THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 2000/29

Moments of the ARMA-EGARCH Model

by

Menelaos Karanasos and J. Kim

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

MOMENTS OF THE ARMA-EGARCH MODEL

M KARANASOS AND J KIM

University of York, Heslington, York, Y010 5DD, UK

June 2000

Abstract

This paper considers the moment structure of the ARMA(r,s)-EGARCH(p,q) model. In particular, we provide the autocorrelation function and any arbitrary moment of the conditional variance/squared errors. In addition, we derive the cross correlations between the process and the conditional variance/squared errors. We also explain our general results using the MA(1)-EGARCH(3,3) and the MA(1)-EGARCH(1,4) models as examples. 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.
JEL Classification: C22.

Address for correspondence: Email: mk16@york.ac.uk, Tel: 01904-433799, Fax: 01904-433759.
We would like to thank K Abadir and M Karanassou for helpful comments and suggestions.

1 INTRODUCTION

One of the principal empirical tool used to model volatility in asset markets has been the ARCH class of models. Following Engle's (1982) pathbreaking idea, several formulations of conditionally heteroscedastic models (e.g. GARCH, Fractional Integrated GARCH, Switching GARCH, Component GARCH) have been introduced in the literature, forming 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[1]. One of the earliest asymmetric GARCH models is the EGARCH (Exponential generalized ARCH) model of Nelson (1991).[2] The EGARCH model is more general than the standard GARCH model in that it allows innovations of different signs to have a differential impact on volatility than does the standard GARCH model. Also, in contrast to the conventional GARCH specification which requires nonnegative coefficients, the EGARCH model by modeling the logarithm of the conditional variance does not impose the nonegativity constraints on the parameter space.

>From the many different functional forms the EGARCH model has become perhaps the most common. In particular, various cases of the EGARCH model have been applied by many researchers. For example, it has been used to examine the interest rates (Brunner and Simon, 1996, Andersen and Lund, 1997, Kavussanos and Alizadeh, 1999, Kim and Sheen, 2000), to model foreign exchange rates (Hu, et al., 1997, Hafner, 1998, Kim, 1998, Lobo and Tufte, 1998), and to analyze stock returns (Koutmos and Booth, 1995, Episcopos, 1996, Reyes, 1996, Hagerud, 1997, Booth, et al., 1997, Donaldson and Kamstra, 1997, Kusi and Pescetto, 1998). The EGARCH model has also been applied to the interest rate futures markets (Tse and Booth, 1996, Tse, 1998).

Various authors have compared the EGARCH model with other alternative conditional heteroscedastic models. For example, Donaldson and Kamstra (1997) compared the EGARCH model with an artificial neural network GARCH model; Hafner (1998) compared the EGARCH model with nonparametric GARCH models; Breidt, et al. (1998) compared long memory stochastic volatility models with EGARCH and FIEGARCH models.

Although the EGARCH model was introduced almost a decade ago and has been widely used in empirical applications, its statistical properties have only recently been examined by researchers. Engle and Ng (1993) artificially nested the GARCH and EGARCH models, estimated this nested specification, and then applied likelihood ratio tests (this approach was also used by Hu, et al., 1997). Hentchel (1995) developed a family of asymmetric GARCH models that nests both the A-PARCH model and the EGARCH model. Deb (1996) examined the finite sample properties of the maximum likelihood and quasi-maximum likelihood estimators of EGARCH(1,1) processes using Monte Carlo methods. Lee and Brorsen (1997) proposed a Cox-type non-nested test of GARCH versus EGARCH models. Pierre (1998) found that, regardless of the assumption made about the conditional error distribution, the EGARCH model is sensitive to the choice of starting values and the degree of computer precision. Andersen, et al. (1999) performed an extensive Monte Carlo study of efficient method of moments estimation of the EGARCH model.

This paper focuses solely on the moment structure of the general ARMA(r,s)-EGARCH(p,q) model. Although the literature on the GARCH/EGARCH models is quite extensive, relative fewer papers have examined the moment structure of models where the conditional volatility is time-dependent. Specifically, some studies have examined the time series properties of the conditional variance (see, for example, Karanasos 1999, 2000) and the relation between the conditional mean and the conditional variance (see, for example, Karanasos and Kim, 2000, and Fountas, et al., 2000).

We contribute to this literature by deriving (i) the autocorrelation function of the conditional variance and of the squared errors, and (ii) the cross correlations between the process and the squared errors. To

[1] The asymmetric response of volatility to positive and negative shocks is well known in the finance literature as the leverage effect of the 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).

[2] Another popular asymmetric GARCH specification is the Asymmetric Power ARCH (A-PARCH) model of Ding, Granger and Engle (1993). The A-PARCH is a general model which includes several other asymmetric models as special cases (e.g. the GJR model, proposed by Glosten, Jagannathan and Runckle, 1993, and the TARCH model, introduced by Zakoian, 1994).

obtain the theoretical results and to carry out the estimation we assume that the error term is drawn from either the normal, or the double exponential or the generalized error distributions.

We should note that, while this research was in progress, Demos (2000) was simultaneously and independently studying the autocorrelation structure of a model that nests both the EGARCH and stochastic volatility specifications. However, he only examines the autocorrelations of the conditional variance/squared errors while we give the autocorrelations between the k_1-th and k_2-th powers (k_1,k_2 > 0) of the conditional variance/squared errors, he assumes normality whereas we also use the double exponential and the generalized error distributions, and finally he restricts his analysis to a white noise specification of the conditional mean[3] while we use a general ARMA(r,s) process.

Also noted that the coefficients of the Wold representation of the conditional mean and variance are needed for the computation of the autocorrelations. In contrast to Demos (2000) we provide exact form solutions which express the coefficients in terms of the parameters of the moving average polynomials and the roots of the autoregressive polynomials.

The rest of the paper is organized as follows. Subsection 2.1 investigates the autocorrelation function of the conditional variance for the EGARCH(1,q) and EGARCH(p,1) models. Subsection 2.2 presents the autocorrelations of the squared errors and provides the cross correlations between the process and its conditional variance for the general ARMA(r,s)-EGARCH(p,q) model. Section 3 discusses the data and presents the empirical results. The final Section concludes the paper and suggests future developments. Proofs are found in the Appendices.

2 EGARCH MODEL

2.1 EGARCH(1,q) Model

Pagan and Schwert (1990), and Donaldson and Kamstra (1997) found that two lag squared errors are required for the S&P 500 EGARCH. Kim and Kon (1994) used an EGARCH(1,3) model for all individual stocks in the Dow Jones Industrial Average. The general EGARCH(1,q) model is given by

$$B(L)\ln(h_t) = \omega + C(L)z_t, \tag{2.1a}$$

$$z_{t-l} = d\frac{\varepsilon_{t-l}}{\sqrt{h_{t-l}}} + \gamma[|\frac{\varepsilon_{t-l}}{\sqrt{h_{t-l}}}| - E|\frac{\varepsilon_{t-l}}{\sqrt{h_{t-l}}}|], \quad \varepsilon_t = e_t h_t^{\frac{1}{2}}, \quad e_{t|t-1} \sim IID(0,1), \tag{2.1b}$$

$$C(L) = \sum_{l=1}^{q} c_l L^l, \tag{2.1c}$$

$$B(L) = 1 - \beta_1 L \tag{2.1d}$$

- *Assumption 1.* The autoregressive parameter β_1 is less than 1 in absolute value (covariance stationarity condition).
- *Assumption 2.* The polynomials B(L) and C(L) have no common left factors other than unimodular ones (irreducibility condition).

Proposition 1 *Under Assumptions 1 and 2 the $(k_1 + k_2)$-th moment of the conditional variance, and the auto covariances/correlations between the k_1-th and k_2-th powers $(k_1, k_2 > 0)$ of the conditional variance, are given by*

[3] Although Demos (2000) examines only the white noise case, he allows the conditional variance to affect the mean with a possibly time varying parameter.

$$
\begin{aligned}
E(h_t^{k_1+k_2}) &= e^{\widetilde{f}_D}\Pi^D_{0,k_1,k_2}, &(2.2a)\\
cov(h_t^{k_1}, h^{k_2}_{t-\underset{(m>0)}{m}}) &= e^{\widetilde{f}_D}[\Pi^D_{m,k_1}\Pi^D_{m,k_1,k_2} - \Pi^D_{0,k_1,0}\Pi^D_{0,0,k_2}], &(2.2b)\\
\rho(h_t^{k_1}, h_{t-m}^{k_2}) &= \frac{[\Pi^D_{m,k_1}\Pi^D_{m,k_1,k_2} - \Pi^D_{0,k_1,0}\Pi^D_{0,0,k_2}]}{\sqrt{[\Pi^D_{0,k_1,k_1} - (\Pi^D_{0,k_1,0})^2][\Pi^D_{0,k_2,k_2} - (\Pi^D_{0,k_2,0})^2]}}, &(2.2c)\\
\widetilde{f}_D &= \frac{(k_1+k_2)[\omega - \gamma f_D \sum_{l=1}^{q} c_l]}{1-\beta_1}, \quad \Pi^D_{0,k_1} = 1 &(2.2d)
\end{aligned}
$$

When the conditional distribution is the normal the $\prod^D_{m,k_1}$, $\prod^D_{m,k_1,k_2}$ and f_D *terms are given by*

$$
\begin{aligned}
\Pi^N_{m,k_1} &= \prod_{i=0}^{m-1}\{\varrho^{\frac{(\gamma+d)^2\xi_{1i}^2}{2}}\frac{1}{2}[1+\varrho^{-2\gamma d\xi_{1i}^2}] + \frac{(\gamma+d)\xi_{1i}}{\sqrt{2\pi}} \times F(1;\frac{3}{2};\frac{(\gamma+d)^2\xi_{1i}^2}{2}) + \\
&\quad + \frac{(\gamma-d)\xi_{1i}}{\sqrt{2\pi}} \times F(1;\frac{3}{2};\frac{(\gamma-d)^2\xi_{1i}^2}{2})\}, &(2.3a)\\
\Pi^N_{m,k_1,k_2} &= \prod_{i=0}^{\infty}\{\varrho^{\frac{(\gamma+d)^2\xi_{2i}^2}{2}}\frac{1}{2}[1+\varrho^{-2\gamma d\xi_{2i}^2}] + \frac{(\gamma+d)\xi_{2i}}{\sqrt{2\pi}} \times F(1;\frac{3}{2};\frac{(\gamma+d)^2\xi_{2i}^2}{2}) + \\
&\quad + \frac{(\gamma-d)\xi_{2i}}{\sqrt{2\pi}} \times F(1;\frac{3}{2};\frac{(\gamma-d)^2\xi_{2i}^2}{2})\}, &(2.3b)\\
f_N &= \sqrt{\frac{2}{\pi}} &(2.3c)\\
\xi_{1i}(k_1) &= k_1\lambda_{1,i+1}, \quad \xi_{2i}(k_1,k_2) = k_2\lambda_{1,i+1} + k_1\lambda_{1,m+i+1}, &(2.3d)\\
\lambda_{1,i} &= \begin{cases} \sum_{n=0}^{i-1} c_{i-n}\beta_1^n, & i \le q \\ \lambda_{1q}\beta_1^{i-q}, & i > q \end{cases} &(2.3e)
\end{aligned}
$$

where F denotes the hypergeometric function (see Abadir, 1999).
The proof of Proposition 1 is given in Appendix A.

2.2 EGARCH(p,1) Model

Nelson (1991) estimated an EGARCH (2,1) model using daily excess returns on the value-weighted CRPS index. Donaldson and Kamstra (1997) found that the optimal EGARCH specification for the TSEC stock index was the 2,1. Pierre (1998) also chose the EGARCH(2,1) model for an equally-weighted index made up of firms on the NYSE/AMEX. The general EGARCH(p,1) model is given by equations (2.1a)-(2.1b), where the $B(L)$ and $C(L)$ polynomials are given by

$$
\begin{aligned}
B(L) &= \prod_{l=1}^{p}(1-\beta_l L), &(2.4a)\\
C(L) &= c_1 L &(2.4b)
\end{aligned}
$$

- *Assumption 3.* All the roots of the autoregressive polynomial $B(L)$ lie outside the unit circle (covariance-stationarity condition).

In what follows we only examine the case where all the roots of the autoregressive polynomial $B(L)$ are distinct.

Proposition 2 *Under Assumption 3 and when the conditional distribution is the normal, the (k_1+k_2)-th moment of the conditional variance, and the auto covariances/correlations between the k_1-th and k_2-th* *powers* $(k_1, k_2 > 0)$ *of the conditional variance are given by equations (2.2a)- (2.2c) and (2.3a)-(2.3b), where the* $\widetilde{f}_D$, ξ_{1i}, *and* ξ_{2i} *terms are given by*

$$\widetilde{f}_D = \frac{(k_1+k_2)[\omega - \gamma f_D c_1]}{\prod_{f=1}^{p}(1-\beta_f)}, \tag{2.5a}$$

$$\xi_{1i}(k_1) = c_1 k_1 \sum_{f=1}^{p} \zeta_f \beta_f^i, \quad \xi_{2i}(k_1,k_2) = c_1 \sum_{f=1}^{p} (k_1 \beta_f^m + k_2)\zeta_f \beta_f^i \tag{2.5b}$$

$$\zeta_f = \frac{\beta_f^{p-1}}{\prod_{\substack{n=1\\ n\neq f}}^{p} (\beta_f - \beta_n)} \tag{2.5c}$$

The proof of Proposition 2 is given in Appendix A.

2.3 ARMA(r,s)-EGARCH(p,q) Model

Donaldson and Kamstra (1997) found that the optimal EGARCH specification for the NIKKEI stock index was the complex 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. The general ARMA(r,s)-EGARCH(p,q) model is

$$\Phi(L) y_t = b + \Theta(L)\varepsilon_t, \quad \varepsilon_t = e_t h_t^{\frac{1}{2}}, \quad e_{t|t-1} \sim IID(0,1) \tag{2.6a}$$

$$\Phi(L) = \prod_{l=1}^{r}(1-\phi_l L), \quad \Theta(L) = 1 + \sum_{l=1}^{s} \theta_l L^l, \quad \varepsilon_t = e_t h_t^{\frac{1}{2}}, \quad e_{t|t-1} \sim IID(0,1) \tag{2.6b}$$

where the h_t is given by (2.1a)-(2.1b) and the B(L) and C(L) are given by (2.4a) and (2.1c), respectively.

- *Assumption 4.* The polynomials $B(L)$ and $C(L)$ have no common left factors other than unimodular ones (irreducibility condition).

Theorem 1a. *Under Assumptions 3 and 4 the* $(k_1 + k_2)$*-th moment of the conditional variance, and the auto covariances/correlations between the* k_1*-th and* k_2*-th powers of the conditional variance,* $(k_1, k_2 > 0)$ *are given by (2.2a)-(2.2c), where* $\widetilde{f}_D$ *is given by*

$$\widetilde{f}_D = \frac{(k_1+k_2)[\omega - \gamma f_D \sum_{l=1}^{q} c_l]}{\prod_{f=1}^{p}(1-\beta_f)} \tag{2.7}$$

When the conditional distribution is the normal, the $\prod_{m,k_1}^{D}$, $\prod_{m,k_1,k_2}^{D}$ *and* f_D *terms are given by (2.3a)-(2.3c), where the* ξ_{1i}, *and* ξ_{2i} *terms are given by*

$$\xi_{1i}(k_1) = k_1 \sum_{f=1}^{p} \zeta_f \lambda_{f,i+1}, \quad \xi_{2i}(k_1,k_2) = \sum_{f=1}^{p} (k_1 \lambda_{f,m+i+1} + k_2 \lambda_{f,i+1})\zeta_f \tag{2.8}$$

and $\lambda_{f,i}$ and ζ_f *are defined in (2.3e) and (2.5c), respectively.*

When the conditional distribution is the generalized error the $\prod^{D}_{m,k_1}$, $\prod^{D}_{m,k_1,k_2}$ and f_D *terms are given by*

$$\prod^{g}_{m,k_1} = \prod_{i=0}^{m-1}\{\sum_{\tau=0}^{\infty}[2^{\frac{1}{v}}\lambda\xi_{1i}]^{\tau}[(\gamma+d)^{\tau}+(\gamma-d)^{\tau}]\frac{\Gamma(\frac{1+\tau}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)}\}, \tag{2.9a}$$

$$\prod^{g}_{m,k_1,k_2} = \prod_{i=0}^{\infty}\{\sum_{\tau=0}^{\infty}[2^{\frac{1}{v}}\lambda\xi_{2i}]^{\tau}[(\gamma+d)^{\tau}+(\gamma-d)^{\tau}]\frac{\Gamma(\frac{1+\tau}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)}\}, \tag{2.9b}$$

$$\lambda = \{2^{\frac{-2}{v}}\Gamma(\frac{1}{v})[\Gamma(\frac{3}{v})]^{-1}\}^{\frac{1}{2}}, \quad f_g = \frac{\Gamma(\frac{2}{v})\lambda 2^{\frac{1}{v}}}{\Gamma(\frac{1}{v})} \tag{2.9c}$$

where ξ_{1i} *and* ξ_{2i} *are given by (2.8); v denotes the degrees of freedom of the generalized error distribution. When* $v > 1$ *the summations in (2.9a) and (2.9b) are finite; when* $v < 1$ *the first summation is finite if and only if* $\xi_{1i}\gamma + |\xi_{1i}d| \leq 0$, *whereas the second summation is finite if and only if* $\xi_{2i}\gamma + |\xi_{2i}d| \leq 0$ *(see Nelson, 1991).*

The above equations, for $v = 1$ *(i.e. the conditional distribution is the double exponential) give*

$$\prod^{d}_{m,k_1} = \prod_{i=0}^{m-1}\frac{2-\sqrt{2}\xi_{1i}\gamma}{2-2\sqrt{2}\xi_{1i}\gamma+\xi_{1i}^2(\gamma^2-d^2)}, \tag{2.10a}$$

$$\prod^{d}_{m,k_1,k_2} = \prod_{i=0}^{\infty}\frac{2-\sqrt{2}\xi_{2i}\gamma}{2-2\sqrt{2}\xi_{2i}\gamma+\xi_{2i}^2(\gamma^2-d^2)}, \tag{2.10b}$$

$$f_d = \frac{1}{\sqrt{2}}, \quad m \neq 0, \tag{2.10c}$$

(2.10a) and (2.10b) hold if and only if $\xi_{1i}\gamma + |\xi_{1i}d| < \sqrt{2}$ and $\xi_{2i}\gamma + |\xi_{2i}d| < \sqrt{2}$, respectively *(see Nelson 1991).*

The proof of Theorem 1a is given in Appendix B.

Theorem 1b. *Under Assumptions 3 and 4 the* $2(k_1+k_2)$*-th moment of the error term, and the auto covariances/correlations between the* k_1*-th and* k_2*-th powers of the error term, are given by*

$$E(\varepsilon_t^{2(k_1+k_2)}) = e^{\widetilde{f}_D}\mu^{D}_{2(k_1+k_2)}\Pi^{D}_{0,k_1,k_2}, \tag{2.11a}$$

$$cov(\varepsilon_t^{2k_1}, \varepsilon_{t-\underset{(m>0)}{m}}^{2k_2}) = e^{\widetilde{f}_D}[\mu^{D}_{2k_1}\Pi^{D}_{m-1,k_1}D_{2k_2,m-1,k_1}\Pi^{D}_{m,k_1,k_2} - \mu^{D}_{2k_1}\mu^{D}_{2k_2}\Pi^{D}_{0,k_1,0}\Pi^{D}_{0,0,k_2}], \tag{2.11b}$$

$$\rho(\varepsilon_t^{2k_1}, \varepsilon_{t-\underset{(m>0)}{m}}^{2k_2}) = \frac{[\mu^{D}_{2k_1}\Pi^{D}_{m-1,k_1}D_{2k_2,m-1,k_1}\Pi^{D}_{m,k_1,k_2} - \mu^{D}_{2k_1}\mu^{D}_{2k_2}\Pi^{D}_{0,k_1,0}\Pi^{D}_{0,0,k_2}]}{\sqrt{[\mu^{D}_{4k_1}\Pi^{D}_{0,2k_1,0} - (\mu^{D}_{2k_1}\Pi^{D}_{0,k_1,0})^2][\mu^{D}_{4k_2}\Pi^{D}_{0,2k_2,0} - (\mu^{D}_{2k_2}\Pi^{D}_{0,k_2,0})^2]}}, \tag{2.11c}$$

$$\widetilde{f}_D = \frac{2(k_1+k_2)[\omega - \gamma f_D\sum_{l=1}^{q}c_l]}{\prod_{f=1}^{p}(1-\beta_f)}, \quad \Pi^{D}_{0,k_1} = 1 \tag{2.11d}$$

where k_1 *and* k_2 *are real positive integers.*

When the conditional distribution is the normal or the generalized error or the double exponential one, the $\prod^{D}_{m,k_1}$, $\prod^{D}_{m,k_1,k_2}$ and f_D *terms are given by (2.3a)-(2.3c), (2.9a)-(2.9c) and (2.10a)-(2.10c),* *respectively, where the* $D^{D}_{2k_2,m-1,k_1}$ and μ^{D}_{2k} *terms are given by*

$$
\begin{aligned}
D^{N}_{2k_2,m-1,k_1} &= \frac{1}{2}\{\frac{\partial}{\partial[\xi_{1,m-1}(\gamma-d)]^{2k_2}}\{e^{\frac{\xi^2_{1,m-1}(\gamma-d)^2}{2}}[1+\Phi(\frac{\xi_{1,m-1}(\gamma-d)}{\sqrt{2}})]\} \\
&\quad +\frac{\partial}{\partial[\xi_{1,m-1}(\gamma+d)]^{2k_2}}\{e^{\frac{\xi^2_{1,m-1}(\gamma+d)^2}{2}}[1+\Phi(\frac{\xi_{1,m-1}(\gamma+d)}{\sqrt{2}})]\}\}, && (2.12a) \\
D^{g}_{2k_2,m-1,k_1} &= 2^{\frac{2k_2}{v}}\lambda^{2k_2}\sum_{\tau=0}^{\infty}(\lambda 2^{\frac{1}{v}}\xi_{1,m-1})^{\tau}[(\gamma+d)^{\tau}+(\gamma-d)^{\tau}]\frac{\Gamma(\frac{\tau+2k_2+1}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)}, && (2.12b) \\
D^{d}_{2k_2,m-1,k_1} &= 2^{-(k_2+1)}\Gamma(2k_2+1)\{F[2k_2+1;\frac{\xi_{1,m-1}(\gamma+d)}{\sqrt{2}}]+F[2k_2+1;\frac{\xi_{1,m-1}(\gamma-d)}{\sqrt{2}})]\} && (2.12c)
\end{aligned}
$$

$$
\mu^{N}_{2k}=\prod_{j=1}^{k}[2k-(2j-1)],\quad \mu^{g}_{2k}=\frac{\Gamma(\frac{1}{v})^{k-1}\Gamma(\frac{2k+1}{v})}{\Gamma(\frac{3}{v})^{k}},\quad \mu^{d}_{2k}=\frac{\Gamma(2k+1)}{2^{k}} \tag{2.13}
$$

In the above expressions Φ *denotes the error function and* $\xi_{1,m-1}(k_1)$ *is given by (2.8). When* $v>1$ *the summation in (2.12b) is finite, whereas when* $v<1$ *the summation is finite if and only if* $\xi_{1,m-1}\gamma+|\xi_{1,m-1}d|\leq 0$. *The right hand side of (2.12c) converges if and only if* $\xi_{1,m-1}\gamma+|\xi_{1,m-1}d|<\sqrt{2}$.

The proof of Theorem 1b is given in Appendix B.

Assumption 5. All the roots of the autoregressive polynomial $\Phi(L)$ lie outside the unit circle.

Assumption 6. The polynomials $\Phi(L)$ and $\Theta(L)$ are left coprime.

In what follows we examine only the case where the roots of the autoregressive polynomial $\Phi(L)$ are distinct.

Theorem 2. *Under Assumptions 3-6 the cross covariances/correlations between the process and the conditional variance/squared errors are given by*

$$
\rho(y_t,h_{t-m}) = \frac{cov(y_t,h_{t-m})}{\sqrt{var(y_t)var(h_t)}} = \begin{cases} \dfrac{\sum\limits_{l=1}^{\infty}\delta_{m+l}\Pi^{D}_{l-1,1}D^{D}_{1,l-1,1}\Pi^{D}_{l,1,\frac{1}{2}}}{[\prod^{D}_{0,1,1}-\prod^{D}_{0,1,0}\prod^{D}_{0,0,1}]\psi} & m\geq 0 \\ \dfrac{\sum\limits_{l=0}^{\infty}\delta_{l}\Pi^{D}_{|m|+l-1,1}D^{D}_{1,|m|+l-1,1}\Pi^{D}_{|m|+l,1,\frac{1}{2}}}{[\prod^{D}_{0,1,1}-\prod^{D}_{0,1,0}\prod^{D}_{0,0,1}]\psi} & m<0 \end{cases},\quad \prod^{D}_{0,1}=1, \tag{2.14a}
$$

$$
\delta_l = \sum_{f=1}^{r}\pi_f s_{fl},\quad \pi_f=\frac{\phi_f^{r-1}}{\prod\limits_{\substack{n=1\\n\neq f}}^{r}(\phi_f-\phi_n)},\quad s_{fl}=\begin{cases}\sum\limits_{n=0}^{l-1}\theta_{l-n}\phi_f^{n}, & l\leq s\\ s_{fs}\phi_f^{l-s}, & l>s\end{cases},\ \delta_0=1 \tag{2.14b}
$$

$$
\psi = \sum_{i=1}^{r}[\sum_{j=0}^{s}\theta_j^2+2\sum_{l=1}^{s}\sum_{j=0}^{s-l}\theta_j\theta_{j+l}\phi_i^{l}]\frac{\pi_i\prod^{D}_{0,1,0}}{\prod\limits_{n=1}^{r}(1-\phi_i\phi_n)}, \tag{2.14c}
$$

$$
p(y_t,\varepsilon^2_{t-m})=\frac{cov(y_t,\varepsilon^2_{t-m})}{\sqrt{var(y_t)var(\varepsilon^2_t)}}=\begin{cases} \dfrac{\sum\limits_{l=1}^{\infty}\delta_{m+l}\Pi^{D}_{l-1,1}D^{D}_{1,l-1,1}\Pi^{D}_{l,1,\frac{1}{2}}}{[\mu_4^{D}\prod^{D}_{0,1,1}-\prod^{D}_{0,1,0}\prod^{D}_{0,0,1}]\psi} & m\geq 0 \\ \dfrac{\sum\limits_{l=0}^{\infty}\delta_{l}\Pi^{D}_{|m|+l-1,1}D^{D}_{1,|m|+l-1,1}\Pi^{D}_{|m|+l,1,\frac{1}{2}}}{[\mu_4^{D}\prod^{D}_{0,1,1}-\prod^{D}_{0,1,0}\prod^{D}_{0,0,1}]\psi} & m<0 \end{cases} \tag{2.15}
$$

When the conditional distribution is the normal, or the generalized error, or the double exponential one, *the* $\Pi^{D}_{l,1}$ *and* $\Pi^{D}_{l,1\frac{1}{2}}$ *terms are given by (2.3a)-(2.3b), (2.9a)-(2.9b), and (2.10a)-(2.10b), respectively, where the* $D^{D}_{1,l,k}$ *term is given by*

$$D^N_{1,l,k} = \frac{1}{2}[\xi_{1l} e^{\frac{\xi^2_{1l}(\gamma+d)^2}{2}}[(\gamma+d) - (\gamma-d)e^{-2\xi^2_{1l}\gamma d}] + \sqrt{\frac{2}{\pi}}\xi^2_{1l}[(\gamma+d)^2 F(1;\frac{3}{2};\frac{(\gamma+d)\xi_{1l}}{\sqrt{2}}) - (\gamma-d)^2 F(1;\frac{3}{2};\frac{(\gamma-d)\xi_{1l}}{\sqrt{2}})], \tag{2.16}$$

$$D^g_{1,l,k} = \sum_{\tau=0}^{\infty}(2^{\frac{1}{v}}\xi_{1l}\lambda)^\tau[(\gamma+d)^\tau - (\gamma-d)^\tau]\frac{\Gamma(\frac{2+\tau}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)}, \tag{2.17a}$$

$$D^d_{1,l,k} = \frac{d\xi_{1l}[\sqrt{2} - \xi_{1l}\gamma]}{\sqrt{2}[1 + \frac{\xi^2_{1l}(\gamma^2-d^2)}{2} - \sqrt{2}\xi_{1l}\gamma]^2}, \quad \text{if } \xi_{1l}\gamma + |\xi_{1l}d| < \sqrt{2} \tag{2.17b}$$

and $\xi_{1l}(k)$ is given by (2.8). When $v > 1$ the summation in (2.17a) is finite, whereas when $v < 1$ the summation is finite if and only if $\xi_{1l}\gamma + |\xi_{1l}d| \leq 0$.

Note that when there is no leverage effect ($d = 0$) the $D^D_{1,l,k}$ is zero and hence the cross correlations between the process and its conditional variance are zero.

The proof of Theorem 2 is given in Appendix C.

3 EMPIRICAL RESULTS

3.1 Data Selections

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. Figure 1 plots the daily returns on the KOSPI, NIKKEI, SE and ST indices.

3.2 Estimation Results

In order to carry out our analysis of stock returns we need to choose a form for the mean equation. Scholes and Williams (1977), Ding, et al. (1993), and Ding and Granger (1996) suggested a MA(1) specification for the mean, Lo and McKinlay (1988), Akgiray (1989) and Nelson (1991) used an AR(1) form, whereas Hentschel (1995) modeled the index return as a white noise process. As a practical matter there is little difference between an AR(1) and a MA(1) model when the AR and the MA coefficients are small and the autocorrelations at lag one are equal, since the higher order autocorrelations die out very quickly in the AR model (Nelson, 1991). We therefore model all the four stock retrns as MA(1) processes. The MA(1) model is

$$y_t = b + (1+\theta L)\varepsilon_t, \quad \varepsilon_t = e_t h_t^{\frac{1}{2}}, \quad e_{t|t-1} \sim IID(0,1) \tag{3.1}$$

To select our 'best' EGARCH specification we begin with low order models (e.g., EGARCH(1,1)) and work upward as required to fit the data. The general EGARCH(p,q) specification that we estimate is

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

Table 2 reports the selected specifications.

The Akaike Information Criterion (AIC) chose high order EGARCH specifications for all indices. When the conditional distribution is the normal, the EGARCH(1,4) specification was chosen in two out of the four indices. For the double exponential conditional distribution, the EGARCH(1,3), specification was chosen for the SE and NIKKEI indices, whereas the EGARCH(2,1) and EGARCH(1,2) specifications were chosen for the KOSPI and the ST indices, respectively. When the conditional distribution is the generalized error, the EGARCH(1,2), EGARCH(2,1) and EGARCH(3,3) specifications were chosen for the SE, ST and KOSPI indices, respectively.

In addition, we use the Likelihood Ratio test to show the performance of the high order models over the simple EGARCH(1,1) model. The tests show the dominance of the high order models with the LR statistics being as shown in Table 3a . These are much bigger than the 5 percent critical values.

For all the stock returns parameter estimation is conducted jointly on a MA(1) mean specification[4] and the appropriate EGARCH model for the conditional variance. Results for the period 1980-1997 are reported in Table 1, which presents parameter estimates along with probability values. The parameters b and θ are the intercept and MA(1) coefficient, respectively, for the return eq. (3.1). The remaining parameters are from the EGARCH model (3.2). Not surprisingly, for all the EGARCH specifications most of the moving average, leverage and autoregressive parameters are significantly different from zero.

For all four indices, parameter estimates are consistent with those generally reported in the literature. In particular, as the persistence measure shows, volatility appears nearly integrated. Persistence is given by the value of the highest root of the autoregressive polynomial and it ranges from .889 to .983 (see Table 3b). 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 2 of Table 3b). In addition, the sum of these products, over all the lagged errors, is negative as well (see column 3 of Table 3b).

3.3 Autocorrelation structure of the estimated models

The EGARCH(1,4), EGARCH(3,3), EGARCH(1,3), EGARCH(2,1) and EGARCH(1,2) specifications are special cases of the EGARCH(4,4) model. The autocorrelations of the conditional variance for the EGARCH(4,4) model are given by (2.2c) where the $\prod_{m,1}^{D}$ and $\prod_{m,1,1}^{D}$ terms, for the normal distribution are given by (2.3a)-(2.3b), and for the generalized error distribution are given by (2.9a)-(2.9b), where the $(\gamma \pm d)\xi_{1i}$ terms are

$$
\begin{aligned}
(\gamma \pm d)\xi_{1i} &= \xi_{1i}^{\pm} = k_1 \sum_{f=1}^{4} \zeta_f \lambda_{f,i+1}^{\pm}, \\
\zeta_f &= \frac{\beta_f^3}{\prod\limits_{\substack{n=1 \\ n\neq f}}^{4} (\beta_f - \beta_n)}, \quad f = 1,2,3,4, \\
\lambda_{f1}^{\pm} &= c_1(1 \pm d_1), \quad \lambda_{f2}^{\pm} = c_2(1 \pm d_2) + c_1(1 \pm d_1)\beta_f, \\
\lambda_{f3}^{\pm} &= c_3(1 \pm d_3) + c_2(1 \pm d_2)\beta_f + c_1(1 \pm d_1)\beta_f^2, \\
\lambda_{f4}^{\pm} &= c_4(1 \pm d_4) + c_3(1 \pm d_3)\beta_f + c_2(1 \pm d_2)\beta_f^2 + c_1(1 \pm d_1)\beta_f^3, \\
\lambda_{fi}^{\pm} &= \lambda_{f4}^{\pm}\beta_f^{i-4}, \quad i \geq 5,
\end{aligned}
$$

and the $(\gamma \pm d)\xi_{2i}$ terms are

$$
(\gamma \pm d)\xi_{2i} = \xi_{2i}^{\pm} = \sum_{f=1}^{4} (k_2 \lambda_{f,i+1}^{\pm} + k_1 \lambda_{f,m+i+1}^{\pm})\zeta_f,
$$

When the conditional distribution is the double exponential one, the $\prod_{m,1}^{D}$ and $\prod_{m,1,1}^{D}$ terms are given

[4]The only exception is the Taiwanese Se weighted index, where the white noise specification is used for the generalized error and the double exponential conditional distributions.

by (2.10a)-(2.10b), where the $\xi_{1i}\gamma$ and $\xi_{1i}d$ terms are

$$\xi_{1i}\gamma = \sum_{f=1}^{4}\zeta_f\lambda_{f,i+1}, \quad ,\lambda_{f,i} = \begin{cases} \sum_{n=0}^{i-1} c_{i-n}\beta_f^n, & i \leq 4 \\ \lambda_{f4}\beta_f^{i-4}, & i > 4 \end{cases},$$

$$\xi_{1i}d = \sum_{f=1}^{4}\zeta_f\widetilde{\lambda}_{f,i+1}, \quad ,\widetilde{\lambda}_{f,i} = \begin{cases} \sum_{n=0}^{i-1} c_{i-n}d_{i-n}\beta_f^n, & i \leq 4 \\ \lambda_{f4}\beta_f^{i-4}, & i > 4 \end{cases}$$

In addition, the $\xi_{2i}\gamma$ and $\xi_{2i}d$ terms are given by

$$\xi_{2i}\gamma = \sum_{f=1}^{4}\zeta_f(\lambda_{f,i+1} + \lambda_{f,m+i+1}), \quad \xi_{2i}d = \sum_{f=1}^{4}\zeta_f(\widetilde{\lambda}_{f,i+1} + \widetilde{\lambda}_{f,i+m+1})$$

Figure 2 plots the theoretical autocorrelations of the estimated conditional variance for all the aforementioned EGARCH models.

4 CONCLUSIONS

In this paper we obtained a complete characterization of the moment structure of the general ARMA(r,s)-GARCH(p,q) model. In particular, we provided the autocorrelation function and any arbitrary moment of the squared errors and the conditional variance. In addition, we derived the cross correlations between the process and the conditional variance/squared errors. To obtain our results we assumed that the error term is drawn from either the normal or the double exponential or the generalized error distributions. The techniques used in this paper can be applied to obtain the moments of more complex EGARCH models like the EGARCH-in-mean model, the Component EGARCH model and the Fractional Integrated EGARCH model. The derivation of the moment structure of these models is left for future research.

REFERENCES

Abadir, K. M., 1999, An introduction to hypergeometric functions for economists, Econometric Reviews, 18, 287-330.

Akgiray, V., 1989, Conditional heteroskedasticity in time series of stock returns: evidence and forecasts, Journal of Business 62, 55-80.

Andersen, T. G., Chung, H. J. and B. E. Sorensen, 1999, Efficient method of moments estimation of a stochastic volatility model: a monte carlo study, Journal of Econometrics, 91, 61-87.

Andersen, T. G. and J. Lund, 1997, Estimating continuous-time stochastic volatility models of the short-term interest rate, Journal of Econometrics, 77, 343-377.

Black, F., 1976, Studies of stock price volatility changes, Proceedings of the 1976 meetings of the Business and Economics Statistics Section, American Statistical Association, 177-181.

Booth, G. G., Martikainen, T. and Y. Tse, 1997, Price and volatility spillovers in Scandinavian stock markets, Journal of Banking and Finance, 21, 811-823.

Breidt, F. J., Carto, N. and P. de Lima, 1998, The detection and estimation of long memory in stochastic volatility, Journal of Econometrics, 83, 325-348.

Brunner, A. D. and D. P. Simon, 1996, Excess returns and risk at the long end of the treasury market: an EGARCH-M approach, The Journal of Financial Research, 14, 3, 443-457.

Deb, P., 1996, Finite sample properties of maximum likelihood and quasi-maximum likelihood estimators of EGARCH models, Econometric Reviews, 15, 51-68.

Demos, A., 2000, The autocorrelation function of conditionally heteroscedastic in mean models, Department of Economics, Athens University of Economics and Business, Discussion Paper, 2000-07.

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., Granger C. W. J. 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, 5, 1749-1778

Episcopos, A., 1996, Stock return volatility and time-varying betas in the Toronto stock exchange, QJBE, 35, 28-38.

Fountas, S., Karanasos, M. and M. Karanassou, 2000, A GARCH model of inflation and inflation uncertainty with simultaneous feedback, Department of Economics and Related Studies, University of York, Discussion Paper, 2000-24.

Glosten, L. R., Jagannathan, R. and D. E. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, Journal of Finance, 5, 1779-1801.

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

Hafner, C. M., 1998, Estimating high-frequency foreign exchange rate volatility with nonparametric ARCH models, Journal of Statistical Planning and Inference, 68, 247-269.

Hagerud, G. E., 1997, Modeling Nordic stock returns with asymmetric GARCH models, Department of Finance, Stockholm School of Economics, Working Paper Series in Economics and Finance No 164.

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

Hu, M. Y., Jiang, C. X. 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, 235-253.

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., 2000, Prediction in ARMA models with GARCH-in-mean effects, forthcoming in the Journal of Time Series Analysis.

Karanasos, M. and J. Kim, 2000, Alternative GARCH in mean models: an application to the Korean stock market, Department of Economics and Related Studies, University of York, Discussion Paper, 2000-25.

Kavussanos, M. G. and A. H. Alizadeh-M, 1999, The expectations hypothesis of the term structure and risk premia in dry bulk shipping freight markets: an EGARCH-M approach, Unpublished Mimeo, Department of Shipping, Trading and Finance, City University Business School.

Kim, D. and S. J. Kon, 1994, Alternative models for the conditional heteroscedasticity of stock returns, Journal of Business, 67, 563-598.

Kim, S. J., 1998, Do Australian and the US macroeconomic news announcements affect the USD/AUD exchange rate? Some evidence from E-GARCH estimations, Journal of Multinational Financial Management, 8, 233-248.

Kim, S. J. and J. Sheen, 2000, International linkages and macroeconomic news effects on interest rate volatility -Australia and the US, Pacific-Basin Finance Journal, 8, 85-113.

Koutmos, G. and G. G. Booth, 1995, Asymmetric volatility transmission in international stock markets, Journal of International Money and Finance, 14, 747-762.

Kusi, J. A. and G. M. Pescetto, 1998, Volatility and volatility spill-overs in emerging markets: the case of the African stock markets, Ekonomia, 2, 171-185.

Lee, J. H. and B. W. Brorsen, 1997, A non-nested test of GARCH vs. EGARCH models, Applied Economics Letters, 4, 765-768.

Lo, A. W. and A. C. MacKinlay, 1988, Stock market prices do not follow random walks: evidence from a simple specification test, Review of Financial Studies, 1, 41-66.

Lobo, B. J. and D. Tufte, 1998, Exchange rate volatility: does politics matter?, Journal of Macroeconomics, 20, 351-365.

Nelson, D. B., 1991, Conditional heteroskedasticity in asset returns: A new approach, Econometrica, 59, 347-370.

Pagan, A. and W. Schwert, 1990, Alternative models for conditional stock volatility, Journal of Econometrics, 45, 267-290.

Pierre, E. F. ST., 1998, Estimating EGARCH-M models: science or art?, The Quarterly Review of Economics and Finance, 38, 167-180.

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

Reyes, M. G., 1996, Index futures trading and stock price volatility: evidence from Denmark and France, Journal of Economics and Finance, 20, 81-88.

Scholes, M. and J. Williams, 1977, Estimating betas from nonsynchronous data, Journal of Financial Economics 5, 309-327.

Tse, Y., 1998, International transmission of information: evidence from the Euroyen and Eurodollar futures markets, Journal of International Money and Finance, 17, 909-929.

Tse, Y. and G. G. Booth, 1996, Common volatility and volatility spillovers between U.S. and Eurodollar interest rates: evidence from the futures market, Journal of Economics and Business, 48, 299-312.

Zakoian, J. M., 1994, Threshold heteroskedastic models, Journal of Economic Dynamics and Control, 18, 931-955.

APPENDIX

A Proof of Propositions 1 and 2

PROOF OF PROPOSITION 1

Using the Wold representation of an ARMA model (as it is given in Karanasos, 2000) and the EGARCH(1,q) conditional variance (eqs 2.1a-2.1d), we have

$$\ln(h_t) = \frac{\omega}{1-\beta_1} + \sum_{l=1}^{\infty} \lambda_{1l} z_{t-l} \Rightarrow \begin{cases} h_t^{k_1} = e^{\frac{\omega k_1}{1-\beta_1}} \times e^{k_1 \sum_{l=1}^{\infty} \lambda_{1l} z_{t-l}} \\ h_{t-m}^{k_2} = e^{\frac{\omega k_2}{1-\beta_1}} \times e^{k_2 \sum_{l=1}^{\infty} \lambda_{1l} z_{t-m-l}} \end{cases} \tag{A.1}$$

where λ_{1l} is defined in (2.3e).
Multiplying $h_t^{k_1}$ by $h_{t-m}^{k_2}$ and taking expectations gives

$$E(h_t^{k_1} h_{t-m}^{k_2}) = \varrho^{\frac{\omega(k_1+k_2)}{1-\beta_1}} \times E(\varrho^{k_1 \sum_{i=1}^{m} \lambda_{1i} z_{t-i}}) \times E(\varrho^{\sum_{i=1}^{\infty} (k_1 \lambda_{1,m+i} + k_2 \lambda_{1i}) z_{t-i-m}}) \tag{A.2}$$

Moreover, from (A.1) it follows that

$$E(h_t^{k_1}) = \varrho^{\frac{\omega k_1}{1-\beta_1}} E(\varrho^{k_1 \sum_{i=1}^{\infty} \lambda_{1i} z_{t-i}}) \tag{A.3}$$

When the conditional distribution is the normal, we use formula 2.3.15 #7 in Prudnikov, et al (Volume 1, 1992), to obtain the expected value of $e_t^k e^{z_t b}$

$$\begin{aligned} E(e_t^k e^{z_t b}) &= \frac{e^{-\gamma b \sqrt{\frac{2}{\pi}}}}{2} \{(-1)^k \frac{\partial}{\partial [b(\gamma - d)]^k} \{ e^{\frac{b^2(\gamma-d)^2}{2}} [1 + \Phi(\frac{b(\gamma-d)}{\sqrt{2}})]\} \\ &+ \frac{\partial}{\partial [b(\gamma+d)]^k} \{ e^{\frac{b^2(\gamma+d)^2}{2}} [1 + \Phi(\frac{b(\gamma+d)}{\sqrt{2}})]\}\} \end{aligned} \tag{A.4}$$

for $k = 0, 1, 2 \ldots$
In the above expression Φ denotes the error function (formula 8.250 #1 in Gradshteyn and Ryzhik, 1994). Using formula 8.253 # 1 in Gradshteyn and Ryzhik (1994), the above expression, for $k = 0$, gives

$$\begin{aligned} E(e^{z_t b}) &= e^{-b\gamma \sqrt{\frac{2}{\pi}}} \{ \frac{1}{2} e^{\frac{b^2(\gamma+d)^2}{2}} [1 + e^{-2\gamma d b^2}] + \\ & \frac{b(\gamma+d)}{\sqrt{2\pi}} F(1; \frac{3}{2}; \frac{b^2(\gamma+d)^2}{2}) + \frac{b(\gamma-d)}{\sqrt{2\pi}} F(1; \frac{3}{2}; \frac{b^2(\gamma-d)^2}{2}) \} \end{aligned} \tag{A.5}$$

where F denotes the hypergeometric function (for an alternative derivation see Theorem A1.1 in Nelson, 1991).
Combining equations (A.2), (A.3) and (A.5), after some algebra, yields (2.2a)-(2.3e). ■

PROOF OF PROPOSITION 2

Using Karanasos's (2000) results and the EGARCH(p,1) conditional variance (eqs. 2.1a-2.1b, 2.4a-2.4b), we have

$$\ln(h_t) = \frac{\omega}{\prod_{f=1}^{p}(1-\beta_f)} + c_1 \sum_{l=0}^{\infty} \sum_{f=1}^{p} \zeta_f \beta_f^l z_{t-1-l} \Rightarrow \begin{cases} h_t^{k_1} = e^{\frac{\omega k_1}{\prod_{f=1}^{p}(1-\beta_f)}} \times e^{k_1 c_1 \sum_{l=0}^{\infty} \sum_{f=1}^{p} \zeta_f \beta_f^l z_{t-1-l}} \\ h_{t-m}^{k_2} = e^{\frac{\omega k_2}{\prod_{f=1}^{p}(1-\beta_f)}} \times e^{k_2 c_1 \sum_{l=0}^{\infty} \sum_{f=1}^{p} \zeta_f \beta_f^l z_{t-1-m-l}} \end{cases}, \tag{A.6}$$

where ζ_f is defined in (2.5c).
Multiplying $\mathrm{h}_t^{k_1}$ by $\mathrm{h}_{t-m}^{k_2}$ and taking expectations yields

$$E(h_t^{k_1} h_{t-m}^{k_2}) = \varrho^{\frac{\omega(k_1+k_2)}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(\varrho^{c_1 k_1 \sum_{i=1}^{m}\sum_{f=1}^{p}\zeta_f \beta_f^{i-1} z_{t-i}}) \times E(\varrho^{c_1 \sum_{i=1}^{\infty}\sum_{f=1}^{p}(k_1\beta_f^m + k_2)\zeta_f \beta_f^{i-1} z_{t-i-m}}) \tag{A.7}$$

Furthermore, from (A.6) it follows that

$$E(h_t^{k_1}) = \varrho^{\frac{\omega(k_1+k_2)}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(\varrho^{c_1 k_1 \sum_{i=1}^{\infty}\sum_{f=1}^{p}\zeta_f \beta_f^{i-1} z_{t-i}}) \tag{A.8}$$

Combining equations (A.7), (A.8), and (A.5), after some algebra, gives (2.2a)-(2.2c), (2.3a)-(2.3c), (2.5a) and (2.5b). ■

B Proof of Theorems 1a and 1b

PROOF OF THEOREM 1a

Using the Wold representation of an ARMA model (as it is given in Karanasos, 2000) we express the k_1-th power of the EGARCH(p,q) conditional variance (eqs. 2.1a-2.1c, 2.4a) as

$$\ln(h_t) = \frac{\omega}{\prod_{f=1}^{p}(1-\beta_f)} + \sum_{l=1}^{\infty}\sum_{f=1}^{p}\zeta_f \lambda_{fl} z_{t-l} \Rightarrow \begin{cases} h_t^{k_1} = e^{\frac{\omega k_1}{\prod_{f=1}^{p}(1-\beta_f)}} \times e^{k_1 \sum_{l=1}^{\infty}\sum_{f=1}^{p}\zeta_f \lambda_{fl} z_{t-l}} \\ h_{t-m}^{k_2} = e^{\frac{\omega k_2}{\prod_{f=1}^{p}(1-\beta_f)}} \times e^{k_2 \sum_{l=1}^{\infty}\sum_{f=1}^{p}\zeta_f \lambda_{fl} z_{t-m-l}} \end{cases} \tag{B.1}$$

Multiplying $\mathrm{h}_t^{k_1}$ by $\mathrm{h}_{t-m}^{k_2}$ and taking expectations yields

$$E(h_t^{k_1} h_{t-m}^{k_2}) = \varrho^{\frac{\omega(k_1+k_2)}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(\varrho^{k_1 \sum_{i=1}^{m}\sum_{f=1}^{p}\zeta_f \lambda_{fi} z_{t-i}}) \times E(\varrho^{\sum_{i=1}^{\infty}\sum_{f=1}^{p}(k_1\lambda_{f,m+i} + k_2\lambda_{fi})\zeta_f z_{t-i-m}}) \tag{B.2}$$

In addition, from (B.1) it follows that

$$E(h_t^{k_1}) = \varrho^{\frac{\omega(k_1+k_2)}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(\varrho^{k_1 \sum_{i=1}^{\infty}\sum_{f=1}^{p}\lambda_{fi}\zeta_f z_{t-i}}) \tag{B.3}$$

Next, when the conditional distribution is the normal, using equations (B.2), (B.3), and (A.5), after some algebra, gives equations (2.2a)-(2.2c), (2.3a)-(2.3c), (2.7) and (2.8).

Moreover, when the conditional distribution is the generalized error, the expected value of $e_t^k \varrho^{z_t b}$ is given by expression A1.5 in Theorem A1.2 in Nelson (1991):

$$E[e_t^k \varrho^{z_t b}] = \varrho^{-b\gamma \frac{\Gamma(\frac{2}{v})\lambda 2^{\frac{1}{v}}}{\Gamma(\frac{1}{v})}} 2^{\frac{k}{v}} \lambda^k \sum_{\tau=0}^{\infty} (\lambda 2^{\frac{1}{v}} b)^{\tau} [(\gamma+d)^{\tau} + (-1)^k (\gamma-d)^{\tau}] \frac{\Gamma(\frac{\tau+k+1}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)} \tag{B.4}$$

where k is a nonnegative integer. 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 equations (B.2), (B.3), and (B.4) for $k=0$, after some algebra, yields equations (2.2a)-(2.2c), and (2.9a)-(2.9c).

Furthermore, the above equation, for $v=1$, gives

$$E[e_t^k \varrho^{z_t b}] = 2^{\frac{-(k+2)}{2}} \varrho^{-b\gamma\frac{1}{\sqrt{2}}} \Gamma(k+1)\{F[k+1; \frac{b(\gamma+d)}{\sqrt{2}}] + (-1)^k F[k+1; \frac{b(\gamma-d)}{\sqrt{2}}]\} \tag{B.5}$$

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

$$E[\varrho^{z_t b}] = \frac{1}{2}\varrho^{-b\gamma\frac{1}{\sqrt{2}}} \frac{2-\sqrt{2}b\gamma}{1-\sqrt{2}b\gamma+\frac{b^2(\gamma^2-d^2)}{2}} \tag{B.6}$$

Finally, (B.5), for $k=1$, gives

$$E[e_t \varrho^{z_t b}] = \varrho^{-b\gamma\frac{1}{\sqrt{2}}} \frac{bd(\sqrt{2}-b\gamma)}{\sqrt{2}[1-\sqrt{2}b\gamma+\frac{b^2(\gamma^2-d^2)}{2}]^2} \tag{B.7}$$

Using equations (B.2), (B.3), and (B.6), after some algebra, gives equations (2.2a)-(2.2c), and (2.10a)-(2.10c).

■

PROOF OF THEOREM 1b

>From (B.1) it follows that the expected value of $\varepsilon_t^{2k_1}\varepsilon_{t-m}^{2k_2}$ is

$$\begin{aligned} E(\varepsilon_t^{2k_1}\varepsilon_{t-m}^{2k_2}) &= E(e_t^{2k_1}h_t^{k_1}e_{t-m}^{2k_2}h_{t-m}^{k_2}) = \varrho^{\frac{\omega(k_1+k_2)}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(e_t^{2k_1}) \times E(\varrho^{k_1\sum_{i=1}^{m-1}\sum_{f=1}^{p}\zeta_f\lambda_{fi}z_{t-i}}) \times \\ &\times E(e_{t-m}^{2k_2} e^{k_1\sum_{f=1}^{p}\zeta_f\lambda_{fn}z_{t-m}}) \times E(\varrho^{\sum_{i=1}^{\infty}\sum_{f=1}^{p}(k_1\lambda_{f,m+i}+k_2\lambda_{fi})\zeta_f z_{t-i-m}}), \quad m>0\ , \end{aligned} \tag{B.8a}$$

$$E(\varepsilon_t^{2(k_1+k_2)}) = \varrho^{\frac{\omega(k_1+k_2)}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(e_t^{2(k_1+k_2)}) \times E(\varrho^{\sum_{i=1}^{\infty}\sum_{f=1}^{p}(k_1\lambda_{f,i}+k_2\lambda_{fi})\zeta_f z_{t-i}}) \tag{B.8b}$$

Next, when the conditional distribution is the normal, combining (B.8a), (B.8b) and (A.5), after some algebra, yields equations (2.11a)-(2.11d), (2.3a)-(2.3c) and (2.12a).

Moreover, when the conditional distribution is the generalized error, using expressions (B.8a), (B.8b) and (B.4), after some algebra, gives (2.11a)-(2.11d), (2.9a)-(2.9c) and (2.12b).

Finally, when the conditional distribution is the double exponential, combining equations (B.8a), (B.8b) and (B.5), after some algebra, yields (2.11a)-(2.11d), (2.10a)-(2.10c) and (2.12c).

■

C Proof of Theorem 2

Using Karanasos's (2000) results, the Wold representation of the ARMA(r,s) process (eq 2.6a) and its conditional variance (eqs 2.1a-2.1c, 2.4a) are

$$y_t = b^* + e_t h_t^{\frac{1}{2}} + \sum_{l=1}^{\infty}\delta_l e_{t-l}h_{t-l}^{\frac{1}{2}}, \quad b^* = \frac{b}{\prod_{l=1}^{r}(1-\phi_l)} \tag{C.9}$$

$$h_{t-m} = e^{\frac{\omega}{\prod_{f=1}^{p}(1-\beta_f)}} \times e^{\sum_{l=1}^{\infty}\sum_{f=1}^{p}\zeta_f\lambda_{fl}z_{t-m-l}} \tag{C.10}$$

where δ_l is defined in (2.14b).

Furthermore, the expected values of $e_{t-m-n}h^{\frac{1}{2}}_{t-m-n}h_{t-m}$ and $e_{t-m-n}h^{\frac{1}{2}}_{t-m-n}h_t$ are

$$
\begin{aligned}
E(e_{t-m-n}h^{\frac{1}{2}}_{t-m-n}h_{t-\underset{m>0}{m}}) &= e^{\frac{1.5\omega}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(e^{\sum_{l=1}^{n-1}\sum_{f=1}^{p}\zeta_f\lambda_{fl}z_{t-m-l}}) \times \\
&\times E(e_{t-m-n}e^{\sum_{f=1}^{p}\zeta_f\lambda_{fn}z_{t-m-n}}) \times E(\varrho^{\sum_{l=1}^{\infty}\sum_{f=1}^{p}(\lambda_{f,n+l}+.5\lambda_{fl})\zeta_f z_{t-l-m-n}}),
\end{aligned} \tag{C.11a}
$$

$$
\begin{aligned}
E(e_{t-m-n}h^{\frac{1}{2}}_{t-m-n}h_t) &= e^{\frac{1.5\omega}{\prod_{f=1}^{p}(1-\beta_f)}} \times E(e^{\sum_{l=1}^{m+n-1}\sum_{f=1}^{p}\zeta_f\lambda_{fl}z_{t-l}}) \times \\
&\times E(e_{t-m-n}e^{\sum_{f=1}^{p}\zeta_f\lambda_{f,m+n}z_{t-m-n}}) \times E(\varrho^{\sum_{l=1}^{\infty}\sum_{f=1}^{p}(\lambda_{f,m+n+l}+.5\lambda_{fl})\zeta_f z_{t-l-m-n}})
\end{aligned} \tag{C.11b}
$$

for $n = 1, \ldots, \infty$.

Note that the variance of y_t is (see Karanasos, 2000)

$$
var(y_t) = \sum_{i=1}^{r}[\sum_{j=0}^{s}\theta_j^2 + 2\sum_{l=1}^{s}\sum_{j=0}^{s-l}\theta_j\theta_{j+l}\phi_i^l]\frac{\pi_i\prod_{0,1,0}^{D}E(h_t)}{\prod_{n=1}^{r}(1-\phi_i\phi_n)} \tag{C.12}
$$

where π_i is given by (2.14b).

Next, when the conditional distribution is the normal, combining equations (C.9)-(C.12), (A.4) for $k = 1$, and (A.5), after some algebra, gives (2.14a)-(2.14c) and (2.16).

Moreover, when the conditional distribution is the generalized error, using equations (C.9)-(C.12), and (B.4), after some algebra, yields (2.14a)-(2.14c) and (2.17a).

Finally, when the conditional distribution is the double exponential, combining equations (C.9)-(C.12), (B.6) and (B.7), after some algebra, gives (2.14a)-(2.14c) and (2.17b).

The derivation of $p(y_t, \varepsilon^2_{t-m})$ is similar to that of $p(y_t, h_{t-m})$.

■