Econometrics Journal (2003), volume **6**, pp. 146–166.

Moments of the ARMA–EGARCH model

M. KARANASOS[1] AND J. KIM

University of York, Heslington, York, UK

Received: February 2003

Summary This paper considers the moment structure of the general ARMA–EGARCH model. In particular, we derive the autocorrelation function of any positive integer power of the squared errors. In addition, we obtain the autocorrelations of the squares of the observed process and cross correlations between the levels and the squares of the observed process. Finally, the practical implications of the results are illustrated empirically using daily data on four East Asia Stock Indices.

Keywords: *Autocorrelations, Exponential GARCH, Stock returns.*

1. INTRODUCTION

One of the principal empirical tools used to model volatility in asset markets has been the ARCH class of models. Following Engle's (1982) ground-breaking idea, several formulations of conditionally heteroscedastic models (e.g. GARCH, Fractional Integrated GARCH, Switching GARCH, Component GARCH) have been introduced in the literature (see, for example, the survey of Bollerslev *et al*. (1994)). These models form an immense ARCH family. Many of the proposed GARCH models include a term that can capture correlation between returns and conditional variance. Models with this feature are often termed asymmetric or leverage volatility models.[2] One of the earliest asymmetric GARCH models is the EGARCH (exponential generalized ARCH) model of Nelson (1991). In contrast to the conventional GARCH specification, which requires non-negative coefficients, the EGARCH model does not impose non-negativity constraints on the parameter space since it models the logarithm of the conditional variance.

Although the literature on the GARCH/EGARCH models is quite extensive, relatively few papers have examined the moment structure of models where the conditional volatility is time dependent. Karanasos (1999) and He and Teräsvirta (1999) derived the autocorrelations of the squared errors for the GARCH(p, q) model, while Karanasos (2001) obtained the autocorrelation function of the observed process for the ARMA-GARCH-in-mean model. Demos (2002) studied the autocorrelation structure of a model that nests both the EGARCH and stochastic volatility specifications. He *et al*. (2002) considered the moment structure of the EGARCH(1,1) model.

[1]Corresponding address: Department of Economics and Related Studies, University of York, York, YO10 5DD, UK. E-mail: `mk16@york.ac.uk`

[2]The asymmetric response of volatility to positive and negative shocks is well known in the finance literature as the leverage effect of stock market returns (Black, 1976). Researchers have found that volatility tends to rise in response to 'bad news' (excess returns lower than expected) and to fall in response to 'good news' (excess returns higher than expected).

This paper focuses solely on the moment structure of the general ARMA(r, s)–EGARCH(p, q) model. It would be useful to know the properties of the autocorrelation function of power-transformed observations when comparing the EGARCH model with the standard GARCH model. In particular, possible differences in the moment structure of these models may shed light on the success of the EGARCH model in applications.

We contribute to the aforementioned literature by deriving: (i) the autocorrelation function of any positive integer power of the squared errors, (ii) the cross correlations between the levels and the squares of both the observed process and the squared errors, and (iii) the autocorrelations of the squared observations. To obtain the theoretical results and to carry out the estimation, we assume that the innovations are drawn from either the normal, double exponential, or generalized error distributions. To facilitate model identification, the results for the autocorrelation function of the power-transformed errors can be applied so that properties of the observed data can be compared with the theoretical properties of the models.

The derivation of the autocorrelations of the fitted power-transformed values and their comparison with the corresponding sample equivalents will help the investigator (a) to decide which is the most appropriate method of estimation (e.g. maximum likelihood estimation (MLE), minimum distant estimator (MDE)) for a specific model, (b) to choose, for a given estimation technique, the model (e.g. asymmetric power ARCH (APARCH), EGARCH) that best replicates certain stylized facts of the data and, (c) in conjunction with the various model selection criteria, to identify the optimal order of the chosen specification.

This paper is organized as follows. Section 2.1 presents the ARMA(r, s)–EGARCH(p, q) process. Section 2.2 investigates the autocorrelation function of any positive integer power of the squared errors for the EGARCH model. Section 2.3 derives the cross correlations between the levels and the squares of the ARMA–EGARCH process. Section 2.4 provides the autocorrelation function of the squared observations. Section 3 discusses the data and presents the empirical results. In the conclusions we suggest future developments. Proofs are found in the appendices.

2. ARMA–EGARCH MODEL

2.1. ARMA(r, s)–EGARCH(p, q) process

Of the many different asymmetric GARCH specifications the EGARCH model has become one of the most common. Here we examine the general ARMA(r, s)–EGARCH(p, q) model. The stochastic process $\{y_t\}$ is assumed to be a causal ARMA(r, s) process satisfying

$$\Phi(L)y_t = b + \Theta(L)\varepsilon_t, \tag{2.1a}$$

where

$$\Phi(L) \equiv \prod_{l=1}^{r}(1 - \phi_l L), \tag{2.1b}$$

$$\Theta(L) \equiv 1 + \sum_{l=1}^{s}\theta_l L^l. \tag{2.1c}$$

Further, let $\{\varepsilon_t\}$ be a real-valued time stochastic process generated by

$$\varepsilon_t = e_t h_t^{\frac{1}{2}}, \tag{2.2}$$

where $\{e_t\}$ is a sequence of independent, identically distributed (*i.i.d.*) random variables with mean zero and variance 1. h_t is positive with probability one and is a measurable function of Σ_{t-1}, which in turn is the sigma-algebra generated by $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\}$. That is, h_t denotes the conditional variance of the errors $\{\varepsilon_t\}$, $(\varepsilon_t|\Sigma_{t-1}) \sim (0, h_t)$. As regards h_t, we assume that it follows an EGARCH(p, q) process

$$B(L)\ln(h_t) = \omega + C(L)z_t, \tag{2.3a}$$

$$z_{t-l} \equiv d\frac{\varepsilon_{t-l}}{\sqrt{h_{t-l}}} + \gamma\left[\left|\frac{\varepsilon_{t-l}}{\sqrt{h_{t-l}}}\right| - \mathsf{E}\left|\frac{\varepsilon_{t-l}}{\sqrt{h_{t-l}}}\right|\right], \tag{2.3b}$$

where

$$C(L) \equiv \sum_{l=1}^{q} c_l L^l, \tag{2.3c}$$

$$B(L) \equiv \prod_{l=1}^{p}(1 - \beta_l L). \tag{2.3d}$$

Various cases of the EGARCH(p, q) model have been applied by researchers. More specifically, Donaldson and Kamstra (1997) found that the optimal EGARCH specification for the NIKKEI stock index was a flexible 3,2. Hu *et al*. (1997) found that in the pre-EMS period, the majority of the European currencies followed an AR(5)–EGARCH(4,4) model.

2.2. Higher-order moments of the squared errors

Although the EGARCH model was introduced over a decade ago and has been widely used in empirical applications, its statistical properties have only recently been examined. Engle and Ng (1993) artificially nested the GARCH and EGARCH models, estimated this nested specification, and then applied likelihood ratio (LR) tests (see also Hu *et al*. (1997)). Hentschel (1995) developed a family of asymmetric GARCH models that nests both the APARCH model and the EGARCH model.

In this section we focus solely on the moment structure of the general EGARCH(p, q) model.

Assumption 1. *All the roots of the autoregressive polynomial $B(L)$ lie outside the unit circle (covariance–stationarity condition).*

Assumption 2. *The polynomials $C(L)$ and $B(L)$ in* (2.3c) *and* (2.3d) *respectively, have no common left factors other than unimodular ones (irreducibility condition). Moreover, $\beta_p, c_q \neq 0$.*

In what follows we examine only the case where all the roots of the autoregressive polynomial $B(L)$ in (2.3d) are distinct. The following proposition establishes the lag-m autocorrelation of $\{\varepsilon_t^{2k}\}$, the kth power of the squared errors: $\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k}) \equiv \text{Corr}\,(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k})$.

Proposition 1. *Let Assumptions* 1 *and* 2 *hold. Suppose further that* $\mathsf{E}(e_t^{4k}) < \infty$, $\mathsf{E}[\exp(2kz_t)] < \infty$ *and* $\mathsf{E}[e_t^{2k}\exp(kz_t)] < \infty\ \forall\, t$, *for any finite positive scalar k. Then the autocorrelation of the* *kth power of the squared error* $\{\varepsilon_t^{2k}\}$*, at lag m* ($m \in \mathsf{N}$)*, has the form*

$$\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k}) = \mathsf{E}(e_t^{2k}) \left\{ \prod_{i=0}^{m-2} \left[\mathsf{E}\left(\exp\left(\xi_{0,\frac{k}{2},i} z_{t-i-1}\right)\right)\right] \mathsf{E}\left[e_{t-m}^{2k} \exp\left(\xi_{0,\frac{k}{2},m-1} z_{t-m}\right)\right] \right.$$
$$\times \prod_{i=0}^{\infty} [\mathsf{E}(\exp(\xi_{m,k,i} z_{t-m-i-1}))] - \mathsf{E}(e_t^{2k}) \left. \left[\prod_{i=0}^{\infty} \left[\mathsf{E}\left(\exp\left(\xi_{0,\frac{k}{2},i} z_{t-i-1}\right)\right)\right] \right]^2 \right\}$$
$$\times \left\{ \mathsf{E}(e_t^{4k}) \prod_{i=0}^{\infty} [\mathsf{E}(\exp(\xi_{0,k,i} z_{t-i-1}))] - [\mathsf{E}(e_t^{2k})]^2 \right.$$
$$\left. \times \left[\prod_{i=0}^{\infty} \left[\mathsf{E}\left(\exp\left(\xi_{0,\frac{k}{2},i} z_{t-i-1}\right)\right)\right] \right]^2 \right\}^{-1}, \tag{2.4a}$$

where

$$\xi_{m,k,i} \equiv k \sum_{f=1}^{p} \zeta_f (\lambda_{f,m+i+1} + \lambda_{f,i+1}), \tag{2.4b}$$

with

$$\lambda_{fi} \equiv \begin{cases} \sum_{n=0}^{i-1} c_{i-n} \beta_f^n, & \text{if } i \le q, \\ \lambda_{fq} \beta_f^{i-q}, & \text{if } i > q, \end{cases} \tag{2.4c}$$

$$\zeta_f \equiv \frac{\beta_f^{p-1}}{\prod_{\substack{n \ne f \\ n=1}}^{p} (\beta_f - \beta_n)}. \tag{2.4d}$$

Note that, when $m = 1$*, the first product term in the right-hand side of* (2.4a) *is replaced by* 1.

Proof. See Appendix A. □

He *et al*. (2002) derived the autocorrelations of positive powers of the absolute-valued errors of the EGARCH(1,1) model.

In the following theorem we provide the autocorrelations of the kth power of the squared errors $\{\varepsilon_t^{2k}\}$ of the EGARCH(p, q) model.

Theorem 1. *Let k be any finite positive integer. Then, when the distribution of* $\{e_t\}$ *is generalized error, the second moment and the autocorrelation function of* $\{\varepsilon_t^{2k}\}$ *are given by*

$$\mathsf{E}(\varepsilon_t^{4k}) = \exp\left(\frac{2k\left[\omega - \gamma \frac{\Gamma\left(\frac{2}{v}\right)\lambda 2^{\frac{1}{v}}}{\Gamma\left(\frac{1}{v}\right)} \sum_{l=1}^{q} c_l \right]}{\prod_{f=1}^{p}(1-\beta_f)} \right) \mu_{4k}^{(g)} B_{0,k}^{(g)}, \tag{2.5a}$$

$$\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k}) = \frac{\mu_{2k}^{(g)}\left[A_{m-1,k}^{(g)} D_{m-1,k}^{(g)} B_{m,k}^{(g)} - \mu_{2k}^{(g)}\left(B_{0,\frac{k}{2}}^{(g)}\right)^2\right]}{\left[\mu_{4k}^{(g)} B_{0,k}^{(g)} - \left(\mu_{2k}^{(g)} B_{0,\frac{k}{2}}^{(g)}\right)^2\right]}, \qquad (A_{0,k}^{(g)} \equiv 1), \quad (2.5b)$$

where

$$A_{m,k}^{(g)} \equiv \prod_{i=0}^{m-1}\left\{\sum_{\tau=0}^{\infty}\left[2^{\frac{1}{v}}\lambda\xi_{0,\frac{k}{2},i}\right]^{\tau}\left[(\gamma+d)^{\tau}+(\gamma-d)^{\tau}\right]\frac{\Gamma\left(\frac{1+\tau}{v}\right)}{2\Gamma\left(\frac{1}{v}\right)\Gamma(1+\tau)}\right\}, \qquad (2.5c)$$

$$B_{m,k}^{(g)} \equiv \prod_{i=0}^{\infty}\left\{\sum_{\tau=0}^{\infty}\left[2^{\frac{1}{v}}\lambda\xi_{m,k,i}\right]^{\tau}\left[(\gamma+d)^{\tau}+(\gamma-d)^{\tau}\right]\frac{\Gamma\left(\frac{1+\tau}{v}\right)}{2\Gamma\left(\frac{1}{v}\right)\Gamma(1+\tau)}\right\}, \qquad (2.5d)$$

and

$$D_{m,k}^{(g)} \equiv 2^{\frac{2k}{v}}\lambda^{2k}\sum_{\tau=0}^{\infty}\left(\lambda 2^{\frac{1}{v}}\xi_{0,\frac{k}{2},m}\right)^{\tau}\left[(\gamma+d)^{\tau}+(\gamma-d)^{\tau}\right]\frac{\Gamma\left(\frac{\tau+2k+1}{v}\right)}{2\Gamma\left(\frac{1}{v}\right)\Gamma(1+\tau)}, \qquad (2.5e)$$

$$\mu_{2k}^{(g)} \equiv \frac{\left[\Gamma\left(\frac{1}{v}\right)\right]^{k-1}\Gamma\left(\frac{2k+1}{v}\right)}{\left[\Gamma\left(\frac{3}{v}\right)\right]^{k}}, \qquad (2.5f)$$

with

$$\lambda \equiv \left\{2^{\frac{-2}{v}}\Gamma\left(\frac{1}{v}\right)\left[\Gamma\left(\frac{3}{v}\right)\right]^{-1}\right\}^{\frac{1}{2}},$$

where $\xi_{m,k,i}$ is defined in Proposition 1, *v are the degrees of freedom of the generalized error distribution and $\Gamma(\cdot)$ is the Gamma function. When $v > 1$, the summations in* (2.5c)–(2.5e) *are finite; when $v < 1$, the three summations are finite if and only if $\xi_{0,\frac{k}{2},i}\gamma + |\xi_{0,\frac{k}{2},i}d| \leq 0$, $\xi_{m,k,i}\gamma + |\xi_{m,k,i}d| \leq 0$ and $\xi_{0,\frac{k}{2},m}\gamma + |\xi_{0,\frac{k}{2},m}d| \leq 0$, respectively* (*see Nelson,* (1991)).

Proof. See Appendix A. □

One of the most widely used models in financial economics to describe a time series, r_t, of the returns from some asset is the martingale process

$$r_t \equiv \varepsilon_t = e_t\sqrt{h_t},$$

where e_t is *i.i.d.* (0, 1) and h_t is a GARCH type process.

In many applications in financial economics, it is not reasonable to assume the normality of e_t, because of the substantial excess kurtosis present in the conditional density of returns. Hence investigators often use MLE assuming some fat-tailed conditional density such as generalized error. Therefore, when it comes to model identification, practitioners in this area may find the results in Theorem 1 quite useful.

In the following proposition, when the innovations are drawn from the double exponential distribution, we provide the autocorrelations of $\{\varepsilon_t^{2k}\}$ for any finite positive integer k.

Proposition 2. *When the distribution of* $\{e_t\}$ *is double exponential, the autocorrelation function of the kth power of the squared errors is given by*

$$\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k}) = \frac{\mu_{2k}^{(d)}\left[A_{m-1,k}^{(d)} D_{m-1,k}^{(d)} B_{m,k}^{(d)} - \mu_{2k}^{(d)}\left(B_{0,\frac{k}{2}}^{(d)}\right)^2\right]}{\left[\mu_{4k}^{(d)} B_{0,k}^{(d)} - \left(\mu_{2k}^{(d)} B_{0,\frac{k}{2}}^{(d)}\right)^2\right]}, \qquad (A_{0,k}^{(d)} \equiv 1), \tag{2.6a}$$

with

$$A_{m,k}^{(d)} \equiv \prod_{i=0}^{m-1} \frac{2 - \sqrt{2}\xi_{0,\frac{k}{2},i}\gamma}{2 - \sqrt{2}\xi_{0,k,i}\gamma + \xi_{0,\frac{k}{2},i}^2(\gamma^2 - d^2)}, \tag{2.6b}$$

$$B_{m,k}^{(d)} \equiv \prod_{i=0}^{\infty} \frac{2 - \sqrt{2}\xi_{m,k,i}\gamma}{2 - 2\sqrt{2}\xi_{m,k,i}\gamma + \xi_{m,k,i}^2(\gamma^2 - d^2)}, \tag{2.6c}$$

and

$$D_{m,k}^{(d)} \equiv 2^{-(k+1)}\Gamma(2k+1) \times \left\{F\left[2k+1; \frac{\xi_{0,\frac{k}{2},m}(\gamma+d)}{\sqrt{2}}\right] + F\left[2k+1; \frac{\xi_{0,\frac{k}{2},m}(\gamma-d)}{\sqrt{2}}\right]\right\}, \tag{2.6d}$$

$$\mu_{2k}^{(d)} \equiv \frac{\Gamma(2k+1)}{2^k}, \tag{2.6e}$$

where $F(\cdot)$ *is the hypergeometric function* (*see Abadir*, (1999)) *and* $\xi_{m,k,i}$ *is defined in Proposition* 1. *Expressions* (2.6b) *and* (2.6c) *hold if and only if* $\xi_{0,\frac{k}{2},i}\gamma + |\xi_{0,\frac{k}{2},i}d| < \sqrt{2}$ *and* $\xi_{m,k,i}\gamma + |\xi_{m,k,i}d| < \sqrt{2}$, *respectively; the right-hand side of* (2.6d) *converges if and only if* $\xi_{0,\frac{k}{2},m}\gamma + |\xi_{0,\frac{k}{2},m}d| < \sqrt{2}$ (*see Nelson*, (1991)).

Proof. See Appendix A. □

Also note that the coefficients of the Wold representation of the kth power of the conditional variance are needed for the computation of the autocorrelations of the kth power of the squared errors (see Demos (2002)). In Appendix A (see equation (A.1)) we provide exact form solutions which express the above coefficients in terms of the parameters of the moving average polynomial and the roots of the autoregressive polynomial.

In the following proposition, when the errors are conditionally normal, we derive the autocorrelation function of the kth power of the squared errors $\{\varepsilon_t^{2k}\}$, $k \in \mathsf{N}$.

Proposition 3. *When the distribution of* $\{e_t\}$ *is normal, the autocorrelations of the kth power of the squared errors are given by*

$$\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k}) = \frac{\mu_{2k}^{(n)}\left[A_{m-1,k}^{(n)} D_{m-1,k}^{(n)} B_{m,k}^{(n)} - \mu_{2k}^{(n)}\left(B_{0,\frac{k}{2}}^{(n)}\right)^2\right]}{\left[\mu_{4k}^{(n)} B_{0,k}^{(n)} - \left(\mu_{2k}^{(n)} B_{0,\frac{k}{2}}^{(n)}\right)^2\right]}, \qquad (A_{0,k}^{(n)} \equiv 1), \tag{2.7a}$$

with

$$A_{m,k}^{(n)} \equiv \prod_{i=0}^{m-1} \left\{ \exp\left(\frac{(\gamma+d)^2 \xi_{0,\frac{k}{2},i}^2}{2} \right) \frac{1}{2}\Big[1+\exp\Big(-2\gamma d \xi_{0,\frac{k}{2},i}^2\Big)\Big] + \frac{(\gamma+d)\xi_{0,\frac{k}{2},i}}{\sqrt{2\pi}} \right.$$
$$\left. \times F\left(1; \frac{3}{2}; \frac{(\gamma+d)^2 \xi_{0,\frac{k}{2},i}^2}{2}\right) + \frac{(\gamma-d)\xi_{0,\frac{k}{2},i}}{\sqrt{2\pi}} \times F\left(1; \frac{3}{2}; \frac{(\gamma-d)^2 \xi_{0,\frac{k}{2},i}^2}{2}\right) \right\}, \tag{2.7b}$$

$$B_{m,k}^{(n)} \equiv \prod_{i=0}^{\infty} \left\{ \exp\left(\frac{(\gamma+d)^2 \xi_{m,k,i}^2}{2} \right) \frac{1}{2}[1+\exp(-2\gamma d \xi_{m,k,i}^2)] + \frac{(\gamma+d)\xi_{m,k,i}}{\sqrt{2\pi}} \right.$$
$$\left. \times F\left(1; \frac{3}{2}; \frac{(\gamma+d)^2 \xi_{m,k,i}^2}{2}\right) + \frac{(\gamma-d)\xi_{m,k,i}}{\sqrt{2\pi}} \times F\left(1; \frac{3}{2}; \frac{(\gamma-d)^2 \xi_{m,k,i}^2}{2}\right) \right\}, \tag{2.7c}$$

and

$$D_{m,k}^{(n)} \equiv \frac{1}{2} \left\{ \frac{\partial}{\partial \left[\xi_{0,\frac{k}{2},m}(\gamma-d)\right]^{2k}} \left\{ \exp\left(\frac{\xi_{0,\frac{k}{2},m}^2 (\gamma-d)^2}{2} \right) \left[1+\Phi\left(\frac{\xi_{0,\frac{k}{2},m}(\gamma-d)}{\sqrt{2}} \right) \right] \right\} \right.$$
$$\left. + \frac{\partial}{\partial \left[\xi_{0,\frac{k}{2},m}(\gamma+d)\right]^{2k}} \left\{ \exp\left(\frac{\xi_{0,\frac{k}{2},m}^2 (\gamma+d)^2}{2} \right) \left[1+\Phi\left(\frac{\xi_{0,\frac{k}{2},m}(\gamma+d)}{\sqrt{2}} \right) \right] \right\} \right\}, \tag{2.7d}$$

$$\mu_{2k}^{(n)} \equiv \prod_{j=1}^{k} [2k-(2j-1)], \tag{2.7e}$$

where ∂ *denotes partial derivative,* $\Phi(\cdot)$ *is the error function of the standard normal distribution and* $\xi_{m,k,i}$ *is given by* (2.4b).

Proof. See Appendix A. □

Several previous articles dealing with financial market data—e.g. Dacorogna *et al*. (1993), Ding *et al*. (1993) and Muller *et al*. (1997)—have commented on the behavior of the autocorrelation function of positive powers of the squared returns, and the desirability of having a model which comes close to replicating certain stylized facts in the data (abstracted from Baillie and Chung (2001)). In this respect, one can apply the results in this section to check whether the EGARCH model can effectively replicate the observed pattern of autocorrelations of power-transformed returns.

2.3. Dynamic asymmetry

In this section we examine the cross correlations between the levels and the squares of the ARMA–EGARCH process in (2.1)–(2.3).

Proposition 4. *Let the distribution of* $\{e_t\}$ *be generalized error and* k *a finite positive integer. Suppose further that* $\mathsf{E}(e_t^{4k}) < \infty$, $\mathsf{E}[e_t^{2k-1}\exp(kz_t)] < \infty$ *and* $\mathsf{E}[\exp(2kz_t)] < \infty \; \forall \, t$*. Then, if Assumptions* 1 *and* 2 *hold, the cross correlations between the* $2k$*th and* $(2k-1)$*th powers of* $\{\varepsilon_t\}$ *are given by*

$$\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k-1}) = \frac{\mu_{2k}^{(g)} A_{m-1,k}^{(g)} D_{m-1,(k)}^{(g)} B_{m,(k)}^{(g)}}{\sqrt{\left[\mu_{4k}^{(g)} B_{0,k}^{(g)} - (B_{0,\frac{k}{2}}^{(g)})^2\right] \mu_{4k-2}^{(g)} B_{0,(\frac{4k-1}{4})}^{(g)}}}, \qquad (m \in \mathsf{N}), \tag{2.8a}$$

with

$$D_{m,(k)}^{(g)} \equiv 2^{\frac{2k-1}{v}} \lambda^{2k-1} \sum_{\tau=0}^{\infty} \left(\lambda 2^{\frac{1}{v}} \varphi_{0,\frac{2k+1}{4},m}\right)^{\tau} \left[(\gamma+d)^{\tau} - (\gamma-d)^{\tau}\right] \frac{\Gamma\left(\frac{\tau+2k}{v}\right)}{2\Gamma\left(\frac{1}{v}\right)\Gamma(1+\tau)}, \tag{2.8b}$$

$$B_{m,(k)}^{(g)} \equiv \prod_{i=0}^{\infty} \left\{ \sum_{\tau=0}^{\infty} \left[2^{\frac{1}{v}} \lambda \varphi_{m,k,i}\right]^{\tau} [(\gamma+d)^{\tau} + (\gamma-d)^{\tau}] \frac{\Gamma\left(\frac{1+\tau}{v}\right)}{2\Gamma\left(\frac{1}{v}\right)\Gamma(1+\tau)} \right\}, \tag{2.8c}$$

and

$$\varphi_{m,k,i} \equiv \sum_{f=1}^{p} (k\lambda_{f,m+i+1} + (k-0.5)\lambda_{f,i+1})\zeta_f,$$

where $A_{m,k}^{(g)}$, $B_{m,k}^{(g)}$ *and* $\mu_{4k}^{(g)}$ *are defined in Theorem* 1*. Note that, when* m *is a negative integer,* $\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k-1}) = 0$.

The proof of Proposition 4 is similar to that of Theorem 1.

Demos (2002) derived the cross correlations between the levels and the squares of an observed series, under the assumption that the mean parameter is time varying and the conditional variance follows a flexible parameterization which nests the autoregressive stochastic volatility and the exponential GARCH specifications.[3] In the same spirit, the following theorem obtains the cross correlations between the levels and the squares of the ARMA(r, s)–EGARCH(p, q) process in (2.1)–(2.3).

Assumption 3. *All the roots of the autoregressive polynomial* $\Phi(L)$ *lie outside the unit circle.*

Assumption 4. *The polynomials* $\Phi(L)$ *and* $\Theta(L)$ *are left coprime.*

Theorem 2. *Let Assumptions* 1–4 *hold. Suppose further that* $\mathsf{E}(e_t^4) < \infty$, $\mathsf{E}[e_t \exp(z_t)] < \infty$ *and* $\mathsf{E}[\exp(2z_t)] < \infty \; \forall \, t$*. Then the cross correlations between the squares and the levels of* $\{y_t\}$ *are given by*

$$\rho(y_t^2, y_{t-m}) = \left(\frac{\eta-1}{\kappa-1}\right)^{\frac{1}{2}} (F_m + H_m), \qquad (m \in \mathsf{N}), \tag{2.9a}$$

[3] Demos called this model time varying parameter generalized stochastic volatility in mean (TVP-GSV-M).

where

$$F_m \equiv \frac{\sum_{i=0}^{j+m-1} \sum_{j=0}^{\infty} \delta_i^2 \delta_j \rho\left(\varepsilon_t^2, \varepsilon_{t-(j+m-i)}\right)}{\left(\sum_{l=0}^{\infty} \delta_l^2\right)^2}, \tag{2.9b}$$

$$H_m \equiv \frac{2\sum_{i=j+m+1}^{\infty} \sum_{j=0}^{\infty} \delta_i \delta_j \delta_{j+m} \rho\left(\varepsilon_t^2, \varepsilon_{t-(i-j-m)}\right)}{\left(\sum_{l=0}^{\infty} \delta_l^2\right)^2}. \tag{2.9c}$$

Furthermore

$$\rho(y_t, y_{t-m}^2) = \left(\frac{\eta-1}{\kappa-1}\right)^{\frac{1}{2}} (K_m + L_m), \qquad (m \in \mathsf{N}), \tag{2.10a}$$

with

$$K_m \equiv \frac{\sum_{i=j+m+1}^{\infty} \sum_{j=0}^{\infty} \delta_i \delta_j^2 \rho\left(\varepsilon_t^2, \varepsilon_{t-(i-j-m)}\right)}{\left(\sum_{l=0}^{\infty} \delta_l^2\right)^2}, \tag{2.10b}$$

$$L_m \equiv \frac{2\sum_{i=j+1}^{\infty} \sum_{j=0}^{\infty} \delta_i \delta_j \delta_{j+m} \rho\left(\varepsilon_t^2, \varepsilon_{t-(i-j)}\right)}{\left(\sum_{l=0}^{\infty} \delta_l^2\right)^2}, \tag{2.10c}$$

where η and κ denote the kurtosis of ε_t and y_t respectively; δ_i is the ith coefficient in the Wold representation of the ARMA(p, q) process in (2.1) (*see equation* (B.2a)). *Note that when the distribution of $\{e_t\}$ is generalized error, $\rho(\varepsilon_t^2, \varepsilon_{t-m})$ is given in Proposition* 4, *and η, κ are given in Proposition* 5 *below.*

Proof. See Appendix B. □

Also observe that when there is no leverage effect ($d = 0$), the $D_{m,(k)}^{(g)}$ term in (2.8b) is zero and hence the cross correlations between the levels and the squares of both the errors and the observed process are zero. Demos (2002), for the TVP-GSV-M model, does not need the asymmetric EGARCH effect to obtain dynamic asymmetry even under the assumption of conditional normality.

2.4. Autocorrelations of the squared observations

In this section, we establish the autocorrelation properties of the squares of the ARMA–EGARCH process in (2.1)–(2.3). Demos (2002) obtained the autocorrelation function of the squares of the observed series for the TVP-GSV-M model.

The result presented in the following proposition, which is a special case of Theorem 2 in Palma and Zevallos (2001), is highly relevant since it helps to identify the nature of the process. By analyzing the autocorrelation function of the squared series it is possible to discard those theoretical models which are incompatible with the data under study.

Proposition 5. *Let Assumptions* 1–4 *hold. Suppose further that* $\mathsf{E}(e_t^4) < \infty$, $\mathsf{E}[\exp(2z_t)] < \infty$ *and* $\mathsf{E}[e_t^2 \exp(z_t)] < \infty\ \forall\ t$. *Then, when the distribution of $\{\varepsilon_t\}$ is generalized error, the autocorrelation function of $\{y_t^2\}$ is given by*

$$\rho(y_t^2, y_{t-m}^2) = \frac{2[\rho(y_t, y_{t-m})]^2}{\kappa-1} + \frac{(\kappa-3)\Gamma_m}{(\kappa-1)} + \frac{\eta-1}{\kappa-1}[G_m + 2\Delta_m - 3\Delta_0\Gamma_m], \qquad (m \in \mathsf{N}), \tag{2.11a}$$

where

$$\Gamma_m \equiv \frac{\sum_{i=0}^{\infty} \delta_i^2 \delta_{i+m}^2}{\sum_{l=0}^{\infty} \delta_l^4}, \tag{2.11b}$$

$$\Delta_m \equiv \frac{\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \delta_i \delta_j \delta_{i+m} \delta_{j+m} \rho\left(\varepsilon_t^2, \varepsilon_{t-(i-j)}^2\right)}{\left(\sum_{l=0}^{\infty} \delta_l^2\right)^2}, \tag{2.11c}$$

$$G_m \equiv \frac{\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \delta_i^2 \delta_j^2 \rho\left(\varepsilon_t^2, \varepsilon_{t-(m+j-i)}^2\right)}{\left(\sum_{l=0}^{\infty} \delta_l^2\right)^2}, \tag{2.11d}$$

and

$$\kappa \equiv 3 - \frac{2\eta \left(\sum_{i=0}^{\infty} \delta_i^4\right)}{\left(\sum_{l=0}^{\infty} \delta_l^2\right)^2} + 3(\eta - 1)\Delta_0, \tag{2.11e}$$

$$\eta \equiv \frac{\mathsf{E}\left(\varepsilon_t^4\right)}{\left[\mathsf{E}\left(\varepsilon_t^2\right)\right]^2}, \tag{2.11f}$$

where $\rho\left(\varepsilon_t^2, \varepsilon_{t-(i-j)}^2\right)$ *and* $\mathsf{E}(\varepsilon_t^4)$ *are given in Theorem* 1*;* δ_i *is defined in Theorem* 2 *and* $\rho(y_t, y_{t-m})$ *is the lag-m autocorrelation of* $\{y_t\}$.

The significance of the above result is that it allows us to establish whether the ARMA–EGARCH model in (2.1)–(2.3) is capable of reproducing key features exhibited by the data. These features include, for example, time series with very little autocorrelation but with strongly dependent squares. Another potential motivation for the derivation of the results in Theorem 2 is that the autocorrelations of the squared process in (2.1) can be used to estimate the ARMA and GARCH parameters in (2.1) and (2.3) respectively. The approach is to use the MDE, which estimates the parameters by minimizing the mahalanobis generalized distance of a vector of sample autocorrelations from the corresponding population autocorrelations (see Baillie and Chung (2001)).

The following proposition provides the lag-m autocorrelation of $\{h_t^k\}$, $k \in \mathsf{R}_+$.

Proposition 6. *Let Assumptions* 1 *and* 2 *hold. Suppose further that* $\mathsf{E}[\exp(2kz_t)] < \infty\ \forall\, t$. *Then, when the distribution of* $\{e_t\}$ *is generalized error, the autocorrelation function of the kth power of the conditional variance is given by* (2.5b) *where now the terms* $D^{(g)}$ *and* $\mu^{(g)}$ *are replaced by* 1*, and* $A_{m-1,k}^{(g)}$ *is replaced by* $A_{m,k}^{(g)}$.

Proof. See Appendix B. □

Demos (2002) derived the autocorrelations of the kth power of the conditional variance for the GSV model under the assumption of conditional normality.

Next, consider a process y_t governed by

$$y_t = \mathsf{E}(y_t \mid \Sigma_{t-1}) + \varepsilon_t.$$

Further, suppose that the conditional mean of y_t, given information through time $t-1$, is

$$\mathsf{E}(y_t \mid \Sigma_{t-1}) = \delta h_t^k, \qquad (k > 0).$$

The results in Proposition 6 can be used to derive the autocorrelation function of the above process. Mean equations of this form have been widely used in empirical studies of time-varying risk premia. Demos (2002), in the TVP-GSV-M model, allowed the conditional variance to affect the mean with a possibly time varying coefficient.

3. EMPIRICAL RESULTS

3.1. Data

We use four daily stock indices: the Korean stock price index (KOSPI), the Japanese Nikkei index (NIKKEI) and the Taiwanese Se weighted index (SE) for the period 1980:01–1997:04, and the Singaporean Straits Times price index (ST) for the period 1985:01–1997:04. The daily observations for each country are extracted from the 'Datastream' database. In each case, the index return is the first difference of log prices without dividends.

3.2. Estimation results

In order to carry out our analysis of stock returns, we have to select a form for the mean equation. Scholes and Williams (1977), Ding *et al*. (1993), and Ding and Granger (1996) suggested an MA(1) specification for the mean. We therefore model the stock returns as MA(1) processes. The MA(1) model is

$$y_t = b + (1 + \theta L)\varepsilon_t. \tag{3.1}$$

To select our 'best' EGARCH specification,[4] we begin with high order models (e.g., EGARCH(4,4)) and follow a 'general to specific' modelling approach to fit the data. The general EGARCH(p, q) specification that we estimate is

$$\varepsilon_t = e_t h_t^{\frac{1}{2}}, \qquad e_t \sim i.i.d.\,(0, 1), \tag{3.2a}$$

$$\prod_{i=1}^{p}(1 - \beta_i L)\ln(h_t) = \omega + \sum_{i=1}^{q} c_i(|e_{t-i}| + d_i e_{t-i}). \tag{3.2b}$$

We estimate EGARCH models of order up to 4,4 for the returns on the four stock indices using three alternative distributions: the normal, double exponential and generalized error. The Akaike information criterion (AIC) (not reported here) chose high order EGARCH specifications for all indices.[5] In addition, we use the LR test to show the performance of the high order models over the EGARCH(1,1) model. The tests (not reported here) show the dominance of the high order models.

For all the stock returns, parameter estimation is conducted jointly on an MA(1) mean specification[6] and the appropriate EGARCH model for the conditional variance. Table 1 reports the results for the period 1980–1997 and presents parameter estimates along with t-statistics.

[4] We define 'best' as the specification chosen by the AIC.

[5] In contrast, in most of the cases, the Schwarz information criterion (SIC) (not reported here) chose the EGARCH(1,1) model.

[6] The only exceptions are the Taiwanese Se weighted index, where the white noise specification is used for the generalized error and double exponential distributions, and the Japanese index where the white noise specification is used when the errors are drawn from the double exponential distribution.

For two out of the four indices the AIC is minimized when the double exponential distribution is used, while for the KOSPI and ST indices, it chooses the generalized error distribution. The parameters b and θ are the intercept and MA(1) coefficient respectively for the return equation (3.1). The remaining parameters are from the EGARCH model (3.2b). Not surprisingly, for all the EGARCH specifications, most of the moving average, leverage and autoregressive parameters are significantly different from zero. The estimated values of degrees of freedom in the generalized error distribution, for the KOSPI and ST indices, are 1.07 and 1.13 respectively. Researchers have found that nontrading periods contribute much less than do trading periods to market variance (see Nelson (1991)). Therefore, the selected specifications reported in Table 1 have been reestimated taking into account the number of nontrading days between day t and $t-1$. That is ω in (3.2b) is replaced by $\omega_t = \omega + \ln(1 + \delta N_t)$. In all cases the estimated δ's were statistically significant and less than unity (the results for these cases are not reported here).

For all four indices, parameter estimates are consistent with those generally reported in the literature. The roots of the autoregressive parts of the conditional variances are reported in the first column of Table 2. In particular, for the KOSPI index volatility appears nearly integrated (the values of the two complex roots are $-0.72 \pm 0.68i$). For the ST index there is one positive and one negative root with values 0.91 and -0.29 respectively. Fiorentini and Sentana (1998) used a measure of persistence of shocks for stationary processes based on the impulse response function, which captures the importance of the deviations of a series from its unperturbed path following a single shock. Accordingly, the persistence of a shock to z_t on $\ln(h_t)$ is $P_\infty[\ln(h_t)|z_t] = \sum_{l=1}^{\infty}\sum_{f=1}^{p}(\zeta_f\lambda_{fl})^2$, where ζ_f and λ_{fl} are defined in Proposition 1. The algebra of this measure is simple and its interpretation straightforward since it is the ratio of the variance of $\ln(h_t)$ to the variance of z_t. The second column of Table 2 reports a measure for the persistence of a (positive) shock to e_t on $\ln(h_t)$. Most noteworthy is the observation that in all EGARCH models, the product of the moving average parameter and the leverage coefficient for the first lagged error is negative (see column 3, Table 2). In addition, the sum of these products, over all the lagged errors, is also negative (see column 4, Table 2).

3.3. Autocorrelation structure of the estimated models

For each of the four indices, Figure 1 plots the estimated theoretical autocorrelations[7] of the squared observations of the 'best' EGARCH model. Specifically, for Korea and Singapore we use the EGARCH(3,3) and EGARCH(2,1) specifications respectively, with innovations that are drawn from the generalized error distribution. Further, for Japan and Taiwan we use the EGARCH(1,3) process with the double exponential distribution. Note that all the 'best' EGARCH models have been estimated without the inclusion of the no-trade dummy (see Figure 1).

The estimated autoregressive coefficient for the SE index is 0.983. As a result the estimated autocorrelations of the squared observations start at 0.1357 for lag one and decrease very slowly. Observe that the autocorrelation at lag 10, 20 and 30 is 0.1068, 0.0824 and 0.0646 respectively. As with the SE index the 'best' EGARCH model for the NIKKEI index is of order 1,3 and has errors that are drawn from the double exponential distribution, but the estimated autoregressive coefficient is lower (0.973). Thus, although the estimated autocorrelations start at 0.1654 for lag one they decrease more rapidly. The autocorrelation at lag 10, 20 and 30 is 0.0767, 0.0499

[7]We used Maple to evaluate the autocorrelations. The codes are available from the authors on request.

Table 1. Parameter estimates for the 'best' EGARCH model.

	KOSPI MA(1)-EGARCH(3,3) (GEN ERROR)	NIKKEI WN-EGARCH(1,3) (DOUBLE EXP)	SE WN-EGARCH(1,3) (DOUBLE EXP)	ST MA(1)-EGARCH(2,1) (GEN ERROR)
b	−0.0002 (1.32)	6E−10 (0.00)	7E−08 (0.00)	0.0002 (1.15)
θ	0.059 (4.57)	—	—	0.200 (11.99)
ω	−2.180 (9.88)	−0.384 (7.07)	−0.273 (6.85)	−1.294 (7.16)
c_1	0.258 (12.93)	0.193 (5.38)	0.145 (3.82)	0.349 (11.05)
c_2	0.374 (12.90)	0.149 (2.95)	0.105 (2.06)	—
c_3	0.254 (12.66)	−0.146 (3.37)	−0.065 (1.65)	—
β_1^*	−0.508 (84.56)	0.973 (206.00)	0.983 (260.60)	0.620 (5.89)
β_2^*	0.387 (40.51)	—	—	0.267 (2.67)
β_3^*	0.949 (157.92)	—	—	—
d_1	−0.124 (2.31)	−1.000 (3.85)	−0.713 (2.88)	−0.196 (3.29)
d_2	−0.099 (1.87)	0.050 (1.84)	0.118 (3.18)	—
d_3	−0.122 (2.28)	−0.637 (2.29)	−1.000 (1.24)	—
υ	1.076 {0.02}	1	1	1.134 {0.02}
AIC	−28073	−30065	−25613	−20859
LL	14049	15041	12815	10438

For each of the four stock indices, Table 1 reports parameter estimates for the 'best' EGARCH model. The general MA(1)–EGARCH(3,3) model is

$$y_t = b + (1 + \theta L)\varepsilon_t,$$
$$\varepsilon_t = \sqrt{h_t} e_t,\ e_t \sim i.i.d.\,(0, 1),$$
$$\left(1 - \sum_{j=1}^{3} \beta_j^* L^j\right) \ln(h_t) = \omega + \sum_{i=1}^{3} c_i (d_i e_{t-i} + |e_{t-i}|).$$

The numbers in parentheses are t-statistics. LL is the maximum log likelihood value. υ are the degrees of freedom of the generalized error distribution. Standard errors are reported in brackets.

and 0.0340 respectively. For the KOSPI and ST indices the distribution of the innovations is generalized error. However the value of the highest root of the autoregressive polynomial for the ST index is 0.913. Therefore, although the autocorrelations start very high, at 0.2257 for lag one, they decrease very quickly. For the KOSPI index, the autocorrelation at lag 10, 20 and 30 is 0.0824, 0.0445 and 0.0261, respectively, whereas for the ST index it is 0.0499, 0.0170 and 0.0064 respectively.

It is useful to uncover the properties of the autocorrelation function of the squared observations, when comparing the EGARCH model with the standard GARCH model family. Possible differences in the moment structure of these models may shed light on the success of the EGARCH model in applications. To facilitate model identification, the results for the autocorrela-

Table 2. Persistence.

	β_j	Persistence	$c_1 d_1$	$\sum_{j=1}^{q} c_j d_j$
KOSPI (GEN ERROR) MA(1)-EGARCH(3,3)	$\beta_1 = 0.95$ $\beta_2 = -0.72 + 0.68i$ $\beta_3 = -0.72 - 0.68i$	0.554	−0.032	−0.100
NIKKEI (DOUBLE EXP) WN-EGARCH(1,3)	$\beta_1 = 0.973$	0.209	−0.193	−0.092
SE (DOUBLE EXP) WN-EGARCH(1,3)	$\beta_1 = 0.983$	0.745	−0.103	−0.026
ST (GEN ERROR) MA(1)−EGARCH(2,1)	$\beta_1 = -0.29$ $\beta_2 = 0.91$	0.298	−0.068	−0.068

The second column of this table reports a measure for the persistence of a (positive) shock to e_t on $\ln(h_t)$.

tions of the power-transformed observations can be applied so that the properties of the observed data can be compared with the theoretical properties of the models. For each of the four stock indices, Figure 1 plots the sample autocorrelations of the squared observations. It also plots the estimated theoretical autocorrelations of the squared observations of the 'best' EGARCH specification and of the GARCH(1,1) model with conditionally normal errors.[8] For all three indices, the autocorrelations of the EGARCH model are closer to the sample autocorrelations than those of the GARCH model.[9] Note that for the KOSPI index, the autocorrelation of the squared observations, at lag 2, 4 and 20 is equal to the corresponding sample autocorrelation. For the ST index, Figure 1(d) plots the autocorrelations of the squared errors of the GARCH(1,1) model with innovations drawn from the generalized error distribution.[10] Observe that these autocorrelations are much higher than those obtained with conditionally normal errors. For the SE index it can be seen that the fitted squared returns from the GARCH model generally have autocorrelations that are substantially too high when compared with the corresponding sample equivalents. In fact, they generally exceed the corresponding sample autocorrelations by a factor of two. In contrast, the EGARCH specification does a good job of replicating the observed pattern of autocorrelations of the squared returns. It generates a model where the autocorrelations of the fitted squared values are relatively 'close' to those of the population equivalents. The autocorrelation of the squared returns, at lag 11, 12, 16 and 26 is equal to the corresponding sample autocorrelation. In other words, the EGARCH model can more accurately reproduce the nature of the sample autocorrelations of squared returns than the GARCH model. Finally, for the four selected specifications, when the no-trade dummy enters in the conditional variance, Figure 2 plots the

[8]The GARCH(1,1) specification that we estimate is $h_t = \omega + a\varepsilon_{t-1}^2 + \beta h_{t-1}$. In order to obtain the estimated theoretical autocorrelations of the squared errors of the above model we use the following formula

$$\rho(\varepsilon_t^2, \varepsilon_{t-k}^2) = \frac{(a+\beta)^k\{1+\beta^2-\beta[(a+\beta)+(a+\beta)^{-1}]\}}{1+\beta^2-2\beta(a+\beta)}.$$

[9]We also estimate a GARCH(1,1) model with conditionally normal errors for the NIKKEI index. The sum of the ARCH and GARCH coefficients is greater than one.

[10]For all indices, we also estimated GARCH(1,1) models with innovations drawn from either the double exponential or t distributions. In all cases, the condition for the existence of either the first moment ($a + \beta < 1$) or the second one ($\beta^2 + 2a\beta + E(e_t^4)a^2 < 1$) was violated.

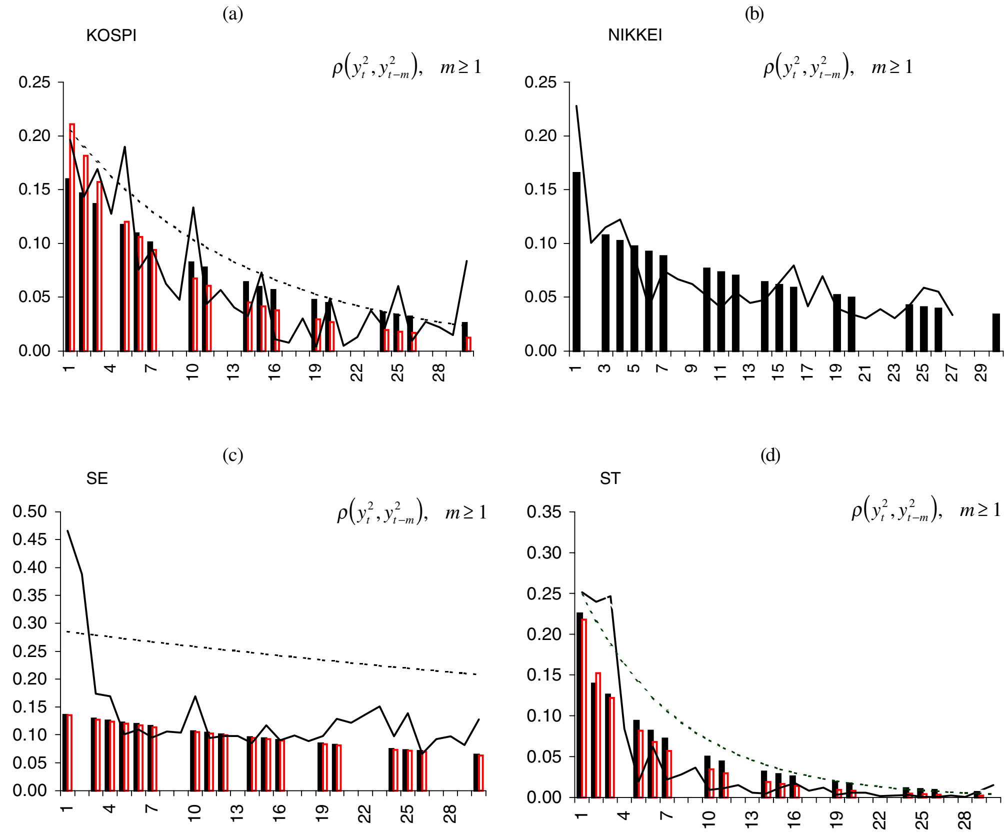

Figure 1. For each of the four indices, Figure 1 plots the sample autocorrelations of the squared observations (solid line). It also plots the estimated theoretical autocorrelations (ETA) of the squared observations for the 'best' EGARCH specification chosen by AIC (dark columns) and by SIC (clear columns). All the 'best' EGARCH models have been estimated without the inclusion of the no-trade dummy. Dotted lines represent the ETA of the squared observations for the GARCH(1,1) model with conditionally normal errors. Moreover, for the ST index, the grey line represents the ETA of the squared observations for the GARCH(1,1) model with errors drawn from the generalized error distribution. Finally, note that for the Nikkei index both information criteria chose the EGARCH(1,3) specification.

estimated theoretical autocorrelations of the squared observations and their corresponding sample equivalents.

4. CONCLUSIONS

In this paper we obtained a complete characterization of the moment structure of the general ARMA(r, s)–EGARCH(p, q) model. In particular, we provided the autocorrelation function of any positive integer power of the squared errors. Additionally, we derived the cross correlations between the levels and squares of the observed process. To obtain our results, we assumed that the error term is drawn from either the normal, double exponential or generalized error distributions.

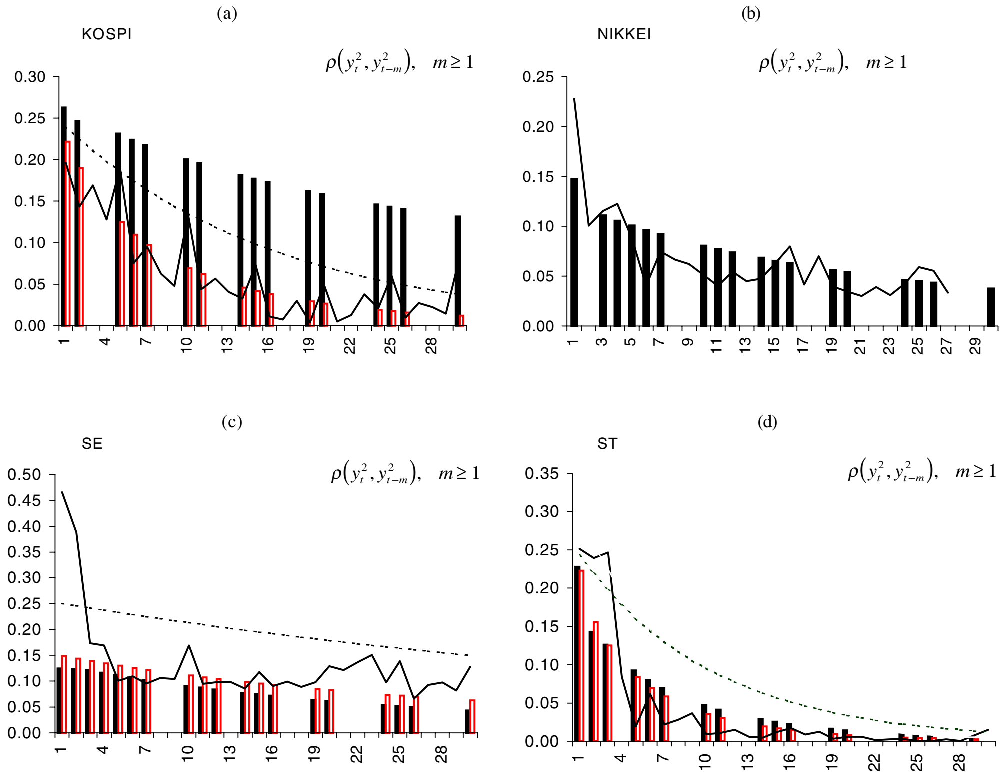

Figure 2. For each of the four indices, Figure 2 plots the sample autocorrelations of the squared observations (solid line). It also plots the ETA of the squared observations for the specifications used in Figure 1. All these EGARCH models have now been estimated with the inclusion of the no-trade dummy. Dotted lines represent the ETA of the squared observations for the GARCH(1,1) model with conditionally normal errors. Moreover, for the ST index, the grey line represents the ETA of the squared observations for the GARCH(1,1) model with errors drawn from the generalized error distribution.

The results of the paper can be used to compare the EGARCH model with the standard GARCH model or the APARCH model. They reveal certain differences in the moment structure between these models. Further, to facilitate model identification, the results for the autocorrelations of the squared observations can be applied so that the properties of the observed data can be compared with the theoretical properties of the models. Finally, the techniques used in this paper can be employed to obtain the moments of more complex EGARCH models, e.g. EGARCH-in-mean, the Component EGARCH, and the Fractional Integrated EGARCH models. The derivation of the moment structure of these models is left for future research.

ACKNOWLEDGEMENTS

We greatly appreciate Neil Shephard and an anonymous referee for their valuable comments, which led to a substantial improvement of the previous version of the article. We would also like to thank Karim Abadir and Marika Karanassou for their helpful suggestions.

REFERENCES

Abadir, K. M. (1999). An introduction to hypergeometric functions for economists. *Econometric Reviews 18*, 287–330.

Baillie, R. and H. Chung (2001). Estimation of GARCH models from the autocorrelations of the squares of a process. *Journal of Time Series Analysis 6*, 631–50.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association*. pp. 177–81. Alexandria VA: American Statistical Association.

Bollerslev, T., R. F. Engle and D. B. Nelson (1994). ARCH models. In R. F. Engle and D. McFadden (eds), *Handbook of Econometrics*, vol. 4. Amsterdam: North Holland.

Dacorogna, M. M., U. Muller, A. Nagler, R. B. Olsen and O. V. Pictet (1993). A geographical model for the daily and weekly seasonal volatility in the FX market. *Journal of International Money and Finance 12*, 413–38.

Demos, A. (2002). Moments and dynamic structure of a time-varying-parameter stochastic volatility in mean model. *The Econometrics Journal 5*, 345–57.

Ding, Z. and C. W. J. Granger (1996). Modelling volatility persistence of speculative returns: a new approach. *Journal of Econometrics 73*, 185–215.

Ding, Z., C. W. J. Ganger and R. F. Engle (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance 1*, 83–106.

Donaldson, R. D. and M. Kamstra (1997). An artificial neural network-GARCH model for international stock return volatility. *Journal of Empirical Finance 4*, 17–46.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica 50*, 987–1007.

Engle, R. F. and V. K. Ng (1993). Measuring and testing the impact of news on volatility. *Journal of Finance 48*, 1749–78.

Fiorentini, G. and E. Sentana (1998). Conditional means of time series processes and time series processes for conditional means. *International Economic Review 39*, 1101–18.

Gradshteyn, I. S. and I. M. Ryzhik (1994). *Table of Integrals, Series, and Products*. London: Academic Press.

He, C. and T. Teräsvirta (1999). Fourth moment structure of the GARCH(p, q) model. *Econometric Theory 15*, 824–46.

He, C., T. Teräsvirta and H. Malmsten (2002). Fourth moment structure of a family of first-order exponential GARCH models. *Econometric Theory 18*, 868–85.

Hentschel, L. (1995). All in the family. Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics 39*, 71–104.

Hu, M. Y., C. X. Jiang and C. Tsoukalas (1997). The European exchange rates before and after the establishment of the European monetary system. *Journal of International Financial Markets, Institutions and Money 7*, 235–53.

Karanasos, M. (1999). The second moment and the autocovariance function of the squared errors of the GARCH model. *Journal of Econometrics 90*, 63–76.

Karanasos, M. (2001). Prediction in ARMA models with GARCH-in-mean effects. *Journal of Time Series Analysis 22*, 555–78.

Muller, U. A., M. M. Dacorogna, R. D. Dave, R. B. Olsen, O. V. Pictet and J. E. Von Weizsacker (1997). Volatilities of different time resolutions-analyzing the dynamics of market components. *Journal of Empirical Finance 4*, 213–39.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica 59*, 347–70.

Palma, W. and M. Zevallos (2001). Analysis of the correlation structure of square time series, unpublished paper, Departmento de Estadística, Pontificia Universidad Católica de Chile.

Prudnikov, A. P., Yu. A. Brychkov and O. I. Marichev (1992). *Integrals and Series*. Glasgow: Bell and Bain Ltd.

Scholes, M. and J. Williams (1977). Estimating betas from nonsynchronous data. *Journal of Financial Economics 5*, 309–27.

APPENDIX A

Proof (Proposition 1). Using the Wold representation of an ARMA model (see Karanasos (2001)), and the EGARCH(p,q) conditional variance (2.3), we have

$$\ln(h_t) = \frac{\omega}{\prod_{f=1}^{p}(1-\beta_f)} + \sum_{l=1}^{\infty}\sum_{f=1}^{p}\zeta_f\lambda_{fl}z_{t-l},$$

where λ_{fl} and ζ_f are defined in (2.4c) and (2.4d) respectively. From the above equation it follows that

$$h_t^k = \exp\left(\frac{\omega k}{\prod_{f=1}^{p}(1-\beta_f)}\right) \times \exp\left(k\sum_{l=1}^{\infty}\sum_{f=1}^{p}\zeta_f\lambda_{fl}z_{t-l}\right), \tag{A.1}$$

or

$$h_{t-m}^k = \exp\left(\frac{\omega k}{\prod_{f=1}^{p}(1-\beta_f)}\right) \times \exp\left(k\sum_{l=1}^{\infty}\sum_{f=1}^{p}\zeta_f\lambda_{fl}z_{t-m-l}\right). \tag{A.2}$$

From (A.1) and (A.2), it follows that the expected value of $\varepsilon_t^{2k}\varepsilon_{t-m}^{2k}$ is

$$\begin{aligned}\mathsf{E}(\varepsilon_t^{2k}\varepsilon_{t-m}^{2k}) = \mathsf{E}(e_t^{2k}h_t^k e_{t-m}^{2k}h_{t-m}^k) &= \exp\left(\frac{2\omega k}{\prod_{f=1}^{p}(1-\beta_f)}\right) \times \mathsf{E}(e_t^{2k}) \\ &\times \mathsf{E}\left[\exp\left(k\sum_{i=1}^{m-1}\sum_{f=1}^{p}\zeta_f\lambda_{fi}z_{t-i}\right)\right] \times \mathsf{E}\left[e_{t-m}^{2k}\exp\left(k\sum_{f=1}^{p}\zeta_f\lambda_{fm}z_{t-m}\right)\right] \\ &\times \mathsf{E}\left[\exp\left(k\sum_{i=1}^{\infty}\sum_{f=1}^{p}(\lambda_{f,m+i}+\lambda_{fi})\zeta_f z_{t-i-m}\right)\right], \qquad (m>0),\end{aligned} \tag{A.3}$$

$$\begin{aligned}\mathsf{E}\left(\varepsilon_t^{4k}\right) &= \exp\left(\frac{2\omega k}{\prod_{f=1}^{p}(1-\beta_f)}\right) \times \mathsf{E}\left(e_t^{4k}\right) \\ &\times \mathsf{E}\left[\exp\left(2k\sum_{i=1}^{\infty}\sum_{f=1}^{p}\lambda_{fi}\zeta_f z_{t-i}\right)\right].\end{aligned} \tag{A.4}$$

The proof of Proposition 1 is completed by inserting (A.3) and (A.4) into $\rho(\varepsilon_t^{2k}, \varepsilon_{t-m}^{2k}) = \frac{\mathsf{E}(\varepsilon_t^{2k},\varepsilon_{t-m}^{2k})-[\mathsf{E}(\varepsilon_t^{2k})]^2}{\mathrm{Var}(\varepsilon_t^{2k})}$. □

Proof (Theorem 1). When the distribution of the innovations is generalized error, the expected value of $e_t^k \exp(z_t b)$ is given by expression A1.5 in Theorem A1.2 (Nelson, 1991):

$$\mathsf{E}[e_t^k \exp(z_t b)] = \exp\left(-b\gamma \frac{\Gamma\left(\frac{2}{v}\right)\lambda 2^{\frac{1}{v}}}{\Gamma\left(\frac{1}{v}\right)}\right) 2^{\frac{k}{v}} \lambda^k \times \sum_{\tau=0}^{\infty} \left(\lambda 2^{\frac{1}{v}} b\right)^{\tau} [(\gamma+d)^{\tau} + (-1)^k (\gamma-d)^{\tau}] \frac{\Gamma\left(\frac{\tau+k+1}{v}\right)}{2\Gamma\left(\frac{1}{v}\right)\Gamma(1+\tau)}, \tag{A.5}$$

where k is a finite non-negative integer and b is a real scalar ($k, b < \infty$). When $v > 1$, the above summation is finite, whereas when $v < 1$, the summation is finite if and only if $b\gamma + |bd| \leq 0$. Using expressions (A.3), (A.4) and (A.5), after some algebra, gives (2.5). □

Proof (Proposition 2). Equation (A.5), for $v = 1$, gives

$$\mathsf{E}[e_t^k \exp(z_t b)] = 2^{\frac{-(k+2)}{2}} \exp\left(-b\gamma \frac{1}{\sqrt{2}}\right) \Gamma(k+1) \times \left\{ F\left[k+1; \frac{b(\gamma+d)}{\sqrt{2}}\right] + (-1)^k F\left[k+1; \frac{b(\gamma-d)}{\sqrt{2}}\right] \right\}. \tag{A.6}$$

The right-hand side of the above expression converges if and only if $b\gamma + |bd| < \sqrt{2}$. In addition, the above equation, for $k = 0$, gives

$$\mathsf{E}[\exp(z_t b)] = \frac{1}{2} \exp\left(-b\gamma \frac{1}{\sqrt{2}}\right) \frac{2 - \sqrt{2} b\gamma}{1 - \sqrt{2} b\gamma + \frac{b^2(\gamma^2 - d^2)}{2}}. \tag{A.7}$$

When the conditional distribution of the errors is double exponential, combining equations (A.3), (A.4), (A.6) and (A.7), after some algebra, yields equation (2.6). □

Proof (Proposition 3). When the errors are conditionally normal, we use formula 2.3.15 #7 in Prudnikov *et al*. (1992, Volume 1), to obtain the expected value of $e_t^k \exp(z_t b)$:

$$\mathsf{E}[e_t^k \exp(z_t b)] = \left\{ (-1)^k \frac{\partial}{\partial [b(\gamma - d)]^k} \left\{ \exp\left(\frac{b^2(\gamma-d)^2}{2}\right) \left[1 + \Phi\left(\frac{b(\gamma-d)}{\sqrt{2}}\right)\right] \right\} + \frac{\partial}{\partial [b(\gamma + d)]^k} \left\{ \exp\left(\frac{b^2(\gamma+d)^2}{2}\right) \left[1 + \Phi\left(\frac{b(\gamma+d)}{\sqrt{2}}\right)\right] \right\} \right\} \frac{\exp\left(-\gamma b \sqrt{\frac{2}{\pi}}\right)}{2}, \tag{A.8}$$

where $\Phi(\cdot)$ is the error function (formula 8.250 #1 in Gradshteyn and Ryzhik (1994)). Note that the above expression is finite. Using formula 8.253 # 1 in Gradshteyn and Ryzhik (1994), equation (A.8) , for $k = 0$, gives

$$\mathsf{E}[\exp(z_t b)] = \exp\left(-\gamma b \sqrt{\frac{2}{\pi}}\right) \left\{ \exp\left(\frac{b^2(\gamma+d)^2}{2}\right) \frac{1}{2}[1 + \exp(-2\gamma d b^2)] + \frac{b(\gamma+d)}{\sqrt{2\pi}} F\left(1; \frac{3}{2}; \frac{b^2(\gamma+d)^2}{2}\right) + \frac{b(\gamma-d)}{\sqrt{2\pi}} F\left(1; \frac{3}{2}; \frac{b^2(\gamma-d)^2}{2}\right) \right\}, \tag{A.9}$$

where $F(\cdot)$ is the hypergeometric function (for an alternative derivation see Theorem A1.1 in Nelson (1991)). Combining (A.3), (A.4), (A.8) and (A.9), after some algebra, yields equation (2.7). □

APPENDIX B

Proof (Theorem 2). $\mathsf{E}(y_t^2, y_{t-m})$ can be written as

$$\mathsf{E}(y_t^2, y_{t-m}) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{l=0}^{\infty} \delta_i \delta_j \delta_{l+m} \mathsf{E}(\varepsilon_{t-i}\varepsilon_{t-j}\varepsilon_{t-l-m}), \tag{B.1}$$

where δ_i, when the roots of the autoregressive polynomial $\Phi(L)$ in (2.1b) are distinct, is given by

$$\delta_l \equiv \sum_{f=1}^{r} \pi_f s_{fl}, \qquad (\delta_0 \equiv 1), \tag{B.2a}$$

with

$$\pi_f \equiv \frac{\phi_f^{r-1}}{\prod_{\substack{n\neq f\\ n=1}}^{r} (\phi_f - \phi_n)}, \tag{B.2b}$$

$$s_{fl} \equiv \begin{cases} \sum\limits_{n=0}^{l-1} \theta_{l-n}\phi_f^n, & \text{if } l \le s, \\ s_{fs}\phi_f^{l-s}, & \text{if } l > s. \end{cases} \tag{B.2c}$$

Next, note that

$$\mathsf{E}(\varepsilon_{t-i}\varepsilon_{t-j}\varepsilon_{t-l-m}) = \begin{cases} \mathsf{E}(\varepsilon_t^2\varepsilon_{t-(l+m-i)}), & \text{if } i = j < l+m, \\ \mathsf{E}(\varepsilon_t^2\varepsilon_{t-(j-l-m)}), & \text{if } i = l+m,\ j > i, \\ \mathsf{E}(\varepsilon_t^2\varepsilon_{t-(i-l-m)}), & \text{if } j = l+m,\ i > j, \\ 0 & \text{otherwise.} \end{cases} \tag{B.3}$$

and

$$\mathsf{E}(\varepsilon_t^2\varepsilon_{t-m}) = \rho(\varepsilon_t^2\varepsilon_{t-m})(\eta - 1)^{\frac{1}{2}}[\mathsf{E}(\varepsilon_t^2)]^{\frac{3}{2}} \qquad (m > 0), \tag{B.4}$$

where η is the kurtosis of ε_t.

Substituting (B.3) and (B.4) into (B.1) yields

$$\mathsf{E}(y_t^2, y_{t-m}) = (\eta - 1)^{\frac{1}{2}}[\mathsf{E}(\varepsilon_t^2)]^{\frac{3}{2}} \times \left\{ \sum_{i=0}^{j+m-1}\sum_{j=0}^{\infty} \delta_i^2\delta_j\rho(\varepsilon_t^2\varepsilon_{t-(j+m-i)}) + 2\sum_{i=j+m+1}^{\infty}\sum_{j=0}^{\infty}\delta_i\delta_j\delta_{j+m}\rho(\varepsilon_t^2\varepsilon_{t-(i-j-m)}) \right\}. \tag{B.5}$$

Further, we have (see Palma and Zevallos (2001))

$$\frac{\mathrm{Var}(y_t^2)}{[\mathsf{E}(\varepsilon_t^2)]^2\left(\sum_{l=0}^{\infty}\delta_l^2\right)} = \kappa - 1, \tag{B.6}$$

where κ is given in (2.11e). The proof of (2.9a) is completed by inserting (B.5) and (B.6) into $\rho(y_t^2, y_{t-m}) = \frac{\mathsf{E}(y_t^2, y_{t-m})}{\sqrt{\mathrm{Var}(y_t^2)\mathsf{E}(y_t^2)}}$. The proof of (2.10a) is like that of (2.9a). □

Proof (Proposition 6). Multiplying (A.1) by (A.2) and taking expectations yields

$$\mathsf{E}(h_t^k h_{t-m}^k) = \exp\left(\frac{2\omega k}{\prod_{f=1}^{p}(1-\beta_f)}\right) \times \mathsf{E}\left[\exp\left(k\sum_{i=1}^{m}\sum_{f=1}^{p}\zeta_f\lambda_{fi}z_{t-i}\right)\right]$$
$$\times \mathsf{E}\left[\exp\left(k\sum_{i=1}^{\infty}\sum_{f=1}^{p}(\lambda_{f,m+i}+\lambda_{fi})\zeta_f z_{t-i-m}\right)\right]. \quad \text{(B.7)}$$

Additionally, from (A.1) it follows that

$$\mathsf{E}(h_t^k) = \exp\left(\frac{\omega k}{\prod_{f=1}^{p}(1-\beta_f)}\right) \times \mathsf{E}\left[\exp\left(k\sum_{i=1}^{\infty}\sum_{f=1}^{p}\lambda_{fi}\zeta_f z_{t-i}\right)\right]. \quad \text{(B.8)}$$

Using equations (B.7), (B.8) and (A.5) for $k = 0$, after some algebra, yields the result in Proposition 6. □