This is an Author's Accepted Manuscript of an article published as: CURTO J.D., PINTO J.C. and TAVARES, G.N. (2009) *Modeling stock markets' volatility using GARCH models with Normal, Student's t and stable Paretian distributions*. **Statistical Papers**, 50(2), 311-321 available online at:
http://dx.doi.org/10.1007/s00362-007-0080-5

myjournal manuscript No.
(will be inserted by the editor)

Modeling stock markets' volatility using GARCH models with Normal, Student's *t* and stable Paretian distributions

José Dias Curto[1], José Castro Pinto[1], Gonçalo Nuno Tavares[2]

[1] ISCTE Business School, Department of Quantitative Methods, Complexo INDEG/ISCTE, Av. Prof. Aníbal Bettencourt, 1600-189 Lisboa, Portugal. Phone: 351 21 7826100. Fax: 351 21 7958605. E-mail: dias.curto@iscte.pt.

[2] Instituto de Engenharia de Sistemas e Computadores - Investigação e Desenvolvimento (INESC-ID), Lisbon and Department of Electrical and Computer Engineering, Instituto Superior Técnico (IST), Lisbon.

Received: date / Revised version: date

Abstract As GARCH models and stable Paretian distributions have been revisited in the recent past with the papers of Hansen and Lunde (2005) and Bidarkota and McCulloch (2004), respectively, in this paper we discuss alternative conditional distributional models for the daily returns of the US, German and Portuguese main stock market indexes, considering ARMA-GARCH models driven by Normal, Student's t and stable Paretian distributed innovations. We find that a GARCH model with stable Paretian innovations fits returns clearly better than the more popular Normal distribution and slightly better than the Student's t distribution. However, the Student's t outperforms the Normal and stable Paretian distributions when the out-of-sample density forecasts are considered.

Key words Non-Gaussian distributions, Conditional heteroskedasticity.

1 Introduction

Since the birth of the modern empirical finance two main approaches have been considered to model the empirical distribution of financial assets returns. The first one, that we name the unconditional approach, admits that stock prices follow a random walk and several models have been proposed to describe the unconditional distribution of financial returns.

However, as the empirical findings suggest the presence of volatility clusters, one might represent this kind of returns behavior using a model where the conditional variance is serially correlated and since the seminal paper of Engle (1982) was published, the second one, that we name the conditional approach, became common in empirical finance.

The unconditional approach is based on the assumption that location and scale parameters are constant, implying that returns are independent and identically distributed (i.i.d.) random variables. The Gaussian distribution was the first to be considered and the normality became one of the most important assumptions in the classical financial models, namely the Portfolio Theory, the Capital Asset Pricing Model (CAPM) and the

Black-Scholes' formula. The Gaussian hypothesis was not seriously questioned until the seminal papers of Mandelbrot (1963) and Fama (1965) were published. Since then, numerous studies have found that the empirical distribution of returns on financial assets exhibit fatter tails and are more peaked around the center than would be predicted by a Gaussian distribution. Thus, alternative distributions possessing such characteristics have been proposed as models for the unconditional distribution of returns. Mandelbrot (1963) attempted to capture the excess of kurtosis by modeling the returns distribution as a member of stable-Lévy or stable Paretian distributions, where the Gaussian distribution is a special case. Other researchers have proposed alternative distributions like the Student's t (Blattberg and Gonedes 1974), the GED: Generalized Error Distribution (Box and Tiao 1962), the Laplace and double Weibull (Mittnik and Rachev 1993).

On the other hand, the conditional approach admits temporal dependencies in the returns series allowing the financial modeling based frequently on past information. Traditionally, serial dependence in time series has been modeled with ARMA (Autoregressive Moving Average) structures, as this class of models provides a good specification for the conditional mean. However, given the homoskedastic nature of the conditional distribution implicit in these models, they are unable to capture the volatility clustering that is common in financial asset returns.

As heteroskedasticity is a common characteristic of returns, and given the importance of predicting volatility in many asset-pricing and portfolio management problems, different solutions for conditional modeling of returns have been proposed in the literature. As the high-frequency financial data exhibits volatility clustering, i. e., large (small) price changes tend to be followed by large (small) changes of either sign (Mandelbrot 1963), the most popular one is the class of Autoregressive Conditional Heteroskedasticity (ARCH) models originally introduced by Engle (1982) and latter Generalized (GARCH) by Bollerslev (1986), possibly in combination with an ARMA specification for the mean equation, referred to as an ARMA-GARCH model.

In its standard form GARCH models assume that the conditional distribution of assets returns is Gaussian. However, for many financial time series, this model specification does not adequately account for leptokurtosis. Thus, several non-Normal alternative distributions have been proposed. Bollerslev (1987) suggests using the Student's t distribution. Nelson (1991) proposed the GED, the Laplace distribution has been employed in Granger and Ding (1995) and Hsieh (1989) used both Student's t and GED as distributional alternative models for innovations. The stable Paretian distributions have been also investigated by Liu and Brorsen (1995), Panorska et al. (1995) and Mittnik et al. (1998, 2003).

In this paper we examine the conditional distribution of daily returns in the US, the German and the Portuguese equity markets, comparing the stable Paretian distribution to the Gaussian and the Student's t distributions for innovations. The Portuguese market is smaller, more recent and less liquid. The German market is representative of intermediary markets, with a longer history but its role on the economy is lower than typical Anglosaxon capital markets. Finally, these markets were compared with the American market, the most liquid and one of the oldest in the World.

The paper is organized as follows. Next section provides a brief description of the GARCH model with Normal, Student's t and stable Paretian innovations. Section 3 describes the returns and presents some preliminary findings. Section 4 discusses the estimation results and compares the goodness-of-fit of the three conditional distributions. Section 5 presents the out-of-sample evaluation results and the final section summarizes our concluding remarks.

2 The models

As it was referred before, ARMA models can be used to explain the conditional mean of the process based on past realizations:

$$r_t = c + \sum_{i=1}^{p} \phi_i r_{t-i} + \sum_{i=1}^{q} \theta_i u_{t-i} + u_t, \tag{1}$$

where u_t is a white noise process and r_t is return at time t.

GARCH models are commonly used to capture the volatility clusters of returns and express the conditional variance as a linear function of past information, allowing the conditional heteroskedasticity of returns. In its standard version, GARCH models are assumed to be driven by normally distributed innovations (Bollerslev, 1986):

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i u_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}^2 \quad \text{and} \quad u_t|\Phi_{t-1} \sim N\left(0;\ \sigma_t^2\right), \qquad (2)$$

where σ_t^2, the scale parameter, represents the conditional variance of the process at time t, $u_t = \varepsilon_t \sigma_t$ and $\Phi_{t-1} = \{u_{t-1},\ u_{t-2},\ ...\}$.

In addition to the normal GARCH model, we also consider two other specifications assuming Student's t [$\varepsilon_t \sim t(v)$, where $t(v)$ refers to the zero-mean t distribution with v degrees of freedom and scale parameter equal to one] and stable Paretian innovations. The parameterization and the properties of the stable Paretian GARCH model are discussed next. The more detailed description of this model is due to its less frequent use to modeling the conditional distribution of financial assets returns.

A process r_t is called a stable Paretian GARCH, $S_{\alpha,\beta}GARCH\,(r, s)$, if it is described by (1) (Liu and Brorsen 1995, Panorska et al. 1995 and Mittnik et al. 1998) and,

$$\sigma_t = \alpha_0 + \sum_{i=1}^{r} \alpha_i \,|\, u_{t-i}| + \sum_{i=1}^{s} \beta_i \sigma_{t-i}, \qquad (3)$$

where σ_t is the scale parameter of the process at time t, $u_t = \sigma_t \varepsilon_t$, and ε_t are i.i.d. realizations of a stable Paretian distributed random variable with $\alpha > 1$: $\varepsilon_t \overset{iid}{\sim} S_{\alpha,\beta}$. $S_{\alpha,\beta}$ represents the standard asymmetric stable Paretian distribution with stable index $\alpha \in (0;\ 2]$, skewness parameter $\beta \in [-1;\ 1]$, zero location parameter and unit scale parameter. From the several notational alternatives of the stable Paretian distribution we select the parameterization suggested by Samorodnitsky and Taqqu (1994) and

Rachev and Mittnik (2000), whereby

$$E\left(e^{itX}\right)=\begin{cases}\exp\left\{i\,t-\left|\,t\right|^{\,\alpha}\left[1-i\beta\tan\frac{\pi\alpha}{2}sign\ (t)\right]\right\}, & \text{if}\quad \alpha\neq 1\\ \exp\left\{i\,t-\left|\,t\right|\,\left[1+i\beta\frac{2}{\pi}sign\ (t)\ln\left|\,t\right|\right]\right\}, & \text{if}\quad \alpha=1\end{cases} \tag{4}$$

is the characteristic function.

The distribution is symmetric about the zero location parameter if $\beta = 0$ and the characteristic exponent α determines the total probability in the extreme tails of the distribution. When $\alpha = 2$, the underlying stable Paretian distribution is the Normal distribution: $N\left(0;\ 2\right)$, with finite moments of all orders. As α decreases from 2 to 0, the tail areas of the stable distribution become increasingly fatter than the Normal. For $\alpha < 2$, the s^{th} absolute moment exists only for $s < \alpha$ and, except for the Normal case, the stable Paretian distribution has infinite variance. Thus, the analog of the variance in this kind of distributions is the scale parameter σ_t. In this paper we assume that $\alpha > 1$, the necessary condition for the existence of the mean.

The $S_{\alpha,\beta}GARCH\left(r,s\right)$ process defined by (1) and (3) with $1 < \alpha < 2$ has a unique strictly stationary solution if $\alpha_i > 0$, $i = 0$, ..., r, $\beta_i > 0$, $i = 1$, ..., s and the measure of volatility persistence, $V_S = \lambda_{\alpha,\beta}\sum\limits_{i=1}^{r}\alpha_i + \sum\limits_{j=1}^{s}\beta_j$, satisfies $V_S \leq 1$ (Rachev and Mittnik 2000), where

$$\lambda_{\alpha,\beta}=E\left|\varepsilon_t\right|=\frac{2}{\pi}\Gamma\left(1-\frac{1}{\alpha}\right)\left(1+\tau_{\alpha,\beta}^2\right)^{\frac{1}{2\alpha}}\cos\left(\frac{1}{\alpha}\arctan\tau_{\alpha,\beta}\right), \tag{5}$$

with $\tau_{\alpha,\beta} = \beta\tan\frac{\alpha\pi}{2}$.

If V_s is strictly less than one, this implies a conditional volatility equation where the impact of a shock dies out over time. If $V_S = 1$, analogous to the ordinary normal GARCH model, we say that r_t is an *inte-*

grated $S_{\alpha,\beta}GARCH\,(r,s)$ process, denoted by $S_{\alpha,\beta}IGARCH\,(r,s)$, which implies non-decaying effects of shocks on the conditional volatility (Engle and Bollerslev 1986). In practice, the estimated volatility persistence, $\hat{V}_S$, tends to be quite close to one for highly volatile series and the integrated model might offer a reasonable data description.

3 Statistical properties of returns

The data consists of daily closing prices of DJIA, DAX and PSI20 indexes, which are main market indexes for the US, German and Portuguese equity markets. These series cover the period from December 31, 1992 to December 31, 2006 yielding 3527, 3565 and 3518 observations, respectively. We analyze the continuously compounded percentage rates of return (not adjusted for dividends) that are calculated by taking the first differences of the logarithm of series (P_t is the closing value for each stock index at time t):

$$r_t = 100 \times \left[\ln\left(P_t\right) - \ln\left(P_{t-1}\right)\right]. \tag{6}$$

Table 1 summarizes the basic statistical properties of the data. The means returns are all positive but close to zero. The returns appear to be somewhat asymmetric as reflected by negative skewness estimates. All three series returns have heavy tails and show strong departure from normality (skewness and kurtosis coefficients are all statistically different from those of the standard Normal distribution which are 0 and 3, respectively). The Jarque-Bera test also clearly rejects the null hypothesis of normality.

Table 1 Summary statistics of returns

Statistics	DJIA	DAX	PSI20
No. of observations	3527	3565	3518
Mean	0.037667	0.040936	0.037439
Median	0.049530	0.103311	0.037463
Maximum	6.154722	7.552676	6.941259
Minimum	-7.454935	-8.882302	-9.589772
Standard deviation	1.001393	1.444509	0.989100
Skewness	-0.227090	-0.259241	-0.630273
Kurtosis	7.748924	6.474062	11.13506
JB[a]	3344.56*	1832.70*	9933.68*
LB Q(10)[b]	15.021	12.77	124.22*
LB $Q^2(10)$[c]	916.56*	2375.90*	849.61*
ARCH(1) LM[d]	110.14*	185.06*	202.42*

* Denotes significant at the 1% level,

[a] JB is the Jarque Bera test for normality,
[b] LB Q(10) is the Ljung-Box test for returns,
[c] LB $Q^2(10)$ is the Ljung-Box test for squared returns,
[d] LM is the Engle's Lagrange Multiplier test for heteroskedasticity.

According to the Ljung-Box statistic for returns, there is no relevant autocorrelation for the DAX and the DJIA indexes. For PSI20, however, returns have statistically significant first order autocorrelation, which can be removed by fitting an autoregressive AR(1) model to this series. The successive returns correlation must be due to the market thinness and non-synchronous trading that is common in many small capital markets.

Even though the series of returns seems to be serially uncorrelated over time, the Ljung-Box statistic for up to tenth order serial correlation of squared returns is highly significant at any level for the three stock indexes, suggesting the presence of strong nonlinear dependence in the data. As non-linear dependence and heavy-tailed unconditional distributions are characteristic of conditionally heteroskedastic data, the Lagrange Multiplier test (Engle, 1982) can be used to formally test the presence of conditional heteroskedasticity and the evidence of ARCH effects. The LM test for a first-order linear ARCH effect (in the last row of table 1) suggests that all stock indexes' returns exhibit ARCH effects, implying that nonlinearities must enter through the variance of the processes (Hsieh, 1989). Such behavior can be captured by incorporating ARCH or GARCH structures in the model, allowing conditional heteroskedasticity by conditioning the volatility of the process on past information. In the next section we use ARMA-GARCH models to describe the conditional distribution of returns.

4 Modeling the empirical distribution of returns

In order to evaluate in-sample and out-of-sample results, the data sample is divided in two parts. One part is used for the in-sample models' estimation; the other part is for the models' performance out-of-sample evaluation. We consider the nine first years of observations to estimate the models (from December 1992 to December 2001, yielding 2268, 2290 and 2226 observations for DJIA, DAX and PSI20, respectively). For the out-of-sample analysis we use the remaining observations comprising the years from 2002 to 2006 (the number of observations is 1259, 1275 and 1292, respectively). In this section we discuss in-sample estimation results.

The ARMA p and q orders ($p = q = 0$ for DJIA, DAX and $p = 1, q = 0$ for PSI20) were determined by the inspection of the sample autocorrelation function (SACF) and the sample partial autocorrelation function (PSACF) of the returns series (not shown in the paper). As is common in financial modeling (see Hansen and Lunde 2005 and Bollerslev et al. 1992) it was found that a GARCH(1,1) specification was adequate to capture the correlation in the squared returns.

The parameters were jointly estimated using conditional ML[1], and it was assumed that the scaled innovations, u_t, are either Normal, Student's t and stable Paretian and σ_t satisfies GARCH recursions (2) and (3), respectively. As there is no analytical expression for the stable Paretian density, the ML estimation is approximate once the density function needs to be approximate. We follow the algorithm of Mittnik et al. (1999) which approximates the stable Paretian density through fast Fourier transform (FFT) of the characteristic function.

To compare unconditional and conditional in-sample fitted models, we employ three likelihood based goodness-of-fit criteria proposed by Mittnik and Paolella (2003). The first is the maximum log-likelihood value obtained from ML estimation (L). This value allows us to judge which model is more likely to have generated the data. The second is the bias-corrected Akaike (Hurvich and Tsai 1989) information criteria (AICC) and the third is the

[1] Estimates are obtained numerically using MATLAB-based custom software.

Schwarz Bayesian criteria (SBC):

$$AICC = -2\log L\left(\hat{\theta}\right) + \frac{2n\,(k+1)}{n-k-2}, \quad SBC = -2\log L\left(\hat{\theta}\right) + \frac{k\ln(n)}{n}, \quad (7)$$

where $\log L\left(\hat{\theta}\right)$ is the maximum log-likelihood value, n is the number of observations and k is the number of parameters. The distribution with a lower value for these information criteria is judged to be preferable. These criteria are also recommended by Sin and White (1996).

To evaluate the goodness-of-fit improvement when the conditional structure is adopted, first we compute the goodness-of-fit measures for the unconditional distributions (not shown in the paper). The Student's t model provides much closer approximation to the empirical unconditional density of PSI20 and DAX returns than Normal and stable Paretian. On the other hand, the stable Paretian model is favored quite strongly for DJIA returns with the largest maximized log-likelihood value and the smallest value for both AICC and SBC information criteria.

The estimation results of the returns' conditional distribution are reported in tables 2, 3 and 4.

These results suggest some important conclusions regarding the unconditional and conditional distributions of returns. All the estimated coefficients are significant at the 5% level. When compared to their unconditional counterparts, the improvement in fit obtained by adopting the conditional models is substantial for each distribution and for all the three stock market indexes. For example, in the stable Paretian case, the maximum log-likelihood value increases from $-3,086.83$ to $-2,913.43$ in the case of DJIA,

Table 2 Maximum likelihood estimates and goodness-of-fit statistics of GARCH(1,1) for DJIA (standard errors are in parentheses)

Estimates	Normal	Student's t	Stable
Intercept c	0.0695 (0.0169)	0.0785 (0.0159)	0.0596 (0.0023)
GARCH parameters			
α_0	0.0115 (0.0024)	0.0079 (0.0028)	0.0061 (0.0013)
α_1	0.0901 (0.0070)	0.0650 (0.0102)	0.0497 (0.0005)
β_1	0.9029 (0.0080)	0.9291 (0.0107)	0.9325 (0.0018)
V[a]	0.9930 (n. a.)	0.9941 (n. a.)	0.9909 (0.0099)
Shape[b]	(n.a.)	6.8566 (0.8779)	1.9252 (0.0089)
Skew[c]	(n.a.)	(n.a.)	-0.9516 (0.0063)
L[d]	-2,969.57	-2,915.51	-2,913.43
AICC[e]	5,951.18	5,849.07	5,842.92
SBC[f]	5,939.16	5,835.04	5,826.88

[a] V is the measure of volatility persistence: $V = \lambda\alpha_1 + \beta_1$ for the Normal and Student's t distributions and $V = \lambda_{\alpha,\beta}\alpha_1 + \beta_1$ for the stable distribution, with $\lambda_{\alpha,\beta}$ given in eq. (5),
[b] "Shape" denotes the degrees of freedom parameter v for the Student's t distribution and stable index α for the stable Paretian distribution,
[c] "Skew" refers to the stable Paretian skewness parameter β,
[d] L refers to the maximum log-likelihood value,
[e] AICC is the bias-corrected Akaike Information Criterion,
[f] SBC is the Schwarz Bayesian Criterion.

from $-3,836,70$ to $-3,625.27$ in the case of DAX and from $-3,128.86$ to $-2,807.73$ in the case of PSI20.

The stable Paretian distribution, by contrast with the unconditional case for DAX and PSI20, achieves the largest likelihood value modeling the conditional distribution of all the stock market indexes' returns, despite the small difference for the Student's t. Consequently, the lower value of both information criteria and the maximum log-likelihood value indicate that the

Table 3 Maximum likelihood estimates and goodness-of-fit statistics of GARCH(1,1) for DAX (standard errors are in parentheses)

Estimates	Normal	Student's t	Stable
Intercept c	0.0729 (0.0226)	0.0867 (0.0218)	0.0679 (0.0210)
GARCH parameters			
α_0	0.0281 (0.0049)	0.0160 (0.0061)	0.0202 (0.0029)
α_1	0.0923 (0.0097)	0.0781 (0.0116)	0.0588 (0.0078)
β_1	0.8939 (0.0106)	0.9154 (0.0126)	0.9058 (0.0108)
V	0.9861 (n. a.)	0.9935 (n. a.)	0.9759 (0.0034)
Shape	(n.a.)	9.6127 (1.5919)	1.8977 (0.0216)
Skew	(n.a.)	(n.a.)	-0.4259 (0.0975)
L	-3,654.83	-3,627.86	-3,625.27
AICC	7,321.70	7,269.76	7,266.60
SBC	7,309.68	7,255.73	7,250.56

See the footnotes on table 2 for the meaning of the estimates.

GARCH(1,1) model with conditional stable Paretian distribution provides a better fit to describe the volatility of returns in the US, German and Portuguese equity markets.

The persistence-of-volatility estimates are very near one showing that conditional models for returns are very close to being integrated. Thus, we also estimated the distributional models with the IGARCH conditions imposed. Not surprisingly, as the persistence measure is close to unity, the IGARCH-restricted parameter estimates and the goodness-of-fit statistics differ very little when compared to non restricted models. This is the reason why the estimation results are not shown in the paper.

Table 4 Maximum likelihood estimates and goodness-of-fit statistics of AR(1)-GARCH(1,1) for PSI20 (standard errors are in parentheses)

Estimates	Normal	Student's t	Stable
Conditional mean			
c	0.0289 (0.0140)	0.0339 (0.0137)	0.0307 (0.0143)
ϕ_1	0.2112 (0.0197)	0.2102 (0.0204)	0.2159 (0.0186)
GARCH parameters			
α_0	0.0152 (0.0027)	0.0140 (0.0036)	0.0149 (0.0029)
α_1	0.2080 (0.0121)	0.2046 (0.0230)	0.1150 (0.0078)
β_1	0.7918 (0.0097)	0.7950 (0.0165)	0.8405 (0.0118)
V	0.9998 (n. a.)	0.9996 (n. a.)	0.9961 (0.0041)
Shape	(n.a.)	6.0374 (0.7418)	1.8553 (0.0216)
Skew	(n.a.)	(n.a.)	-0.0919 (0.0975)
L	-2,870.55	-2,809.92	-2,807.73
AICC	5,753.14	5,627.86	5,625.49
SBC	5,741.12	5,619.85	5,615.47

See the footnotes on table 2 for the meaning of the estimates.

5 Out-of-sample density forecasts

As we are testing non-Gaussian distributions for the underlying process of returns, we support the out-of-sample analysis in density forecasts (Poon and Granger, 2003). Thus, rather predicting first and second moments as is common in the GARCH literature, we focus on density predictions and the overall density forecasting performance of competing models can be compared by evaluating their conditional densities at the future observed values (Mittnik and Paolella, 2003):

$$\hat{f}_{t+1|t}(r_{t+1}) = f\left[\frac{r_{t+1} - \mu\left(\hat{\theta}_t\right)}{\sigma_{t+1}\left(\hat{\theta}_t\right)}|r_t, r_{t-1}, \ldots\right], \tag{8}$$

where $\hat{f}_{t+1|t}(\cdot)$ is the one-step-ahead density forecast, $\hat{\theta}_t$ refers to the estimated parameters based on the sample information up to and including period t and σ_{t+1} results from (2) and (3) by using $\hat{\theta}_t$. We re-estimate (via ML estimation) the model parameters at each step, as is common in actual applications. Thus, 1258, 1274 and 1291 observations are used for DJIA, DAX and PSI20 out-of-sample density forecasts evaluation, respectively. A model will fare well in the comparison among competing models if realization r_{t+1} is near the mode of $\hat{f}_{t+1|t}(r_{t+1})$ and if the mode of the conditional density is more peaked (Mittnik and Paolella, 2003).

Table 5 presents the means, standard deviations and medians of the estimated density values $\hat{f}_{t+1|t}(r_{t+1})$. For all the three stock market indexes (except for DAX when the median is considered) the best performance is achieved by the Student's t distribution when both central tendency measures are used. Notice that this is contrary to the model selection based on the goodness-of-fit measures where the stable Paretian distribution had better in-sample performance to describe returns data.

6 Conclusions

In this paper we have discussed alternative unconditional and conditional distributional models for the daily returns of DJIA, DAX and PSI20 stock market indexes considering AR-GARCH models driven by Normal, Student's t and stable Paretian distributed innovations.

Table 5 Comparison of overall predictive performance

	Distribution	DJIA	DAX	PSI20
Mean	Normal	0.3597	0.2562	0.4462
	Student's t	0.3733	0.2651	0.4776
	Stable	0.2962	0.2339	0.3499
St. deviation	Normal	0.1839	0.1471	0.2239
	Student's t	0.2101	0.1651	0.2726
	Stable	0.1193	0.1110	0.1156
Median	Normal	0.3571	0.2477	0.4419
	Student's t	0.3603	0.2509	0.4593
	Stable	0.3228	0.2566	0.3747

The entries represent average predictive likelihood values.

The behavior of returns has the common characteristics of high-frequency financial time series. First, they exhibit a considerable level of excess kurtosis. Second, nevertheless the small autocorrelation of returns, the serial dependence in the square returns is clearly not rejected pointing towards the existence of volatility clusters. Finally, the volatility of returns tends to be highly persistent.

When compared to their unconditional counterparts, the improvement in fit obtained by the conditional models is substantial for each distribution and for all the three stock market indexes.

By contrast with the unconditional case, where the Student's t achieves the best fit for DAX and PSI20 returns, the stable Paretian distribution achieves the largest likelihood value modeling all the three conditional distributions of returns, despite the small difference for the Student's t. However, in the out-of-sample analysis the best performance is also achieved by the Student's t distribution for all the three stock market indexes.

Acknowledgements

The authors would like to thank the editor and the referee whose thoughtful comments led to substantial improvements in the paper.

References

1. Akaike H (1978) Time series analysis and control through parametric models. Applied Time Series Analysis. In: D. F. FINDLEY (Ed.), Academic Press, New York
2. Bidarkota PV, McCulloch JH (2004) Testing for persistence in stock returns with GARCH-stable shocks. Quantitative Finance 4, 256-265
3. Blattberg RC, Gonedes NJ (1974) A comparison of the Stable and Student distributions as statistical methods for stock prices. Journal of Business 47, 244-280
4. Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327
5. Bollerslev T (1987) A conditional heteroskedastic time series model for speculative prices and rates of return. Review of Economics and Statistics 69, 542-547
6. Bollerslev T, Chou RY, Kroner KF (1992) Arch modeling in finance. Journal of Econometrics 52, 5-59
7. Box GEP, Tiao GC (1962) A further look at robustness via Bayes theorem. Biometrika 49, 419-432

8. Engle RF (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4), 987-1006
9. Engle RF, Bollerslev T (1986) Modelling the persistence of conditional variances. Econometric Reviews 5, 1-50
10. Fama EF (1965) The behavior of stock prices. Journal of Business 38, 34-105
11. Granger C, Ding Z (1995) Some properties of absolute return, an alternative measure of risk. Annales d'Economie et de Statistique 40, 67-91
12. Hansen PR, Lunde A (2005) A forecast comparison of volatility models: does anything beat a GARCH(1,1)? Journal of Applied Econometrics 20, 873-889
13. Hsieh, DA (1989) The statistical properties of daily foreign exchange rates: 1974-1983. Journal of International Economics 24, 129-145
14. Hurvich CM, Tsai CL 1989 Regression and time series model selection in small samples. Biometrika 76, 297-307
15. Jarque CM, Bera AK (1987) A test for normality of observations and regression residuals. International Statistical Review 55 (2), 163-172
16. Ljung GM, Box GEP (1978) On measure of lack of fit in time series models. Biometrika 65, 297-303
17. Liu SM, Brorsen BW (1995) Maximum likelihood estimation of a GARCH-stable model. Journal of Applied Econometrics 10, 273-285
18. Mandelbrot B (1963) The variation of certain speculative prices. Journal of Business 36, 394-419
19. Mittnik S, Rachev S T (1993) Modeling asset returns with alternative stable models. Econometric Reviews 12, 261-330
20. Mittnik S, Paolella MS, Rachev ST (1998) Unconditional and conditional distributional models for the Nikkei Index. Asia-Pacific Financial Markets 5, 99-1281

21. Mittnik S, Rachev ST, Doganoglu T, Chenyao D (1999) Maximum likelihood estimation of stable Paretian models. Mathematical and Computer Modeling 29, 275-293
22. Mittnik S, Paolella MS. (2003) Prediction of financial downside risk with heavy-tailed conditional distributions. In: Handbook of Heavy Tailed Distributions in Finance, Volume 1, Chapter 9, Elsevier North-Holland.
23. Nelson DB (1991) Conditional heteroskedasticity in asset returns: a new approach, Econometrica 59(2), 347-370
24. Panorska A, Mittnik S, Rachev ST (1995) Stable ARCH models for financial time series. Applied Mathematical Letters 8, 33-37
25. Poon S, Granger C (2003) Forecasting volatility in financial markets: a review. Journal of Economic Literature 41, 478-539
26. Rachev S, Mittnik S (2000) Stable Paretian models in finance. John Wiley & Sons, New York
27. Samorodnitsky G, Taqqu MS (1994) Stable non-Gaussian random processes. Stochastic models with infinite variance. Chapman and Hall, London
28. Schwarz G (1978) Estimating the dimension of a model. Annals of Statistics 6, 461-64
29. Sin CY, White H (1996) Information criteria for selecting possibly misspecified parametric models. Journal of Econometrics 71, 207-225