Fat Tails and Asymmetry in Financial Volatility Models

Peter Verhoeven[a] and Michael McAleer[b]

[a]School of Economics and Finance, Curtin University of Technology, Bentley, WA 6102, Australia

[b]Department of Economics, University of Western Australia, Crawley, WA 6009, Australia

Abstract

Although the GARCH model has been quite successful in capturing important empirical aspects of financial data, particularly for the symmetric effects of volatility, it has had far less success in capturing the effects of extreme observations, outliers and skewness in returns. This paper examines the GARCH model under various non-normal error distributions in order to evaluate skewness and leptokurtosis. The empirical results show that GARCH models estimated using asymmetric leptokurtic distributions are superior to their counterparts estimated under normality, in terms of: (i) capturing skewness and leptokurtosis; (ii) the maximized log-likelihood values; and (iii) isolating the ARCH and GARCH parameter estimates from the adverse effects of outliers. Overall, the flexible asymmetric Student-t distribution performs best in terms of capturing the non-normal aspects of the data.

Keywords: Asymmetric volatility, Conditional non-normality, Skewness, Leptokurtosis, Outliers, Location parameter.

1. Introduction

In finance, knowledge of the stochastic process underlying asset returns is essential for making correct investment decisions as it provides essential information about the riskiness of assets. The key underlying assumption of most financial models, such as the Black-Scholes option pricing model (BSM) and the capital asset pricing model (CAPM), is that the (logarithmic) returns are independently and identically distributed (*i.i.d.*) normal variates.

Unfortunately, the i.i.d. assumption about empirical returns underlying these financial models is typically not satisfied as financial returns are not, in general, normally distributed. Even if the underlying returns were normally distributed, the returns of portfolios that use dynamic strategies or include options on these assets will not be [13]. Furthermore, investors view upside and downside risks differently, with a preference for positively skewed returns, implying that more than the first two moments of returns may be priced in equilibrium. In short, the basic assumption of normality is highly suspect.

A pervasive temporal feature of asset price movements is the conditional dependency in the second moment. Such conditional dependency has been widely established through generalised autoregressive conditional heteroskedasticity (GARCH(p,q)) models [1]. Although these models generate reasonable amounts of excess kurtosis for high frequency (e.g. daily or weekly) data, the empirical distribution of returns conditioned on the current level of volatility is not normally distributed, as is frequently assumed. Since the suggestion of the Student-t distribution in [1], there has been an endless search beyond the normality assumption to model the empirical distribution of conditional returns.

The main purpose of this paper is to examine alternative probability density functions (pdf) for conditional returns, with particular emphasis on how well these capture fat tails and asymmetry. Such an empirical examination is important for finance practitioners as the shape of the conditional distribution also affects the means and variances, and hence the upside and downside probabilities of returns.

This paper is organised into six sections. Section 2 presents the specification of the mean-variance model. Section 3 discusses some important issues with maximum likelihood estimation, while Section 4 describes the alternative probability density functions used in the paper. The data and empirical results are discussed in Section 5, with special attention on leptokurtosis, asymmetry and the location parameter. Section 6 concludes the paper.

2. Specification of the mean-variance model

Consider the following specification of a GARCH(1,1) model, where the conditional mean (or log-returns) is given by an AR(1) process:

$$y_t = \mu + \varphi y_{t-1} + \varepsilon_t , \quad |\varphi| < 1 \tag{1}$$

where $\varepsilon_t \sim f\left(0, \sqrt{h_t}\right)$, and the conditional variance of the residuals is given by a general GARCH(1,1) process:

$$h_t = \omega + G(\varepsilon_{t-1})\varepsilon_{t-1}^2 + \beta h_{t-1} . \tag{2}$$

The conditional variance of ε_t can be used to obtain the normalized (or standardized) error, $\eta_t = \frac{\varepsilon_t}{\sqrt{h_t}}$, which is assumed to be i.i.d. $f(0,1,\theta)$. Sufficient conditions for positivity of the conditional variance are $\omega > 0$, $G(\varepsilon_{t-1}) \geq 0$, and $\beta \geq 0$, where $G(\varepsilon_{t-1})$ is a response function which models the effect of lagged shocks, ε_{t-1}, on the conditional volatility. When $G(\varepsilon_{t-1})$ is a constant (α), equation (2) reduces to the standard GARCH(1,1) model [1].

Although the simple GARCH(1,1) model is usually a good starting point when modelling financial returns, there is substantial evidence that suggests that time-varying asymmetry is a major component of volatility dynamics [8]. Hence, to avoid misspecification of the conditional variance equation, we include a leverage term, namely GJR [4]. Now, $G(\varepsilon_{t-1}) = \alpha + \gamma I(\varepsilon_{t-1} < 0)$, where I(.) is an indicator function which equals 1 when $\varepsilon_{t-1} < 0$, and zero otherwise. In this model, good news (or positive shocks, $\varepsilon_t > 0$) have an impact of $\alpha\varepsilon_t^2 \geq 0$ on volatility, while bad news (or negative shocks, $\varepsilon_t < 0$) have an impact of $(\alpha + \gamma)\varepsilon_t^2 \geq 0$. The leverage term usually arises when the unconditional returns are skewed, resulting in a positive (negative) γ estimate when the returns are negatively (positively) skewed, on average.[1]

For the GJR-GARCH(1,1) model [4], assuming that the standardized residuals are symmetrically distributed, the second moment regularity condition is $\frac{1}{2}\gamma+\alpha+\beta<1$ [15]. It follows that the unconditional variance implied by the GJR-GARCH(1,1) model is $E[h_t]=E[(\varepsilon_t-E[\varepsilon_t])^2]=\dfrac{\omega}{1-\left(\dfrac{\gamma}{2}+\alpha+\beta\right)}$. The necessary and sufficient condition for the existence of the unconditional fourth moment of the distribution of ε_t is $\beta^2+2\alpha\beta+\beta\gamma+k\alpha\gamma+k\alpha^2+\frac{k}{2}\gamma^2<1$, where k is the kurtosis of η_t. When the fourth moment condition is satisfied, the kurtosis for the unconditional distribution of ε_t becomes $K=k\dfrac{1-(\beta^2+2\alpha\beta+\beta\gamma+\alpha\gamma+\alpha^2+\frac{1}{4}\gamma^2)}{1-(\beta^2+2\alpha\beta+\beta\gamma+k\alpha\gamma+k\alpha^2+\frac{k}{2}\gamma^2)}$. Hence, using a conditional distribution with low kurtosis, high kurtosis of the unconditional distribution can be captured from variations in volatility over time. Furthermore, additional kurtosis may be induced when the conditional distribution is assumed to be leptokurtic ($k > 3$).

3. Maximum likelihood estimation (MLE)

The most common technique of estimating mean-variance models is by maximising the log-likelihood function, in which the returns are generated by a specific mean-variance model with an assumed pdf, f_{yt}. There are three properties of an estimator that are important, consistency, efficiency, and asymptotic normality. Although the property of efficiency is not always attained in large samples, the property of consistency is always required.

Consistency of the (quasi-) maximum likelihood estimator, or (Q)MLE, of the parameters of the mean and variance equations requires the expected (Q)MLE to have a unique maximum

[1] In this model, skewness in returns can be partly captured by the correlation between the covariance of returns with the level of volatility.

at the true value of the parameters. Proving consistency and asymptotic normality of the (Q)MLE of dependent processes, such as for GARCH-type models, is not always straightforward. Sufficient conditions for the consistency and asymptotic normality of the QMLE of the GARCH(1,1) model have been established under the following set of assumptions [2,11,15,16,22]: (i) normality; (ii) correct specification of the mean and variance equation;[2] (iii) strict stationarity of η_t; and (iv) various additional moment conditions on ε_t.

Efficiency of the (Q)MLE rests on the assumption of the true underlying pdf. For estimation purposes, if the normal pdf is not the true conditional density, the resulting maximum likelihood estimates are quasi-maximum likelihood estimates (QMLE). Although the asymptotic standard errors can be estimated consistently using QMLE, subject to the regularity conditions, the penalty is that they will not attain the Cramer-Rao bound. The loss of efficiency is directly related to the divergence of the true conditional distribution from the assumed normality, and is much greater for the QMLE of the variance than the mean equation [14]. Furthermore, fat tails without skewness are less serious [5].

Consistency is preserved for the normal distribution, under the above regularity conditions, but may be problematic when non-normality is used, and may yield inconsistent estimates of some moments when the assumed pdf is false [6].

Newey and Steigerwald [18] show that the identification condition can still hold for a non-Gaussian QMLE of the relative scale parameter if either (i) the conditional mean is identically zero; or (ii) both the assumed (theoretical) and true (empirical) pdf's are unimodal and conditionally symmetric about zero. When the symmetry condition is not satisfied, the correct specification of the conditional mean and variance is no longer sufficient to ensure consistency of the QMLE as the mean and variance are not necessarily the natural location and scale parameters, respectively. They show that an additional location parameter (ξ) is necessary to identify (and anchor) the location of the distribution in order to satisfy the identification condition for consistency. The location parameter accounts for asymmetry of η_t, that is, the discrepancy between the conditional mean and mode, and can be introduced in either the: (i) conditional mean equation; or (ii) distribution function. Thus, the conditional

[2] Hence, consistency may fail due to omitted dependency.

mean expressed in terms of the conditional location equation becomes[3] $y_t = \mu + \varphi y_{t-1} + \sqrt{h_t}(\eta_t - \xi)$, so that[4] $\eta_t = \frac{\varepsilon_t}{\sqrt{h_t}} + \xi$ and mode[ε_t] = 0. If there is heavy mass in the negative tail, making the distribution negatively skewed, the mean will be pulled in the direction of the skewness (to the left) and will no longer coincide with the natural location (mode) of the distribution. As a result, the location parameter is required to adjust for the discrepancy between the mean and the mode, thereby accounting for the skewness.

4. Alternative probability density functions

There are three parameters that define a pdf, namely (i) location; (ii) scale; and (iii) shape. The location parameter (mean, median, or mode) specifies the abscissa (x-axis) locations of the range of values. As the *location* parameter is the midpoint for symmetric distributions, as it shifts, the pdf shifts without a change in shape. The *scale* parameter (variance) measures the spread or variability of a pdf, and sets the scale (unit) of measurement of the values in the range of the pdf. As the scale changes, the pdf compresses (expands), but retains its shape. The *shape* parameter (skewness and kurtosis[5]) determines how the variation is distributed about the location, and determines the form of a distribution within the general family of distributions. As the shape parameter changes, the properties of the pdf change.

In general, the desirable properties of a pdf are that: (i) it must be sufficiently flexible so as to generate a range of shapes; (ii) the shape parameters must explain the skewness and kurtosis that may be encountered in finance; and (iii) it must be estimable. A trade-off for using pdf's with more flexible shapes is that it increases the impact of sampling errors on the parameter estimates.

[3] Newey and Steigerwald [18] point out that to retain the mean as the conditional mean of y_t, it must be assumed that $E[\lambda\eta_t - \xi] = \lambda E[\eta_t] - \xi = 0$, that is, $E[\eta_t] = \xi/\lambda$.
[4] Hence, it must now be assumed that $E[\eta_t] = E[\varepsilon_t/\sqrt{h_t} + \xi] = \xi$ and $E[\varepsilon_t \xi] = 0$.
[5] It is important to note that kurtosis is both a measure of peakedness and fat tails of the distribution.

In this paper, we estimate the mean-variance model using standardized pdf's[6] with a variety of shapes, all of which have been used previously in the finance literature. These are the: (i) asymmetric Student-t distribution (asStudent-t) [7]; (ii) asymmetric generalised error distribution (asGED) [21]; (iii) the asymmetric generalised t-distribution (asGTD) [20]; (iv) Gram-Charlier (Type A) distribution [9,12]; and (v) Pearson Type IV distribution [17,19];

Both the symmetric Student-t distribution and GED have one parameter (ν) to capture leptokurtosis. The implied kurtosis of the Student-t distribution is $k = \frac{6}{\nu - 4} + 3$ for all $\nu > 4$, while the implied kurtosis of the GED is $k = \frac{\Gamma(\frac{1}{\nu})\Gamma(\frac{5}{\nu})}{\Gamma(\frac{3}{\nu})^2}$. It follows that the Student-t distribution is leptokurtic when $4 < \nu \leq 25$, whereas the GED is leptokurtic when $1 < \nu < 2$. The greatest amount of kurtosis that can be generated by the GED is 6 (the Laplace distribution), which is twice the implied kurtosis of the normal distribution, and (two-thirds) less than can be captured by the Student-t distribution. Compared with the Student-t distribution, for equivalent kurtosis (k = 6), the GED is substantially more peaked (0.71 versus 0.47), with a higher pdf between 0-0.3σ and 2-6σ, but with substantially thinners tails (outside $\pm 6\sigma$). Hence, although the GED distribution may be better able to capture peaks, it is far worse for capturing fat tails.

A distribution that provides more flexibility than the Student-t distribution or GED is the GTD. This distribution has two parameters to control leptokurtosis, providing flexibility in the tails (ν) as well as in the peakedness (r). The GTD nests seven other well-known distributions, including the Student-t distribution (when r = 2) and GED ($\nu = \infty$). The kurtosis implied by the GED is $k = \frac{\mathrm{B}(\frac{5}{r}, \frac{\nu - 4}{r})\mathrm{B}(\frac{1}{r}, \frac{\nu}{r})}{\mathrm{B}(\frac{3}{r}, \frac{\nu - 2}{r})^2}$. Compared with the Student-t distribution (when ν = 6), reducing r from 2 to 1 has the effect of increasing the peakedness of the distribution (the height increases from 0.47 to 0.95), as well as fattening the tails (outside

[6] The mathematical expressions for these standardised probability density functions are given in the endnotes of Table I (see also [10]).

$\pm 2.8\sigma$). It does this by decreasing the density within 0.5-2.8σ. For both the GED and GTD, asymmetry is introduced by adding indicator functions.

The Gram-Charlier (Type A) distribution is an approximate pdf generated by a simple polynomial expansion of the normal density function, where the skewness (s) and kurtosis (k) appear directly as parameters. At an implied kurtosis of 6, the Gram-Charlier distribution is substantially more peaked than the Student-t distribution, but less peaked than GED. It puts a higher density around 2-4σ, rather than in the tails, which is approximately normal.

Of the asymmetric distributions discussed above, the Pearson Type IV distribution has the most flexible shape, covering a very large area in the β_1-β_2-plane. For this distribution, δ and r are measures of skewness and kurtosis around the mean, respectively. The necessary condition for the Type IV distribution is $0 < q < 1$, where $q = \frac{s^{2(k+3)^2}}{4(2k-3s^2-6)(4k-3s^2)}$ [10].[7] Type IV distribution nests the Student-t distribution with δ=0 and $3 < k < 9$. Furthermore, the possible area of the beta points for the asStudent-t distribution lies entirely within the Type IV area. The Type IV distribution can be more leptokurtic than the Student-t distribution when $3 < r < 4$. For the Type IV distribution, the implied skewness is determined by δ when r is small, and by δ and r when the distribution is closer to the boundary of the Type IV region. When r is high, the flexibility of skewness about the mean is low and the kurtosis becomes independent of skewness. The minimum value of the square of skewness is zero when $\delta = 0$ (Type VII (symmetric) distribution), and its maximum value is 32 when $r \to 3$ and $\delta^2 \to \infty$. When the magnitude of δ reaches a certain value (the transition line of Type V), r must decrease in order to allow for any further increase in absolute skewness. Thus, the peakedness and fatness of the tails are adjusted to accommodate any further increase in skewness. When $\delta = 0$, $k = 3 + \frac{6}{r-3}$. Hence, the Type IV distribution converges to the normal distribution when $r \to \infty$. The kurtosis increases exponentially when $r \to 3$, resulting in higher density very far away (> 10σ) in the tails. As a result, the peakedness does not change substantially.

[7] This condition must be met for the imaginary roots of the quadratic function to exist.

5. Data and empirical results

5.1 Data

The data consist of the daily close-to-close logarithmic returns of the National Association of Securities Dealers Automated Quotation (NASDAQ) Composite Index (IXIC), the Australian All Ordinaries Index (AOI), and Kuala Lumpur SE Composite Index (KLSE), from 1990-2000. A large sample size (2500) issued to reduce the effects of sampling errors on the estimates. The algorithm used for maximum likelihood estimation is Newton-Raphson, and all the optimisation routines are coded in GAUSS.

5.2 Empirical results

The parameter estimates and some diagnostics for the AR(1)-GJR-GARCH(1,1) model under various pdf's are reported in Table 1 for IXIC.[8]

5.2.1 Leptokurtosis

For all series, the specification of the mean-variance model under conditional normality captures much, but not all, of the excess kurtosis of the unconditional returns. Compared with the kurtosis of the unconditional returns, which varies from 8.11 for AOI to 31.13 for KLSE, the kurtosis of the conditional returns is substantially lower, varying over a relatively narrow range from 4.47 to 5.98.

Additional leptokurtosis can be captured when the conditional returns are assumed to be leptokurtic. For the symmetric distributions, the greatest amount of kurtosis is implied by the Student-t distribution (up to 8.00), followed by GTD (up to 4.39), and GED (up to 4.36). The fit of the model, measured in terms of the MLL values, can be substantially improved using leptokurtic distributions. Overall, the symmetric GTD provides the best fit, which can be attributed to its greater flexibility, as it has two parameters to control the shape of the distribution, one for peakedness and the other for fat tails, both of which are measures of

[8] Results for the two other returns series are available from the authors upon request.

kurtosis. This also implies that there may be some trade-off in capturing fat tails and peakedness for the Student-t distribution.

Allowing greater flexibility in the shape of the distribution by introducing asymmetry results in a greater degree of implied kurtosis. The kurtosis of the asymmetric leptokurtic distributions implied from the estimated parameters ranges from 3.48 for the Gram-Charlier distribution, to 8.22 for the Type IV distribution, and 8.81 for the asStudent-t distribution. Overall, Type IV and asStudent-t distributions best capture leptokurtosis, leaving the least amount of unexplained residual kurtosis.

As expected, capturing fat tails through a leptokurtic distribution substantially decreases the demands on the volatility model, as is evidenced by the significantly decreased QMLE of α and γ, and increased volatility persistence, as measured by β.

5.2.2 Asymmetry

Based on the Lagrange multiplier test statistic for asymmetry (LM(A)) [3], the presence of time-varying non-linearity beyond ARCH effects (at the 5% significance level) is rejected for all indices, except IXIC. For this series, there is evidence of a sign bias, as well as of a positive and a negative size bias in the conditional returns. The sign of the t-ratio indicates that negative shocks cause a larger increase in volatility than positive shocks. Modelling sign asymmetries by including a leverage term in the volatility model reduces the LM(A) statistics only slightly. Furthermore, despite the fact that the LM(A) statistic is insignificant for both AOI and KLSE, the γ estimate is statistically significant for these series.[9] This implies that the leverage term does not fully accommodate the impacts of positive and negative shocks on volatility, as captured by the LM(A) test [3].[10]

[9] The LM(A) statistic [3], which tests for unconditional skewness (and kurtosis), rejects the presence of non-linearities beyond ARCH effects in the conditional returns. The finding of a significant QMLE of γ in all series is a result of conditional skewness. This is further demonstrated by the fact that for IXIC, when the unconditional skewness of the conditional returns is captured by the Type IV distribution, the LM(A) statistic becomes insignificant.

[10] This is because the four moments that are specified in the GJR-GARCH(1,1) model are inflexible, that is, although the model captures $S^{2/3}$ and $K^{1/2}$, it does not capture all aspects of asymmetry and excess kurtosis.

Allowing time-varying asymmetries to be captured by the leverage term in the volatility model does little to reduce the empirical skewness of the conditional returns. Only for AOI is the skewness of the conditional returns smaller when the leverage term is included (-0.30 versus -0.36). This may be related to that fact that, for this series, the α estimate is insignificant. Consistent with these findings, when we allow for non-time-varying asymmetry in the conditional distribution, the leverage term is little affected. Adding the leverage term to the GARCH(1,1)-Type IV model does little to improve the maximized likelihood function [19], concluding that accommodating skewness through the Type IV distribution means there is "nothing" left to capture through the conditional variance. When asymmetries are captured through modelling the time-varying third moment, the statistical significance of the leverage term is substantially reduced [17]. This implies that the leverage term in the variance equation only partially captures the time-varying third moment.

Inference on skewness and kurtosis are highly dependent in all asymmetric distributions, in particular, the Type IV and asStudent-t distributions: the larger (smaller) the skewness, the larger (smaller) the kurtosis.[11] This is consistent with the empirical results for the asStudent-t distribution, in that the implied kurtosis of the distribution substantially increases when skewness is introduced. For KLSE, this results in overshooting of kurtosis, as well as skewness.

For both asGED and asGTD, where asymmetry is introduced by adding indicator functions, there appears to be some trade-off between capturing excess kurtosis and excess skewness. For example, when asymmetry is introduced into the GTD for IXIC, the residual skewness decreases from -0.64 to -0.32, whereas the residual kurtosis increases from 1.37 to 1.95. For these distributions, the fit of the model improves substantially when asymmetry is accommodated, implying that skewness may be more important than kurtosis.

An important feature that is consistent across the returns series is that the flexible Type IV distribution best describes the data, both in terms of having the lowest sum of the absolute residual skewness and residual kurtosis, and in terms of having the highest MLL values. The normal distribution performs the worst. Furthermore, the QMLE estimates of the mean-

[11] The reverse is not necessarily true.

variance model converge when the Type IV or asStudent-t distribution is used, implying that asset returns may be distributed as asStudent-t.

5.2.3 Location parameter

If the symmetry condition is not satisfied, that is, when the assumed or true conditional returns are skewed, correct specification of the mean and variance equation may not suffice to ensure consistency because the mean is not the natural location parameter of the error distribution [18]. Consequently, a location parameter has been included to provide a diagnostic test of whether the symmetry condition is satisfied [7].

For the three returns series (IXIC, KLSE, and AOI), the location parameter is negative and (marginally) significant when the volatility model is estimated under conditional symmetry and when the leverage term is not included. As the unconditional returns for these series are negatively skewed, the mean of the distribution will be shifted to the left. It follows that a positive correction is necessary for consistency as the mean is now no longer the natural location of the distribution. Including the leverage term in the variance equation changes the sign of the estimated location parameter: for the negatively skewed returns series (IXIC and AOI), the location parameter becomes positive, while the location parameter becomes negative for the positively skewed returns series (KLSE). Furthermore, for the negatively (positively) skewed series, the sign asymmetry (Asym) increases (decreases). These findings imply that the location parameter may account for some of the time-varying skewness.

The location parameter is generally found to be significant when the variance equations are estimated under conditional asymmetry. For the Type IV and asStudent-t distributions, the location parameter is positive when the conditional returns are negatively skewed (IXIC), and negative when the conditional returns are positively skewed (KLSE). The significance of the location parameter is consistent with expectations as both the location (ξ) and skewness (δ) parameters shift the location of the distribution relative to the origin, though in opposite directions. When the distribution is negatively skewed, δ is negative, shifting the location of the distribution to the left, relative to the origin. As a result, the location parameter will be positive, shifting the distribution back to the origin. The total effect will be that the shape of the distribution is changed. If a location parameter is not included, skewness around the natural location parameter (median or mode) is estimated, which is generally insignificant.[12]

6. Conclusion

The empirical results of this paper show there are benefits to estimating conditional mean-variance models using conditional non-time-varying asymmetric leptokurtic distributions instead of a normal distribution. In particular, asymmetric distributions are capable of capturing outlying observations, which cannot be adequately captured by a time-varying conditional variance. As expected, the benefits of estimating GARCH models using asymmetric leptokurtic distributions are more substantial for highly volatile series, which have a higher degree of non-normality. The results also show that constant skewness of the unconditional returns does not help in capturing time-varying asymmetries. Hence, a more rewarding direction for future research may be to accommodate time variations in the third moment.

Acknowledgments

The first author wishes to acknowledge the financial support of the Curtin Business School Small Grant Scheme. The second author is most grateful for the financial support of the Australian Research Council.

[12] Without including a location parameter, the skewness of the conditional returns implied by the parameter estimates is generally very small.

References

[1] T. Bollerslev, T., Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics, 31 (1986), 307-327.

[2] T. Bollerslev, J.M. Wooldridge, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, Econometric Reviews, 11 (1992), 143-172.

[3] R.F. Engle, V.K. Ng, Measuring and testing the impact of news on volatility, Journal of Finance, 48 (1993), 1749-1778.

[4] L.R. Glosten, R. Jagannathan, D. Runkle, On the relation between the expected value and the volatility of the nominal excess return on stocks, Journal of Finance, 48 (1993), 1779-1801.

[5] G. Gonzalez-Rivera, F.C. Drost, Efficiency comparisons of maximum-likelihood-based estimators in GARCH models, Journal of Econometrics, 93 (1999), 93-111.

[6] G. Gonzalez-Rivera, J. Racine, On the recovery of efficiency losses: Maximum and quasi-maximum likelihood estimators, Working Paper, Department of Economics, University of California, 1995.

[7] C.R. Harvey, A. Siddique, Autoregressive conditional skewness, Journal of Financial and Quantitative Analysis, 34 (1999), 465-487.

[8] D.A. Hsieh, Chaos and nonlinear dynamics: Application to financial markets, Journal of Finance, 46 (1991), 1839-1877.

[9] R. Jarrow, A. Rudd, Approximate option valuation for arbitrary stochastic processes, Journal of Financial Economics, 10 (1982), 347-369.

[10] M.G. Kendall, A. Stuart 1969, The advanced theory of statistics, Vol. 1, 3rd edition, Charles Griffin & Co. Ltd, London

[11] S.W. Lee, B.E. Hansen, Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator, Econometric Theory, 10 (1994), 29-52.

[12] T.K.Y. Lee, Y.K. Tse, Term structure of interest rates in the Singapore Asian dollar market, Journal of Applied Econometrics, 6 (1991), 143-152.

[13] H.E. Leland, Beyond mean-variance: Performance measurement in a nonsymmetrical world, Financial Analysts Journal, 55 (1999), 27-36.

[14] D.X. Li, H.J. Turtle, Semiparametric ARCH models: An estimating function approach, Journal of Business & Economic Statistics, 18 (2000), 174-186.

[15] S. Ling and M. McAleer, Stationarity and the existence of moments of a family of GARCH processes, Journal of Econometrics, 106 (2002b) 109-117.

[16] R.L Lumsdaine, Consistency and asymptotic normality of the quasi-Maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, Econometrica, 64 (1996), 575-596.

[17] Y. Nagahara, The PDF and CF of Pearson type IV distributions and the ML estimation of the parameters, Statistics & Probability Letters, 43 (1999), 251-264.

[18] W.K. Newey, D.G. Steigerwald, Asymptotic bias for quasi-maximum-likelihood estimators in conditional heteroskedasticity models, Econometrica, 65 (1996), 587-599.

[19] G. Premaratne, A.K. Bera, Modeling asymmetry and excess kurtosis in stock return data, Working Paper 00-123, Department of Economics, University of Illinois, Champaign, 2001.

[20] P. Theodossiou, Financial data and the skewed generalized-t distribution, Management Science, 44 (1998), 1650-1661.

[21] P. Theodossiou, Skewed generalised error distribution of financial assets and option pricing, Working Paper, School of Business, Rutgers University, New Jersey, 2001.

[22] A.A. Weiss, Asymptotic theory for ARCH models: estimation and testing, Econometric Theory, 2 (1986), 107-131.

Table 1. QMLE of the AR(1)-GJR-GARCH(1,1) model for IXIC

	Normal	Student-t	asStudent-t	GTD	asGTD	GED	asGED	Type IV	Gram-Charlier
μ	1.040e-4	1.540e-05	-4.440e-05	-1.160e-05	3..310e-4	1.060e-4	2.560e-4	-4.540e-05	-1.940e-4
φ	1.678e-1*	1.621e-1*	1.464e-1*	1.616e-1*	1..538e-1*	1.628e-1*	1.525e-1*	1.443e-1*	1.407e-1*
ω	3.517e-6*	2.195e-6*	2.062e-6*	2.108e-6*	1.651e-6*	2.722e-6*	2.523e-6*	2.033e-6*	2.788e-6*
α	6.346e-2*	4.677e-1*	5.158e-2*	4.575e-2*	5.508e-2*	5.387e-2*	5.878e-2*	5.213e-2*	6.137e-2*
γ	1.216e-1*	1.132e-1*	1.054e-1*	1.127e-1*	1.042e-1*	1.152e-1*	1.055e-1*	1.026e-1*	1.057e-1*
β	8.556e-1*	8.867e-1*	8.880e-1*	8.891e-1*	8.812e-1*	8.739e-1*	8.735e-1*	8.885e-1*	8.743e-1*
1/ν		1.270e-1*	1.226e-1*	1.590e-1*	1.584e-1*				
(ν-2)						-5.369e-1*	-4.67e-1*		
r				2.242*	2.297*			8.177*	
δ			-1.039*		-0.142*		-1.35e-1*	-2.917*	
ξ	5.210e-2	9.990e-2	1.032*	9.944e-2*	2.564e-1*	9.310e-2	2.377e-1*	9.710e-1*	7.77e-2
s									-0.395*
k									0.666*
Asym	0.343	0.292	0.329	0.289	0.346	0.319	0.358	0.337	0.367
MLL	3.144	3.162	3.171	3.163	3.169	3.158	3.166	3.172	3.162
***Diagnostics*:**									
Mean	0.052	0.062	0.07	0.065	0.039	0.052	0.026	1.042	0.085
Residual Mean	0.000	-0.038	0.187	-0.034	0.004	-0.034	0.232	0.971	0.010
Standard deviation	1.000	1.003	1.006	1.002	1.000	1.002	1.005	1.002	0.992
Skewness	-0.614	-0.636	-0.644	-0.638	-0.643	-0.626	-0.617	-0.645	-0.633
Residual Skewness	-0.614	-0.636	-0.093	-0.638	-0.322	-0.626	-0.061	-0.062	-0.239
Kurtosis	4.977	5.097	5.088	5.107	5.105	5.039	5.001	5.085	4.987
Residual Kurtosis	1.977	0.548	0.14	1.37	1.949	1.174	1.453	0.318	1.321

Table 1 (continued)

LM(N)	563*	626*	626*	631*	633*	596*	575*	625*	578*
Q(12)	7.477	8.127	7.901	6.837	7.613	7.545	7.251	8.112	12.993
Q(12)^2	5.861	5.812	6.196	5.834	6.031	5.830	5.968	7.158	6.595
Sign	3.972*	4.148*	4.395*	4.248*	4.147*	4.019*	3.769*	1.835*	4.542*
Size(+)	-2.854*	-2.867*	-3.082*	-2.882*	-2.964*	-2.854*	-2.776*	-1.411	-3.23*
Size(-)	-2.049*	-2.501*	-2.564*	-2.543*	-2.509*	-2.314*	-1.892	-1.978*	-2.397*
LM(A)	16.19*	17.17*	19.37*	18.21*	17.18*	16.25*	14.71*	4.65	21.05*

* denotes significance at the 5% level based on heteroskedasticity-consistent standard errors. Asym is measured as $(\alpha+\gamma)/\alpha$. LM(N) is the LM statistic for normality of η_t ($LM(N)=N(S^2/6 + (K-3)^2/24)$, which is asymptotically χ^2 distributed with two degrees of freedom under the null hypothesis of normality). Q(12) is the Ljung-Box test statistic for serial correlation in η_t with 12 lags. $Q(12)^2$ is the Ljung-Box test statistic for an ARCH process based on η_t^2. Under the null hypotheses of uncorrelated and conditionally homoskedastic errors, respectively, the test statistics are asymptotically χ^2 distributed with 12 degrees of freedom. The t-ratios for the coefficients b_1, b_2, and b_3 are the sign bias, the positive size bias, and the negative size bias test statistics, respectively, as suggested in [3], and are based on the following auxiliary regression model: $\eta_t^2 = a + b_1 S_t^- + b_2 S_t^+\eta_{t-1} + b_3 S_t^-\eta_{t-1} + \nu_t$. Under the null hypothesis that $b_1 = b_2 = b_3 = 0$, the joint test statistic for asymmetry (LM(A)) is asymptotically χ^2 distributed with 3 degrees of freedom. The density functions are as follows: for the normal distribution: $f(.) = \frac{1}{\sqrt{2\pi h_t}}\exp(-\frac{1}{2}\eta_t^2)$; asGED: $f(.) = \frac{\nu}{2\lambda\sqrt{h_t}\Gamma(\frac{1}{\nu})}\exp(-\{\frac{|\eta_t|}{\lambda(1+I(\eta_t)\delta}\}^{\nu})$; asStudent-t distribution: $f(.) = \exp(-\frac{1}{2}\delta^2)\frac{\Gamma(\frac{\nu+1}{2}}{\sqrt{\{\pi h_t(\nu-2)\}}\Gamma(\frac{\nu}{2})}(1+\frac{\eta_t^2}{\nu-2})^{-\frac{\nu+1}{2}}\sum_{j=0}^{\infty}\frac{\Gamma(\frac{\nu+j+1)}{2}}{j!\Gamma(\frac{\nu+1}{2})}(\frac{\eta_t\delta\sqrt{2}}{\lambda\sqrt{\nu+\frac{\eta_t^2}{\lambda^2}}})^j$; asymmetric GTD: $f(.) = c(1+\frac{r}{\nu-2}\{\frac{|\eta_t|}{\lambda(1+I(\eta_t)\delta}\}^r)^{-\frac{\nu+1}{r}}$; Gram-Charlier distribution: $f(.) = f(.)_N\left\{1+\frac{1}{6}s(\eta_t^3-3\eta_t)+\frac{k-3}{24}(\eta_t^4-6\eta_t^2+3)\right\}$; and Pearson Type IV distribution: $f(.)=\frac{1}{C\lambda\sqrt{h_t}}(1+\frac{\eta_t^2}{\lambda^2})^{-\frac{r=2}{2}}\exp(\delta\arctan(\frac{\eta_t}{\lambda})$.