The Effect of Innovation Assumptions on Asymmetric GARCH Models for Volatility Forecasting

Diego Acuña[1(✉)], Héctor Allende-Cid[2], and Héctor Allende[3]

[1] Universidad Técnica Federico Santa María, Valparaíso, Chile
diego.acuna@usm.cl
[2] Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile
hector.allende@ucv.cl
[3] Universidad Adolfo Ibáñez, Viña Del Mar, Chile
hallende@uai.cl

Abstract. The modelling and forecasting of volatility in Time Series has been receiving great attention from researchers over the past years. In this topic, GARCH models are one of the most popular models. In this work, the effects of choosing different distribution families for the innovation process on asymmetric GARCH models are investigated. In particular, we compare A-PARCH models for the IBM stock data with Normal, Student's t, Generalized Error, skew Student's t and Pearson type-IV distributions. The main findings indicate that distributions with skewness have better performance than non-skewed distributions and that the Pearson IV distribution arises as a great candidate for the innovation process on asymmetric models.

Keywords: Financial markets · GARCH models · Asymmetry · Innovation processes

1 Introduction

In financial markets, the volatility of an asset is considered to be a metric of the risk associated with the asset itself, so its estimation is crucial in pricing models and in Value-at-Risk (VaR) calculations. Due to the large amount of research in this topic, stylized facts about the volatility of financial assets have emerged and been confirmed over the years [6], such as the mean reversion property, persistence and the asymmetric impact of innovations on the volatility. It should be expected that a good volatility model must be able to capture these stylized facts. One of the most successful systems for volatility modelling is the GARCH methodology developed by Engle et. al. [2,5] which is able to encompass many of those stylized facts. Nevertheless, in the original formulation of GARCH models, the asymmetry effect was not addressed. As a workaround,

Funded by CONICYT-PCHA/Magíster Nacional/2014 - 22141595.

A. Pardo and J. Kittler (Eds.): CIARP 2015, LNCS 9423, pp. 527–534, 2015.
DOI: 10.1007/978-3-319-25751-8_63

many extensions to this system have been proposed, being one of the most successful ones the asymmetric power ARCH model (A-PARCH) of Ding et. al. [4]. This model outperforms other classical GARCH[1] formulations for equity markets where asymmetry is a necessity [9]. Due to this reason, it is going to be the preferred volatility formulation in this work.

Since the first works of Mandelbrot [13] and Fama [7] the distribution of financial asset returns have been known to be non-Gaussian and negatively skewed [10] with heavy tails, but there is few research on the probability density function (pdf) scheme of the innovations of a GARCH process with the ability to capture the properties described before. The Normal distribution has been widely used since the first work on ARCH models [5], but it obviously fails on accounting fat tails. Also, the Student's t [2] and the generalized error distribution (GED) [15] have received some attention mainly because they are more flexible than the Normal distribution regarding fat tails and skewness. More recently, Stavroyiannis et. al. [18] proposed the use of the Pearson type-IV distribution (see [14] for a detailed study on the distribution) due to its flexibility to approximate pdfs with fat tails [12], obtaining good results for standard GARCH models compared to Student's t and skewed Student's t (s-Student) distributions in the sense of Fernandez and Steel [8].

The aim of this paper is to make a formal study on the effect of the selection of the family distribution used in the innovation part of a GARCH model for the modelling of asymmetric process. In particular, it is of interest the analysis of Normal, Student's t, GED and Pearson IV distributions on A-PARCH models, regarding the performance of the model in tasks of forecasting volatility. The results should give researchers guidelines to the correct specification of a volatility model for a financial process. The remainder of this paper is organized as follows: Section 2 explains the A-PARCH model and the distributions used for the innovation process. In Section 3, the methodology of the study and the data used for experimentation is presented. Section 4 reports the results and the analysis of the study. Finally, in Section 5 we give some concluding remarks.

2 A-PARCH Model and the Innovation Process

Lets P_t denote an asset price at time t and its continuously compounded return over the period $t-1$ to t as $r_t = ln(P_t) - ln(P_{t-1})$. Following Engle's formulation [6] we can define both the conditional mean and variance as:

$$m_t = E_{t-1}[r_t] \tag{1}$$

$$h_t = E_{t-1}[(r_t - m_t)^2] \tag{2}$$

where $E_{t-1}[u]$ is the conditional expectation of u given the information set $\mathcal{F}$ at time $t-1$ (sometimes denoted as $E[u|\mathcal{F}_{t-1}]$). The return process R_t can be defined as:

[1] The term GARCH is generally used in literature (and in this work) to refer to the entire family of GARCH models when no conflict exist.

$$R_t = m_t + \sqrt{h_t}\epsilon_t, \text{ where } E_{t-1}[\epsilon_t] = 0 \text{ and } Var_{t-1}[\epsilon_t] = 1 \tag{3}$$

As a general assumption, $\{\epsilon_t\} \sim i.i.d.\, F()$ for some distribution function F. The GARCH methodology focuses on providing an expression for the conditional variance h_t. As stated before, in this work h_t will follow the formulation of an A-PARCH(p,q) process:

$$\begin{aligned} h_t^\delta &= \alpha_0 + \sum_{i=1}^{p} \alpha_i (|R_{t-i}| - \gamma_i R_{t-i})^\delta + \sum_{j=1}^{q} \beta_j h_{t-j}^\delta, \quad \text{where} \\ \alpha_0 &> 0, \delta \geq 0, \\ \alpha_i &\geq 0, i = 1, \ldots, p, \\ -1 &< \gamma_i < 1, i = 1, \ldots, p, \\ \beta_j &\geq 0, j = 1, \ldots, q. \end{aligned} \tag{4}$$

This model imposes a Box-Cox transformation on h_t with order δ. The asymmetry is handled by the parameter γ. If $\gamma = 0$ there is no asymmetry and with $\delta = 2$ the model behaves as a standard GARCH model. Since we are interested on the asymmetry, in general γ will not be equal to zero.

2.1 The Innovation Process

Engle [5] proposed the use of the standard Normal distribution for the specification of ϵ_t. As explained before, the distribution of the returns tends to have fat tails, so the use of a Normal distribution is a strong assumption that needs to be revisited. Bollerslev [2] and Nelson [15] showed that the innovations with Student's t and GED distributions obtained better results than the Normal one. Later, several studies reported good results using skewed versions of Normal and Student's t distributions in contrast of the non-skewed versions on asymmetric models [1,16]. Recently, Stavroyiannis et. al. [18] used the Pearson type-IV distribution in a standard GARCH model outperforming skewed Student's t versions of the innovation process. So, it is of interest to perform a general study of those distributions for the innovation process on asymmetric models. As seen on equation (3) one of the requirements over the distribution used for ϵ is being specified as a zero mean and one variance process ($0-1$ from now on). Next, we show how to get a $0-1$ version for non-obvious distributions.

Generalized Error Distribution. The density of a $0-1$ GED random variable z is given by [15]:

$$\begin{aligned} f(z) &= \frac{\nu \cdot exp[-\frac{1}{2}|\frac{z}{\lambda}|^\nu]}{\lambda \cdot 2^{(1+1/\nu)}\Gamma(\frac{1}{\nu})}, \quad \text{where} \\ \lambda &= \left[2^{\frac{-2}{\nu}}\frac{\Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{3}{\nu}\right)}\right]^{\frac{1}{2}}, \quad \text{for } -\infty < z < \infty \text{ and } 0 < \nu \leq \infty. \end{aligned} \tag{5}$$

where $\Gamma(\cdot)$ is the gamma function and ν is a tail-thickness parameter. Note that when $\nu = 2$, z behaves as the standard Normal distribution.

Skewed Distributions by Inverse Scale Factors. It is possible to introduce skewness into unimodal and symmetric distributions by using inverse scale factors in the positive and negative orthant [8]. Briefly, the procedure is as follows: given a skew parameter ξ, the density of a random variable z can be represented as:

$$f(z|\xi) = \frac{2}{\xi + \xi^{-1}} \left[f(\xi z) H(-z) + f(\xi^{-1} z) H(z) \right] \tag{6}$$

where $\xi \in \mathbb{R}^+$ and $H(\cdot)$ is the Heaviside function. The mean and variance are defined as:

$$\begin{aligned} E[z] &= M_1(\xi - \xi^{-1}) \\ V[z] &= (M_2 - M_1^2)(\xi^2 - \xi^{-2}) + 2M_1^2 - M_2 \end{aligned} \tag{7}$$

where $M_r = 2\int_0^\infty z_r f(z) dz$.

It is possible to standardize skewed versions of the Normal, Student's t and GED distributions using the conditions given above.

Pearson Type-IV Distribution. A normalized version of the Pearson type-IV distribution is given by [18] using a modern form for the distribution obtained by Nagahara [14]. For a random variable z:

$$\begin{aligned} f(x) &= \frac{\hat{\sigma} \cdot \Gamma(\frac{m+1}{2})}{\sqrt{\pi} \cdot \Gamma(\frac{m}{2})} \left| \frac{\Gamma(\frac{m+1}{2} + \mathbf{i}\frac{\nu}{2})}{\Gamma(\frac{m+1}{2})} \right|^2 \frac{1}{(1+x^2)^{\frac{m+1}{2}}} \exp(-\nu \cdot \tan^{-1} x), \text{ where} \\ x &= \hat{\sigma} z + \hat{\mu} \\ \hat{\mu} &= -\frac{\nu}{m-1} \quad \text{and} \\ \hat{\sigma}^2 &= \frac{1}{m-2} \left[1 + \frac{\nu^2}{(m-1)^2} \right]. \end{aligned} \tag{8}$$

for $m > 1/2$, m controls the kurtosis and ν the asymmetry of the distribution.

3 Methodology

3.1 Data Description

The data set consist of continuously compounded returns of IBM stocks, where the estimation period spans from January 1, 1990 to January 1, 2015 (6300 observations) and the out-of-sample evaluation period spans from January 2, 2015 to May 12, 2015 (90 observations). On Figure 1 two plots of the stock price and return value for IBM are shown.

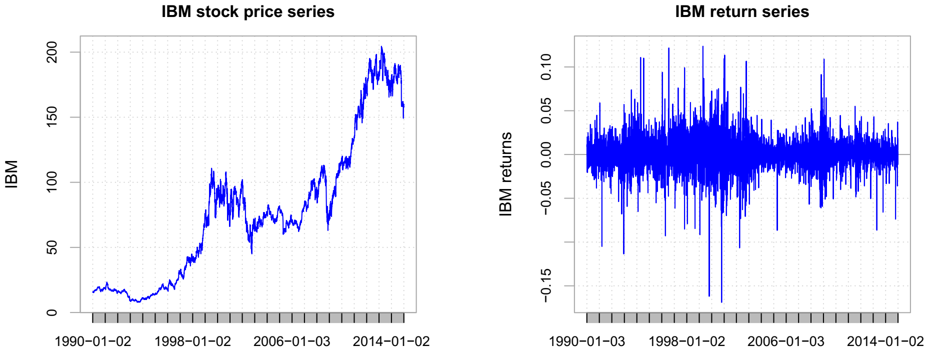

Fig. 1. IBM stock and return values.

3.2 The Model and Performance Evaluation

From a previous study of IBM return series and its autocorrelation and partial autocorrelation function, an A-PARCH(1,1) model was used for the conditional variance:

$$h_t^{\delta} = \alpha_0 + \alpha_1(|R_{t-1}| - \gamma_1 R_{t-1})^{\delta} + \beta_1 h_{t-1}^{\delta} \tag{9}$$

The five distributions (Normal, Student-t, GED, Skewed Student-t and Pearson type-IV) used for the innovation process ϵ_t have mean 0 and variance 1. The shape parameter for Student-t, GED, Skewed Student-t and Pearson type-IV distributions is equal to ν. The skew parameter for the Skewed Student-t distribution is ξ and for the Pearson type-IV is m. For the estimation phase, the preferred algorithm is maximum likelihood as usual in GARCH literature. For the evaluation of the models, an out-of-sample evaluation is going to be used with the second period of the data. Since there is no a preferred loss function for GARCH models [9], we are going to use the following performance measures:

$$MSE = n^{-1} \sum_{t=1}^{n} (\sigma_t - h_t)^2 \tag{10}$$

$$MAE = n_{-1} \sum_{t=1}^{n} |\sigma_t - h_t| \tag{11}$$

$$Logloss = -n^{-1} \sum_{t=1}^{n} (\sigma_t ln(h_t) + (1 - \sigma_t) ln(1 - h_t)) \tag{12}$$

where n is the sample size. As we need a measure for the volatility (σ_t in each loss function), a historical estimator for it is going to be needed. In particular, the Garman-Klass estimator [11] has shown to be very efficient, provided that there are no destabilizing large values [17] and it will be used in this study. The expression for the Garman-Klass estimator consists of:

$$\hat{\sigma}_t^2 = \frac{1}{2}\left(ln\frac{H_t}{L_t}\right) - (2 \cdot ln2 - 1)\left(\frac{C_t}{O_t}\right) \tag{13}$$

where O_k and C_t are the opening and close price in day t, and, H_t and L_t are the highest and lowest price of the asset in study in time t.

4 Experimentation and Analysis

The parameters estimation for the A-PARCH(1,1) model are given in Table 1 grouped by the distribution used on the innovation process. Also, below each value, in parenthesis we give the standard error for the estimated value.

Table 1. Parameters estimated for the volatility model

Innovation	μ	ω	α_1	γ_1
Normal	3.042e-04 (9.706e-05)	1.018e-03 (1.452e-04)	7.856e-02 (5.902e-03)	4.735e-01 (4.922e-02)
Student	2.380e-04 (1.119e-04)	5.872e-04 (1.368e-04)	7.509e-02 (7.371e-03)	0.4154 (6.166e-02)
GED	2.055e-04 (1.342e-04)	7.439e-04 (1.533e-04)	7.594e-02 (7.290e-03)	0.4431 (6.173e-02)
s-Student	2.565e-04 (1.461e-04)	5.934e-04 (1.388e-04)	7.53553e-02 (7.390e-03)	0.4142 (6.196e-02)
Pearson IV	2.838e-04 (1.697e-04)	5.917e-04 (1.387e-04)	7.544e-02 (7.399e-03)	0.4123 (6.193e-02)
Innovation	β_1	δ	shape	skew
Normal	9.279e-01 (5.466e-03)	6.406e-01 (8.009e-02)	- -	- -
Student	0.9344 (6.848e-03)	0.6932 (0.1050)	5.1893 (0.3170)	- -
GED	0.9319 (6.8529e-03)	0.6737 (0.1012)	1.2744 (2.771e-02)	- -
s-Student	0.9342 (6.8692e-03)	0.6913 (0.1050)	5.1859 (0.3164)	1.0156 (1.808e-02)
Pearson IV	0.9341 (6.8844e-03)	0.6914 (0.1055)	5.1835 (0.3161)	-0.0874 (0.1203)

It is interesting to note that, for the Student's t, skewed Student's t and the Pearson IV distributions the shape parameter is practically the same. The difference between those three distribution is given by the skewness: the Student's t is symmetric, where the skewed Student's t and Pearson IV are asymmetric distributions. In Table 2, we show the performance measures (loss function) used.

For all cases from the out-of-sample evaluation the Pearson IV distribution gave the best results. For the log-likelihood of the estimation phase, the skewed Student's t distribution obtained the best result. Those results show that the skewness for the innovation process plays a fundamental role in the volatility modelling.

Table 2. Loss functions and Log Likelihood for each model estimated (in bold each value which is the best measure for the selected loss function)

Innovation	MSE	MAE	Log Loss	Log Likelihood
Normal	7.4849e-06	2.19496e-03	6.44927e-02	-8849.456
Student	3.6194e-06	1.42859e-03	6.43533e-02	17541.49
GED	3.6519e-06	1.43351e-03	6.43545e-02	-8583.458
s-Student	3.6167e-06	1.42848e-03	6.43531e-02	**17541.91**
Pearson IV	**3.5970e-06**	**1.42641e-03**	**6.43524e-02**	17541.83

For the forecast made (in table 2) we can observe that the best distributions for the innovation process were the Pearson type-IV followed by the skewed Student's t distribution. A corrected Diebold-Mariano test [3] for checking the statistical significance of the better forecast obtained from Pearson IV was made. The results are presented in table 3 where it can be seen that for the MSE, it is clear that the Pearson IV distribution has statistically different forecast accuracy than the skewed Student's t and because it has lower MSE in this case we choose the Pearson IV innovation. For the MAE loss function we cannot reject the null hypothesis so statistically the two forecast have the same accuracy.

Table 3. Diebold Mariano test statistic (DM) for forecast comparison using skewed Student's t and Pearson IV distributions.

Innovation used in forecast	Test result	MSE	MAE
s-Student vs Pearson IV	DM	2.3176	0.6521
	p-value	0.0277	0.5195

5 Conclusion

The election of the distributions used for the disturbances in GARCH asymmetric volatility modelling has shown to be an important phase on the construction, estimation and forecasting in volatility financial markets models. In particular, we have concluded that the model with skewed distributions outperforms the model with non-skewed distributions in the forecast of the IBM stock time series in terms of the performance measures MSE, MAE and Log Loss.

We tested the introduction of the Pearson type-IV distribution for the disturbances on an A-PARCH model. The results obtained in IBM stock time series are encouraging (similar than for standard GARCH models [18]). The proposal was also validated with the S&P-500 time series obtaining similar results (available upon request). Future work will deal with the validation of this formulation with other well-known time series.

References

1. Alberg, D., Shalit, H., Yosef, R.: Estimating stock market volatility using asymmetric GARCH models. Applied Financial Economics **18**(15), 1201–1208 (2008)
2. Bollerslev, T.: Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics **31**, 307–327 (1986)
3. Diebold, F.X., Mariano, R.S.: Comparing predictive accuracy. Journal of Business & Economic Statistics **20**(1), 134–144 (2002)
4. Ding, Z., Granger, C.W.J., Engle, R.F.: A long memory property of stock market returns and a new model. Journal of Empirical Finance **1**(1), 83–106 (1993)
5. Engle, R.F.: Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the Econometric Society **50**(4), 987–1007 (1982)
6. Engle, R.F., Patton, A.J.: What good is a volatility model. Quantitative Finance **1**(2), 237–245 (2001)
7. Fama, E.F.: The Behavior of Stock-Market Prices **38**(1), 34–105 (1965)
8. Fernandez, C., Steel, M.F.J.: On Bayesian Modeling of Fat Tails and Skewness. Journal of the American Statistical Association **93**(441), 359–371 (1998)
9. Hansen, P.R., Lunde, A.: A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? Journal of Applied Econometrics **20**(7), 873–889 (2005)
10. Jondeau, E., Poon, S.H., Rockinger, M.: Financial modeling under non-Gaussian distributions. Springer Science & Business Media (2007)
11. Klass, M.J., Mark, B.G.: On the Estimation of Security Price Volatility from Historical Data. The Journal of Business **53**, 67–78 (1980)
12. Magdalinos, M.A., Mitsopoulos, G.P.: Conditional heteroskedasticity models with pearson family disturbances. The Refinement of Econometric Estimation and Test Procedures, Finite Sample and Asymptotic Analysis, 1–33 (2007)
13. Mandelbrot, B.: The Variation of Certain Speculative Prices (1963)
14. Nagahara, Y.: The PDF and CF of Pearson type IV distributions and the ML estimation of the parameters. Statistics & Probability Letters **43**(3), 251–264 (1999)
15. Nelson, D.B.: Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the Econometric Society **59**, 347–370 (1991)
16. Peters, J.P.: Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (Skewed) Student-t densities. Preprint, University of Liege, Belgium (2001)
17. Poon, S.H.: A practical guide to forecasting financial market volatility. Wiley Finance (2005)
18. Stavroyiannis, S., Makris, I., Nikolaidis, V., Zarangas, L.: Econometric modeling and value-at-risk using the Pearson type-IV distribution. International Review of Financial Analysis **22**, 10–17 (2012)