The GARCH models: A study on the volatility of gold returns

Junhyuck Song
Université Paris Nanterre*

September, 2019

Abstract

Gold has been considered as a safe-haven asset. In particular, gold has been used as a tool for investors to hedge against market turbulence that presents high volatility and uncertainty of risky assets. Market turbulence periods can be characterized as a regime change where investors expect to generate a constant return of risky assets in low volatility regime. In high volatility regime, on the other hand, buying gold as a safer asset to reduce potential losses. Especially for buy-and-hold investors need to recognize these regime changes beforehand. MS-GARCH model that allows to switch from one regime to another by unobservable Markov-chain process provide more flexible setup to recognize this regime changes on the gold market. This paper implements different GARCH-type models on the volatility in gold returns in order to compare its forecasting power.

1 Introduction

Gold has been considered as a safe-haven asset. Thanks to this fact, gold has been used as a tool for investors to hedge against market turbulence that presents high volatility and uncertainty of risky assets. For investors, this implies that future volatility analysis is a crucial part to include gold in their portfolio to reduce potential losses. Unarguably, the most popular model classes in volatility analysis is that of GARCH [28] models generalized from ARCH [24]. However, fitting GARCH models with financial data often exhibits a high persistence of the conditional variance and models usually have poor prediction power for out-of-sample. Diebold (1986)[13] Lamoureux et al (1990)[9] argue that this nearly integrated behavior of the conditional variance is due to the stationary conditions imposed on the model that can not fully estimate the structural changes in the variance process. Fitting a GARCH model on a sample which shows structural changes in the unconditional variance can also create integrated GARCH effect[30]. Aside from this potential bias of persistence parameters in classic GARCH model, gold returns exhibit strong leverage effects which make volatility forecasting more inaccurate[23]. To counter this problem, several approaches have been proposed. The first approach consists of using conditional asymmetric models such as TGARCH, GJR-GARCH and EGARCH with imposed asymmetric parameters in the equation of the conditional variance. The second approach is Markov Regime-Switching (MS) approach. Several configurations based on Markov-Switching ARCH and GARCH models have been proposed.

*The university dose not endorse any of the opinions expressed therein

Schwert (1989)[16] and Gray (1996)[26] use MS-GARCH framework by allowing two regimes depending on the level of volatility. Hamilton and al (1994)[17] introduce MS-ARCH framework to avoid the problem of path dependence due to the recursive nature of the conditional variance of GARCH model. Marcucci (2005)[18] uses GARCH, EGARCH and GJR-GARCH to compare the ability to forecast S&P 100 volatility of returns. Statistical inferences of MS-GARCH for conditional mean and variance switch and Markov chain properties are studied by Bauwens, Preminger and Rombouts (2009)[5]. Statistical conditions of models are given by Abramson and Cohen (2007)[2]. De la Torre-Torres, Galeana-Figueroa and Alvarez-Garcia (2019)[3] use MS-GARCH framework to test an investment strategy on commodity markets. As gold returns can highly be categorized by regime switches where investors expect a constant return of risky assets in a low volatility market regime and shift to safer assets in a high volatility market regime, the advantage of modeling the volatility in gold returns is to be able to track down regime change points. MS-GARCH framework is proven to provide better performance when forecasting future volatility because the model allows to capture the variance dynamics.

For estimating method, computation by the likelihood function requires of taking all possible regime paths. The problem of path dependence renders the likelihood method computationally very expensive as the number of possible paths grows in an exponential way and this makes maximum likelihood estimation makes difficult in many commercial and freeware packages. Several alternative estimation has been proposed. Hamilton et al (1994) use MS-ARCH framework to avoid the path dependence problem. Klaassen (2002)[14] uses GARCH framework and avoid the path dependency problem by taking the conditional variance in each regime into a single variance. Bauwens, Preminger and Rombouts (2009) propose Bayesian Markov Chain Monte Carlo (MCMC) estimation by using Gibbs sampler to sample jointly the whole regime path history including the whole regime path in the parameter space.

The objective of this paper is to show the performance of volatility forecasting using different GARCH models. In particular, this paper attempt to capture the leverage effects in gold returns depending on regime changes which are strongly influenced by market conditions. These points of market regime switching can be beneficial for investors who want to hedge against upcoming high volatility risk. Section 2 introduce briefly symmetric and asymmetric GARCH models. Section 3 focus on its estimation method without explaining the details of algorithms generally used in GARCH fitting. Section 4 presents the implementation of GARCH models and Section 5 concludes.

2 GARCH Models

2.1 GARCH

Let y_t is the price of gold at time t, their returns are defined by (1). Generalized ARCH model where the conditional variance depends on previous own lags is given by the equation (2) where ϵ_t is the market shock. To grantee positivity of parameters in the conditional variance, the parameters should be $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$. The persistence of shocks to conditional volatility is measured by $\alpha + \beta$ must be smaller than 1 in order to assure the conditional variance to be stationary. The conditions of regularity of GARCH model is given by Ling and McAleer (2002)[21].

$$y_t = logS_t - logS_{t-1} \tag{1}$$

$$y_t = \mu + \epsilon_t \tag{2}$$

$$\epsilon_t = \sigma_t z_t \tag{3}$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{4}$$

h_t of equation (3) is often specified as normal distributions or Student distributions in order to capture fat tail effects and excess kurtosis of financial returns.

2.2 The Inverse Leverage Effect of Gold Returns

The leverage refers to the relationship between financial assets and both implied and realized volatility. Contrary to stock market volatility where the volatility is negatively correlated with price, gold market exhibits the inverse leverage effect. Since gold is considered as a safe-haven asset where investors seek to hedge against potential losses when they anticipate market turbulence, the volatility in gold returns is positively correlated with upward price movement. This fact illustrates asymmetrical volatility of gold returns. To measure this asymmetric of the conditional volatility, GJR-GARCH, TGARCH, EGARCH and APARCH specifications are used.

2.2.1 Asymmetric GARCH Models

The Threshold GARCH (TGARCH) model was introduced by Zakoian (1994) in order to measure asymmetric effects of volatility. The conditional variance equation is defined by indicated function taking values when given conditions are satisfied. if γ asymmetric coefficient is significant, there is asymmetric effect. TGARCH model dose not impose positivity restrictions on the volatility coefficients so $sigma_t$ is no longer assumed to be positive. The conditions for existence of moments are given by He and Teräsvirta (1999)[7].

The GJR-GARCH model developed by Glosten, Jagnnathan et Runkle (1993)[19]. Asymmetric effect on conditional volatility is measured by the term $I\gamma(\epsilon_{t-1} < 0)$ where ϵ_{t-1} is 1 when it is negative otherwise it is 0. When the coefficient γ is significant, positive and negative shocks of equal magnitude have different impact on conditional volatility. The coefficient of good news is captured by $\alpha + \gamma$ and bad new is captured by γ. The existence of moments and stationarity conditions are analyzed by Ling and McAleer (2002)[21] for GJR-GARCH(1, 1) model.

The EGARCH model proposed by Nelson (1991)[4] able to capture volatility leverage. Developed to compensate a few limitations of GARCH model, Nelson summarize these limitations as follow: *GARCH (1, 1) model is unable to include asymmetric leverage effect, Non negativity constraints on parameters are too limitative and the conditional moments of GARCH may explode even if the process is strictly stationary and ergodic.* The term $\theta_1((|z_{t-1}|) - E(|z_{t-1}|))$ denotes the magnitude effect and $\theta_2 z_t$ as the sign effect. if θ_2 is negative, that means negative shocks induce higher volatility than positive shocks of the same magnitude.

The asymmetric power of ARCH (APARCH) model proposed by Ding, Granger and Engle (1993)[31]. Parameter δ ($\delta > 0$) assure the normality of the conditional standard deviation. Coefficient γ where γ is between -1 and 1, measures the asymmetric effect. Here, positive or negative significant coefficient means fitted series has positive or negative asymmetric effect.

2.2.2 GARCH Model Fitting by MLE

The most common method of estimating GARCH models is the method of maximum likelihood estimation (MLE). The term of GARCH model (3) plays crucial role in determining the fitness so the predictive power of GARCH models. Here, we consider only the two most popular distribution choices which are normal, t-student. Given (3), the log-likelihood function L with the normal distribution,

$$logL = \Sigma_{t=1}^{T} l_t = -\frac{1}{2}\Sigma_{t=1}^{T}(ln(2\pi) + ln(\sigma_t^2)) - \frac{1}{2}\Sigma_{t=1}^{T}\frac{\epsilon_t^2}{\sigma_t^2} \tag{5}$$

GARCH fitting by the maximum likelihood estimation is well documented in the literature. First of all, maximizing the log likelihood needs to initiate starting values for the model parameters μ, ω, $\alpha_i, i \in (1,..,q)$ and $\beta_i, i \in (1,...,p)$ needs to be initialized along with ϵ_t^2 and σ_t^2[8]. General choices for μ is the mean of y_t as the starting value. For ω is usually set to unconditional variance of y_t and zero values are assigned to the conditional variance parameters. Finally, for σ_t^2 the mean of the residuals resulting from a regression of y_t on a constant is usually used. The most popular method to maximize the log-likelihood is using a Newton-Raphons algorithm as below.

$$\hat{\theta_{n+1}} = \hat{\theta_n} - \lambda_n H(\hat{\theta_t})^{-1} s(\hat{\theta_n}) \tag{6}$$

θ_n is the vector of model parameters to be estimated at iteration n, λ_n is a scalar step-length parameter and $s(\theta_n)$ and $H(\theta_n)$ denote respectively the gradient vector and Hessian matrix of the log-likelihood at iteration n. $H(\theta_n)$ is given by the equation (7). The scalar-length parameter is chosen by satisfying $lnL(\theta_{n+1}) \geq lnL(\theta_n)$. The Berndt-Hall-Hall-Hausman (BHHH) algorithm is often used to approximate the Hessian matrix using first derivative. Analytic solution of GARCH estimation is provided by Fiorentini et al (1996)[15].

$$-H(\theta) = \Sigma_t^T \frac{\partial l_t}{\partial \theta}\frac{\partial l_t}{\partial \theta'} \tag{7}$$

The BHHH algorithm is similar to the Newton-Raphson algorithm but the only difference consist of using the Hessian matrix of second derivatives $H(\hat{\theta_n})$. The Hessian matrix is approximated by

$$-H(\hat{\theta}_i \approx B(\hat{\theta}_i) = \Sigma_{t=1}^{T} G(\hat{\theta}_i)G'(\hat{\theta}_i) = \Sigma_{t=1}^{T}\frac{\partial l_t}{\partial \theta_i}\frac{\partial l_t}{\partial \theta_i'} \tag{8}$$

Table 1: Conditional variance and Stationary conditions on different asymmetric GARCH models

Model	Conditional variance	Stationary conditions
TGARCH	$\omega + \alpha I(\epsilon_{t-1} \geq 0) + \gamma I(\epsilon_{t-1} < 0)\epsilon_{t-1} + \beta\sigma_{t-1}$	$\omega, \beta > 0$
EGARCH	$ln(\sigma_t^2) = \omega + \theta_1(\|z_{t-1}\| - E(\|z_{t-1}\|) + \theta_2 z_{t-1} + \beta ln(\sigma_{t-1}^2)$	$\|\beta\|$
GJR-GARCH	$\omega + \alpha\epsilon_{t-1}^2 + \gamma I(\epsilon_{t-1} < 0)\epsilon_{t-1}^2 + \beta\epsilon_{t-1}^2$	$\omega > 0,\ \alpha + \gamma \deg 0 and \beta > 0$
APARCH	$\sigma_t^\delta = \omega + \alpha(\|\epsilon_{t-1} - \gamma\epsilon_{t-1})^\delta + \beta\sigma_{t-1}^\delta$	$\delta, \omega, \alpha, \beta > 0$

The advantage of BHHH algorithm is increased computational speed by not calculating the Hessian matrix at each iteration for each time step. However, it requires more iteration if the log-likelihood function is far from its maximum.

The ML estimates are consistent and asymptotically Normally distributed and an estimate of the asymptotic covariance matrix of the ML estimates is constructed from the Hessian matrix where the optimization algorithm produced. However, the appropriate regularity conditions were verified for some GARCH models.[25][20][29][11][12]

2.3 MS-GARCH

2.3.1 Model

Let s_t be an ergodic Markov chain with transition probabilities $\eta = P(s_t = i|s_{t-1} = j)$ and invariant probability is denoted as π_i. The MS-GARCH model is given by

$$y_t = \mu_{s_t} + \sigma_t z_t \tag{9}$$

$$\sigma_t^2 = \omega_{s_t} + \alpha_{s_t}\epsilon_{t-1}^2 + \beta_{s_t}\sigma_{t-1}^2 \tag{10}$$

Conditions on σ_t^2's positivity is imposed the same as GARCH model. Conditions for weak stationarity and existence of moments for the Markov-Switching GARCH (p, q) model with zero means μ_{s_t} are given by Franq and Zakoian (2005)[6]. As the process $Z_t = (y_t, \sigma_t h_t, s_t)'$ follows Markov chain, mild regularity conditions the existence of moments and geometrically ergodicity are given by Meyn and Tweedie (1993) and Chan (1993)[27]. Before defining a Markov chain process on MS-GARCH, Bauwens, Preminger and Rombouts (2007)[5] impose five assumptions,

C1. The error term z_t is i.i.d with a continuous density centered on zero with $E(|z_t|^\delta)$ for $\delta > 0$.

C2. $\alpha_i > 0$, $\beta_i > 0$ and $\eta_{ii} \in \{1, ..., n\}$.

C3. $\Sigma_{i=1}^T \pi_{ii} E(log(\alpha_i z_i^2 + \beta_i)) < 0$

C4. $E(|z_t^2|^k) < \infty$ for $k \geq 1$.

C5. $\rho(\Omega) < 1$

Where Ω is defined as (n, n) matrix

$$\Omega = \begin{pmatrix} E(\alpha_i z_i^2 + \beta_i)^k \eta_{11} & ... & E(\alpha_i z_i^2 + \beta_i)^k \eta_{1n} \\ \vdots & & \vdots \\ E(\alpha_i z_i^2 + \beta_i)^k \eta_{n1} & ... & E(\alpha_i z_i^2 + \beta_i)^k \eta_{nn} \end{pmatrix}$$

For the first assumption, for $\delta = 1$, for $\delta > 1$, the data is estimated by the conditional scaling factor. Second assumption states the conditions for strict positivity of the variance and the discrete Markov chain is irreductible and ergodic. Third assumption assures that at least one of the regimes is stable.

Under $C1-C3$ assumptions y_t is geometrically ergodic and if Markov chain is initiated from its stationary distribution, the process is strictly stationary and $\beta-mixing$. The geometric ergodicity ensures a unique stationary probability measure for the existence of the process as well as the chain process. Meyn and Tweedie [27] gives the properties of this Markov chain that satisfies these conditions with conventional limit theorems for any given initiating value given the existence of suitable moments. The regular mixing which defines the exponential decaying rate of the mixing numbers is given by Doukhan, Massart and Rio(1994)[22]. The matrix Ω is defined to analyze the existence of higher order moments. Under the assumptions C1-C2 and C4-C5, the process is geometrically ergodic and $E(|z_t^2|^k)<\infty$ for $k\geq 1$, where expectations are taken from the stationary distribution.

2.3.2 Bayesian Inference

This section's Bayesian inference and Sample method are well documented in the paper of Bauwens et al [5]. Due to the path dependence of the conditional variance and its recursion nature, the estimation of MSGARCH models by the maximum likelihood estimation is computationally expensive. Most of approaches use Bayesian inference which allows to treat the latent state variables as parameters of model and to estimate the likelihood function given unconditional states. To estimate the parameters of MS-GARCH model, Gibbs sampling algorithm that allows to sample from the full conditional posterior densities by blocking each parameters given by θ,μ,η. For the choice of the standard normal distribution for z_t. Let $s_t=(s_1,s2)$ takes two states, Y_t is the vector of $(y_1,y_2,...,y_t)$. For model parameters, let $\eta=(\eta_{11},\eta_{12},\eta_{21},\eta_{22})'$, respectively active first state, transition probability from first state to second state, transition probability from second state to first state and active second state. $\mu=(\mu_1,\mu_2)'$ and model parameters $\theta=(\omega_k,\alpha_k,\beta_k)'$ for $k=1,2$. For the joint density of y_t given the parameters and the past information, the joint density is:

$$f(y_t,s_t|\mu,\theta,\eta,Y_{t-1},S_{t-1})=f(y_t|s_t,\mu,\eta,Y_{t-1},S_{t-1})f(s_t|\mu,Y_{t-1},S_{t-1}). \tag{11}$$

And the conditional density of y_t is set to Gaussian density

$$f(y_t|s_t,\mu,\theta,Y_{t-1},S_{t-1})=\frac{1}{\sqrt{2\pi\sigma_t^2}}\exp(-\frac{(y_t-\mu_{s_t})^2}{2\sigma_t^2}) \tag{12}$$

with the marginal density of s_t is defined by

$$f(s_t|s_t,\mu,Y_{t-1},S_{t-1})=f(s_t|\mu,s_{t-1})=\eta_{s_{s_t s_{t-1}}} \tag{13}$$

Where, $\eta_{11}+\eta 21=1$, $\eta_{12}+\eta_{22}=1$, $0<\eta_{11}<1$ and $0<\eta_{22}<1$. The joint density of $y=(y_1,y_2,...,y_T)$ and $S=(s_1,s_2,...,s_T)$ is obtained by taking the product of two densities (11) and (12) as below

$$f(y,S|\mu,\theta,\eta)\propto\prod_{t=1}^{T}\sigma^{-1}exp(\frac{(y_t-\mu_{s_t})^2}{2\sigma_t^2}\eta_{s_t s_{t-1}} \tag{14}$$

2.3.3 Sampling s_t

The conditions imposed $\eta_{11} + \eta 21 = 1$, $\eta_{12} + \eta_{22} = 1$ ensure Markov chain does not depend on historical values but to sample s_t, conditions on previous and posterior to the present state due to the path dependence of the conditional variances. The full conditional mass function of state t is given by

$$\varphi(s_t|S,\mu,\theta,y) \propto \eta_{1,s_{t-1}}^{2-s_t}\eta_{2,s_{t-1}}^{s_t-1}\eta_{1,s_t}^{2-s_t+1}\eta_{2,s_t}^{s_{t+1}-1}\prod_{j=t}^{T}\sigma_j^{-1}exp-\frac{(y_j-\mu_{s_j})^2}{2\sigma_j^2} \tag{15}$$

2.3.4 Sampling η

Given a prior density $\pi(\eta)$

$$\varphi(\eta|S,\mu,\theta,y) \propto \pi(\eta)\prod_{t=1}^{T}\eta_{s_t s_{t-1}} \tag{16}$$

η dose not depend on $\mu\eta$ and y.

2.3.5 Sampling θ

Given a prior density $\pi(\eta)$

$$\varphi(s_t|S,\mu,\eta,y) \propto \pi(\theta)\prod_{t=1}^{T}\sigma_t^{-1}exp-\frac{(y_j-\mu_{s_j})^2}{2\sigma_j^2} \tag{17}$$

where the models parameters do not depend on η. Bauwen et al uses the griddy-Gibbs sampler to sample θ. Let (r) be the step of iteration, the algorithm loops as follow:

1. With the equation (16), compute $\kappa(\omega_1|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)$ where κ is the kernel of the conditional posterior density of w_1 μ is sampled at iteration r over a grid $(w_1^1, w_1^2, ..., w_G)$ to get the vector $G_\kappa = (\kappa_1, ..., \kappa_G)$.

2. Compute $G_f = (0, f_2, ..., f_G)$ by a deterministic integration where f_i is

$$f_i = \int_{w_1^1}^{w_1^i}\kappa(\omega_1|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)dw_1 \tag{18}$$

3. Generate u $U(0, f_G)$ and invert $\kappa(\omega_1|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)$ by numerical interpolation to get a draw ω_1^{r+1} $\varphi(\omega_1|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)$

4. Repeat previous steps for $\varphi(\beta_1|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)$, $\varphi(\alpha_1|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)$, $\varphi(\omega_2|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)$, $\varphi(\alpha_2|S^{(r)},\beta_1^{(r)},\alpha_1^{(r)},\theta_2^{(r)},\mu^{(r)},y)$ and so on.

2.3.6 Sampling μ

Given a prior density $\pi(\eta)$

$$\varphi(\mu|S,\theta,\eta,y) \propto \pi(\mu)\prod_{t=1}^{T}\sigma_t^{-1} exp - \frac{(y_j-\mu_{s_j})^2}{2\sigma_j^2} \tag{19}$$

3 Model Fitting

3.1 Data

Gold data downloaded from Federal Reserve Economic Data (Fred) St. Louis is the gold fixing price on the London bullion market based in U.S dollar[1]. Total number of observation is 13427 starting from on 1 April 1968 to on 28 June 2019 where training set has 10365 observations. For the period of our data, gold returns are overall positive but compare to other assets, gold's kurtosis is very high, meaning that gold returns have a lot of extreme values. Positive skewness in gold returns signify investors can expect small losses frequently and a few large gains.

Figure 1: Gold price and log returns of in-sample data

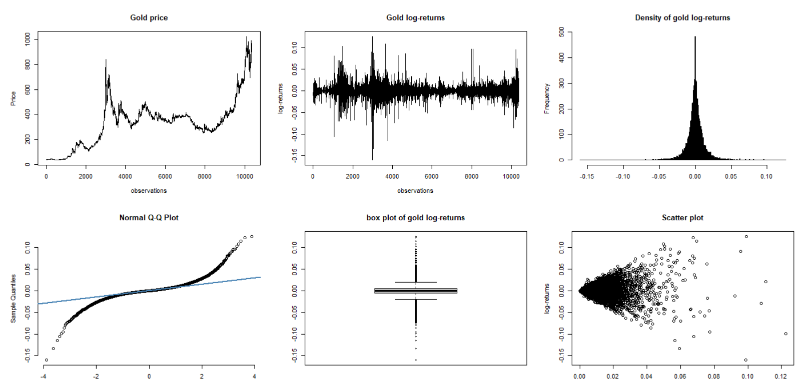


Figure 1 shows non-normality in gold returns and the traditional relationship between returns and volatility. As volatility increases, positive and negative returns increases. In our data, gold returns are positively skewed with kurtosis equal to 13.70.

3.2 GARCH Models Estimation

Data is fitted by using *R* software with *rugarch* package. This package offers flexible choices concerning optimization problems. Contrary to the general steps of GARCH fitting, the main optimization method is the augmented Lagrange solver *solnp* of Ye (1997). Statistical significance

[1]Ticker for this data is available: GOLDAMGBD228NLBM, Two data are available on FRED statistic for gold fixing price in London Bullion market. London time 10:30 am and 3:00 pm

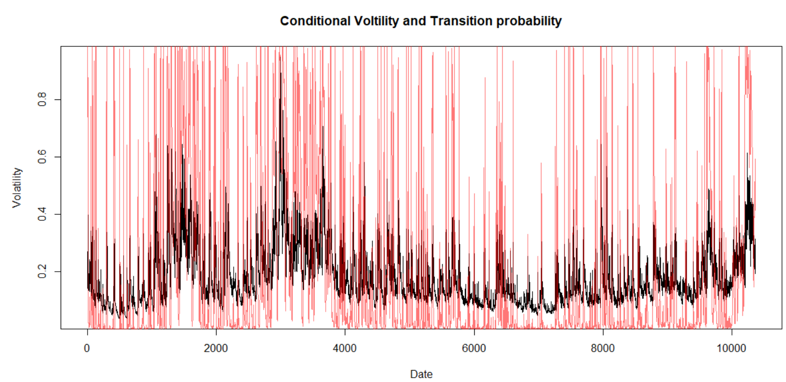
is provided by the robust standard errors based on the method of White (1982), which produces asymptotically valid confidence intervals by calculating the covariance (V) of the parameter θ as follow, Ghalanos (2016)[1]. The choice of GARCH model package is based on the robustness of statistical accuracy since GARCH fitting involving non-linear optimization[8].

$$\hat{V} = -L''(\hat{\theta})\Sigma_{i=1}^{n} g_i(x_i|\hat{\theta})^T g_i(x_i|\hat{\theta})L''(\hat{\theta}) \tag{20}$$

The fitted models are presented in the table 2. For our training data, GARCH coefficient β was not statistically significant but ARCH coefficient is statistically significant. The model selection is based on information criteria AIC and BIC. For z_t, normal distribution and Student-t distribution are used. Unfortunately, GARCH (1, 1) with normal and Student-t distribution model coefficients are not statistically significant. The coefficients of ARCH(∞) which is GARCH(1, 1) are significant for both distributions but Student-t distribution gives better fitted models than normal distribution. To take a parsimonious model, ARCH(2) has to be chosen among symmetric ARCH-GARCH type models. As shown in table 2, the persistence parameter $\alpha + \beta$ is very close to 1 for all GARCH-type models. This shows nearly integrated behavior of volatility as we discussed above.

The coefficient γ being the asymmetric parameter, the sign of γ TGARCH is negative and significant. Since good news has an impact on α_1 and bad news has an impact on $\alpha_1 + \gamma$, bad news trigger more volatility on gold market which confirm stylized facts of gold market behavior. Furthermore, this is confirmed by other asymmetric GARCH-type models. For example, the coefficient γ of EGARCH with normal and Student-t distribution has a positive sign and significant. That means bad news would have greater impact on volatility. Among asymmetric GARCH models, EGARCH exhibits the best fit.

Figure 2: Fitted Conditional Volatility and Transition Probability



Red line represents transition probability. As the graph shows, high volatility is associated with second regime.

As MS-GARCH estimation method is done by MCMC, information criterion is based on "Deviance Information Criterion" (DIC). Different combinations of models are fitted to find the best model. Generally, when GARCH(1, 1) model is included, unconditional transition probability shows

Table 2: Fitted symmetric and asymmetric GARCH models

Model	μ	ω	α_1	α_2	β_1	γ	δ	AIC	BIC
ARCH (2, 0)	0.00016	0.000068	0.322121	0.314225	-			-6.1929	-6.1901
(Norm)	(0.15283)	(0.0000)	(0.0000)	(0.0000)	-				
GARCH (1, 1)	0.000052	0.00001	0.104040	-	0.89496	-		-6.4004	- 6.3976
(Norm)	(0.5518)	(0.92795)	(0.39041)	-	(0.0000)	-			
EGARCH(1, 1)	0.000114	-0.196844	0.028515	-	0.975943	0.29053	-	-6.4078	-6.4043
(Norm)	(0.402088)	(0.000013)	(0.034163)	-	(0.0000)	(0.000)	-	-	
GJRGARCH(1, 1)	0.000121	0.000001	0.119562	-	0.897649	-0.036425	-	-6.4036	-6.4001
(Norm)	(0.17548)	(0.94051)	(0.46840)	-	(0.0000)	(0.18978)	-	-	
APARCH (1, 1)	0.000108	0.000003	0.110720	-	0.895742	-0.096177	1.774511	-6.4041	-6.4000
(Norm)	(0.71792)	(0.94512)	(0.31983)	-	(0.0000)	- (0.029311)	(0.253036)	-	-
TGARCH (1, 1)	0.000135	0.000168	0.143752	-	0.883326	-0.118309	-	-6.3972	-6.3937
(Norm)	(0.206494)	(0.003389)	-	-	(0.00000)	(0.00000)	-	-	
ARCH (2, 0)	0.000119	0.000056	0.513333	0.485666	-	-	-	-6.4172	- 6.4137
(Std)	(0.073207)	(0.00000)	(0.00000)	(0.00000)	-	- -	-	-	-
GARCH (1, 1)	0.000025	0.000001	0.115521	-	0.883479	-	-	-6.5558	-6.5523
(Std)	(0.67266)	(0.99354)	(0.92709)	-	- (0.46311)	-	-	-	
EGARCH (1, 1)	0.000062	-0.125559	0.039255	-	- 0.986418	0.296159	-	-6.5677	-6.5635
(Std)	(0.2187)	(0.00000)	(0.000036)	-	- (0.00000)	(0.00000)	-	-	
GJRGARCH (1, 1)	0.000051	0.000001	0.137904	-	0.886196	-0.050199	-	-6.5585	-6.5543
(Std)	(0.77561)	(0.98788)	(0.85430)	-	(0.16414)	(0.74252)	-	-	
APARCH (1, 1)	0.000059	0.000019	0.157318	-	0.883520	-0.134241	1.282758	-6.5657	-6.5608
(Std)	(0.259013)	(0.219955)	(0.00000)	-	(0.00000)	(0.000077)	(0.00000)	-	-
TGARCH (1, 1)	0.000058	0.000085	0.157419	-	0.884699	-0.151030	-	-6.5645	-6.5603
(Std)	(0.281643)	(0.005526)	(0.00000)	-	(0.00000)	(0.000126)	-	-	-

Table 3: MS-GARCH model fit

Model	Unconditional probability	transition matrix	DIC
ARCH(1)	0.5677	0.9872, 0.0128	-67145.248
ARCH(1)	0.4323	0.0168, 0.9832	
GARCH(1, 1)	0.8004	0.9946, 0.0054	-67978.29
GARCH(1, 1)	0.1996	0.0216, 0.9784	
ARCH(1)	0.0545	0.8775, 0.1225	-65750.84
GARCH(1, 1)	0.9455	0.0071, 0.9929	
ARCH(1)	0.6944	0.9767, 0.0233	-65589.3717
EGARCH(1, 1)	0.3056	0.0529, 0.9471	
TGARCH(1, 1)	0.7028	0.9758, 0.0244	-67876.79
TGARCH(1, 1)	0.2972	0.0578, 0.9422	
EGARCH(1, 1)	0.6939	0.9738, 0.0262	-67929.40
EGARCH(1, 1)	0.3061	0.0593, 0.9407	
TGARCH(1, 1)	0.6829	0.9728, 0.0272	-67769.07
EGARCH(1, 1)	0.3171	0.0587, 0.9413	

it is hard to switch from one to another once one regime is activated. Other models have similar unconditional transition probability about 70% to 30%. For MS-GARCH, two EGARCH regimes show the best fit. The fitted model is shown in the table 4 where Mean is the posterior mean, SD is Standard deviation, SE correspond to the naive standard error of the mean, TSSE is time series standard error based on an estimated of the spectral density at zero and RNE is the relative numerical efficiency defined as $(SD/TSSE)^2$[10].

Table 4: MS-GARCH fit parameters

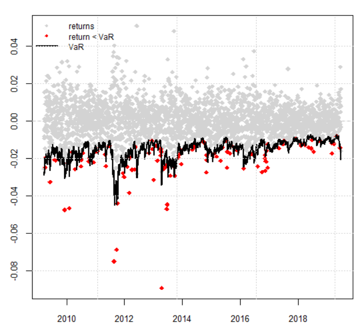
	Mean	SD	SE	TSSE	RNE
μ_1	-0.2132	0.0056	0.0002	0.0007	0.0587
α_1	0.1552	0.0120	0.0004	0.0005	0.5172
θ_1	0.0534	0.0072	0.0002	0.0004	0.3392
β_1	0.9796	0.0004	0.0000	0.0001	0.0677
μ_2	-0.8576	0.1356	0.0043	0.0198	0.0469
α_2	0.2221	0.0261	0.0008	0.0020	0.1727
θ_2	0.0492	0.0182	0.0006	0.0012	0.2457
β_2	0.8792	0.0186	0.0006	0.0027	0.0467

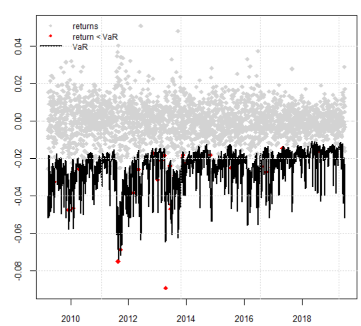
Annualized volatility of each regimes is respectively 0.091 and 0.467 where first regime represents low volatility regime and second represents high volatility regime. As we can show in Figure 2

4 Backtesting

Back-testing of GARCH models are done by calculating Value-at-Risk in order to compare the predicted losses from the actual realized losses. The model chosen for back-testing is EGARCH which showed the best fit in our in-sample period. For MS-GARCH back-testing, two regimes with EGARCH will be used. The estimated model has 13065 observations whose parameters will be used in out-sample that has 2592 observations. For the distribution choice, t-distribution is used for both EGARCH and MS-GARCH.

Figure 3: VaR 0.01% of EGARCH(1, 1) and MSGARCH





Left graph is VaR at 0.01 from EGARCH(1, 1) and right graph is VaR at 0.01 from MSGARCH where 2 regimes of EGARCH(1,1), EGARCH(1,1) are used. Red points correspond to the violation where the actual realized returns are smaller than VaR.

The result shows MS-GARCH model is clearly better than single EGARCH(1, 1). However,

Table 5: Fitted GARCH models with volatility forecasting metrics

Model	AIC	BIC	$HQIN$	MAE	$RMSE$	$MASE$	$MAPE$
ARCH(2) $\mathcal{N}$	-6.1687	-6.1659	-6.1677	0.008	0.0117	25.627	0.9926
GARCH(1, 1) $\mathcal{N}$	-6.4004	-6.3976	-6.3995	0.001	0.0014	3.3451	0.1197
EGARCH(1, 1) $\mathcal{N}$	-6.4078	-6.4043	-6.4066	0.0016	0.002	5.2697	0.1948
GJRGARCH(1, 1) $\mathcal{N}$	-6.4036	-6.4001	-6.4024	0.001	0.0014	3.3591	0.1203
APARCH(1, 1) $\mathcal{N}$	-6.4041	-6.4000	-6.4027	0.0015	0.0020	5.021	0.1811
TGARCH(1, 1) $\mathcal{N}$	-6.3972	-6.3937	-6.3960	0.016	0.0019	5.0482	0.1869
ARCH(2) $\mathcal{T}$	-6.4172	-6.4137	-6.4160	0.0037	0.005	12.046	0.4542
GARCH(1, 1) $\mathcal{T}$	-6.5558	-6.5523	-6.5546	0.001	0.0014	3.1752	0.1069
EGARCH(1, 1) $\mathcal{T}$	-6.5677	-6.5635	-6.5663	0.0017	0.002	5.3173	0.193
GJRGARCH(1, 1) $\mathcal{T}$	-6.5585	-6.5543	-6.5571	0.001	0.0014	3.2497	0.1098
APARCH(1, 1) $\mathcal{T}$	-6.5657	-6.5608	-6.5641	0.0015	0.0019	4.843	0.1723
TGARCH(1, 1) $\mathcal{T}$	-6.5645	-6.5603	-6.5631	0.017	0.0021	5.4014	0.1945
MSGARCH:($\mathcal{T}$, $\mathcal{T}$)	-	-	-	0.011	0.0182	1.139	inf

MS-GARCH VaR tends to overestimate risks. This is shown where market volatility is relatively low between 2014 and 2016, 2017 and 2018.

5 Conclusion

Among different GARCH models, asymmetric GARCH models show the best fit for the volatility of gold returns. The particularity of the volatility of gold returns has an inverse leverage effect where the asset becomes more volatile when there is upward pressure. This is thanks to the fact that investors consider gold as a safer asset when risky assets experience market turbulence. As it is shown in Figure 2, high volatility regime correspond to volatile period of risky assets. Backtesting result show that allowing regime changes outperform single regime models. Only issue with MS-GARCH estimation is that it is computationally expensive. This might render its uses when refitting the model is frequently needed such as calculating implied volatility and option pricing.

References

[1] Ghalanos A. "Package "Rugarch"". In: *CRAN* (2018), pp. 1–108.

[2] A Abramson and Cohen I. "On the Stationarity of Markov-Switching GARCH Process". In: *Econometric Theory* 23 (2007), pp. 485–500.

[3] D la Torre-Torres et al. "A Test of Using Markov-Switching GARCH models in Oil and Natural Gas Trading". In: *Energies* 1 (2019), pp. 1–24.

[4] Nelson D B. "Conditional heteroskedasticity in asset returns: A new approach". In: *Econometrica* 59(2) (1991), pp. 347–370.

[5] Rombouts J Bauwens L Preminger A. "Theory and Inference for a Markov-Switching GARCH model". In: *Working paper 07-33, Centre Interuniversitaire sur le Risque, les politiques économiques et l'emploi* (2007), pp. 1–25.

[6] Francq C and J.-M Zakoian. "Comments on the paper by Minxian Yang: "Some Properties of vector autoregressive processes with Markov-Switching coefficients"". In: *Econometric Theory* 18 (), pp. 815–818.

[7] He C and Teräsvirta T. "Properties of Moments of a Family of GARCH Processes". In: *Journal of Econometrics* 92 (1992), pp. 173–192.

[8] Hill C and Mccuillough B.D. "On the Accuracy of GARCH Estimation in R packages". In: *Econometric Reseearch in Finance* 4 (2019), pp. 133–155.

[9] Lamoureux C and W Lastrapes. "Heteroskedasticity in Stock Return Data: Volume versus GARCH effects". In: *Journal of Finance* (1990), pp. 221–229.

[10] Ardia D, Bluteau K, and Boudt K. "Markov-Switching GARCH models in R: The MSGARCH Package". In: *Journal of Statistical Software* 91(4) (2019), pp. 1–38.

[11] Kristensen D and Rahbek A. "Asymptotics of the QMLE for a class of ARCH(q) Models". In: *Econometric Theory* 21 (2005), pp. 946–961.

[12] Straumann D. "Estimation in Conditionally Volatility Heteroskeastic Time Series Models". In: *Lecture Notes in Statistics, Springer, Berlin* (2005).

[13] Diebold F. "Comment on Modeling the Persistence of Conditional Variances". In: *Econometric Reviews* (1986), pp. 51–56.

[14] Klaassen F. "Improving GARCH volatility Forecasts with Regime-Switching GARCH". In: *Empirical Economics* 27 (2002), pp. 363–394.

[15] Fiorentini G, Calzolari G, and Panattoni L. "Analytic Derivatives and The Computation of GARCH Estimates". In: *Journal of Applied Econometrics* 11, 4 (1996), pp. 399–417.

[16] Schwert G. "Why Does Stock Market Volatility Change Over Time?" In: *Journal of Finance* (1989), pp. 1115–1153.

[17] Hamilton J and R Susmel. "Autoregressive Conditional Heteroskedasticity and Changes in Regime". In: *Journal of Econometrics* 64 (), pp. 307–333.

[18] marcucci J. "Forecasting Stock Market Volatility with Regime-Switching GARCH models". In: *Studies in Nonlinear Dynamics and Econometrics* 9 (2005), pp. 1–55.

[19] Glosten L, Jogannathan R, and Engle R. "A long Memory property of Stock Market Returns and A new model". In: *Journal of Empirical Finance* 1 (1993), pp. 83–106.

[20] Lee and Hansen B.E. "Asymptotic Theory for the GARCH(1, 1) Quasi-Maximum Likelihood Estimator". In: *Econometric Theory* (1993), pp. 29–52.

[21] McAleer J M and Ling S. "Necessary and Sufficient Moment of conditions for the GARCH(r, s) and Asymmetric Power (r, s) Models". In: *Econometric Theory* 18(03) (2002), pp. 722–729.

[22] Doukhan P, Massart P, and Rio E. "The Functional Central Limit Theorem for Strongly Mixing Processes". In: *Annales de l'I.H.P, Probabilités et Statistiques* 30 (1994), pp. 63–82.

[23] Giot P and Laurent S. "Value-at-Risk for Long and Short Trading Positions". In: *Journal of Applied Econometrics* (2003), pp. 641–663.

[24] Engle R. "Autoregressive Conditional Heteroskedasticity with Estimates of United Kingdom inlfation". In: *Econometrica* (1982), pp. 987–1008.

[25] Lumsdaine R.L. "Consistency and Asymptotic Theory for the GARCH 1, 1), Quasi-Maximum Likelihood Estimator in IGARCH(1, 1) and Covariance Stationary GARCH(1, 1) models". In: *Econometrica* 64 (1992), pp. 575–596.

[26] Gray S. "Modeling the conditional distribution of interest rates as a regime-switching process". In: *Journal of Financial Economics* (), pp. 27–62.

[27] Meyn S and R Tweedie. "Markov Chains and Stochastic Stability". In: *London, Springer Verlag* (1993).

[28] Bollerslev T. "Generalized Autoregressive Conditional Heteroskedasticity". In: *Journal of Econometrics* (1986), pp. 307–327.

[29] Jensen T and Rahbek A. "Asymptotic Normality of the QML Estimator of ARCH in the Nonstationary Case". In: *Econometrica* 72(2) (2004), pp. 641–646.

[30] Mikosch T and C Starica. "Nonstationarities in Financial Time Series, The Long-Range Dependence, and the IGARCH Effects". In: *Review of Economics ans Statistics* (), pp. 378–390.

[31] Ding Z, Granger C, and Engle R. "A Long Memory Property of Stock Market Returns and A New Model". In: *Journal of Empirical Finance* 1 (1993), pp. 83–106.