A joint Initiative of Ludwig-Maximilians-Universität and Ifo Institute for Economic Research

CESifo Working Papers

AN ALTERNATIVE CONDITIONAL ASYMMETRY SPECIFICATION FOR STOCK RETURNS*

Kurt Brännäs
Niklas Nordman

CESifo Working Paper No. 448

April 2001

CESifo
Center for Economic Studies & Ifo Institute for Economic Research
Poschingerstr. 5, 81679 Munich, Germany
Tel.: +49 (89) 9224-1410
Fax: +49 (89) 9224-1409
e-mail: office@CESifo.de

* A previous version of the paper has gained from comments made during a seminar at the Economics Department, Umeå University, and from the comments of Jan De Gooijer and Xavier de Luna. The paper was completed while the first author enjoyed the kind hospitality of CESifo, Munich.

CESifo *Working Paper No. 448*
April 2001

AN ALTERNATIVE CONDITIONAL ASYMMETRY SPECIFICATION FOR STOCK RETURNS

Abstract

The paper advances the log-generalized gamma distribution as a suitable generator of conditional skewness. Based on the NYSE composite daily returns an asMA-asQGARCH model along with skewness dynamics is estimated. The results indicate a skewness that varies between sizeable negative skewness and almost symmetry. The conditional variance and skewness measures are negatively correlated.

JEL Classification: C22, C51, C52, C53, G14.

Keywords: Time series, finance, nonlinearity, skewness, gamma, estimation, NYSE.

Kurt Brännäs
Department of Economics
Umeå University
SE-90187 Umeå
Sweden
Kurt.Brannas@econ.umu.se

Niklas Nordman
Department of Economics
Umeå University
SE-90187 Umeå
Sweden

1. Introduction

Harvey and Siddique (1999) recently proposed a modelling approach for conditional skewness in financial time series and also presented some empirical results. While there is no dispute about the existence of skewness in returns, there appears to be a shortage of well-established economic explanations for this phenomenon. Chen, Hong and Stein (2000) summarize previous explanations and advance their suggestion that it is investor heterogeneity that plays a central role. The heterogeneity is due to differences in opinion about fundamental value and to some investors facing short-sales constraints. Harvey and Siddique's (1999) focus is empirical and they use the noncentral t-distribution for modelling. This distribution therefore potentially also accounts for fat-tailedness. Computationally their approach is, however, quite demanding. Recently, Premaratne and Bera (2000) advanced the Pearson type IV distribution, that closely resembles the noncentral t-distribution, as a computationally simpler alternative.

We propose and study the usefulness of an alternative distribution – the log generalized gamma – that allows for skewness and other features in the distribution through a single distribution characteristic q (e.g., Stacey, 1962, Prentice, 1974, Farewell and Prentice, 1977). The normal and extreme-value distributions are among the special cases. Numerically this specification is simpler than the parametric Harvey and Siddique (1999) model or seminonparametrically formulated models (e.g., Brunner, 1992). In addition, we focus directly on the basic density characteristic and then only indirectly on skewness. This is in contrast to Harvey and Siddique's approach which starts from a skewness model and has to solve a pair of nonlinear equations at each time point and at each iterative step in order to get to the basic parameters required for conditional maximum likelihood estimation. If our idea of a more direct modelling was applied to the noncentral t-distribution considerable gains in computational time relatively theirs would likely be obtained.

Empirically we study the NYSE composite daily returns, January 2, 1981 – December 31, 1999 ($T = 4956$) and partly re-use the study of Brännäs and De Gooijer (2000).

2. Model and Estimation

The present modelling exercise extends the one of Brännäs and De Gooijer (2000, hereafter BDG) by adding a more flexible density that allows for skewness in innovations. Their model has as conditional mean the asMA specification $E(y_t|Y_{t-1}) = e_t = \theta_0 + \sum_{i=1}^{q} \theta_i^+ u_{t-i}^+ + \sum_{i=1}^{q} \theta_i^- u_{t-i}^-$, where $u_t^+ = \max(0, u_t)$, $u_t^- = \min(u_t, 0)$, and $Y_t = (y_1, \ldots, y_t)$ is the information set (Wecker, 1981, Brännäs and De Gooijer, 1994). The conditional variance is an asQGARCH specification $V(y_t|Y_{t-1}) = h_t^2 = \alpha_0 + \sum_{i=1}^{Q}(\alpha_i^+ u_{t-i}^+ + \alpha_i^- u_{t-i}^-) + \sum_{i=1}^{p} \beta_i u_{t-i}^2 + \sum_{i=1}^{p} \gamma_i h_{t-i}^2$ (BDG). These conditional moments catch shocks u_t asymmetrically. One may view the conditional mean as containing not a full risk measure h_t^2 but an unrestricted reduced form of such a measure. The suggested approach does not depend on the precise nature of these moment specifications but can be varied according to context. Let $v_t = y_t - E(y_t|Y_{t-1})$ be the zero mean prediction error. In the model $u_t = v_t = \varepsilon_t h_t$, where $h_t > 0$ is the conditional standard deviation. The random variable ε_t has zero mean, unit variance, and is assumed to be conditionally independent of h_t.

We think of ε_t as a standardization of a log-generalized gamma (LGG) distributed variable (cf. the Mathematical Appendix for details). This family of densities has a characteristic $q \in (-\infty, \infty)$, which indicates the member. For some q values well-known distributions arise, e.g., $q = 0$ corresponds to a normal distribution, $q = 1$ to an the extreme-value distribution, and $q = -1$ to the reciprocal extreme-value distribution. The LGG density is negatively skewed for positive q values, while negative q corresponds to positive skewness. Symmetry only emerges for $q = 0$. The LGG distribution to be used for estimation is

$$f(u) = \begin{cases} \sigma h^{-1}|q|(q^{-2})^{q^{-2}} \exp[q^{-2}(w - \exp(w))]/\Gamma(q^{-2}), & q \neq 0 \\ (2\pi h^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}u^2/h^2), & q = 0 \end{cases}, \tag{1}$$

where $w = q\sigma u/h + q\mu$, $\mu = q^{-1}\psi(q^{-2}) - q^{-1}\ln q^{-2}$, and $\sigma^2 = q^{-2}\psi'(q^{-2})$ with $\psi(x) = \partial \ln \Gamma(x)/\partial x$ and $\psi'(x) = \partial\psi(x)/\partial x$. The skewness of ε is $s = \sigma^{-3}q^{-3}\psi''(q^{-2})$. Here, $\psi^{(k)}(x) = \partial^k \ln \Gamma(x)/\partial x^k$ are ψ-functions and $\Gamma(.)$ is the gamma function (Abramowitz and Stegun, 1965).

By estimating the unknown parameter q from data, evidence of an asymmetric ε_t distribution may be found. Note, however, that even for a symmetric ε_t distribution the marginal distribution of y_t is generally asymmetric due to the nonlinearities in the model.

The main idea of this paper is to add a conditional skewness specification indirectly in terms of q. Using the structures of the first two conditional moments we may write

$$q_t = \kappa_0 + \sum_{i=1}^{r_1}(\kappa_i^+ u_{t-i}^+ + \kappa_i^- u_{t-i}^-) + \sum_{i=1}^{r_2} \mu_i u_{t-i}^2 + \sum_{i=1}^{r_3} \nu_i q_{t-i}.$$

With estimates of q_t, estimates of the skewness $s_t = \sigma^{-3}q_t^{-3}\psi''(q_t^{-2})$ can easily be calculated. Since we have estimates of q_t we can also obtain conditional kurtosis and higher conditional moments or graph the conditional density for any t.

Given the density function in (1), conditional maximum likelihood estimation is quite straightforward. For the empirical results given in the next section, the

RATS 5.0 package was employed. The required ψ-functions were evaluated in terms of numerical differencing of $\ln\Gamma(x)$ rather than by approximating functions (cf. Abramowitz and Stegun, 1965). To estimate the parameters both the SIMPLEX and BFGS routines were employed.

3. Results

The estimation results are based on the NYSE composite daily returns (defined as $y_t = 100[\ln(I_t) - \ln(I_{t-1})]$, where I_t is the price index), January 2, 1981 – December 31, 1999 (4956 observations, source: Datastream). Table 1 summarizes the main estimation results. The first part reproduces estimation results for the asMA-asQGARCH model obtained under the normal distribution assumption for ε_t, cf. BDG, and is included for comparison. In terms of conditional mean estimates a notable change between the specifications is the increase in the constant term, θ_0, as q is introduced and allowed to vary. This is partly due to the constant threshold of zero in forming the $\{u_t^+\}$ and $\{u_t^-\}$ sequences. Allowing for skewness by the use of q there may be larger probabilities for negative u_t values. Depending on signs of parameter estimates this may then have an enhancing effect on the constant term. The estimates of the low u_t^+ lags appear to become smaller, while those of the u_{t-3}^+ lag becomes larger. The changes for the u_t^- lags are smaller. There are no sign changes in neither $\hat{e}_t$ nor in $\hat{h}_t^2$. For the latter all estimates except for the persistence parameter become smaller. Using the response time measure of BDG we find very small differences between the three models. On average all measures respond to news within the day.

In the asMA-asQGARCH model with a constant q, we get $\hat{q} = 0.223$ or the skewness measure $\hat{s} = -0.23$. A rather similar $\hat{q}$ estimate was obtained for a pure asMA model. In both cases a normality assumption on ε_t can be rejected. For the full model we note the very small effect of lagged q_{t-1}. This indicates that the distribution of ε_t can be expected to shift shape very quickly. In Figure 1 we plot the effects of u_{t-1} on the conditional variance (h_t^2) and on the skewness parameter (q_t) for the full model. Introducing a variable q function indicates less and a more symmetric impact on h_t^2 of the lagged u_{t-1} values. The effect on q_t of u_{t-1} is quite asymmetric and at its smallest for $u_{t-1} \approx 2.8$. Figure 2 plots the conditional skewness against the observations and reveals that skewness is smallest for small values on y_t, while for larger absolute values on y_t there will always be negative skewness.

There is roughly a linear negative relationship between the skewness $\hat{s}_t$ and $\hat{q}_t$ ($\hat{s}_t = 0.199 - 1.96\hat{q}_t$, $R^2 = 0.98$). This suggest two things. First, we should expect to get analogous lag structures whether we model as here or in terms of a model for s_t (from which q_t would be solved for estimation purposes, cf. the Harvey and Siddique

Table 1: Parameter estimates for the conditional mean ($\hat{e}_t$), the conditional variance ($\hat{h}_t^2$) and conditional skewness ($\hat{q}_t$) models. Standard errors in parentheses

Model	Cond. Moment	Estimate
asMA-	$\hat{e}_t$	$\underset{(0.009)}{0.004} + \underset{(0.014)}{0.109}\,\hat{u}^{+}_{t-1} + \underset{(0.012)}{0.056}\,\hat{u}^{+}_{t-2} - \underset{(0.014)}{0.045}\,\hat{u}^{+}_{t-3}$
		$+ \underset{(0.014)}{0.033}\,\hat{u}^{-}_{t-1} - \underset{(0.014)}{0.051}\,\hat{u}^{-}_{t-4}$
asQGARCH	$\hat{h}_t^2$	$\underset{(0.002)}{0.010} - \underset{(0.006)}{0.032}\,\hat{u}^{+}_{t-1} - \underset{(0.006)}{0.096}\,\hat{u}^{-}_{t-1} + \underset{(0.002)}{0.059}\,\hat{u}^{2}_{t-1} + \underset{(0.002)}{0.905}\,\hat{h}^{2}_{t-1}$
(ε_t normal)		$\ln L = -5802.9$, $\mathrm{LB}_{10}(\hat{u}_t/\hat{h}_t) = 5.90$, $\mathrm{LB}_{10}(\hat{u}_t^2/\hat{h}_t^2) = 6.23$,
		$\hat{\sigma}_u^2 = 0.827$, Skewness $= -0.68$, Kurtosis $= 6.93$
asMA-	$\hat{e}_t$	$\underset{(0.017)}{0.021} + \underset{(0.027)}{0.098}\,\hat{u}^{+}_{t-1} + \underset{(0.025)}{0.032}\,\hat{u}^{+}_{t-2} - \underset{(0.024)}{0.070}\,\hat{u}^{+}_{t-3}$
		$+ \underset{(0.025)}{0.037}\,\hat{u}^{-}_{t-1} - \underset{(0.027)}{0.047}\,\hat{u}^{-}_{t-4}$
asQGARCH(q)	$\hat{h}_t^2$	$\underset{(0.002)}{0.000} - \underset{(0.010)}{0.002}\,\hat{u}^{+}_{t-1} - \underset{(0.029)}{0.079}\,\hat{u}^{-}_{t-1} + \underset{(0.012)}{0.033}\,\hat{u}^{2}_{t-1} + \underset{(0.003)}{0.935}\,\hat{h}^{2}_{t-1}$
	$\hat{q}$	$\underset{(0.005)}{0.223}$
		$\ln L = -5741.3$, $\mathrm{LB}_{10}(\hat{u}_t/\hat{h}_t) = 7.00$, $\mathrm{LB}_{10}(\hat{u}_t^2/\hat{h}_t^2) = 12.45$,
		$\hat{\sigma}_u^2 = 0.825$, Skewness $= -0.85$, Kurtosis $= 9.01$
Full model	$\hat{e}_t$	$\underset{(0.002)}{0.025} + \underset{(0.011)}{0.099}\,\hat{u}^{+}_{t-1} + \underset{(0.010)}{0.027}\,\hat{u}^{+}_{t-2} - \underset{(0.013)}{0.066}\,\hat{u}^{+}_{t-3}$
		$+ \underset{(0.004)}{0.042}\,\hat{u}^{-}_{t-1} - \underset{(0.002)}{0.044}\,\hat{u}^{-}_{t-4}$
	$\hat{h}_t^2$	$\underset{(0.000)}{0.003} - \underset{(0.005)}{0.012}\,\hat{u}^{+}_{t-1} - \underset{(0.016)}{0.066}\,\hat{u}^{-}_{t-1} + \underset{(0.009)}{0.041}\,\hat{u}^{2}_{t-1} + \underset{(0.003)}{0.932}\,\hat{h}^{2}_{t-1}$
	$\hat{q}_t$	$\underset{(0.008)}{0.172} - \underset{(0.005)}{0.108}\,\hat{u}^{+}_{t-1} + \underset{(0.009)}{0.027}\,\hat{u}^{-}_{t-1} + \underset{(0.007)}{0.023}\,\hat{u}^{2}_{t-1} + \underset{(0.018)}{0.073}\,\hat{q}_{t-1}$
		$\ln L = -5726.5$, $\mathrm{LB}_{10}(\hat{u}_t/\hat{h}_t) = 7.05$, $\mathrm{LB}_{10}(\hat{u}_t^2/\hat{h}_t^2) = 10.69$,
		$\hat{\sigma}_u^2 = 0.825$, Skewness $= -0.82$, Kurtosis $= 8.54$

Notes: $\mathrm{LB}_{10}(u_t)$ is the Ljung-Box test against serial correlation in u_t. $\ln L$ is the log-likelihood function value.

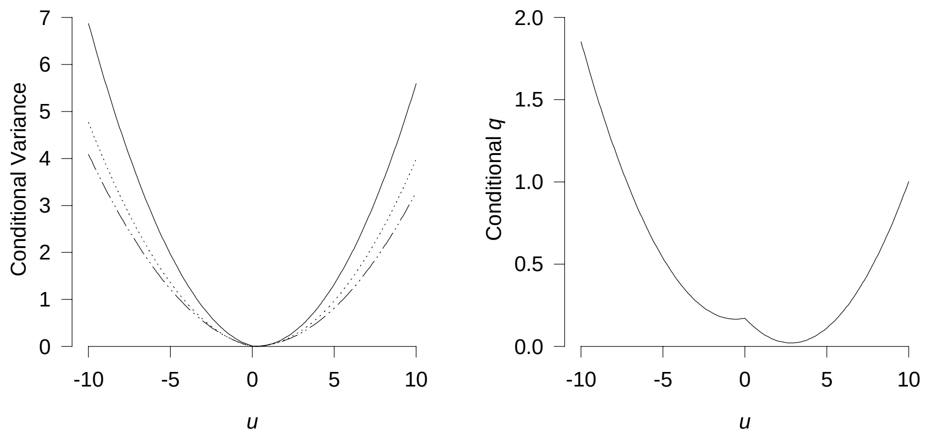

Figure 1: Effects of u_{t-1} on conditional variance (left, for $h_{t-1}^2 = 0$, solid line asMA-asQGARCH, dotted line asMA-asQGARCH(q), dot-dashed line Full model) and conditional skewness (right, for $q_{t-1} = 0$).

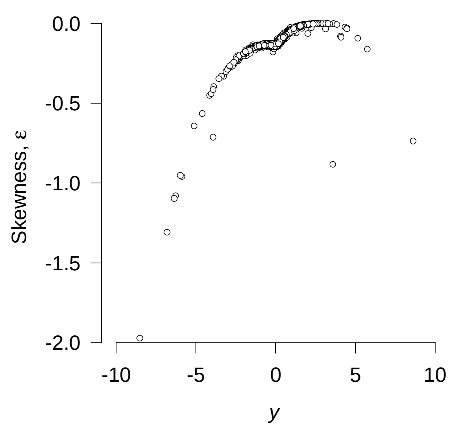

Figure 2: Skewness measure $\hat{s}_t$ versus observations y_t. (Note that one far outlying observation $y_t = -21.5$ is not included. For that observation the skewness measure has its smallest value $\hat{s}_t = -17.4$ and the conditional variance its largest value $\hat{h}_t^2 = 22.8$.)

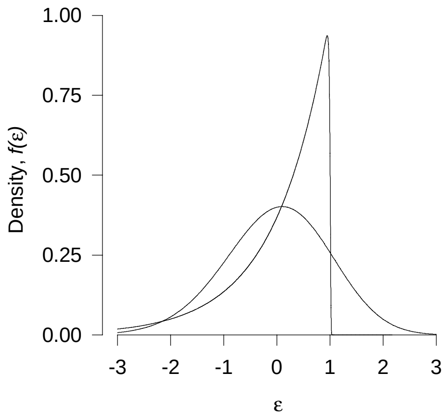

Figure 3: The standarized density of ε_t for $q = 0.2$ (the bell shaped curve) and $q = 8.52$.

(1999) approach). Second, using the relationship between $\hat{s}_t$ and $\hat{q}_t$ we see that the estimated effects for $\hat{q}_t$ will have changed signs and be doubled in size for $\hat{s}_t$. On comparison with Harvey and Siddique (1999) we find a negative but smaller effect of the skewness in the previous period. In our case, we find no smaller persistence in the conditional variance.

Figure 3 shows the standardized ε_t density for two selected values on q. For the black Monday of October 1987 ($y_t = -21.5$) the estimated $\hat{q}_t = 8.52$, while most $\hat{q}_t$s are between 0.05 and 0.3. Hence, the extreme shape for the $\hat{q}_t = 8.52$ case only arises once. The density for $\hat{q}_t = 0.2$ is very close to the shape of a standardized normal density function.

4. Discussion

The present approach is based on a direct modelling of a distributional characteristic (q_t) rather than on the skewness (s_t), which is a function of the characteristic. We could apply the same approach to the noncentral t-distribution used by Harvey and Siddique (1999). That distribution contains two distribution characteristics, and we could model both or the main determinant of skewness, i.e. the noncentrality parameter. In a future paper the details of such an approach will be given and compared to the results reported here.

The parameters were estimated by the use of a general purpose econometric program package, so that application to other series should be straightforward. Our experience with respect to the estimates of parameters was positive on all accounts. Only for the full model did we have to re-estimate multiple times before we managed to get reasonable estimates of the standard errors. The use of analytical derivatives for the covariance matrix could presumably reduce this problem.

Mathematical Appendix

Let the random variable V have the log generalized gamma density

$$f_V(v) = \Gamma^{-1}(k)\exp(kv - e^v)$$

(cf. Prentice, 1974). As $k \to \infty$, V is normally distributed. To obtain the normal density for a finite parameter value, Prentice advocates transformation and re-parametrization. Set $W = k^{1/2}(V - \ln k)$ so that $f_W(w) = k^{-1/2}\exp(k^{1/2}w + k\ln k - ke^{k^{-1/2}w})/\Gamma(k)$ and re-parametrize with $q = k^{-1/2}$. This gives the density $f_W(w) = |q|(q^{-2})^{q^{-2}}\exp[q^{-2}(qw - e^{qw})]/\Gamma(q^{-2})$ for which the normal arises as $q \to 0$.

Since V has the cumulant generating function $\ln\Gamma(q^{-2}+\theta) - \ln\Gamma(q^{-2})$ the first moments are $\psi(q^{-2}), \psi'(q^{-2})$ and $\psi''(q^{-2})$. Using the result on cumulant generating functions of Cramér (1945, p. 187) for linear transformations we easily obtain the moments for the random variable W: $\mu = \mathrm{E}(W) = q^{-1}\psi(q^{-2}) - q^{-1}\ln q^{-2}, \sigma^2 = \mathrm{Var}(W) = q^{-2}\psi'(q^{-2})$, and skewness $s = q^{-3}\psi''(q^{-2})$.

To obtain zero mean and unit variance, set $\varepsilon = (W-\mu)/\sigma$ to obtain the density $f(\varepsilon) = \sigma|q|(q^{-2})^{q^{-2}} \times \exp[q^{-2}(q\sigma\varepsilon + q\mu - e^{q\sigma\varepsilon + q\mu})]/\Gamma(q^{-2})$. In addition, we need the density for the innovation $u = \varepsilon h \in (-\infty, \infty)$:

$$f(u) = \frac{\sigma h^{-1}|q|(q^{-2})^{q^{-2}}}{\Gamma(q^{-2})}\exp[q^{-2}(q\sigma u/h + q\mu - e^{q\sigma u/h + q\mu})].$$

This density forms the basis for conditional maximum likelihood estimation of the parameters generating q, u, and h.

References

Abramowitz, M. and Stegun, I.A. (1965). *Handbook of Mathematical Functions.* Dover, New York.

Brännäs, K. and De Gooijer, J.G. (1994). Autoregressive-Asymmetric Moving Average Models for Business Cycle Data. *Journal of Forecasting* **13**, 529-544.

Brännäs, K. and De Gooijer, J.G. (2000). Asymmetries in Conditional Mean and Variance: Modelling Stock Returns by asMA-asQGARCH. Umeå Economic Studies 535.

Brunner, A.D. (1992). Conditional Asymmetries in Real GNP: A Seminonparametric Approach. *Journal of Business and Economic Statistics* **10**, 65-72.

Chen, J., Hong, H. and Stein, J.C. (2000). Forecasting Crashes: Trading Volume, Past Returns and Conditional Skewness in Stock Prices. NBER Working Paper 7687.

Cramér, H. (1945). *Mathematical Methods of Statistics.* Almqvist & Wiksells, Uppsala.

Farewell, V.T. and Prentice, R.L. (1977). A Study of Distributional Shape in Life Testing. *Technometrics* **19**, 69-75.

Harvey, C.R. and Siddique, A. (1999). Autoregressive Conditional Skewness. *Journal of Financial and Quantitative Analysis* **34**, 465-487.

Premaratne, G. and Bera, A.K. (2000). Modeling Asymmetry and Excess Kurtosis in Stock Return Data. Working Paper 00-123, Department of Economics, University of Illinois, Champaign.

Prentice, R.L. (1974). A Log Gamma Model and its Maximum Likelihood Estimation. *Biometrika* **61**, 539-544.

Stacey, E.W. (1962). A Generalization of the Gamma Distribution. *Annals of Mathematical Statistics* **33**, 1187-1192.

Wecker, W.E. (1981). Asymmetric Time Series. *Journal of the American Statistical Association* **76**, 16-21.