

A FLEXIBLE PARAMETRIC GARCH MODEL WITH AN APPLICATION TO EXCHANGE RATES

KAI-LI WANG,^{a*} CHRISTOPHER FAWSON,^b CHRISTOPHER B. BARRETT^c AND JAMES B. MCDONALD^d

^a *Department of International Trade, Tam Kang University, Taiwan*

^b *Department of Economics, 3530 Old Main Hill, Utah State University, Logan, UT 84322*

^c *Department of Applied Economics and Management, 351 Warren Hall, Cornell University, Ithaca, NY 14853-7801, USA*

^d *Department of Economics, 130, FOB, Brigham Young University, Provo, Utah 84602, USA*

SUMMARY

Many asset prices, including exchange rates, exhibit periods of stability punctuated by infrequent, substantial, often one-sided adjustments. Statistically, this generates empirical distributions of exchange rate changes that exhibit high peaks, long tails, and skewness. This paper introduces a GARCH model, with a flexible parametric error distribution based on the exponential generalized beta (EGB) family of distributions. Applied to daily US dollar exchange rate data for six major currencies, evidence based on a comparison of actual and predicted higher-order moments and goodness-of-fit tests favours the GARCH-EGB2 model over more conventional GARCH-*t* and EGARCH-*t* model alternatives, particularly for exchange rate data characterized by skewness. Copyright © 2001 John Wiley & Sons, Ltd.

1. INTRODUCTION

Contemporary modelling of exchange rate time series makes widespread use of generalized autoregressive conditional heteroskedastic (GARCH) models.¹ GARCH models have been shown not only to capture volatility clustering but also to accommodate some of the leptokurtosis (i.e. thick tails) commonly found in exchange rate time series. However, GARCH models with conditionally normal errors generally fail to sufficiently capture the leptokurtosis evident in asset returns (Bollerslev, 1987; Baillie and Bollerslev, 1989; Hsieh, 1989; Baillie and DeGennaro, 1990; Wang *et al.*, 1996). The increased attention focused on distributional properties (particularly tail thickness), when estimating exchange rates models (Boothe and Glassman, 1987; Koedijk *et al.*, 1992; Loretan and Phillips, 1994; Huisman *et al.*, 1998), has led to the widespread adoption of non-normal conditional error distributions, most commonly the Student-*t* (Bollerslev, 1987; Bollerslev *et al.*, 1994). The Student-*t* distribution models thicker tails than the normal, but does not permit skewness. With more than \$1.5 trillion traded daily in global currency markets, specification issues that affect estimates of potentially time-varying, higher-order central moments have significant practical implications for exchange rate risk management.

Economic theories of exchange rate determination offer two likely explanations for the empirical regularity of fat-tailed exchange rate returns.² The first is the overshooting of floating nominal

* Correspondence to: Kai-Li Wang, Department of International Trade, Tam-Kang University, Taiwan.
E-mail: kalwang@mail.tku.edu.tw

¹ Bollerslev *et al.* (1992) offer a good survey.

² See Taylor (1995) and Obstfeld and Rogoff (1996) for excellent, formal treatments of exchange rate determination models.

exchange rates associated with monetary or fiscal shocks in the presence of sticky prices (Dornbusch, 1976). The other is speculative attacks against fixed exchange rates (Krugman, 1979). Both models imply infrequent, extraordinarily sharp movements in exchange rates that are likely to appear as long tails in a distribution of differenced exchange rates. In addition, sticky prices in floating rate regimes, and especially fixed exchange rates, also generate modal daily exchange rate changes near zero (Obstfeld and Rogoff, 1996). The implication is that exchange rate changes are concentrated near the mean (high peakedness) but are likely to have long tails.³ As such, the choice of a conditional distribution should accommodate both long tails and high peakedness in the exchange rate series. Commonly used leptokurtic distributions, such as the Student-*t*, are not sufficiently flexible to capture both the high peakedness and the fat-tailed properties of exchange rate returns.

Moreover, skewness might also be important in exchange rate return series that exhibit episodes of sharp depreciation (appreciation) not offset by subsequent sharp appreciation (depreciation). Two reasons for skewness are: first, permanent shocks that lead to changes in the equilibrium exchange rate may be asymmetric; rapid improvements in Japanese productivity over the past thirty years is such an example; and second, speculative attacks against a currency tend to be one-sided. The 1992–3 European, 1994 Mexican, and 1997–8 East Asian currency crises are recent examples of such episodes. Since significant skewness is observed in exchange rate series that have experienced speculative attacks or other adverse shocks (Boothe and Glassman, 1987; Hsieh, 1988; Peruga, 1988; Huisman *et al.*, 1998), estimation methods that accommodate skewness are needed.⁴

Given the problems associated with quasi-maximum likelihood GARCH estimation (Pagan and Sabau, 1987; Lee and Hansen, 1994; Deb, 1996), incomplete accommodation of the statistical characteristics of exchange rates may yield inaccurate estimates of exchange rate dynamics. Common stylized facts associated with exchange rate distributions include: clustering, possible skewness, thick tails, and peakedness. The linear GARCH model can model clustering and as, noted above, combined with a Student-*t* distribution can also capture some kurtosis characteristics, but not any skewness in the data. This paper introduces a GARCH-EGB2 model, based on the exponential generalized beta distribution of the second kind (EGB2), which can accommodate all four of these features. The GARCH-EGB2 model is then applied to six different data sets and performs very well.

2. THE GARCH-EGB2 MODEL

McDonald (1984, 1991) introduced the exponential generalized beta distribution of the second kind (EGB2); a flexible distribution that is able to accommodate not only thick tails but also asymmetry. The EGB2 distribution includes many other well-known distributions as special or limiting cases and has been useful in applications characterized by nonnormal errors (McDonald, 1984; 1993). Of

³ An alternative perspective, provided by Friedman (1953), is to recognize that profit-maximizing speculators stabilize transitory shocks to the exchange rate and accelerate movement in response to permanent shocks. Thus, if transitory shocks are far more common than permanent shocks, the empirical distribution of exchange rate changes will be high-peaked and long-tailed.

⁴ Hansen (1994) uses a modified Student-*t* distribution to accommodate skewness but is unable to simultaneously accommodate high peakedness. Moreover, Hansen's model depends on appropriate *ex ante* lag selection. Huisman *et al.* (1998) demonstrate that the GARCH-*t* model does an adequate job of capturing the leptokurtosis inherent in exchange rate returns but they do not explore its ability to satisfactorily capture the high peakedness or asymmetry.

particular interest, the four parameter EGB2 distribution is sufficiently flexible to model peakedness and skewness commonly observed in high-frequency data. The EGB2 distribution is defined by the probability density function (pdf):

$$\text{EGB2}(z; \delta, \sigma, p, q) = \frac{e^{\frac{p(z-\delta)}{\sigma}}}{|\sigma|B(p, q) \left(1 + e^{\frac{z-\delta}{\sigma}}\right)^{p+q}}$$

where δ is a location parameter that affects the mean of the distribution, σ reflects the scale of the density function, and p and q are shape parameters that together determine the skewness and kurtosis of the distribution. The EGB2 converges in distribution to the normal when $p = q$ approaches infinity. It is symmetric for $p = q$ and is positively (negatively) skewed for $p > q$ ($p < q$) for $\sigma > 0$; the skewness results reverse for $\sigma < 0$. The EGB2 can accommodate coefficient of skewness values between -2 and 2 and coefficient of kurtosis values up to 9 (McDonald, 1991). These constraints are sufficiently flexible for most data series and will not constrain our investigation of exchange rate return data in this paper. Although McDonald and Xu (1995) introduce another more general EGB distribution, which contains the EGB2 as a special case, preliminary results of estimating the EGB distribution suggest most financial data are adequately modelled by the EGB2 distribution. Indeed, in this spirit, McDonald and Xu (1995, p. 134) found that the exponential generalized beta distribution of the second kind (EGB2) accommodates possibly thick-tailed and skewed error distributions and provides a systematic basis for partially adaptive estimation in regression and time-series models. Therefore, we employ the EGB2 specification in our study.

Quasi-maximum likelihood GARCH estimation has been shown to have poor finite sample properties if the data-generating process is not correctly specified (Pagan and Sabau, 1987; Lee and Hansen, 1994; Deb 1996). To address this potential shortcoming, we propose a general autoregressive integrated moving average [ARIMA(m, d, n)] specification with GARCH(1,1) conditional variance based on the EGB2 distribution.⁵ This approach allows us to account for most of the characteristics observed in empirical financial distributions, including first-order serial correlation, time-varying conditional variance, asymmetry, thick tails, and high peakedness. Denoting a time-series dependent variable as y_t , the general form of this model is given by:

$$\begin{aligned}\phi_m(L)y_t &= \mu + \theta_n(L)\epsilon_t \\ \epsilon_t &= h_t^{0.5} z_t \\ E(\epsilon_t^2 | \psi_{t-1}) &= h_t = w + \alpha \epsilon_{t-1}^2 + \beta h_{t-1} \\ \epsilon_t | \psi_{t-1} &\sim D(0, h_t, \eta)\end{aligned}$$

where the $\phi_m(L)$ and $\theta_n(L)$ are polynomials in the lag operator (L) of order m and n , respectively, and $w, \alpha, \beta > 0$ to ensure strictly positive conditional variance. The errors, ϵ_t , follow the assumed zero mean conditional density function (D) with variance h_t , and the parameter vector η . The latter are shape parameters, $\eta = \{p, q\}$ under EGB2, $\eta = \{v\}$ under the Student- t distribution, and

⁵ The GARCH(1,1) specification we employ is generally excellent for a wide range of financial data (Bollerslev *et al.*, 1992).

η is the empty set under the normal distribution. $\{z_t\}$ should have zero mean and unit variance and be i.i.d. if the model is correctly specified. To achieve efficiency, we jointly estimate the conditional mean and conditional variance equations with the conditional distribution by full information maximum-likelihood (FIML) estimation using the GAUSS Constrained Maximum-Likelihood (CML) module.

For the standardized EGB2 distribution with shape parameters p and q , the GARCH-EGB2 log-likelihood function is:

$$\log L = T[\log(\sqrt{\Omega}) - \log(B(p, q)) + p\Delta] + \sum \left[p \left(\frac{\sqrt{\Omega}\varepsilon_t}{\sqrt{h_t}} \right) - 0.5 \log(h_t) - (p + q) \log \left(1 + \exp \frac{\sqrt{\Omega}\varepsilon_t}{\sqrt{h_t}} + \Delta \right) \right]$$

where

$$\begin{aligned} \Delta &= \psi(p) - \psi(q) \\ \Omega &= \psi'(p) + \psi'(q) \end{aligned}$$

and $\psi(p)$ and $\psi'(p)$ represent digamma and trigamma functions, respectively. We show the detailed parameterization of the GARCH-EGB2 model in the Appendix.

For the Student- t distribution with ν degrees of freedom, the GARCH- t log-likelihood function (as presented by Bollerslev, 1987) is:

$$\begin{aligned} \log L = T & \left[\log \Gamma \left(\frac{\nu + 1}{2} \right) - \log \Gamma \left(\frac{\nu}{2} \right) - 0.5 \log(\pi(\nu - 2)) \right] \\ & - 0.5 \sum \left[\log h_t + (\nu + 1) \log \left(1 + \frac{\varepsilon_t^2}{h_t(\nu - 2)} \right) \right] \end{aligned}$$

where Γ denotes the gamma function.

Of particular concern are potential skewness and high peakedness that international finance theory suggests are regular features of exchange rate return series. The Student- t is symmetric and so cannot accommodate skewness in the underlying series, while the EGB2 distribution readily captures skewness. The EGB2 also achieves higher peaks than the Student- t . Consider the Student- t and symmetric EGB2($p = q$), both with unit variance and zero mean. Thus, with $p = q$ in the EGB2, each distribution has only one free parameter: ν for the t and p for the EGB2. The kurtosis parameter for the t distribution is then given by $3[(\nu - 2)/(\nu - 4)]$ for $\nu \geq 4$ while the same parameter for the EGB2 is given by $[\psi'''(p)/(2(\psi'(p))^2)] + 3$. Simple plots of the two density functions, selecting ν and p to equalize kurtosis between the two distributions, shows that increasing kurtosis makes the EGB2 more peaked than the Student- t distribution. For kurtosis values equal to three, the reference value for the normal, the two densities are essentially indistinguishable. Once kurtosis exceeds 4, the higher peak of the EGB2 becomes evident and the differences increase with kurtosis.⁶ Since exchange rate series typically exhibit kurtosis in the 4–6

⁶ Density plots showing this difference are available on the *Journal of Applied Econometrics Data Archive* (<http://www.econ.queensu.ca/jae>).

range (see Section 3, below), the peakedness difference applies and seems relevant to specification decisions.

The linear GARCH model does not capture the asymmetric second moment (a so-called leverage effect) since the conditional variance is only linked to past conditional variance and squared innovations, and hence prohibits an asymmetric response in the conditional variance to positive and negative errors (Black, 1976; Christie, 1982; French *et al.*, 1987). This limitation led to the introduction of a non-linear Exponential GARCH (EGARCH) specification by Nelson (1991) in which the asymmetrical behaviour of exchange rate returns is modelled as an asymmetric, non-linear specification of the conditional variance process and a symmetric distribution (such as the Student- t) for the conditional error. To provide a more complete assessment of the relative strength of the GARCH-EGB2 model, we also consider the EGARCH- t specification that accommodates a leverage effect (asymmetry), volatility clustering, and leptokurtosis while still relying on a symmetric distribution.⁷ Our findings suggest that where asymmetry exists in the underlying returns, as distinct from asymmetry in their conditional variance, a linear GARCH model based on the more flexible EGB2 distribution clearly outperforms the non-linear EGARCH model based on the symmetric t distribution. Refining the conditional variance specification is advantageous only when it does not come at the cost of sacrificing important features of the conditional mean distribution. Put bluntly, accommodating skewness in the conditional first moment seems more important than accommodating it in the conditional second moment.

The log-likelihood function for the EGARCH- t model is the same as that for the GARCH- t model with the exception that the conditional variance h_t is specified as:

$$\log(h_t) = w + \sum_{i=1}^a \alpha_i g(z_{t-i}) + \sum_{i=1}^p \beta_i \log(h_{t-i})$$

where

$$g(z_t) = \theta z_t + \gamma[|z_t| - E|z_t|]$$

$$E|z_t| = \sqrt{v}\Gamma[0.5(v-1)]/[\sqrt{\pi}\Gamma(0.5v)]$$

Since we are estimating an EGARCH(1,1) model, the parameter α_1 can be set equal to one.

When estimating the GARCH- t , EGARCH- t , and GARCH-EGB2 models, all parameters were jointly estimated using the constrained maximum likelihood routine in Gauss and were found to be invariant to starting points and algorithm adopted.⁸ The main difference between the estimations was the number of iterations required for convergence, with the Gaussian GARCH model converging most rapidly followed by the GARCH- t , the GARCH-EGB2 and then the non-linear EGARCH- t model.

In addition to non-negativity constraints on the parameter space to ensure positive conditional variance in linear GARCH models, regardless of the underlying distributional assumption, the shape parameters, η , also need to be constrained. In the case of the EGB2 distribution, p and q must be positive, while in estimation based on the Student- t distribution, v must be greater than two, which is necessary for the standardized t distribution to be defined. In practice, these constraints are easy to impose.

⁷ This suggestion was made by one of the referees.

⁸ The tolerance level in estimation was set at 0.00001.

The GARCH- t , EGARCH- t , and GARCH-EGB2 are non-nested models, thus comparisons are difficult. The GARCH-EGB2 maintains an identical specification relative to the GARCH- t up to the non-nested choice of distribution. So, although one cannot make formal comparisons based on likelihood values, model performance across multiple indicators nonetheless provides reasonably direct evidence on the benefits of moving to the more flexible distributional specification. By adding just one extra shape parameter, relative to the GARCH- t model, the GARCH-EGB2 model is able to account not only for the first, second, and fourth moments of the conditional distribution of the dependent variable, as do popular Gaussian GARCH and GARCH- t models, it is also able to accommodate flexibility in the third moment and peakedness. As we will show, this enables the GARCH-EGB2 model to consistently outperform the GARCH- t model.

Comparisons between the EGARCH- t and GARCH-EGB2 models are somewhat more difficult because of different specifications for both the error distribution and for the conditional variance. However, some insights can still be gained by comparing these two models' performance across multiple indicators as well.

3. DATA

We use daily noon spot US dollar exchange rate data (\$/local currency) for the German deutsche mark (DM), British pound, Japanese yen (¥), French franc (FF), Belgian franc (BF), and Italian lira (IL) over the period 1 January 1985 to 21 November 1996 (3016 observations per series) to demonstrate the properties of the GARCH-EGB2 estimators. Data were obtained from the Exchange Rate Service of the Pacific Data Center at the University of British Columbia. To achieve stationarity, we transform the nominal exchange rate data by taking the first-difference of the logarithm for each exchange rate series:⁹

$$R_{i,t} = \ln[S_{i,t}/S_{i,t-1}] * 100$$

where $S_{i,t}$ equals the nominal spot foreign exchange rate of currency i at period t , expressed as US \$/currency $_i$, $R_{i,t}$ equals the percentage change in the nominal exchange rate of currency i at period t , and $R_{i,t} > 0$ ($R_{i,t} < 0$) indicates currency appreciation (depreciation).¹⁰

Table I presents descriptive statistics for each exchange rate series ($R_{i,t}$). All six currencies exhibit leptokurtosis, a coefficient of kurtosis (KUR) significantly in excess of the normal distribution's reference value of three, and the yen, pound, and lira all show significant skewness. The observed skewness may be attributable to permanent structural shocks that led to the yen's dramatic appreciation over the sample period and to the autumn 1992 speculative attacks that knocked the pound and lira of the European monetary system's exchange rate mechanism (ERM).¹¹

⁹ The data and unit root test results are reported in Appendix 2 on the *Journal of Applied Econometrics Data Archive* (<http://www.econ.queensu.ca/jae>).

¹⁰ Since we do not adjust for weekends or holidays, $R_{i,t}$ reflects exchange rate changes between two successive trading days.

¹¹ The existence of significant structural shocks raises the possibility that one might be able to fit the data better using multiple models over sample subperiods. The strategy of looking for break points, however, implies either the use of a flexible distribution through a mixture model with endogenous switching points—thereby reinforcing our core point that more flexible distributions are necessary for modeling exchange rate dynamics—or the use of multiple models absent explicitly transition dynamics, in which case parameter stability is sacrificed for fit.

Table I. Descriptive statistics for percentage change in the (\$/currency) nominal exchange rate for the German deutsche mark (DM), British pound, Japanese yen (¥), French franc (FF), Belgian franc (BF), and Italian lira (IL) over the period 1 January 1985 to 21 November 1996

	SK	KUR	$f_{0.75} - f_{0.25}$	$f_{0.6} - f_{0.4}$	JB	$Q(30)$	$Q^2(30)$
DM	-0.037 (0.045)	5.1 (0.09)	1.13	0.41	566.73* [0.00]	30.03 [0.46]	391.36 [0.00]
£	-0.12 (0.045)	5.2 (0.09)	1.08	0.40	604.92* [0.00]	37.98 [0.15]	451.30 [0.00]
¥	0.286 (0.045)	6.1 (0.09)	1.02	0.38	1282.9* [0.00]	37.80 [0.16]	233.98 [0.00]
FF	0.02 (0.045)	6.0 (0.09)	1.14	0.39	511.89* [0.00]	41.16 [0.08]	391.67 [0.00]
BF	0.024 (0.045)	5.0 (0.09)	1.12	0.41	521.53* [0.00]	41.69 [0.08]	370.85 [0.00]
IL	-0.616 (0.045)	8.8 (0.09)	1.14	0.41	4377.65* [0.00]	33.39 [0.31]	641.75 [0.00]

Notes: SK = coefficient of skewness. $E(R_{i,t} - \mu)^3/\sigma^3$, where μ is the mean and σ is the standard deviation.)
 KUR denotes the kurtosis coefficient. $(E(R_{i,t} - \mu)^4/\sigma^4)$, where μ is the mean and σ is the standard deviation.)
 The asymptotic standard errors of SK and KUR are reported in parentheses and computed as $(6/T)^{0.5}$ and $(24/T)^{0.5}$, respectively.

$f_{\alpha_1} - f_{\alpha_2}$ denotes the inter-percentile range. Thus, $f_{0.75} - f_{0.25}$ represents the distance between the values of the random variable at which the cumulative distribution function equals 0.75 and the value at which the cumulative distribution equals 0.25. Hence, 50% of the observations lie within this range around the median. The lower the value of $f_{\alpha_1} - f_{\alpha_2}$, the higher the peakedness of a unimodal distribution.

JB denotes the Jarque–Bera Test for normality defined by $T[\text{SK}^2/6 + (\text{KURT} - 3)^2/24]$ which is asymptotically distributed as $\chi^2(2)$. As a benchmark, the 1% critical value equals 9.21.

Q and Q^2 represent the Ljung–Box test statistics for up to 30th-order serial correlation for each exchange rate series and its square, respectively. Similar results are obtained for different orders. P -values against the null hypothesis of white noise are reported in brackets.

*denotes statistical significance at the 1% level. P -values are reported in brackets.

As we will see in Section 4, the GARCH-EGB2 model is especially appealing for currencies such as these three, which exhibit significantly skewed percentage change distributions.

Let $f_{\alpha_1} - f_{\alpha_2}$ denote the inter-percentile range corresponding to the cumulative probabilities α_1 and α_2 ($f_{0.75} - f_{0.25}$ is represents the distance between the values at which the cumulative distribution function equals 0.75 and the value at which the cumulative distribution equals 0.25) Given α_1 and α_2 , on opposite sides of the median, the lower the value of $f_{\alpha_1} - f_{\alpha_2}$, the higher the peakedness of a unimodal distribution. Across all six standardized exchange rates, the value $f_{0.75} - f_{0.25}$ is uniformly less than 1.36, the reference range corresponding to the standard normal distribution. The unconditional distributions of these exchange rates have higher peaks than does a normal distribution around the central 50% of probability mass. The high peakedness is corroborated as well over the narrower interval $f_{0.6} - f_{0.4}$, for which all exchange rates' ranges are less than 0.5, the inter-percentile value of the standard normal over its central 20% of probability mass.

Given skewness, fat tails, and high peakedness, it is not surprising that the null hypothesis of normality is strongly rejected by the Jarque–Bera (JB) asymptotic test for each exchange rate. Table I also presents Ljung–Box test statistics indicating significant autocorrelation in $R_{i,t}$ at a lag of 30 trading days ($Q(30)$), and significant volatility clustering at the same lag ($Q^2(30)$). In summary, the series' descriptive statistics suggest the unconditional distributions of daily exchange

rate changes deviate considerably from the traditional Gaussian assumption. These results are consistent with previous empirical findings and economic theory (Boothe and Glassman, 1987; Hsieh, 1988; Wang *et al.*, 1996; Husiman *et al.*, 1998).

4. EMPIRICAL RESULTS

We began estimation by identifying and estimating a common ARMA process for the stationary $R_{i,t}$. First, Box–Jenkins techniques were used to reduce the set of prospective ARMA specifications. Next, we further narrowed the pool of possible models to those having a p -value for the Ljung–Box portmanteau $Q(30)$ statistic of greater than 0.3, a significance level selected to support a reasonable assumption of white noise. Finally, we chose the ARMA specification having the lowest Schwarz Bayesian criterion (SBC) value from among the candidate models having passed the Box–Jenkins and $Q(30)$ screens. In other words, the Ljung–Box Q statistic was used to identify a few possible models and then the information criterion (SBC) selected the final ARMA specification for the conditional mean equation.

Five different models were considered: homoscedastic normal (HN), Gaussian GARCH (GARCH), Student- t GARCH (GARCH- t), exponential Student- t GARCH (EGARCH- t), and a GARCH-EGB2 model. In each specification, the parameters in the ARIMA model were jointly estimated with the GARCH conditional variance and distributional parameters for each currency using maximum likelihood procedures. Table II reports the Ljung–Box Portmanteau statistics of squared standardized residuals (z_t) for all currencies for each of the five estimated models and clearly indicates the elimination of serial correlation in the conditional variance.

Parameter estimates and associated standard errors of each model fit to each exchange rate series are reported in Appendix 3 on the JAE Data Archive. White (1982) showed that if the model is correctly specified, conventional standard errors and White robust standard errors will be stochastically the same. Our results for the GARCH-EGB2 specification routinely yield nearly identical standard error estimates by either method, providing informal evidence that the GARCH-EGB2 captures important stylized facts.

Table II. Ljung–Box $Q^2(30)$ tests for serially correlated conditional variance for homoscedastic normal (HN), Gaussian GARCH (GARCH), GARCH- t , and GARCH-EGB2 model specifications

	DM	£	¥	FF	BF	IL
HN	403.20 [0.00]	447.50 [0.00]	239.53 [0.00]	404.10 [0.00]	370.75 [0.00]	620.03 [0.00]
GARCH	29.28 [0.50]	24.36 [0.76]	28.72 [0.53]	25.50 [0.70]	35.82 [0.21]	23.67 [0.79]
GARCH- t	28.71 [0.53]	24.46 [0.75]	29.06 [0.52]	25.41 [0.71]	34.73 [0.25]	26.22 [0.66]
EGARCH- t	28.35 [0.55]	23.24 [0.81]	30.03 [0.46]	23.95 [0.77]	32.08 [0.36]	31.19 [0.41]
GARCH-EGB2	28.96 [0.52]	24.64 [0.74]	29.52 [0.49]	25.50 [0.70]	35.01 [0.24]	25.03 [0.72]

Notes: The figure in brackets is the p -value of the Ljung–Box $Q^2(30)$ test against the null hypothesis of no serial correlation.

The estimated parameters of the conditional variance, between the GARCH-EGB2 and GARCH- t , are close to each other for the non-skewed distributions, the German mark, French franc, and Belgian franc. However, the results exhibit greater differences for currencies associated with skewed distributions, the British pound, Japanese yen, and Italian lira. This fact is not surprising and emphasizes the importance of being able to specify a distribution which accommodates the underlying data characteristics.

While all the GARCH models appear to successfully model second-order serial correlation, the issue of non-normality remains. To assess the relative descriptive power of the GARCH-EGB2 model over competing models we utilize two diagnostics: (1) a comparison of actual and predicted higher-order moments; and (2) a comparison of goodness of fit statistics across model specifications.¹²

4.1. A Comparison of Actual and Predicted Higher-order Moments

Table III reports estimated sample skewness and kurtosis coefficients and corresponding predicted moments of standardized residuals, along with estimated standard errors, for each data series and estimated model. A comparison of the results for the homoscedastic normal (HN) and the Gaussian GARCH model (GARCH) suggests that skewness and excess kurtosis of the standardized residuals persist for most currencies in the Gaussian GARCH models. For all currencies except the Japanese yen, the Gaussian GARCH leptokurtic characteristics ($m_4^{\text{Gaussian-GARCH}}$) are smaller than those for the homoscedastic model (m_4^{HN}) and they both remain greater than the theoretical kurtosis coefficient of a normal distribution.¹³ This result confirms the finding that the Gaussian assumption is not sufficiently flexible to fully account for leptokurtosis in exchange rate data and is consistent with previously referenced literature that employs the Student- t conditional error distribution to account for leptokurtosis.

Predicted and observed conditional kurtosis values for the non-Gaussian models are reported in Table III in the sections labelled GARCH- t , EGARCH- t , and GARCH-EGB2.¹⁴ In contrast with the Gaussian model results, the predicted kurtosis coefficients ($\phi_4^{\text{GARCH-}t}$ and $\phi_4^{\text{EGARCH-}t}$) of the Student- t distribution models are all greater than the coefficients ($m_4^{\text{GARCH-}t}$ and $m_4^{\text{EGARCH-}t}$) calculated from the standardized residuals. The implication is that GARCH modelling based on the leptokurtic Student- t distribution tends to overestimate the fourth moment.

A comparison of the predicted ($\phi_4^{\text{GARCH-EGB2}}$) and observed kurtosis coefficient ($m_4^{\text{GARCH-EGB2}}$) for the GARCH-EGB2 model suggests a much closer correspondence, with absolute differences

¹² In all diagnostics, the predicted values are constructed from the estimates of distribution parameters which are jointly estimated with the conditional mean and variance equations.

¹³ Milhøj (1987), MaCurdy and Morgan (1987), and Hsieh (1989) also found the Gaussian GARCH model can reduce some degree of leptokurtosis compared to the Gaussian homoscedasticity model. The 'theoretical' or predicted skewness and kurtosis in these models are given by 0 and 3, respectively with traditional sample estimates of standard errors, calculated by $(6/T)^{0.5} = 0.045$ and $(24/T)^{0.5} = 0.089$ for $T = 3016$, being reported in parentheses in Table III. The strict validity of the standard errors is conditional on independent and identically distributed observations.

¹⁴ Predicted skewness and kurtosis coefficients ($\phi_4^{\text{GARCH-}t}$ in the case of GARCH- t $\phi_4^{\text{EGARCH-}t}$ for EGARCH- t and $\phi_4^{\text{GARCH-EGB2}}$ for GARCH-EGB2) are calculated from estimated values for the shape parameters (ν in the case of GARCH- t and EGARCH- t , p and q for GARCH-EGB2). These formulas are given as follows: the kurtosis for the Student- t and EGB2, respectively, are given by $3(\nu - 2)/(\nu - 4)$ for $\nu > 4$ and $\{[\psi'''(p) + \psi'''(q)] + 3[\psi'(p) + \psi'(q)]^2\}/[\psi'(p) + \psi'(q)]^2$; and the predicted skewness coefficient of EGB2 distribution is $[\psi''(p) - \psi''(q)]/[\psi'(p) + \psi'(q)]^{1.5}$. Two different standard errors are reported for the non-Gaussian models: (1) the traditional standard errors defined in footnote 14 and (2) standard errors obtained using delta methods, denoted by $(\)^{\text{delta}}$. The details of the derivations are contained in Appendix 4 on the JAE Data Archive. Observed kurtosis coefficients ($m_4^{\text{GARCH-}t}$, $m_4^{\text{EGARCH-}t}$, and $m_4^{\text{GARCH-EGB2}}$) for all series are calculated directly from each model's standardized residuals.

Table III. Skewness and kurtosis statistics of sample standardized residuals and predicted values

	DM	£	¥	FF	BF	IL
HN						
m_3^{HN}	-0.034 (0.034) (0.045)	-0.100 (0.100) (0.045)*	0.281 (0.281) (0.045)*	0.026 (0.026) (0.045)*	0.036 (0.036) (0.045)*	-0.586 (0.586) (0.045)*
m_4^{HN}	5.054 (2.054) (0.089)*	5.053 (2.053) (0.089)*	6.099 (3.099) (0.089)*	4.943 (1.943) (0.089)*	4.966 (1.966) (0.089)*	8.550 (5.550) (0.089)*
GARCH						
$m_3^{\text{Gaussian-GARCH}}$	0.075 (0.075) (0.045)	-0.110 (0.110) (0.045)*	0.464 (0.464) (0.045)*	0.094 (0.094) (0.045)*	0.114 (0.114) (0.045)*	-0.118 (0.118) (0.045)*
$m_4^{\text{Gaussian-GARCH}}$	4.419 (1.419) (0.089)*	4.365 (1.365) (0.089)*	6.154 (3.154) (0.089)*	4.350 (1.350) (0.089)*	4.402 (1.402) (0.089)*	4.745 (1.745) (0.089)*
GARCH-t						
$m_3^{\text{GARCH-t}}$	0.083 (0.083) (0.045)	-0.108 (0.108) (0.045)*	0.502 (0.502) (0.045)*	0.103 (0.103) (0.045)*	0.127 (0.127) (0.045)*	-0.160 (0.160) (0.045)*
$m_4^{\text{GARCH-t}}$	4.451 (1.416) (0.089)*	4.345 (1.959) (0.089)*	6.373 (21.627) (0.089)*	4.398 (1.317) (0.089)*	4.462 (1.430) (0.089)*	5.172 (1.189) (0.089)*
$\phi_4^{\text{GARCH-t}}$	5.867 (0.921) ^{delta}	6.304 (1.155) ^{delta}	28 (36.563) ^{delta}	5.715 (0.857) ^{delta}	5.892 (0.948) ^{delta}	6.361 (1.211) ^{delta}
EGARCH-t						
$m_3^{\text{EGARCH-t}}$	0.080 (0.080) (0.045)	-0.114 (0.114) (0.045)*	0.455 (0.455) (0.045)*	0.114 (0.114) (0.045)*	0.127 (0.127) (0.045)*	-0.161 (0.161) (0.045)*
$m_4^{\text{EGARCH-t}}$	4.374 (1.399) (0.089)*	4.347 (2.108) (0.089)*	6.316 (20.304) (0.089)*	4.364 (1.273) (0.089)*	4.412 (1.353) (0.089)*	5.265 (1.644) (0.089)*
$\phi_4^{\text{EGARCH-t}}$	5.773 (0.883) ^{delta}	6.482 (1.249) ^{delta}	26.62 (32.736) ^{delta}	5.637 (0.825) ^{delta}	5.765 (0.888) ^{delta}	6.909 (1.535) ^{delta}
GARCH-EGB2						
$m_3^{\text{EGARCH-EGB2}}$	0.083 (0.088) (0.005)	-0.115 (-0.071) (0.044)	0.502 (0.326) (0.176)	0.102 (0.087) (0.015)	0.125 (0.076) (0.049)	-0.148 (-0.025) (0.123)
$\phi_3^{\text{EGARCH-EGB2}}$	(0.045) (0.419) ^{delta}	(0.045) (0.475) ^{delta}	(0.045)* (0.584) ^{delta}	(0.045) (0.414) ^{delta}	(0.045) (0.433) ^{delta}	(0.045)* (0.608) ^{delta}
$m_4^{\text{EGARCH-EGB2}}$	4.451 (0.133) (0.089)	4.395 (0.383) (0.089)*	6.355 (0.999) (0.089)*	4.394 (0.147) (0.089)	4.457 (0.144) (0.089)	5.065 (0.156) (0.089)
$\phi_4^{\text{EGARCH-EGB2}}$	4.584 (0.189) ^{delta}	4.778 (0.195) ^{delta}	5.356 (0.575) ^{delta}	4.541 (0.190) ^{delta}	4.601 (0.193) ^{delta}	4.909 (0.227) ^{delta}

Notes: m_3 is the coefficient of skewness of the standardized residuals from the estimated model.

m_4 is the coefficient of kurtosis of the standardized residuals from the estimated model.

The predicted skewness coefficient for the Gaussian and student- t is 0 (i.e. $\phi_3^{\text{OLS}} = \phi_3^{\text{Gaussian-GARCH}} = \phi_3^{\text{GARCH-t}} = \phi_3^{\text{EGARCH-t}} = 0$), and the predicted kurtosis coefficient for the Gaussian is 3.0 (i.e. $\phi_4^{\text{OLS}} = \phi_4^{\text{Gaussian-GARCH}} = 3$).

The asymptotic standard error of the coefficients of skewness and kurtosis are reported in parentheses and computed as $(6/T)^{0.5} = 0.045$ and $(24/T)^{0.5} = 0.089$, respectively; $()^{\text{delta}}$ indicates the standard error of skewness and kurtosis calculated by standard delta approach.

ϕ_4^t is the predicted kurtosis coefficient of Student- t distribution = $3(\nu-2)/(\nu-4)$, $\nu > 4$.

ϕ_3^{EGB2} is predicted skewness coefficient of EGB2 distribution = $[\psi''(p) - \psi''(q)]/[\psi'(p) + \psi'(q)]^{1.5}$.

ϕ_4^{EGB2} is predicted kurtosis coefficient of EGB2 distribution = $\{[\psi''(p) + \psi''(q)] + 3[\psi'(p) + \psi'(q)]^2\}/\{\psi'(p) + \psi'(q)\}^2$.

The $()$ reports the absolute difference between observed and predicted coefficients and * denotes the absolute difference is more than two times the standard errors of the skewness and kurtosis coefficients.

being smaller (sometimes considerably smaller) than for other models for every currency. Based on the traditional estimate of the standard error of the sample kurtosis, the observed differences between the sample and predicted kurtosis are statistically significant for all currencies for the GARCH- t and EGARCH- t models and for the pound and yen for the GARCH-EGB2. However, because of the non-linear relationship between the distributional parameters and predicted kurtosis, corresponding standard errors of predicted kurtosis, based on the delta method, are also calculated and reported for the (E)GARCH- t and GARCH-EGB2 models. The delta-standard errors for the (E)GARCH- t and EGB2 models tend to increase by a factor of at least ten and two, respectively, with none of observed differences between the observed and predicted kurtosis coefficients being statistically significant. However, regardless of the statistical significance, the close agreement between the EGB2 observed and predicted kurtosis estimates is impressive. For example, the absolute difference between the predicted and observed kurtosis coefficients for the British pound and Japanese yen are (0.383 for pound, 0.999 for yen) for the EGB2 specification and (1.959 for pound, 21.67 for yen) for the GARCH- t and (2.108 for pound, 20.304 for yen) for the EGARCH- t model.

Evidence of asymmetry in the distribution of standardized residuals is also explored through a comparison of observed and predicted coefficients of skewness, reported in Table III. The Student- t is symmetric with a predicted skewness coefficient of zero. The standard error of the EGB2 predicted skewness is calculated by the delta method and is reported in Table III. The observed coefficients of skewness ($m_3^{\text{Gaussian-GARCH}}$, $m_3^{\text{GARCH-}t}$, and $m_3^{\text{EGARCH-}t}$) for Gaussian GARCH, GARCH- t , and EGARCH- t models, across all currencies except the German DM, are statistically significant as judged by the traditional standard error. According to the delta standard error, the null hypothesis of no difference between the observed and predicted skewness coefficient can't be rejected for any currencies for GARCH-EGB2 models. However, based on traditional standard errors, the hypothesis of no difference between observed and predicted skewness coefficients is rejected for the Japanese yen and the Italian lira for all models. However, the difference (0.176 for yen, 0.123 for lira) GARCH-EGB2 model is far less than for competing symmetric distribution models (0.502 for yen, 0.160 for lira in the GARCH- t model and 0.455 for yen and 0.161 for lira in the EGARCH- t model). The yen case reinforces the relative strength of the GARCH-EGB2 approach. In particular, the highly skewed distribution of yen standardized residuals highlights the potential that estimated parameters are likely to be distorted under an inappropriate distribution assumption and reinforces the value of a model specification having sufficient flexibility to adapt to the distributional properties of the data. In general, the more flexible EGB2 distribution appears to be more adept at capturing properties of the higher order moments of exchange rate series while the traditional GARCH model, under Gaussian and Student- t assumptions, appears to yield systematic under or overestimation of third and fourth moments.

4.2. Goodness of Fit Statistics

Boothe and Glassman (1987) presented empirical evidence suggesting that nonnested distribution comparisons based on log-likelihood values frequently lead to spurious conclusions. Greene (1997, p. 162) also cautions against the inappropriate use of likelihood function comparisons when distributional assumptions differ across the models. Consequently, we use the χ^2 goodness of fit (GoF) statistic to compare differences between observed frequencies of standardized residuals and theoretically predicted frequencies based upon estimated distribution shape parameters. The

chi-square goodness of fit statistic is calculated by:

$$\text{GoF} = \sum_{i=1}^k \frac{(f_i - F_i)^2}{F_i}$$

where f_i is the observed count frequency of actual standardized residuals in the i th data class (interval), F_i is the predicted count frequency derived from the estimated values for the distribution shape parameters, and k is the number of data intervals used in the distributional comparisons. GoF has an asymptotic chi-squared distribution with degrees of freedom equal to the number of intervals minus the number of estimated (distribution) parameters minus one. The null hypothesis tested by the GoF statistic is that the observed and predicted distribution functions are identical.

Results for the χ^2 -test for goodness of fit test, using 40 (equal probability) intervals are reported in Table IV. Given the large sample size used in estimation (3016 observations), it is not surprising that test statistics suggest rejection of the hypothesis that the residuals are drawn from the assumed distribution at conventional levels of significance. This result is consistent with other large sample applications (Kloek and Van Dijk, 1978; McDonald, 1984).

However, if one looks at the relative magnitude of calculated statistics across model specifications, the evidence clearly favors the GARCH-EGB2 specification over the alternative GARCH- t for all cases, and particularly so for the case of skewed data series (£, ¥, IL). The GARCH-EGB2 has a smaller GoF statistic than the EGARCH- t for the skewed series. Direct accommodation of skewness in the series seems to dominate accommodation of potential asymmetry in series' conditional variance. For the DM, FF, and BF, with approximately symmetric distributions, the EGARCH- t has a smaller GoF statistic than both the GARCH- t and GARCH-EGB2 models. So the more flexible conditional variance specification of the EGARCH model only appears of value when the underlying series is reasonably symmetric. Indeed in two of the three skewed series, the EGARCH- t performs even more poorly than the GARCH- t , which is always dominated by the GARCH-EGB2 model. Further analysis of goodness of fit tests under different interval size assumptions reveals that our results are very robust to the number of intervals chosen.

Table V reports log-likelihood values for each model specification and data series. The best model based on a GoF criterion agrees with the best model obtained for a log-likelihood criterion except for the case of the model for the Japanese yen where a GARCH- t is the worst, based on log-likelihood and the EGARCH- t is worst based on GoF. The Schwarz criterion (SC) and Akaike information criterion (AIC), which take account of the number of parameters, yield rankings identical to those based on the GoF criterion.

Table IV. χ^2 goodness of fit test statistics for GARCH- t , EGARCH- t , and GARCH-EGB2 model specifications

Currency	GARCH- t	GARCH-EGB2	EGARCH- t
DM	116.8	115.3	95.8
£	94.4	83.4	88.4
¥	111.0	87.0	118.2
FF	94.4	93.6	92.8
BF	100.3	99.9	93.9
IL	113.7	99.8	114.1

Table V. Log-likelihood values for the GARCH- t , EGARCH- t , and GARCH-EGB2 model specifications

Currency	GARCH- t	GARCH-EGB2	EGARCH- t
DM	-3163.5	-3161.3	-3158.2
£	-2991.0	-2987.8	-2989.0
¥	-2970.2	-2963.7	-2969.5
FF	-3024.2	-3022.1	-3018.8
BF	-3137.1	-3134.8	-3132.0
IL	-3005.8	-3002.0	-3006.7

5. CONCLUSIONS

Although GARCH modelling based on normal or Student- t conditional distributions has proved useful in capturing the volatility clustering and leptokurtosis commonly present in asset price series, we have demonstrated the difficulty these models have in accommodating other commonly observed characteristics of high-frequency exchange rate data, notably high peakedness and skewness. Since economic theory suggests these are potentially important statistical characteristics of the underlying series, we have proposed a GARCH model based on the more flexible EGB2 distribution. The GARCH-EGB2 specification can model thick-tailed, high-peaked, and asymmetrically distributed data as well as volatility clustering. These attractive properties make it useful in empirical estimation of financial markets in which the specification of the distribution is vitally important. An application to daily logarithmic changes in six major exchange rates over ten years shows that the GARCH-EGB2 model consistently outperforms the commonly employed GARCH- t specifications. With skewed data series, the GARCH-EGB2 also outperformed the EGARCH- t , in spite of the latter's more flexible conditional variance specification. The GARCH-EGB2 appears to be a promising specification to accommodate high peakedness and thick tails in data series characterized by skewness and volatility clustering.

APPENDIX: PARAMETERIZATION OF THE STANDARDIZED EGB2

Following the traditional definition of a GARCH process, suppose that:

$$\varepsilon_t = h_t^{0.5} Z_t \quad (\text{A1})$$

where $\{\varepsilon_t\}$ is the error term sequence from the conditional mean equation and $\{Z_t\}$ is an i.i.d. sequence with zero mean and unit variance. Let h_t evolve according to a GARCH(1,1) process:

$$h_t = \alpha_0 + \alpha_1 h_{t-1} + \beta_1 \varepsilon_t^2 \quad (\text{A2})$$

The error ε_t from an EGB2 density is given by:

$$\text{EGB2}(\varepsilon; \delta, \sigma, p, q) = \frac{\exp\left(\frac{n(\varepsilon-\delta)}{\sigma}\right)}{|\sigma| B(p, q) \left(1 + e^{\frac{\varepsilon-\delta}{\sigma}}\right)^{p+q}} \quad (\text{A3})$$

The standardized residual, Z_t , follows an EGB2 distribution with zero mean and unit variance, therefore:

$$\text{Var}(z) = \sigma^2(\psi'(p) + \psi'(q)) = 1 \quad (\text{A4})$$

$$E(z) = \delta + \sigma[\psi(p) - \psi(q)] = 0 \quad (\text{A5})$$

where $\psi()$, $\psi'()$ are digamma and trigamma functions, respectively (Davis, 1935).

Solving for σ and δ in terms of Δ and Ω where:

$$\Delta = \psi(p) - \psi(q) \quad (\text{A6})$$

and

$$\Omega = \psi'(p) + \psi'(q) \quad (\text{A7})$$

results in:

$$\sigma = \sqrt{\frac{1}{\psi'(p) + \psi'(q)}} = \sqrt{\frac{1}{\Omega}} \quad (\text{A8})$$

$$\delta = -\sigma[\psi(p) - \psi(q)] = -\Delta \sqrt{\frac{1}{\Omega}} \quad (\text{A9})$$

Substituting the expressions for δ and σ back into the EGB2 distribution yields an EGB2 density function with zero mean and unit variance as:

$$\text{EGB2}(z; p, q) = \frac{\sqrt{\Omega} \exp\left(p \left(z + \frac{\Delta}{\sqrt{\Omega}}\right) \sqrt{\Omega}\right)}{B(p, q) \left(1 + \exp\left(\left(z + \frac{\Delta}{\sqrt{\Omega}}\right) \sqrt{\Omega}\right)\right)^{p+q}} \quad (\text{A10})$$

According to assumption (A1):

$$z_t = \frac{\varepsilon_t}{\sqrt{h_t}} \quad (\text{A11})$$

Changing the variable from z to ε as follows ($dz = d\varepsilon/h^{0.5}$):

$$\text{EGB2}(\varepsilon; h, p, q) = \frac{\sqrt{\Omega} \exp\left(p \left(\frac{\varepsilon}{\sqrt{h}} + \frac{\Delta}{\sqrt{\Omega}}\right) \sqrt{\Omega}\right)}{\sqrt{h} B(p, q) \left(1 + \exp\left(\left(\frac{\varepsilon}{\sqrt{h}} + \frac{\Delta}{\sqrt{\Omega}}\right) \sqrt{\Omega}\right)\right)^{p+q}} \quad (\text{A12})$$

Algebraic manipulation then yields:

$$\text{EGB2}(\varepsilon; h, p, q) = \frac{\sqrt{\Omega} \exp\left(p \left(\frac{\sqrt{\Omega}}{\sqrt{h}} \varepsilon + \Delta\right)\right)}{\sqrt{h} B(p, q) \left(1 + \exp\left(\frac{\sqrt{\Omega}}{\sqrt{h}} \varepsilon + \Delta\right)\right)^{p+q}} \quad (\text{A13})$$

ACKNOWLEDGEMENTS

We thank Basudeb Biswas, John Geweke, Terry Glover, Grant McQueen, and two anonymous referees for helpful comments, and Ron Schoenberg for programming advice. The remaining errors are ours.

REFERENCES

- Baillie RT, Bollerslev T. 1989. The message in daily exchange rate: a conditional-variance tale. *Journal of Business and Economic Statistics* **7**: 297–305.
- Baillie RT, DeGennaro RP. 1990. The impact of delivery terms on stock return volatility. *Journal of Financial Service Research* **3**: 55–76.
- Black F. 1976. Studies of stock market volatility changes. *1976 Proceedings of the American Statistical Association*, Business and Economic Statistics Section. 177–181.
- Bollerslev T. 1987. A conditional heteroskedastic time series model for speculative price and rate of return. *Review of Economics and Statistics* **9**: 542–547.
- Bollerslev T, Chou RY, Kroner KF. 1992. ARCH modeling in finance. *Journal of Econometrics* **52**: 5–59.
- Bollerslev T, Engle RF, Nelson DB. 1994. ARCH models. In *Handbook of Econometrics* (Vol. IV), Engle RF, McFadden DL (eds). Elsevier Science: Amsterdam.
- Boothe P, Glassman D. 1987. The statistical distribution of exchange rates. *Journal of International Economics* **22**: 297–319.
- Christie AA. 1982. The stochastic behavior of common stock variance: value, leverage and interest rates effects. *Journal of Financial Economics* **10**: 407–432.
- Davis HT. 1935. *Tables of the Higher Mathematical Functions* (Vols I and II). Principia Press: Bloomington, IND.
- Deb P. 1996. Finite sample properties of maximum likelihood and quasi-maximum likelihood estimators of EGARCH models. *Econometric Reviews* **15**: 51–68.
- Dornbusch R. 1976. Expectations and exchange rate dynamics. *Journal of Political Economy* **84**: 1161–1176.
- Engle RF. 1982. Autoregressive conditional heteroskedasticity with the estimates of the variance of U.K. inflation. *Econometrica* **50**: 987–1008.
- Exchange Rate Service of the Pacific Data Center at the University of British Columbia (<http://pacific.commerce.ubc.ca/xr/>).
- French KR, Schwert GW, Stambaugh RF. 1987. Expected stock returns and volatility. *Journal of Financial Economics* **19**: 3–29.
- Friedman M. 1953. The case for flexible exchange rates. *Essays in Positive Economics*. University of Chicago Press: Chicago IL.
- Greene W. 1997. *Econometric Analysis* (3rd edn). Prentice Hall: Englewood Cliffs, NJ.
- Hansen BE. 1994. Autoregressive conditional density estimation. *International Economic Review* **35**: 705–730.
- Hsieh DA. 1988. The statistical properties of daily foreign exchange rates: 1974–1983. *Journal of International Economics* **24**: 129–145.
- Hsieh DA. 1989. Modeling heteroskedasticity in daily foreign-exchange rates. *Journal of Business and Economic Statistics* **7**: 307–317.
- Huisman R, Koedijk K, Kool C, Palm F. 1998. The fat-tailedness of FX returns. Mimeo, Maastricht University.
- Kloek T, Van Dijk HK. 1978. Efficient estimation of income distribution parameters. *Journal of Econometrics* **8**: 856–860.
- Koedijk KG, Stork PA, de Vries CG. 1992. Differences between foreign exchange rate regimes: the view from the tails. *Journal of International Money and Finance* **11**: 462–473.
- Krugman PR. 1979. A model of balance of payments crises. *Journal of Money, Credit and Banking* **11**: 311–325.
- Lee S, Hansen BE. 1994. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory* **10**: 29–52.

- Loretan M, Phillips PCB. 1994. Testing the covariance stationarity of heavy-tailed time series: an overview of the theory with applications to several financial data series. *Journal of Empirical Finance* **21**: 248.
- MaCurdy TH, Morgan L. 1987. Tests of the martingale hypothesis for foreign currency futures with time varying volatility. *International Journal of Forecasting* **3**: 131–148.
- McDonald JB. 1984. Some generalized functions for the size distribution of income. *Econometrica* **52**(3): 647–663.
- McDonald JB. 1991. Parametric models for partially adaptive estimation with skewed and leptokurtic residuals. *Economics Letters* **37**: 273–8.
- McDonald JB, White SB. 1993. A comparison of some robust, adaptive, and partially adaptive estimators of regression models. *Econometric Reviews* **12**: 103–124.
- McDonald JB, Xu YJ. 1995. A generalization of the beta distribution with applications. *Journal of Econometrics* **66**: 133–152.
- Milhøj A. 1987. A conditional variance model for daily deviations of an exchange rate. *Journal of Business and Economic Statistics* **5**: 99–103.
- Nelson D. 1991. Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**(2): 347–370.
- Obstfeld M, Rogoff K. 1996. *Foundations of International Macroeconomics*. MIT Press: Cambridge MA.
- Pagan AR, Sabau H. 1987. On the inconsistency of the MLE in certain heteroskedasticity regression models. Mimeo, University of Rochester.
- Peruga R. 1988. The Distributional Properties of Exchange Rate Changes Under a Peso Problem. Ph D dissertation, University of California, San Diego, Dept of Economics.
- Taylor MP. 1995. The economics of exchange rates. *Journal of Economic Literature* **33**: 13–47.
- Wang KF, Barrett CB, Fawson C. 1996. Estimating East Asian exchange rates at different frequencies. Utah State University Economic Research Institute Study Paper #96-43, Logan, Utah.
- White H. 1982. Maximum likelihood estimation of misspecified models. *Econometrica* **50**: 1–25.