INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES

HRMARS MANAGEMENT ACADEMIC RESEARCH SOCIETY
EXPLORING INTELLECTUAL CAPITAL
www.hrmars.com
ISSN: 2225-8329

Improving the Forecasting Power of Volatility Models

Ahmed Bensaida

To Link this Article: http://dx.doi.org/10.6007/IJARAFMS/v2-i3/9944 DOI:10.6007/IJARAFMS /v2-i3/9944

Received: *16* July *2012,* ***Revised:*** *11* August 2012*,* ***Accepted:*** *24* August 2012

Published Online: 18 September 2012

In-Text Citation: (Bensaida, 2012)

To Cite this Article: Bensaida, A. (2012). Improving the Forecasting Power of Volatility Models. *International Journal of Academic Research in Accounting Finance and Management Sciences*, *2*(3), 43–57.

Vol. 2, No. 3, 2012, Pg. 43 - 57

http://hrmars.com/index.php/pages/detail/IJARAFMS	JOURNAL HOMEPAGE

INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN ACCOUNTING, FINANCE AND MANAGEMENT SCIENCES
www.hrmars.com
ISSN: 2225-8329

Improving the Forecasting Power of Volatility Models

Ahmed Bensaida
Faculty of Economics and Management of Mahdia, University of Monastir, Sidi Massaoud, Hiboun 5111, Mahdia, TUNISIA
Email: ahmedbensaida@yahoo.com

Abstract
Volatility models have been extensively used in risk modeling especially GARCH models under the normal distribution. Although they generate highly significant coefficient estimates, these models are known to have poor forecasting power. It is therefore interesting to develop a different approach of risk modeling to improve forecasting results. By using the generalized t-distribution in modeling the changes in the distribution of stock index returns, the results show a significant improvement in the forecasting power. Moreover, Monte Carlo simulations have confirmed that the index returns are better explained by ARCH-type models.
Keywords: Generalized t, GARCH, Forecast, Index Return

Introduction

A trader is always faced by the risk of price fluctuation when buying or selling a given stock. In response, financial intermediaries have developed many hedging strategies to protect their positions against risk. Nevertheless, these strategies depend on the expected future volatility of the asset which is usually forecasted by GARCH models. Recent studies have shown that GARCH models have poor forecasting power and suggested the use of intra-day observations to increase forecasting efficiency. Since intra-day observations are not readily available, there is a need to modify the model properties concerning the default choice of normal distribution.

Existing literature dealing with ARCH models focuses on volatility components, and ignores the real distribution of the returns, which is an important factor for model maximum likelihood estimation. Although it is well known that the distribution of a given financial time series has thicker tails than the normal, the use of the normal distribution with ARCH-type models offer a fatter-tail conditional distribution.[1] That's why researchers did not give much interest on the used distribution,

[1] If we have $u_t = \varepsilon_t \sqrt{h_t}$, and the conditional distribution of ε_t is assumed to be time invariant with a finite fourth moment, it follows by Jensen's inequality that: $E\left(u_t^4\right) = E\left(\varepsilon_t^4\right)E\left(h_t^2\right) \geq E\left(\varepsilon_t^4\right)E\left(h_t\right)^2 = E\left(\varepsilon_t^4\right)E\left(u_t^2\right)^2$. Given a standardized normal ε_t, the unconditional distribution for u_t is therefore leptokurtic.

and have focused their efforts in search for new forms of the volatility equation inside the ARCH-type model to capture newly discovered behavior. Zhang et al. (2006) for example, have developed the Mixture GARCH, which offers thicker tails than those of the associated GARCH models regardless of the used distribution. However, it is still preferable to capture the tail thickness by the estimated distribution and not by the variance equation because the tails of the conditional normal distribution are not thick enough to describe the process and the distribution is not fully adaptable to the type of data.

This paper is divided into 6 sections: Section one describes the sample, section two describes the methodology, section three describes the generalized *t*- distribution, section four presents our results, section five is concerned with forecasting, section six is devoted to simulations, and we conclude in section seven.

Sample

The sample consists of three daily closing stock indexes: S&P 500 ${}_{S}S_t$, Nikkei 225 ${}_{N}S_t$, and CAC 40 ${}_{C}S_t$, starting from January 1st, 1996 and ending on September 15th, 2006 (10 years, 8 months and 15 days of daily observations). Data are collected from the Yahoo finance web page. The sample is thereafter divided into two sub-samples, the first ten years (until December 31st, 2005) are used for estimation and the rest is used for out-of-sample forecasting.

Let's denote S_t the spot price of a stock index, its return is computed as follow: $r_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$; hence, providing ($T$–1) observations. The return is then analyzed for linear and non-linear dependencies. Linear dependency can be detected by the study of the autocorrelation function (ACF), and partial autocorrelation function (PACF) in a way to determine the ARIMA process that can fit the observations. The degree of integration in an ARIMA (p, d, q) process can be determined by applying the GPH test developed by Geweke & Porter-Hudack (1983). All index returns are linearly independent.

The linearly whitened residuals are next tested for non-linear dependency through the BDS test developed by Geweke & Porter-Hudack (1983) and improved by Kanzler (1999).[i] Non-linear dependency is caused by non-stationarity, chaotic behavior or stochastic behavior.

Non-stationarity implies a change of the behavior of the returns over a long time period. Changes in the economy can affect such change. The non-stationarity can be caused by structural changes: technological and financial innovations, policy changes, war …etc. Chow breakpoint test shows that the returns are stable over the period of study.

Chaotic behavior is detected by the newly developed test based on the Lyapunov exponent (BenSaida A., 2007. Using the Lyapunov Exponent as a Practical Test for Noisy Chaos. Retrieved from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=970074.). The last possibility of non-linear dependency is that returns are stochastic as verified by the Lagrange multiplier or ARCH test.

2. Methodology of research

Despite the extensive work on ARCH models, the GARCH (1, 1) is still the favorite model chosen in the majority of cases. Such choice seems to be rather arbitrary. Moreover, no consistent work has yet been done on the true distribution of the risk of a given asset. Indeed, the normal distribution remains a mechanical choice in the studies. It is worth noting that some have studied the tail-fatness usually observed in financial time series data, and hence have suggested other fatter-tail distributions. Bollerslev (1987) for example has suggested the standardized *t*-distribution to model American stock price indexes and DEM/USD and GBP/USD exchange rates under the GARCH (1, 1) model and found a relative improvement over the normal distribution. However, the fact that the standardized *t*-distribution is fully adaptable is debatable and a more detailed study on the true distribution is needed.

The GARCH (*p*, *q*) model is defined as:

$$\begin{cases} r_t = y_t\varsigma + u_t \\ u_t = \varepsilon_t\sqrt{h_t}\ ;\varepsilon_t \to NID(0,1) \\ h_t = \alpha_0 + \sum_{j=1}^{p}\beta_j h_{t-j} + \sum_{i=1}^{q}\alpha_i u_{t-i}^2 \end{cases} \tag{1}$$

u_t has zero mean and ε_t are Normally and Independently Distributed (IID) with zero mean and unitary variance, ε_t are serially uncorrelated and are independent from r_t. The exogenous variable y_t may contain past realizations of r_t. Usually, when modeling time series, the dependent variable r_t is first centered to have “zero” mean, then the residuals from the regression are modeled using ARCH specification. y_t is a matrix of exogenous variables affecting the endogenous variable r_t (including autoregressive and moving average ARMA), and ς is a vector of coefficients. Non-negativity and stationarity conditions state that all coefficients must be positive: $\alpha_0 > 0$, $\alpha_i \geq 0$ (i = 1, …, q), and $\beta_j \geq 0$ (j = 1, …, p), and $\sum_{i=1}^{\max(p,q)}(\alpha_i + \beta_i) < 1$.

To estimate ARCH models we need to maximize the log-likelihood function derived from the used distribution. A quick comparison of the true dispersion represented by the non-parametric distribution along with the normal distribution of the Nikkei 225 return clearly shows that the normal choice is not adequate.

Normality tests have rejected the hypothesis of normal returns for all studied indexes (Table 1).

Table 1. Normality tests results

Normality tests		${}_Sr_t$	${}_Nr_t$	${}_Cr_t$
Jarque-Bera	*JB*	920	386	724
	p-			
Kolomogorov-	*K*	0.0	0.0	0.0
	p-	4.0	0.0	5.1
Shapiro-Wilk	*S*		0.9	0.9
	p-	3.6	3.1	2.5

* Lilliefors *p*-value < 0.01.

The estimation is carried on the 3 returns using GARCH, IGARCH, EGARCH and APARCH models. Orders *p* and *q* will be varied from (1, 1) up to (5, 5) for ${}_Sr_t$, and up to (6, 6) for ${}_Nr_t$, and ${}_Cr_t$ as suggested by the ACF and PACF of the squared residuals. Volatility in-mean specification, according to which an asset with a higher perceived risk would pay a higher return on average, and holiday effect, according to which the information stream continues even during weekends and holidays, are also tested.

Model selection is based on the Schwarz Information Criteria. Indeed, Liew and Chong (2005) have found that the Schwarz Information Criterion "SIC" identifies the true ARCH-type model better than any other information criteria.

Maximum likelihood is based on the BHHH method because this algorithm is known to be faster in execution. Sometimes, the BHHH algorithm do not reach convergence after a long number of iterations, in this case the Marquardt algorithm is used. The Marquardt method modifies the BHHH algorithm by adding a correction matrix or ridge factor to the Hessian approximation. The ridge correction handles numerical problems when the outer product is near singular and may improve the convergence rate. As above, the Marquardt algorithm pushes the updated parameter values in the direction of the gradient. In conclusion, the BHHH algorithm and the Marquardt algorithm are complementary; failure of one method to reach convergence may be cured by the other method.

The Generalized T Distribution

The GTD has the following form:

$$f\left(x,\eta,\psi,b\right)=\frac{\eta}{2b.\mathrm{B}\left(\dfrac{1}{\eta},\psi\right)\left[1+\dfrac{|x|^{\eta}}{b^{\eta}}\right]^{\psi+\frac{1}{\eta}}} \qquad (2)$$

Where $\eta > 0$, $\psi > 0$, and $b > 0$. *B(.)* is the Beta function.

An important characteristic of the GTD that it nests both the standard *t*- distribution when η=2, the degree of freedom becomes 2ψ; and the Generalized Error Distribution when ψ tends to infinity, in this case the GED has η degree of freedom. When both conditions are met, *i.e.*, η=2 and ψ tends to infinity, the GTD becomes the normal distribution.

The GTD is a symmetric function; its mean equals zero. The reason of choosing a symmetric function is quite simple to explain: the purpose of risk modeling is to determine its behavior and to give a reasonable forecast of future realizations.

Nevertheless, the fact that past realizations of stock index returns have shown a large probability for negative changes compared to positive changes, does not imply that future realizations of stock index returns will have the same gap of probability between positive values and negative values.[2] Consequently, and by taking in fact that the future is uncertain, the assumption that the stock index returns have the same chance to increase as to decrease is assumed. Therefore, a symmetric probability distribution is the best guess for an uncertain future. And if a non-symmetric distribution was assumed, a strong hypothesis for the uncertain future concerning the movement of

[2] A test for skewness has suggested that the S&P 500 return and Nikkei 225 return distributions are symmetric, the CAC 40 return distribution is skewed to the left, and the BVMT return distribution is skewed to the right.

the stock index returns is made, which is the assertion that the last would tend to move to one way more than to another; and this statement is skeptical.

The next step is to determine under which conditions the variance of GTD equals one. The only condition under which GTD has unitary variance is to set:

$$b=\sqrt{\frac{\Gamma\left(\frac{1}{\eta}\right)\Gamma(\psi)}{\Gamma\left(\frac{3}{\eta}\right)\Gamma\left(\psi-\frac{2}{\eta}\right)}}\ ,\ \eta.\psi>2 \tag{3}$$

The skewness of the GTD is zero. Its kurtosis is computed using the formula: The kurtosis of this distribution is calculated using the formula: $Kurtosis=\kappa=\frac{\int_{-\infty}^{+\infty}f(x)\left[x-E(x)\right]^4dx}{Var(x)^2}$. Knowing that *E(x)* = 0 and *Var(x)* = 1, we obtain: $\kappa=\int_{-\infty}^{\infty}x^4f(x)dx=2\int_0^{\infty}x^4f(x)dx$. To solve this integral, an integration rule proved by Gradshteyn&Ryzhik (2007), p. 341, § 3.241.4, is used: $\int_0^{+\infty}\frac{x^p}{\left(1+a.x^n\right)^k}dx=\frac{\Gamma\left(\frac{p+1}{n}\right)\Gamma\left(-\frac{p-k.n+1}{n}\right)}{n.a^{\frac{p+1}{n}}\Gamma(k)}$, under the conditions that $n>0$, $p>-1$, and $(k.n-p)>1$. In the present case, the kurtosis of GTD is:

$$\kappa_{GTD}=\frac{\Gamma\left(\frac{1}{\eta}\right)\Gamma\left(\frac{5}{\eta}\right)\Gamma(\psi)\Gamma\left(\psi-\frac{4}{\eta}\right)}{\Gamma\left(\frac{3}{\eta}\right)^2\Gamma\left(\psi-\frac{2}{\eta}\right)^2} \tag{4}$$

This kurtosis is defined when $\eta.\psi>4$ and it is useful to compare it with the sample kurtosis after estimating the distribution parameters. As η and ψ increase, the kurtosis decreases toward zero; and conversely, as the product $\eta.\psi$ goes toward 4, the kurtosis increases exponentially to reach infinity because $\lim_{x\to0^+}\Gamma(x)=+\infty$.

Since the generalized *t*-distribution is governed by two shape parameters, it becomes inevitable that the GTD can have a large variety of shapes.

The log-likelihood to be maximized is:

$$L(\theta)=T\log(\eta)-T\log\left(2b.B\left(\frac{1}{\eta},\psi\right)\right)-\left(\psi+\frac{1}{\eta}\right)\sum_{t=1}^{T}\log\left[1+\frac{|u_t|^\eta}{b^\eta h_t^{\frac{\eta}{2}}}\right]-\frac{1}{2}\sum_{t=1}^{T}\log(h_t) \tag{5}$$

For EGARCH model specification, the expectation of the absolute value of ε_t under the generalized *t*-distribution is given by:

$$E|\varepsilon_t| = \frac{\Gamma\left(\frac{2}{\eta}\right)\Gamma\left(\psi - \frac{1}{\eta}\right)}{\sqrt{\Gamma\left(\frac{1}{\eta}\right)\Gamma\left(\frac{3}{\eta}\right)\Gamma(\psi)\Gamma\left(\psi - \frac{2}{\eta}\right)}} \tag{6}$$

For the APARCH model, the stationarity condition is given by:

$\sum_{i=1}^{q}\alpha_i E\left[\left(|\varepsilon_t| - \gamma_i \varepsilon_t\right)^d\right] + \sum_{j=1}^{p}\beta_j < 1$, where:

$$E\left[\left(|\varepsilon_t| - \gamma_i \varepsilon_t\right)^d\right] = \frac{\Gamma\left(\psi - \frac{d}{\eta}\right)\left[\Gamma\left(\frac{1}{\eta}\right)\Gamma(\psi)\right]^{\frac{d}{2}-1}}{\left[\Gamma\left(\frac{3}{\eta}\right)\Gamma\left(\psi - \frac{2}{\eta}\right)\right]^{\frac{d}{2}}}\left[(1-\gamma_i)^d + (1+\gamma_i)^d\right]\frac{\Gamma\left(\frac{d+1}{\eta}\right)}{2}$$

Under the condition that: $\eta.\psi > max\ (d, 2)$. Proof is provided in the appendix (1).

Results

For comparison purpose, all models are estimated under the Gaussian errors (Table 2) and GTD (Table 3).

Table 2. Results summary under Normal errors

	$_Sr_t$	$_Nr_t$	$_Cr_t$
Model type	EGARCH	EGARCH	EGARCH
Orders (p, q)	(1, 3)	(1, 1)	(1, 1)
SIC	-15934.15	-14177.84	-15092.04

Table 3. Results summary under GTD errors

	$_Sr_t$	$_Nr_t$	$_Cr_t$
Model type	EGARCH	EGARCH	EGARCH
Orders (p, q)	(1, 3)	(1, 1)	(1, 1)
η	1.99	1.83	2.07
ψ	8.11	8.64	7.56*
SIC	-15936.44	-14200.03	-15088.72

* Not significant at 5% significance level.

The models described above perform well relatively to other same-type models. However, the question now is whether these models are consistent or not. In other words, do they capture the effect generated by the volatility of the stock index returns? For this task, a specification test is needed. When specifying ARCH type models, the errors ε_t are assumed to be independently and identically distributed IID. Therefore, it seems reasonable to use the BDS test as a specification test by applying it to the fitted residuals from the concerned model, *i.e.*, test the null hypothesis that $\varepsilon_t = \frac{u_t}{\sqrt{h_t}}$ is IID. This test has a good power for testing misspecification of ARCH-type models. Unexpected result was found for the S&P 500. The respective model is inconsistent under Gaussian and GTD errors. Did the estimation go wrong? If so, then why the estimated EGARCH model for the Nikkei 225 returns is consistent? It is possible to say that the EGARCH model is not perfectly adaptable to fit the S&P 500 return. To verify this hypothesis, the same consistency test is carried on the second best model for the S&P 500 return, which is the IGARCH model (Table 4).

Table 4. Consistent models for the S&P 500 return

	$_{S}r_t$	$_{S}r_t$
Model type	IGARCH	IGARCH
Orders (*p*, *q*)	(1, 1)	(1, 1)
Distribution	Normal	GTD
η	-	1.99
ψ	-	4.79
SIC	-15808.44	-15851.68

All estimated models under the GTD have outperformed those estimated under the normal distribution, except for the CAC 40 return due to insignificant parameter ψ. Moreover, we notice that the GTD parameter η is close to 2, in this case the GTD is nested by the standard *t*-distribution STD. Wald coefficient test and the log-likelihood ratio test have both accepted the null hypothesis $\eta = 2$ for all models. The volatility in-mean specification did not improve any of the estimated models. The Nikkei 225 return is affected by the holiday effect (Table 5).

Table 5. Results summary under STD errors

	$_{S}r_t$	$_{N}r_t$	$_{C}r_t$
Model type	IGARCH	EGARCH*	EGARCH
Orders (*p*, *q*)	(1, 1)	(1, 1)	(1, 1)
Degree of	9.3905	11.064	18.837
SIC	-15859.50	-14207.09	-15096.45

* This model includes the holiday coefficient.

Next, a standard efficiency test is conducted. This test was conducted frequently in the literature Pagan & Schwert (1990) and it consists of estimating the following model using OLS:

$$r_t^2 = a + b \cdot h_t + v_t \tag{7}$$

If the model is correctly specified and if indeed: $Var\left(r_t \mid \underline{r_{t-1}}\right) = h_t$ (the conditional volatility of the index return equals h_t), one should expect to have "*a*" and "*b*" equal zero and unity respectively. Of course, in practice the values for h_t are subject to estimation error, resulting in a standard errors-in-variables problem and a downward bias in the regression estimate for *b*. The use of such test is justified to the extent that the squared returns provide an unbiased estimator of the underlying latent volatility. The joint null hypothesis {a = 0 and b = 1} is rejected for the S&P 500. However, H_0 is accepted when conducting the test on volatility derived from GARCH models instead of IGARCH.

The R^2 is often interpreted as a simple gauge of the degree of predictability in the volatility process; and hence of the potential economic significance of the volatility forecasts. Its use as a guide to the accuracy of the volatility forecasts, however, is problematic. As discussed in Anderson & Bollerslev (1998), the realized squared returns are poor estimators of daily volatility due to the large idiosyncratic component in daily returns. Consequently, it's insignificant to interpret R^2 unless we have a benchmark for the expected value under the hypothesis of correct model specification. The (true or theoretical) population R^2 from the OLS regression under H_0 equals $R^2_{H_0} = \frac{Var(h_t)}{Var(r_t^2)}$, ht is obtained from the estimated model (Table 6).

Table 6. Theoretical vs. reported R2

Stock index	S&P 500	Nikkei 225	CAC 40
Theoretical	0.14	0.10	0.20
Reported $R2$	0.09	0.09	0.20

This form of R^2 can be written as a function of ARCH model parameters (in case of GARCH (1, 1) model, the value of R^2 is given by: $R^2 = \frac{\alpha_1^2}{1-\beta_1^2-2\alpha_1\beta_1}$), since the volatility h_t is a non-linear function of r_t. $Var(h_t)$ is a function of r_t because it is derived from the estimated GARCH model and it's a direct result given by the ARCH-type models.

Besides, with the estimated volatility ht, the population value of R^2 is below this upper bound. Therefore, a low R^2 is not an anomaly, yet a direct implication of ARCH models. Without a doubt, low R^2 largely reflects the inherent noise in the daily squared returns as a measure for the underlying latent volatility factor.

The main handicap in this procedure is that we are trying to compare volatility to simple daily squared returns, in other words the evaluation method is not adapted to the type of data. Indeed, the daily volatility cannot be represented by the simple square of the observed daily return because the variability in one day is the result of the return's change over the whole day. Anderson & Bollerslev (1998) have demonstrated that the volatility can explain much better the daily cumulative 5-minute squared returns (or continuous return) represented by the ex-post daily sample variance, *i.e.*, $\frac{\sum_{one\ day} r_{t+dt}^2}{number\ of\ observation\ in\ one\ day}$ (288 observations for each day). Moreover, if the time interval of the returns goes smaller than 5 minutes, the forecast becomes better because the ex-post daily sample variance approaches the true daily sample variance.

Forecasting

The one-step-ahead volatility forecasts are computed based on the estimated model for each stock index return relative to the out-of-sample period.

For the IGARCH(1, 1) model, the one-step-ahead volatility forecast $ht+1^*$ is computed as:

$$h_{t+1}^* = E_t\left(h_{t+1}\right) = h_{t+1} = \alpha_0 + \alpha_1 u_t^2 + \beta_1 h_t \tag{8}$$

For the EGARCH(1, 1) model, the one-step-ahead volatility forecast is:

$$h_{t+1}^* = e^{\left[\alpha_0 + \beta_1 \log(h_t) + \alpha_1\left(\zeta \frac{u_t}{\sqrt{h_t}} + \left|\frac{u_t}{\sqrt{h_t}}\right| - E\left|\varepsilon_t\right|\right)\right]} \tag{9}$$

The out-of-sample realized squared returns (from January 1st, 2006 until September 15th, 2006) are once again regressed against a constant and the one-step-ahead volatility forecasts. The obtained coefficients of multiple determinations R^2 are reported in Table 7.

Table 7. Out-of-sample R^2

Stock index		S&P 500	Nikkei 225	CAC 40
Estimated model	Theoretical	0.065	0.100	0.193
	Reported R^2	**0.013**	**0.035**	**0.104**
Same model under Normal	Theoretical	0.072	0.110	0.190
	Reported R^2	**0.007**	**0.028**	**0.098**
Rate of improvement		**85%**	**25%**	**6%**

The above results, however, does not satisfy our expectation although the forecasting power was improved under the generalized *t*-distribution. Researchers who studied ARCH models usually accept the idea that the poor forecasting power of these models is due to their type. The coefficient of determination is computed for an out-of- sample of 8 months and half, which is a long period. So to check the effect of the forecast horizon on the forecasting power of our models, R^2 is computed for the first month, and each time we increase the sample by one month and compute the R^2 again until we reach the end of the out-of-sample (Table 8).

Table 8: Out-of-sample R^2 fluctuation

*R*2 fluctuation	S&P 500	Nikkei 225	CAC 40
1/1/06 → 31/1/06	0.299	0.003	0.096
1/1/06 → 28/2/06	0.260	0.002	0.0003
1/1/06 → 31/3/06	0.068	0.027	0.0005
1/1/06 → 30/4/06	0.068	0.049	0.003
1/1/06 → 31/5/06	0.003	0.033	0.178
1/1/06 → 30/6/06	0.016	0.016	0.114
1/1/06 → 31/7/06	0.013	0.022	0.103
1/1/06 → 31/8/06	0.012	0.030	0.099
1/1/06 → 15/9/06	0.013	0.035	0.104

Except for the S&P 500 model, the forecast horizon does not have a significant effect on the coefficient of determination on the short run. An ARCH-type process is stochastic; and the volatility is generated from the return itself and not from other stochastic exogenous variable. Hence, the predictability of the ARCH-type models is weak but improved under the generalized *t*-distribution.

Simulations

The objective of the conducted simulations is to verify whether ARCH models are really adequate to describe index returns or not, so a special care is given to the random number generator *RNG* which is the core part of Monte Carlo simulations. The *RNG* used in these simulations is based on the Mersenne-Twister algorithm developed by Matsumoto & Nishimura (1998) to generate uniformly distributed random numbers with a huge period of 219937-1. Marsaglia's (2000) "*ziggurat method*" could next be applied on the uniform random numbers to obtain normally or any other distribution random numbers. The ziggurat method consists of generating random points *(x, y)* uniformly distributed in the plane, and rejects any of them that do not fall under the curve of the desired probability density function; the remaining *x*'s form the desired distribution random numbers.

Besides the ziggurat method, another more powerful method is used here. It is based on the inverse cumulative distribution function. The cumulative distribution function or *CDF* of any probability distribution is a continuous ascending function which accepts any real *x* and steadily increases from 0 to 1. Denote *F(x)* the *CDF* of the desired distribution. The idea behind using the inverse *CDF* or $F^{-1}(y)$, with $F^{-1}(y)$: [0, 1] →]-∞, +∞ [, is that if we generate uniformly distributed random numbers on the interval [0, 1], the transformed numbers through $F^{-1}(y)$ are randomly distributed on the interval]-∞, +∞[, they correspond to the used distribution random numbers. The *CDF* of the generalized *t*-distribution is given by:

$$F\left(x\right)_{GTD}=\frac{1}{2}+\frac{S\left(x\right)}{2}\left[I_{\frac{|x|^{\eta}}{|x|^{\eta}+b^{\eta}}}\left(\frac{1}{\eta},\psi\right)\right] \tag{10}$$

Where: *S(x)* is the sign function, and $I_z(a, b)$ is the *regularized incomplete beta function* that satisfies: $I_z\left(a,b\right)=\frac{1}{B\left(a,b\right)}B_z\left(a,b\right)$, with: $\mathrm{B}_z\left(a,b\right)=\int_0^z t^{a-1}\left(1-t\right)^{b-1}dt$ is the *incomplete beta* function. Note that $I_z\left(a,b\right)=1-I_{1-z}\left(b,a\right)$. Proof is provided in appendix (2).

$F^{-1}(y)$ can be derived only when knowing the inverse of the regularized incomplete beta function. Fortunately, some algorithms are designed to find solutions of this special function. Since $I_z(a, b)$ is monotone, it is still possible to find *z* that satisfies $s = I_z(a, b)$, in this case $z = I_s^{-1}(a,b)$. One good algorithm is the Newton's method.

The starting value which forms the state of the *RNG*, called *seed*, is set by the clock of the computer at the time the program was run.

For each stock index return model, the estimated coefficients and distribution are used to generate 10,000 paths or realizations; each path has the same sample size as the corresponding studied return. Afterward, the coefficients are re-estimated under the same model type and distribution for all simulated paths to check the consistence of the abovementioned methodology. The Wald coefficients and log-likelihood ratio tests are next applied on the re-estimated coefficients for every path to compare similarities between the original and the re-estimated models. The rates of acceptance of the null hypothesis that the models for the simulated paths are the same as the initial model are presented in Table 9.

Table 9. Simulated paths acceptance rate

*H*0	S&P 500	Nikkei 225*	CAC 40
Wald test	80.8%	92.7%	92.4%
LL ratio test	77.6%	96.0%	94.9%

* The rate of acceptance is non-including the holiday effect. When including the holiday effect it becomes 60% and 63%.

Conclusions

Risk modeling has known an impressive development since the first ARCH paper appeared. The trade-off between risk and return, where risk is generally measured by the volatility, is a decisive element in financial theories. In fact, accurate measures and good forecasts of future volatility are critical for the implantation and evaluation of asset pricing theories and hedging strategies. Hence, a thorough understanding of the determinants of the volatility process is crucial for issues for the functioning of markets.

However, it is still believed that the normal distribution is the best choice to use with ARCH-type models because it can explain the behavior of stock index risk and because it's the easiest one to model the volatility of any given financial asset, although all performed normality tests have strongly rejected the hypothesis of normal returns.

This is out of surprise, because it is well known that the distribution of a given financial time series has thicker tails than the normal, and the use of the normal distribution with ARCH-type models offer a fatter-tail conditional distribution. That's why researchers did not give much interest on the used distribution, and have focused their efforts in search for new forms of the volatility equation inside the ARCH-type model to capture newly discovered behavior. However, the tails of the conditional normal distribution are not thick enough to describe the process, and the distribution is not fully adaptable to the type of data.

The generalized *t*-distribution GTD was found to outperform the normal distribution in modeling the stock index volatility. It nests many other distributions, and it is more powerful in approximating the process's behavior. Hence, it is preferable to avoid the normal choice of the normal density and to choose a more adapted one.

Although the forecasting power represented by the R^2 of the ex-post model $r_t^2 = a + b \cdot h_t + v_t$ is rather due to the nature of ARCH models and to the idiosyncratic components in daily returns, in the present study the forecasting power was improved by using daily returns with different distribution. Moreover, except for the S&P 500 model, the forecast horizon does not have a significant effect on the forecasting power on the short run.

Monte Carlo simulations have confirmed the common belief that stock index returns are better explained by ARCH-type models more than any other model.

References

Anderson, G. T., and Bollerslev, T. (1998). Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review*, 39, pp. 855-905.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The Review of Economics and Statistics*, 69, pp. 542-547.

Brock, W. A., Dechert, W. D., Scheinkman, J. A., and LeBaron, B. (1996). A Test for independence based on the correlation dimension. *Economic Reviews*, 15, pp. 197-235.

Geweke, J., and Porter-Hudak, S. (1983). The Estimation and Application of Long Memory Time Series Models. *Journal of Time Series Analysis*, 4, pp. 221-238.

Gradshteyn, I. S., and Ryzhik, I. M. (2007). *Table of integrals, series, and products*. 7th edition, New York: Alain Jeffrey & Daniel Zwillinger.

Liew, V. K., and Chong, T. T. L. (2005). Autoregressive lag length selection criteria in the presence of ARCH errors. *Economics Bulletin*, 3, pp. 1-5.

Marsaglia, G., and Tsang, W. (2000). The ziggurat method for generating random variables. *Journal of statistical software*, 5, pp. 1-7.

Matsumoto, M., and Nishimura, T. (1998). Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudorandom Number Generator. *ACM Transactions on Modeling and Computer Simulation*, 8, pp. 3-30.

Pagan, A. R., and Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, 45, pp. 267-290.

Zhang, Z., Li, K. W., and Yuen, K. C. (2006). On a mixture GARCH time-series model. *Journal of Time Series Analysis,* 27, pp. 577-597.

Appendices

1. Stationarity condition of APARCH model under GTD

The stationarity condition of the APARCH model is: $\sum_{i=1}^{q}\alpha_i E\left[\left(|\varepsilon_t|-\gamma_i\varepsilon_t\right)^d\right]+\sum_{j=1}^{p}\beta_j<1$

The quantity $E\left[\left(|\varepsilon_t|-\gamma_i\varepsilon_t\right)^d\right]$ is computed as follow:

$$E\left[\left(|\varepsilon_t|-\gamma_i\varepsilon_t\right)^d\right]=E\left[\left(\varepsilon_t-\gamma_i\varepsilon_t\right)^d\right]_{\varepsilon_t=0}^{\infty}+E\left[\left(-\varepsilon_t-\gamma_i\varepsilon_t\right)^d\right]_{\varepsilon_t=-\infty}^{0}$$

$$=E\left[\varepsilon_t^d\left(1-\gamma_i\right)^d\right]_{\varepsilon_t=0}^{\infty}+E\left[\left(-\varepsilon_t\right)^d\left(1+\gamma_i\right)^d\right]_{\varepsilon_t=-\infty}^{0}$$

$$=\left(1-\gamma_i\right)^d E\left[|\varepsilon_t|^d\right]_{\varepsilon_t=0}^{\infty}+\left(1+\gamma_i\right)^d E\left[|\varepsilon_t|^d\right]_{\varepsilon_t=-\infty}^{0}$$

The GTD is symmetric with mean zero, hence: $E\left(|\varepsilon_t|^d\right)_{\varepsilon_t=0}^{\infty}=E\left(|\varepsilon_t|^d\right)_{\varepsilon_t=-\infty}^{0}=\frac{1}{2}E\left(|\varepsilon_t|^d\right)_{\varepsilon_t=-\infty}^{\varepsilon_t=+\infty}$, therefore: $E\left[\left(|\varepsilon_t|-\gamma_i\varepsilon_t\right)^d\right]=\frac{1}{2}E\left(|\varepsilon_t|^d\right)\left[\left(1+\gamma_i\right)^d+\left(1-\gamma_i\right)^d\right]$

$$E\left(|\varepsilon_t|^d\right)=E\left[\left(-\varepsilon_t\right)^d\right]_{\varepsilon_t=-\infty}^{0}+E\left[\varepsilon_t^d\right]_{\varepsilon_t=0}^{\infty}$$

$$=\int_{-\infty}^{0}(-x)^d f(x)dx+\int_0^{\infty}x^d f(x)dx$$

$$=2\int_0^{\infty}x^d f(x)dx$$

Using Gradshteyn & Ryzhik (2007), p. 341, § 3.241.4:

$\int_0^{+\infty}\frac{x^p}{\left(1+a.x^n\right)^k}dx=\frac{\Gamma\left(\frac{p+1}{n}\right)\Gamma\left(-\frac{p-k.n+1}{n}\right)}{n.a^{\frac{p+1}{n}}\Gamma(k)}$, under the conditions that $n>0$, $p>-1$ and $(k.n-p)>1$, we obtain:

$$\int_0^{\infty}x^d f(x)dx=\frac{\eta}{2b.B\left(\frac{1}{\eta},\psi\right)}\int_0^{+\infty}\frac{x^d}{\left[1+\frac{x^\eta}{b^\eta}\right]^{\psi+\frac{1}{\eta}}}dx=\frac{b^d\Gamma\left(\frac{d+1}{\eta}\right)\Gamma\left(\psi-\frac{d}{\eta}\right)}{2\Gamma\left(\frac{1}{\eta}\right)\Gamma(\psi)}$$

$$=\frac{\Gamma\left(\psi-\frac{d}{\eta}\right)\left[\Gamma\left(\frac{1}{\eta}\right)\Gamma(\psi)\right]^{\frac{d}{2}-1}}{\left[\Gamma\left(\frac{3}{\eta}\right)\Gamma\left(\psi-\frac{2}{\eta}\right)\right]^{\frac{d}{2}}}\frac{\Gamma\left(\frac{d+1}{\eta}\right)}{2}$$

Therefore,

$$E\left[\left(|\varepsilon_t|-\gamma_i\varepsilon_t\right)^d\right]_{GTD}=\frac{\Gamma\left(\psi-\frac{d}{\eta}\right)\left[\Gamma\left(\frac{1}{\eta}\right)\Gamma(\psi)\right]^{\frac{d}{2}-1}}{\left[\Gamma\left(\frac{3}{\eta}\right)\Gamma\left(\psi-\frac{2}{\eta}\right)\right]^{\frac{d}{2}}}\left[\left(1-\gamma_i\right)^d+\left(1+\gamma_i\right)^d\right]\frac{\Gamma\left(\frac{d+1}{\eta}\right)}{2}$$

2. Cumulative distribution function of the GTD

The *CDF* of the generalized *t*-distribution GTD is given by:

$$F(x)_{GTD} = \int_{-\infty}^{x} \frac{\eta}{2b.B\left(\frac{1}{\eta},\psi\right)\left[1+\frac{|t|^{\eta}}{b^{\eta}}\right]^{\psi+\frac{1}{\eta}}} dt = \frac{1}{2} + S(x)\frac{\eta}{2b.B\left(\frac{1}{\eta},\psi\right)}\int_{0}^{|x|}\frac{1}{\left[1+\frac{t^{\eta}}{b^{\eta}}\right]^{\psi+\frac{1}{\eta}}}dt$$

Where: *S(x)* is the sign function: $S(x)=\begin{cases} 1 \text{ if } x>0 \\ -1 \text{ if } x<0 \\ 0 \text{ if } x=0 \end{cases}$. Note that: $F(0)_{GTD}=\frac{1}{2}$.

Let's first compute the integral $H=\int_0^x \frac{1}{\left(1+\frac{1}{a}t^n\right)^k}dt$ when $x > 0$.

By posing $X=\frac{t^n}{t^n+a}$, we obtain: $H=\int_0^{\frac{x^n}{x^n+a}}(1-X)^k\, dt$, now replace *dt* with *dX*. Knowing that $t=a^{\frac{1}{n}}\left(\frac{1}{\frac{1}{X}-1}\right)^{\frac{1}{n}}$, we obtain $dt=a^{\frac{1}{n}}\frac{1}{n}X^{-2}\left(\frac{1}{X}-1\right)^{-\frac{1}{n}-1}dX$.

Hence;

$$H=\frac{a^{\frac{1}{n}}}{n}\int_0^{\frac{x^n}{x^n+a}}(1-X)^k X^{-2}\left(\frac{1}{X}-1\right)^{-\frac{1}{n}-1}dX=\frac{a^{\frac{1}{n}}}{n}\int_0^{\frac{x^n}{x^n+a}}X^{\frac{1}{n}-1}(1-X)^{k-\frac{1}{n}-1}dX=\frac{a^{\frac{1}{n}}}{n}B_{\frac{x^n}{x^n+a}}\left(\frac{1}{n},k-\frac{1}{n}\right)$$

Replacing this result into the *CDF*, we obtain:

$$F(x)_{GTD}=\frac{1}{2}+S(x)\frac{1}{2B\left(\frac{1}{\eta},\psi\right)}B_{\frac{|x|^{\eta}}{|x|^{\eta}+b^{\eta}}}\left(\frac{1}{\eta},\psi\right)=\frac{1}{2}+\frac{S(x)}{2}I_{\frac{|x|^{\eta}}{|x|^{\eta}+b^{\eta}}}\left(\frac{1}{\eta},\psi\right)$$

$I_z(a, b)$ is the *regularized incomplete beta function* that satisfies: $I_z(a,b)=\frac{1}{B(a,b)}B_z(a,b)$, with: $\mathrm{B}_z(a,b)=\int_0^z t^{a-1}(1-t)^{b-1}dt$ is the *incomplete beta function*, with $a > 0$ and $b > 0$.

---

[i] Kanzler (1999). Very fast and correctly sized estimation of the BDS statistic. *Unpublished manuscript*, Department of Economics, University of Oxford. Retrieved from: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=151669.