
<td>-0.047458</td>
</tr>
<tr>
<td>Maximum</td>
<td>0.11356</td>
<td>0.10957</td>
<td>0.04908</td>
<td>0.04181</td>
</tr>
<tr>
<td>Number of observation</td>
<td>648</td>
<td>3453</td>
<td>122,379</td>
<td>1,136,971</td>
</tr>
<tr>
<td>Jarque–Bera statistic</td>
<td>1833*</td>
<td>16.203*</td>
<td>2.25e+07*</td>
<td>7.73e+09*</td>
</tr>
</tbody>
</table>

* The null “data are normally distributed” is rejected at 5% confidence level.

Please cite this article as: BenSaïda, A., The frequency of regime switching in financial market volatility, J. Empir. Finance (2015), http://dx.doi.org/10.1016/j.jempfin.2015.03.005
5. Results

Generally, when dealing with high-frequency financial data, microstructure noises must be taken into consideration, since they contaminate the volatility measures. Indeed, as argued by Anderson et al. (2005), realized volatility, which is the sum of intraday squared returns, is a consistent yet a noisy estimator of the integrated (actual) volatility. The mean of the noise becomes more negligible as the frequency increases, and filtering the integrated volatility from the realized volatility will have a smaller impact on its quality as the time scale over which the returns are calculated decreases. A practical investigation of this statement is conducted by Bai et al. (2004), where they studied the effect of the frequency of financial intraday data on daily volatility estimates, using the mean square error (MSE) as a gauge for unconditional variance estimators, and found that the minimum MSE occurs at 10-minute interval and that the MSE is monotonically increasing for larger or shorter sampling intervals. In addition, since our objective is not to approximate the daily integrated volatility via the intraday realized volatility, a small interest is given to microstructure effect as argued by Hansen & Lunde (2005).

Table 3 summarizes the results for a single regime model, Tables 4 and 5 present the results for a two-regime model under the SGT and the skewed t, respectively. Asymptotic t-statistics of the coefficients are between parentheses. The comparison between the single-state and the two-state models is performed through the Schwarz information criterion (SIC), a lower SIC means a better model. For each frequency, the elapsed estimation time is reported in the last row of every table. All estimations are performed using a 2.40 GHz Intel quad-core i7 computer with 8GB of RAM.

Under one regime, all estimations are successful for all frequencies, and all coefficients are statistically significant at 5% confidence level. Moreover, estimations for single-regime models did not consume large computer time, roughly 117 s for a sample size larger than one million observations. The flexibility of the SGT, represented by the shape parameters $\eta$ and $\phi$, and the skewness parameter $\lambda$, allows the conditional distribution to fit the data under different frequencies. Nevertheless, the sample skewnesses for intraday returns are positive, and the respective estimated As are negative. In fact, Peiró (1999) and Harris & Küçüközmen (2001) have noticed that the sample skewness is severely distorted when the data are highly leptokurtic; therefore, it doesn't have a well-established distribution, which generates inference about its true value. In addition, removing outliers from the data considerably affects the sample skewness. For example, removing outliers that exceed, in absolute value, twelve standard deviations from the 10-minute data (47 out

---

6 The stock market is 6.5 hours trading and not 24 hours trading like the FX market.
7 Also called Bayesian criterion. $SIC = -2 \ln(\text{LLF}) + n \ln(T)$, where $n$ is the number of estimated coefficients.

Fig. 2. S&P 500 index prices for all frequencies.
of 122,379 observations) and 1-minute data (642 out of 1,136,971 observations) reduces the sample skewnesses to −0.1162 and −0.0275, respectively. So differences between fitted and sample skewness coefficients for highly leptokurtic samples are likely to be of limited significance.

Estimations under two regimes are successful for all frequencies, since each likelihood function did converge to a solution without any problem of tractability. Furthermore, the estimation time is judged extremely convenient relative to the MCMC and Gibbs sampler technique employed by Augustyniak (2014) among others, where it took 30 min to estimate a data set of 5000 observations on a computer with higher performance, against only 7 s with the approach presented in this paper, and for the same simulated sample size. Indeed, the Gibbs sampler technique is an iterative simulation-based algorithm, it is computationally intensive because the conditional variance at every step must be recalculated for each simulated state sequence whenever the parameters change (Bauwens et al., 2015).

Table 3
Single regime GARCH(1,1) under the SGT.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>One week</th>
<th>One day</th>
<th>10-minute</th>
<th>1-minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>0.0047056</td>
<td>0.0022599</td>
<td>4.8483e-05</td>
<td>3.2488e-06</td>
</tr>
<tr>
<td>($39.8535)^*$</td>
<td>($211.827)^*$</td>
<td>($40.4206)^*$</td>
<td>($35.6957)^*$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>0.00021972</td>
<td>2.4965e-05</td>
<td>2e-07</td>
<td>3.0431e-08</td>
</tr>
<tr>
<td>($3.5865)^*$</td>
<td>($5.3848)^*$</td>
<td>($44.8323)^*$</td>
<td>($253.093)^*$</td>
<td></td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.35953</td>
<td>0.12145</td>
<td>0.10714</td>
<td>0.24005</td>
</tr>
<tr>
<td>($5.1507)^*$</td>
<td>($5.4388)^*$</td>
<td>($46.1539)^*$</td>
<td>($674.697)^*$</td>
<td></td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.64047</td>
<td>0.87855</td>
<td>0.89286</td>
<td>0.75994</td>
</tr>
<tr>
<td>($13.4861)^*$</td>
<td>($55.201)^*$</td>
<td>($503.985)^*$</td>
<td>($1031.07)^*$</td>
<td></td>
</tr>
<tr>
<td>$\eta_1$</td>
<td>0.62195(10.8109)$^*$</td>
<td>0.484(25.8399)$^*$</td>
<td>0.77895(168.979)$^*$</td>
<td>0.71927(636.695)$^*$</td>
</tr>
<tr>
<td>$\phi$</td>
<td>21.0967</td>
<td>10.1178</td>
<td>4.736</td>
<td>5.8829</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.26836</td>
<td>0.15725</td>
<td>0.2602</td>
<td>0.23442</td>
</tr>
<tr>
<td>Maximum log-likelihood</td>
<td>1710.930</td>
<td>11935.038</td>
<td>672248.45</td>
<td>77799.93</td>
</tr>
<tr>
<td>Schwarz criterion</td>
<td>-3376.543</td>
<td>-23813.047</td>
<td>-134414.9</td>
<td>-15559.889</td>
</tr>
<tr>
<td>SIC per observation</td>
<td>-5.211</td>
<td>-6.896</td>
<td>-10.986</td>
<td>-13.685</td>
</tr>
<tr>
<td>Elapsed estimation time</td>
<td>0.45 s</td>
<td>0.83 s</td>
<td>8.00 s</td>
<td>117.4 s</td>
</tr>
</tbody>
</table>

Note: asymptotic t-statistics of the estimated coefficients are between parentheses.  
* Significant at 5% confidence level.

Please cite this article as: BenSaïda, A., The frequency of regime switching in financial market volatility, J. Empir. Finance (2015), http://dx.doi.org/10.1016/j.jempfin.2015.03.005
of the ef
of being in the high stress regime 1 at time volatility state since its conditional risk is greater than the conditional risk of the normal stress regime 2. The smoothed probabilities are inferred using the backward algorithm (Klaassen, 2002). The massive flexibility in adapting shapes to different volatility states.

The conditional distributions of the index returns have different shapes and are far from the normal. In fact, the SGT shape parameter \( \beta \) is negative in both regimes, so the distributions are skewed to the left. Differences between the two-regime distributions are barely noticeable for weekly data and become more accentuated as the frequency of the returns increases. This is due to the improved ability into detecting the volatility dynamics for sufficiently large sample size. Fig. 4 is a visual ascertainment of the efficiency of the SGT compared to a low persistence model. The skewness parameter \( \lambda \) is less than one in both regimes, which indicates a rapidity of regime switching or low persistence. In fact, for weekly data the expected duration in state 1 is 1.22 days, while for 10-minute and 1-minute frequencies it is 16.03 days and 7.43 days, respectively. The unconditional probability of regime 1 is 0.63913, which is significantly greater than 0.5 at 5% confidence level. The maximum log-likelihood is 1743.95 for one week, 12,258.29 for one day, 679,875.6 for 10-minute, and 7,865,049 for 1-minute frequencies. The elapsed estimation time is 2.3 s, 5.7 s, 11.093 s, and 13.8349 s, respectively.

Table 4
Switching regime GARCH(1,1) under the SGT.

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>One week</th>
<th>One day</th>
<th>10-minute</th>
<th>1-minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regime 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c )</td>
<td>(-0.01401)</td>
<td>0.0029788</td>
<td>(-3.3682e-05)</td>
<td>2.4564e-06</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>0.00014928</td>
<td>3.4162e-05</td>
<td>1.6405e-06</td>
<td>5.175e-08</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>(1.5818)</td>
<td>(3.0113)</td>
<td>(21.0022)</td>
<td>(119.635)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.71384</td>
<td>0.92674</td>
<td>0.89618</td>
<td>0.83433</td>
</tr>
<tr>
<td>( \eta )</td>
<td>7.902</td>
<td>5.5109</td>
<td>2.9871</td>
<td>2.1287</td>
</tr>
<tr>
<td>( \psi )</td>
<td>(7307.16)</td>
<td>(27.0187)</td>
<td>(3865.86)</td>
<td>(5605.87)</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>(-0.36149)</td>
<td>(-0.18148)</td>
<td>(-0.32363)</td>
<td>(-0.24999)</td>
</tr>
<tr>
<td>Expected duration</td>
<td>1.22</td>
<td>2.14</td>
<td>16.03</td>
<td>4.70</td>
</tr>
<tr>
<td>Regime 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( c )</td>
<td>0.0067123</td>
<td>(-0.0038853)</td>
<td>6.9806e-05</td>
<td>(-3.7755e-07)</td>
</tr>
<tr>
<td>( \alpha_0 )</td>
<td>(38.4259)</td>
<td>(219.03)</td>
<td>(66.474)</td>
<td>(-2.2658)</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>9.8307e-05</td>
<td>7.7805e-05</td>
<td>4.7806e-07</td>
<td>3.0431e-08</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>(3.7943)</td>
<td>(3.1727)</td>
<td>(13.144)</td>
<td>(101.23)</td>
</tr>
<tr>
<td>( \eta )</td>
<td>0.27255</td>
<td>0.24782</td>
<td>0.06018</td>
<td>0.07613</td>
</tr>
<tr>
<td>( \psi )</td>
<td>(5.6076)</td>
<td>(55.699)</td>
<td>(12.6285)</td>
<td>(110.755)</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>(-0.5252)</td>
<td>(-12.7596)</td>
<td>(-143.567)</td>
<td>(-477.794)</td>
</tr>
<tr>
<td>Expected duration</td>
<td>2.99</td>
<td>1.14</td>
<td>17.68</td>
<td>7.43</td>
</tr>
<tr>
<td>Transition matrix</td>
<td>( P_{0,1} = p )</td>
<td>( P_{0,2} = q )</td>
<td>( P_{1,2} = P_{2,1} )</td>
<td></td>
</tr>
<tr>
<td>( P_{0,1} )</td>
<td>0.5336</td>
<td>0.9376</td>
<td>0.4756</td>
<td>0.3874</td>
</tr>
<tr>
<td>( P_{0,2} )</td>
<td>0.6519</td>
<td>0.9434</td>
<td>0.3874</td>
<td>0.3874</td>
</tr>
<tr>
<td>( P_{1,2} )</td>
<td>0.1265</td>
<td>0.9434</td>
<td>0.3874</td>
<td>0.3874</td>
</tr>
<tr>
<td>( P_{2,1} )</td>
<td>0.1265</td>
<td>0.9434</td>
<td>0.3874</td>
<td>0.3874</td>
</tr>
<tr>
<td>Unconditional probability ( \pi )</td>
<td>0.082513</td>
<td>0.92674</td>
<td>0.82513</td>
<td>0.92674</td>
</tr>
<tr>
<td>Maximum log-likelihood</td>
<td>1743.95</td>
<td>12,258.29</td>
<td>679,875.6</td>
<td>7,865,049</td>
</tr>
<tr>
<td>Schwarz criterion</td>
<td>(-3371.38)</td>
<td>(-24,369.93)</td>
<td>(-13,595,540.3)</td>
<td>(-15,729,848)</td>
</tr>
<tr>
<td>SIC per observation</td>
<td>5.2027</td>
<td>11.1093</td>
<td>10.975</td>
<td>10.975</td>
</tr>
<tr>
<td>Elapsed estimation time</td>
<td>2.3 s</td>
<td>11.093 s</td>
<td>91.5 s</td>
<td>14.8 min</td>
</tr>
</tbody>
</table>

Note: asymptotic t-statistics of the estimated coefficients are between parentheses.

* Significant at 5% confidence level.

** Significant at 10% confidence level.

The transition probabilities are statistically significant at 5% confidence level. When comparing with the single-regime estimation (Table 5), volatility switches regimes for all studied frequencies. For weekly returns, however, the single regime GARCH outperforms the switching regime based on the SIC. Although the maximum log-likelihood is greater for the two-state model, the gain in efficiency is not enough to compensate the loss due to the increased number of estimated coefficients.

The conditional distributions of the index returns have different shapes and are far from the normal. In fact, the SGT shape parameters \( \eta \) and \( \psi \) are both different from one regime to another, and the parameter \( \eta \) is less than one in both regimes, which indicates a highly leptokurtic distribution as shown in Fig. 4, a plot of the unconditional distributions of both regimes compared to the normal. The skewness parameter \( \lambda \) is negative in both regimes, so the distributions are skewed to the left. Differences between the two-regime distributions are barely noticeable for weekly data and become more accentuated as the frequency of the returns increases.

Fig. 5 illustrates the conditional risk, represented by the conditional standard deviations, in both regimes. The smoothed probabilities of being in the high stress regime 1 at time \( t \), \( Pr(s_t = 1 | T) \), given the information available at time \( T \) are plotted in Fig. 5. These probabilities are inferred using the backward–forward algorithm (Klaassen, 2002). The massive fluctuation of these probabilities suggests that the model exhibits rapid risk shifts between high and normal stress periods in the market. The estimated transition probabilities confirm the rapidity of regime switching or low persistence. In fact, for weekly data the expected duration in state 1 is significantly longer than in state 2.

We have also estimated the GJR(1,1) model of Glosten et al. (1993). Although the leverage effect is positive and significant in both regimes, the GARCH(1,1) model outperforms the GJR model according to the SIC. Alexander & Lazar (2009) have also noticed that the leverage effect has no influence when a skewed distribution is used.

2014; Billio et al., in press).
Note: asymptotic t-statistics of the estimated coefficients are between parentheses. $d_f$ stands for degrees of freedom.

* Significant at 5% confidence level.
** Significant at 10% confidence level.

$(1 - p)^{-1} = 1.22$ week against $(1 - q)^{-1} = 2.99$ weeks in state 2; for daily data, state 1 lasts 2.14 days and state 2 lasts 1.14 day; and for 1-minute data, the regime duration is 4.7 min in state 1 and 7.43 min in state 2. The low persistence of the different states has been also noticed by Klaassen (2002) and Wilfing (2009), where the conditional variance is governed more by smooth transitions between different states, which supports the presumption that both regimes and GARCH terms can be important.

For 10-minute returns, however, the estimated transition probabilities are slightly important: $p = 0.9376$ and $q = 0.9434$, which means that once the market holds 10 min in the high volatility state, then on average 93.76% of the time it will remain in that state for the next 10 min. Similarly, if the market holds 10 min in the low volatility state, it will remain in that state for the next 10 min 94.34% of the time, on average. This implies that both regimes are relatively persistent: the high stress regime 1 persists 2.67 h and the normal stress regime 2 persists 2.95 h (1 h is considered long compared to 10 min). On the other hand, the smoothed probabilities of the 10-minute returns (Fig. 6c) show clear domination of the high stress regime 1 (periods of remarkably high probabilities) during two distinctive events, (i) the preparation for and invasion of Iraq in the fall of 2002 until the spring of 2003, (ii) the beginning of "credit crunch" and the subprime mortgage crisis in December 2007 followed by the global financial crisis which reaped from the bankruptcy of Lehman Brothers in September 2008 and the collapse of many large financial institutions until the second quarter of 2009. These periods are characterized by smoothed probabilities of the high volatility regime close to one, so in Fig. 6c we observe clear gaps near zero in the probability plot.

Next, we estimate a GARCH(1,1) switching regime model under Hansen’s (1994) skewed $t$ distribution, since it is preferred to the normal and even to the asymmetric normal mixture distribution (Alexander & Lazar, 2009; Haas, 2010; Wilfing, 2009). Results are presented in Table 5, where $d_f$ stands for the distribution’s degrees of freedom and $\lambda$ is the skewness coefficient.

![Image with Table 5](http://dx.doi.org/10.1016/j.jempfin.2015.03.005)
Fig. 4. Unconditional SGT density function (PDF) for both regimes vs. $N(0,1)$.

Fig. 5. Conditional standard deviations for both regimes.
Estimation times for the skewed $t$ are slightly shorter than those under the skewed generalized $t$, partly because the initial coefficients’ values are carefully chosen to avoid invalid convergence. Based on the SIC, the SGT remarkably outperforms the skewed $t$; even the single regime GARCH under the SGT (Table 3) is better than the two-regime GARCH under the skewed $t$ (except for daily returns). The gain in efficiency and the ability to detect different regimes for multiple frequencies, even intra-daily, make the SGT an acclaimed choice for switching regime GARCH-type models.

6. Conclusion

In this paper, we propose and motivate the use of the highly flexible skewed generalized $t$ distribution as an acclaimed choice for switching regime GARCH models. The introduced approach, supported by a simulation study, generates tractable maximum log-likelihood estimates. Additionally, extensive simulations support the fact that switching regime GARCH models are suitable to capture the different states of volatility in financial markets. For this purpose, a closed form distribution function and a quantile function of the SGT, which are crucial in generating pseudorandom numbers, are developed.

Empirical study is carried on the S&P 500 index for weekly, daily, 10-minute, and 1-minute data, the Markov-switching regime GARCH is preferred to the classical single-regime GARCH for daily and intra-daily returns. For weekly returns, although the switching regime model provides significant coefficient estimates, the single regime GARCH remains a favorite choice. This is mainly due to two reasons; (i) first, the gain in efficiency by the maximum log-likelihood is not enough to compensate the loss caused by increased number of estimated parameters, (ii) and second, the massive fluctuation of the smoothed probabilities for all frequencies suggests that the market exhibits rapid risk shifts between high and normal stress periods, which indicates that weekly data cannot flawlessly capture the quick switching in volatility states.

The skewed generalized $t$ remarkably outperforms the skewed $t$, and improves the efficiency to detect diverse regimes without problems of inaccurate convergence. Its high flexibility allows the detection of the different distribution’s shapes for each regime without any substantial loss in computation time. Moreover, compared to the Markov chain Monte Carlo and Gibbs sampler technique, the gain in estimation time is extremely convenient.

This paper offers many suggestions for future researches. First, we can extend the MS-GARCH to include the APARCH model (Ding et al., 1993) under the SGT, and compare the gain in efficiency, yet we need to construct a proper stationarity condition to include the distribution’s specific parameters. BenSaïda (2012) has indeed constructed the stationarity condition for the APARCH model under the non-skewed version GTD and has found a superior forecasting power. Second, the derived closed-form distribution and quantile functions of the SGT can be implemented to improve the performance of the Value-at-Risk analysis, as argued by Harris & Küçüközmen (2001) and Bali & Theodossiou (2007). Third, we can include the SGT into a switching regime dependence structure analysis via a

Fig. 6. Smoothed probabilities of being in high volatility regime 1.
Copula-based Markov approach. And finally, the SGT exists only in a univariate form and apply it to a multivariate switching regime GARCH model by considering Haas et al.'s (2004) representation, since it is not path dependent. Hence, estimation of a multivariate model can be done by a direct maximization of the log-likelihood.

**Appendix A. Proof of Theorem 1**

The CDF of the skewed generalized t (SGT) distribution is given by:

$$F(x) = \int_{-\infty}^{x} \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left[ 1 + \frac{x^2 - \mu^2}{(1-x^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dx$$

$$F(x) = \int_{-\infty}^{\mu} \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left[ 1 + \frac{(x-\mu)^2}{(1-x^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dx + \int_{\mu}^{x} \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left[ 1 + \frac{(x-\mu)^2}{(1-x^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dx$$

Pose $y = t - \mu$.

$$F(x) = \int_{-\infty}^{0} \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left[ 1 + \frac{(-y)^2}{(1+y^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dy + \int_{0}^{x} \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left[ 1 + \frac{y^2}{(1+y^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dy$$

where: $\text{sgn}(\cdot)$ stands for the sign function.

Following Gradshteyn & Ryzhik (2007, p. 322), Section 3.241.4, $\int_{0}^{\infty} x^{p-1} e^{-ax^2} dx = \frac{1}{a^{p/2}} B \left( \frac{p+1}{2}, \frac{n-p+1}{2} \right)$, under the conditions that $-1 < p < n q - 1$ and $a \neq 0$, we have $F(Y = 0) = F(\mu) = 1 - \frac{\lambda}{2}$.

$$F(\mu) = \int_{0}^{0} \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left[ 1 + \frac{(-\mu)^2}{(1+\mu^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dy = 1 - \int_{0}^{\infty} \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left[ 1 + \frac{y^2}{(1+y^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dy$$

$$= 1 - \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right)} \left( 1 + \lambda \right) \alpha \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) - \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right)} B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) - \frac{1}{2} = 1 - \frac{\lambda}{2}$$

Case when $(x - \mu) \geq 0$

$$F(x) = \frac{1 - \lambda}{2} + \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right)} \int_{0}^{x-\mu} \left[ 1 + \frac{y^2}{(1+y^2)^{\alpha/2}} \right]^{\frac{\alpha+1}{2}} dy.$$

Let's first compute the integral $H = \int_{0}^{x} \frac{1}{(1+b t^2)^k} dt$ when $x > 0$.

By posing $X = \frac{t}{\sqrt{b}}$, we obtain: $H = \int_{0}^{\frac{1}{\sqrt{b}}} (1-X)^k dX$, now replace $dt$ with $dX$. Knowing that $t = a \left( \frac{1}{x^2 - 1} \right)^{\frac{1}{2}}$, we obtain $dt = a \frac{1}{n} X^{-2} \left( \frac{x}{x^2 - 1} \right)^{\frac{1}{2}} - \frac{1}{2} dX$. Hence;

$$H = \frac{a}{n} \int_{0}^{\frac{1}{\sqrt{b}}} (1-X)^k X^{-2} \left( \frac{x}{x^2 - 1} \right)^{\frac{1}{2}} dX = \frac{a}{n} \int_{0}^{\frac{1}{\sqrt{b}}} X^{\frac{1}{2}} X^{-2} (1-X)^k - \frac{1}{2} dX = \frac{a}{n} B \left( \frac{1}{n}; \frac{1}{n} - \frac{1}{2} \right).$$

Replacing $H$ into the CDF, we obtain:

$$F(x) = \frac{1 - \lambda}{2} + \frac{\eta}{2\sigma^2 \theta B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right)} B \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right) = \frac{1 - \lambda}{2} + \frac{\lambda + 1}{2} \left( \frac{1}{\eta}; \frac{\psi}{\eta} \right).$$

$I(a, b)$ is the regularized incomplete beta function, with $a > 0$ and $b > 0$.  

Please cite this article as: BenSaida, A., The frequency of regime switching in financial market volatility, J. Empir. Finance (2015), http://dx.doi.org/10.1016/j.jempfin.2015.03.005
Case when \((x - \mu) < 0\)

\[
F(x) = \frac{1 - \lambda}{2} + \frac{\eta(x - \mu)}{2\sigma \theta B(\frac{1}{\eta}; \frac{1}{\eta})} \int_0^{x-\mu} \frac{1}{1 + \left(\frac{\psi}{\eta}\right)^n} dy.
\]

Let’s first compute the integral \(H = \int_0^x \frac{1}{1 + \left(\frac{-t}{\eta}\right)^n} dt\) when \(x < 0\). By posing \(X = \left(\frac{-t}{\eta}\right)^n\), we obtain: \(H = \int_0^{-x} (1 - X)^{\frac{1}{n}} dx\), now replace \(dt\) with \(dX\). Knowing that \(t = \frac{-\psi}{\eta}\), we obtain \(dt = -\frac{\eta}{\psi} X^{-\frac{1}{n}} dX\). Hence;

\[
H = -\frac{\psi}{\eta} \int_0^{-x} (1 - X)^{-\frac{1}{n}} X^{-\frac{1}{n}} dX = -\frac{\psi}{\eta} \int_0^{-x} X^{-\frac{1}{n}} (1 - X)^{-\frac{1}{n}} dX = -\frac{\psi}{\eta} B_{\frac{1}{n}, 1 - \frac{1}{n}}(1, k - 1).
\]

Then,

\[
F(x) = \frac{1 - \lambda}{2} - \frac{(1 - \lambda)}{2B(1; \frac{1}{\eta})} \left(\frac{1}{\eta}\right)^n \int_0^{-x} (1 - X)^{-\frac{1}{n}} dX = \frac{1 - \lambda}{2} + \frac{\lambda - 1}{2} \int_0^{-\psi} \left(\frac{1}{\eta}\right)^n \frac{1}{1 + \left(\frac{\psi}{\eta}\right)^n} d\psi.
\]

Finally,

\[
F(x) = \frac{1 - \lambda}{2} + \frac{\lambda - 1}{2} \int_0^{-\psi} \left(\frac{1}{\eta}\right)^n \frac{1}{1 + \left(\frac{\psi}{\eta}\right)^n} d\psi.
\]

**Remark.** The CDF of the SGT can also be written as:

\[
F(x) = \frac{1 - \lambda}{2} + \frac{\eta(x - \mu)}{2\sigma \theta B(1; \frac{1}{\eta})} \left(\frac{1}{\eta}\right)^n \int_0^{\frac{|x - \mu|}{\psi}} \frac{1}{1 + \left(\frac{\psi}{\eta}\right)^n} d\psi.
\]

where \(\text{E}_{i}^n\) is the Gauss hypergeometric function, by using both of the following transformation rules: \(B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}\), \(\text{E}_{i}^n\) by Abramowitz & Stegun (1972, p. 945), Section 26.5.23, and the Pfaff’s second transformation \(\text{E}_{i}^n(a, b; c) = (1-z)^{\frac{-b}{a}}\text{E}_{i}^n(b, c-a; c; \frac{z}{1-z})\) by Abramowitz & Stegun (1972, p. 559), Section 15.3.5. Nevertheless, this representation is avoided since only a very few software packages can compute the numerical approximation of the Gauss hypergeometric function; moreover, its inverse is highly expensive in computer time.

**Appendix B. Proof of Theorem 2**

If we have \(F(x) = z\), then the inverse CDF is simply \(x = F^{-1}(z)\).

Case when \((x - \mu) \geq 0\)

\[
F(x) = z = \frac{1 - \lambda}{2} + \frac{\lambda + 1}{2} \int_0^{\frac{|x - \mu|}{\psi}} \left(\frac{1}{\eta}\right)^n \frac{1}{1 + \left(\frac{\psi}{\eta}\right)^n} d\psi.
\]

which yields:

\[
\frac{|x - \mu|}{\psi} = \frac{1}{\lambda + \frac{1}{\lambda} \frac{|x - \mu|}{\psi}} \frac{1}{\eta}\frac{1}{\eta}
\]

or:

\[
F^{-1}(z) = \frac{(1 + \lambda)\sigma \theta}{\left[\frac{1}{\lambda + \frac{1}{\lambda} (1 - \frac{|x - \mu|}{\psi})} - 1\right]^{\frac{1}{\eta}}} + \mu.
\]

Note that in this case \(F(x) \geq F(\mu) = \frac{1 - \lambda}{2};\) consequently, \((z - \frac{1 - \lambda}{2}) \geq 0\).
Case when $\eta > 0$

$$F(x) = \frac{1}{\eta} \lambda \frac{\lambda - 1}{2} - \frac{1}{\eta} \left( \frac{1}{\sqrt{\pi}} \right)$$

which yields:

$$F^{-1}(z) = \frac{\sqrt{2\mu - z}}{\sqrt{2\mu}} + \mu.$$

Note that in this case $F(x) < F(\mu) = \frac{1}{2}$; consequently, $(z - \frac{1}{2}) < 0$.

Finally,

$$F^{-1}(z) = \mu + \frac{\sigma \theta}{\lambda} \left[ \text{sgn} \left( \frac{z - \frac{1}{2}}{\lambda} + \mu \right) + \lambda \right] \left( \frac{1}{\sqrt{\pi}} \right)$$

References


Please cite this article as: BenSaïda, A., The frequency of regime switching in financial market volatility. J. Empir. Finance (2015), http://dx.doi.org/10.1016/j.jempfin.2015.03.005

Please cite this article as: BenSaïda, A., The frequency of regime switching in financial market volatility, J. Empir. Finance (2015), http://dx.doi.org/10.1016/j.jempfin.2015.03.005