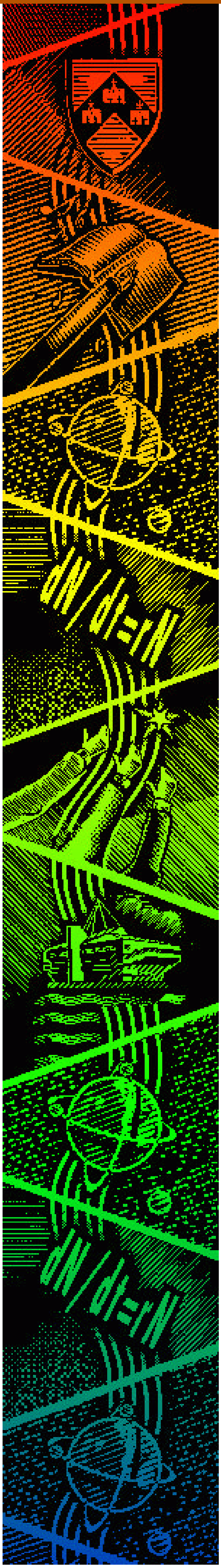

THE UNIVERSITY of York

Discussion Papers in Economics

No. 2000/25

Alternative GARCH in Mean Models: An Application to the Korean Stock Market

by

Menelaos Karanasos and J Kim

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

ALTERNATIVE GARCH IN MEAN MODELS:
AN APPLICATION TO THE KOREAN STOCK MARKET

M KARANASOS AND J KIM

University of York, Heslington, York, YO10 5DD, UK

Abstract

The purpose of this paper is the theoretical and empirical comparison of alternative GARCH-in-mean models. We examine three GARCH specifications: Bollerslev's (1986) GARCH model, Taylor (1986) - Schwert's (1989) GARCH model, and Nelson's (1991) Exponential GARCH model. In addition, we employ four of the most common forms in which the time-varying variance enters the specification of the mean to determine the risk premium: the quadratic, the linear, the logarithmic and the square root one. For all the aforementioned models we give the auto/cross correlations of the process and its conditional variance. The practical implications of the results are illustrated empirically using daily data on the Korean Stock Price Index (KOSPI).

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

.

1 INTRODUCTION

One of the most common mean equations for excess asset returns is the GARCH-in-mean (GARCH-M) model of Engle, Lilien and Robins (1987). Mean equations of this form have been widely used in empirical studies of time-varying risk-premia, in the term structure (Engle and Ng, 1993, Hurn, McDonald and Moody, 1995, Brunner and Simon, 1996, Kavussanos and Alizadeh, 1999), in forward and future prices of commodities (Hall, 1991, Moosa and Loughani, 1994), in industrial production (Caporale and McKierman, 1996) and especially in stock returns (Black and Fraser 1995, Fraser 1996, Hansson and Hordahl, 1997, Elyasiani and Mansur, 1998, Ellul, 1999).

In particular, the empirical literature has attempted to characterize the intertemporal relation between risk and return on the stock market. However, the reported findings are conflicting. For example, French, Schwert and Stambaugh (1987) and Campbell and Hentschel (1992) conclude that the data are consistent with a positive relation between conditional expected excess return and conditional variance, whereas Fama and Schwert (1977), Campbell (1987), Breen, Glosten and Jagannathan (1989), Turner, Starz and Nelson (1989), Nelson (1991), Pagan and Hong (1991) and Glosten, Jagannathan and Runckle (1993) find a negative relation. Although there is a large body of empirical research on the trade off between risk and return there have been relatively fewer theoretical advancements. This paper contributes to both the theoretical and empirical developments in this research area.

In order to carry out our analysis of stock returns we need to choose a form for the mean and variance equations. Scholes and Williams (1977) suggested a MA(1) specification for the mean, Lo and McKinlay (1988) and Nelson (1991) used an AR(1) form, while Hentschel (1995) modeled the index return as a white noise process. Here, we adopt the ARMA(1,1) form which includes the white noise, MA(1), and AR(1) specifications as special cases. For the variance equation we examine three alternative GARCH processes: (i) Bollerslev's GARCH(1,1) model, (ii) Taylor/Schwert's model where the conditional standard deviation ($h_t^{\frac{1}{2}}$) follows a GARCH(1,1) process and (iii) Nelson's Exponential GARCH(1,1) [EGARCH(1,1)] model where the logarithm of the conditional variance [$\ln(h_t)$] follows an ARMA(1,1) process.

Furthermore, we need to choose the form in which the time varying variance enters the specification of the mean to determine the risk premium. This is a matter of empirical evidence. Although a number of authors assumed linearity between the conditional variance and mean return (see, for example, Nelson, 1991, Glosten, Jagannathan and Runckle, 1993, and Hentschel, 1995) the theoretical justification for this linear relationship is meagre, since the required excess return on a portfolio is linear in its conditional variance only under very special circumstances. In many other applications the square root of the conditional variance ($\sqrt{h_t}$) has been used (see for example Domowitz and Hakkio, 1985, Bollerslev, Engle and Wooldridge, 1988, Kavussanos and Alizadeh, 1999 and Ellul, 1999). However, Engle, Lilien and Robins (1987) found that the logarithm of the conditional variance [$\ln(h_t)$] worked better in their estimation of the time varying risk premia in the term structure, whereas Kroner and Lastrapes (1993) used the square of the conditional variance (h_t^2) as a regressor in their preferred model. Following the above discussion, we model the relationship between the expected return and the time varying risk premium by assuming (i) linearity, i.e. we use h_t, $h_t^{\frac{1}{2}}$, and $\ln(h_t)$ as a regressor in Bollerslev's, Taylor/Schwert's and Nelson's models, respectively, and (ii) nonlinearity, i.e. we use h_t^2, h_t, and h_t^k in Bollerlsev's, Taylor/Schwert's and Nelson's[1] models, respectively.

To obtain the theoretical results and to carry out the estimation we need to make a distributional assumption for the error term[2]. In the light of empirical evidence of fat-tail errors, several authors (Bollerslev, 1987, Nelson, 1991, and Sentana, 1995) have chosen distributions such as the Student-t distribution or the Generalized Error Distribution[3]. We should note that for the GARCH models our theoretical results hold under any distributional assumption, whereas for the EGARCH model we assume that the error term is drawn from either the Normal, or the Double Exponential or the Generalized Error Distribution.

The rest of the paper is organized as follows. In Section 2 we present the theoretical results for the above GARCH-M models. First, for the GARCH specifications we give their ARMA representations, their autocorrelation functions and the condition for the existence of their second moments[4]. Second, we obtain the ARMA

[1] In the case of the EGARCH(1,1) model with the k-th power of the conditonal variance in the mean equation we only examine a MA(1) mean specification.

[2] In addition to the potential gains in efficiency, the exact form of the error distribution also plays an important role in several applications of the GARCH model such as option pricing and the construction of optimal forecast error intervals; see Engle and Mustafa (1992) and Baillie and Bollerlsev (1992).

[3] Other parametric densities that have been considered in the estimation of GARCH models include the normal-Poison mixture distribution in Jorion (1988), the power exponential distibution in Baillie and Bollerslev (1989), the normal -lognormal mixture distribution in Hsieh (1989) (see also, for an excellent analysis of a variety of non-normal distributions, Paolella, 1999).

[4] The moment structure of GARCH models is a topic that has recently attracted plenty of attention. Karanasos (1999) examined

representations and the autocorrelation functions of the stock returns. Finally, we give the cross correlations between the stock return and the GARCH specifications[5].

In Section 3 we proceed with the estimation of models from the GARCH-in-mean family in order to take into account the serial correlation, the GARCH effects and the time varying risk premium observed in our time series data. Our empirical analysis is based on daily data for the price of the Korean Stock Index. For the estimation of the GARCH(1,1) specification and the in-mean effect we use three alternative functional forms: the logarithmic, the linear and the square root one[6]. All our estimated models suggest a positive and statistically significant relation between risk and return.

2 GARCH-IN-MEAN MODELS

2.1 GARCH-in-mean-models

GARCH models have been applied in modeling the relation between conditional variance and asset risk premia. In what follows, we examine an ARMA(1,1) form for the mean equation which includes both the AR(1) and MA(1) forms as special cases. The AR(1) term allows for the autocorrelation induced by discontinuous trading in the stocks making up an index (Scholes and Williams, 1977, Lo and Mckinlay, 1988). In addition, for the variance equation, we assume that either the conditional variance is a linear function of the lagged squared error as in Bollerslev's GARCH(1,1) model or that the conditional standard deviation is a linear function of the lagged absolute error as in Taylor/Schwert's GARCH(1,1)[7] model. In the latter model we also include an asymmetric term[8]. Finally, in Bollerslev's model we include the variance in the mean equation, whereas in Taylor/Schwert's model we use the standard deviation as a regressor.

- **Model 1**. Consider the ARMA(1,1)-GARCH(1,1)-in-mean model

$$(1-\phi L)r_t = b + (1-\theta L)\varepsilon_t + \varsigma h_t^{\frac{\delta_i}{2}}, \;\; \varepsilon_t = e_t h_t^{\frac{1}{2}}, \;\; e_{t|t-1} \sim IID(0,1) \tag{2.1}$$

$$h_t^{\frac{\delta_i}{2}} = \omega + af(\varepsilon_{t-1}) + \beta h_{t-1}^{\frac{\delta_i}{2}}, \;\; f(\varepsilon_{t-1}) = [|\varepsilon_{t-1}| + d_i\varepsilon_{t-1}]^{\delta_i} = h_{t-1}^{\frac{\delta_i}{2}}[|e_{t-1}| + d_i e_{t-1}]^{\delta_i}, \tag{2.2a}$$

$$\delta_i = \begin{cases} 2 & \text{if} \quad i=a \\ 1 & \text{if} \quad i=b \end{cases}, \quad d_i = \begin{cases} 0 & \text{if} \quad i=a \\ d & \text{if} \quad i=b \end{cases}, \;\; \omega, a, \beta \geq 0, \;\; -1 < d < 1 \tag{2.2b}$$

Note that when i=a ($\delta_a = 2, \; d_a = 0$) we have Bollerslev's GARCH-in-mean model (Model 1a), and when i=b ($\delta_b = 1, \; d_b = d$) we have Taylor/Schwert's asymmetric GARCH-in-mean model (Model 1b).

In these models the $\frac{\delta_i}{2}$-th power of the conditional variance follows a GARCH(1,1) process and it also enters the specification for the mean. In other words, in these models of volatility feedback the long-horizon forecasts are linearly related to short-horizon forecasts of some volatility measure and the expected stock return is linear in this same measure.

the moment structure of the squared errors for the standard GARCH model (see also, He and Teräsvirta, 1999). Karanasos (2000a) derived the auto/cross covariance functions of the aggregate variance and the component variances for the N component GARCH(n,n) model. K (2000b) and Fountas, Karanasos and Karanassou (2000) derived expressions for the autocovariances of the conditional variance and obtained the cross covariances between the process and its conditional variance for the ARMA-GARCH-in-mean and the ARMA-GARCH-in-mean-level models, respectively.

[5] While we were writing this paper, the work by He, Teräsvirta and Malmsten (1999), hereafter HTM, came to our attention. HTM examined only the conditional variance in the context of the simple EGARCH(1,1) model whereas we derive results for both the conditional variance and mean in the context of an EGARCH(1,1)-in-mean model. We should also note that our research has been contacted independently of HTM.

[6] Taylor(1986) and Schwert(1989) first suggested ARCH models for the conditional standard deviation. Nelson and Foster (1994) show that a GARCH extension of the Taylor/Schwert ARCH model is a consistent estimator of the conditional variance of near diffusion processes. In the context of a power GARCH model, Ding, Granger and Engle (1993) using daily data on the S&P 500 index estimate a power of 1.43 which was significantly different from 1 or 2; on the other hand, Hentschel (1995) using U.S stock return data estimate a power of 1.131 and found it to be insignificantly different from 1. Although in our empirical analysis we estimate a power GARCH model in our theoretical analysis we restrict our attention to either Bollerlsev's or Taylor/Schwert's models.

[7] Our choice of the first-order model is motivated by the fact that is the most widely applied GARCH model.

[8] 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).

Proposition 1. *The ARMA(1,1) representation of the* $\frac{\delta_i}{2}$-th power of the conditional variance $(h_t^{\frac{\delta_i}{2}})$ *is*

$$\begin{aligned}(1-\beta_1 L)h_t^{\frac{\delta_i}{2}} &= \omega + a v_{1,t-1}, \ \ \beta_1 = a k_1 + \beta, \ \ v_{1,t-1} = f(\varepsilon_{t-1}) - k_1 h_{t-1}^{\frac{\delta_i}{2}}, & (2.3a)\\ k_1 &= E[f(e_{t-1})] = E[|e_{t-1}| + d_i e_{t-1}]^{\delta_i} & (2.3b)\end{aligned}$$

The stationarity condition for $h_t^{\frac{\delta_i}{2}}$ *is* $\beta_1 < 1$.
In addition, the univariate ARMA representation of the stock return (r_t) *is*

$$(1-\phi L)(1-\beta_1 L)r_t = [b(1-\beta_1) + \varsigma\omega] + (1-\theta L)(1-\beta_1 L)\varepsilon_t + \varsigma a v_{1t-1} \qquad (2.4)$$

The stationarity condition for r_t *is* $\beta_1, |\phi| < 1$. *In what follows we examine only the case where* $\phi \neq \beta_1$.
Moreover, the auto/cross covariances/correlations of the stock return (r_t*) and the* $\frac{\delta_i}{2}$*-th power of its conditional variance (*$h_t^{\frac{\delta_i}{2}}$*) are*

$$\begin{aligned}\rho_m(r_t) &= \frac{\gamma_{rm}}{\gamma_{r0}} = \frac{cov(r_t, r_{t-m})}{var(r_t)}, & (2.5a)\\ cov(r_t, r_{t-m}) &= \gamma_{rm} = \frac{\phi^{|m|}}{1-\phi^2}[1+\theta^2-\theta(\phi+\phi^{-1})]E(h_t) + \\ & \frac{1}{(1-\phi\beta_1)(\phi-\beta_1)}[\frac{\phi^{|m|+1}}{1-\phi^2} - \frac{\beta_1^{|m|+1}}{1-\beta_1^2}]\varsigma^2 a^2[k_2-k_1^2]E(h_t^{\delta_i}) + \\ & [\frac{\phi^{|m|}[(1+\phi^2)(1+\beta_1\theta) - 2\phi(\theta+\beta_1)]}{(1-\phi^2)(1-\phi\beta_1)(\phi-\beta_1)} + \\ & + \frac{\beta_1^{|m|+1}(-\theta+\beta_1^{-1})}{(\beta_1-\phi)(1-\beta_1\phi)}]E(h_t^{\frac{\delta_i}{2}+\frac{1}{2}})\varsigma a\lambda_{i1}, & (2.5b)\\ k_2 &= E[f(e_{t-1})^2] = E[|e_{t-1}| + d_i e_{t-1}]^{2\delta_i}, \ \ \lambda_{ij} = E[f(e_t)^j e_t], & (2.5c)\\ var(r_t) &= \frac{1}{1-\phi^2}[1+\theta^2-2\theta\phi]E(h_t) + \frac{(1+\phi\beta_1)\varsigma^2 a^2[k_2-k_1^2]E(h_t^{\delta_i})}{(1-\phi\beta_1)(1-\phi^2)(1-\beta_1^2)} + \\ & + \frac{2(\phi-\theta)\varsigma a\lambda_{i1}E(h_t^{\delta_i+\frac{1}{2}})}{(1-\phi^2)(1-\phi\beta_1)}, & (2.5d)\end{aligned}$$

$$\begin{aligned}cov(h_t^{\frac{\delta_i}{2}}, h_{t-m}^{\frac{\delta_i}{2}}) &= \gamma_{hm} = \frac{\beta_1^{|m|}}{1-\beta_1^2}a^2[k_2-k_1^2]E(h_t^{\delta_i}), \ \ \rho_m(h_t^{\frac{\delta_i}{2}}) = \beta_1^{|m|}, & (2.6)\\ \rho(r_t, h_{t-m}^{\frac{\delta_i}{2}}) &= \begin{cases}\frac{\gamma^+_{rh,m}}{\sqrt{\gamma_{r0}\times\gamma_{h0}}} & \text{if} \ \ m \geq 0\\ \frac{\gamma^-_{rh,m}}{\sqrt{\gamma_{r0}\times\gamma_{h0}}} & \text{if} \ \ m < 0\end{cases}, \\ cov(r_t, h_{t-m}^{\frac{\delta_i}{2}}) &= \begin{cases}\gamma^+_{rh,\underset{(m\geq 0)}{m}} = [\frac{\phi^{1+m}}{(\phi-\beta_1)(1-\phi\beta_1)} + \frac{\beta_1^{1+m}}{(\beta_1-\phi)[1-\beta_1^2]}]\varsigma a^2[k_2-k_1^2]E(h_t^{\delta_i}) + \frac{\phi^m(\phi-\theta)a\lambda_{i1}E(h_t^{\frac{\delta_i}{2}+\frac{1}{2}})}{(1-\phi\beta_1)}\\ \gamma^-_{rh,m} = \frac{\beta_1^{|m|}\varsigma a^2[k_2-k_1^2]}{(1-\beta_1\phi)[1-\beta_1^2]}E(h_t^{\delta_i}) + \frac{a\lambda_{i1}\beta_1^{|m|}(\beta_1^{-1}-\theta)}{(1-\phi\beta_1)}E(h_t^{\frac{\delta_i}{2}+\frac{1}{2}}), \ m<0\end{cases}\end{aligned}$$

(2.7)

Note that when i=a, $\lambda_{a1} = 0$ *and thus the last terms in the right hand side of equations (2.5b), (2.5d) and (2.7) are 0; when i=b,* $E(h_t^{\frac{\delta_b}{2}+\frac{1}{2}}) = E(h_t)$.
Furthermore, the j-th moment of $h_t^{\frac{\delta_i}{2}}$ *is given by*

$$\begin{aligned}E(h_t^{\frac{\delta_i}{2}j}) &= \frac{\sum_{n=0}^{j-1}\binom{j}{n}\omega^{j-n}E(h_t^{\frac{\delta_i}{2}n})\beta_n}{1-\beta_j}, & (2.8a)\\ \beta_n &= \sum_{l=0}^{n}\binom{n}{l}\beta^{n-l}a^l k_l, \ \ k_l = E[f(e_t)^l] & (2.8b)\end{aligned}$$

Hence, the condition for the existence of the j-th moment of $h_t^{\frac{\delta_i}{2}}$ *is* $\beta_1, \ldots \beta_j < 1$. *In particular, the first and second moments of the* $h_t^{\frac{\delta_i}{2}}$ *are*

$$E(h_t^{\frac{\delta_i}{2}}) = \frac{\omega}{1-\beta_1}, \quad E(h_t^{\delta_i}) = \frac{[E(h_t^{\frac{\delta_i}{2}})]^2}{1 - \frac{a^2}{1-\beta_1^2}[k_2 - k_1^2]} \tag{2.9}$$

The proof of Proposition 1 is given in Appendix A.

The theoretical justification for including the ςh_t term in the mean equation of Bollerslev's GARCH(1,1) model is meagre[9], since the required excess return on a portfolio is linear in its conditional variance only under very special circumstances (Nelson, 1991)[10]. Therefore, in what follows we include the ςh_t^2 and ςh_t terms in the mean equation of Bollerslev's and Taylor/Schwert's models, respectively.

- **Model 2** Consider the ARMA(1,1)-GARCH(1,1)-in-mean model

$$(1-\phi L)r_t = b + (1-\theta L)\varepsilon_t + \varsigma h_t^{\delta_i}, \ \varepsilon_t = e_t h_t^{\frac{1}{2}}, \ e_{t|t-1} \sim IID(0,1), \tag{2.10}$$

$$h_t^{\frac{\delta_i}{2}} = \omega + af(\varepsilon_{t-1}) + \beta h_{t-1}^{\frac{\delta_i}{2}}, \ f(\varepsilon_{t-1}) = [|\varepsilon_{t-1}| + d_i\varepsilon_{t-1}]^{\delta_i} = h_{t-1}^{\frac{\delta_i}{2}}[|e_{t-1}| + d_i e_{t-1}]^{\delta_i}, \tag{2.11a}$$

$$\delta_i = \begin{cases} 2 & \text{if} \quad i=a \\ 1 & \text{if} \quad i=b \end{cases}, \ d_i = \begin{cases} 0 & \text{if} \quad i=a \\ d & \text{if} \quad i=b \end{cases}, \ \omega, a, \beta \geq 0, \ -1 < d < 1 \tag{2.11b}$$

As in model 1, when i=a ($\delta_i = 2, \ d_i = 0$), we have Bollerslev's GARCH-in-mean model (model 2a), and when i=b ($\delta_i = 1, \ d_i = d$) we have Taylor/Schwert's asymmetric GARCH-in-mean model (model 2b).

In these models the $\frac{\delta_i}{2}$-th power of the conditional variance follows a GARCH(1,1) model whereas the δ_i-th power of the conditional variance enters the specification for the mean. In other words, in these models of volatility feedback the long-horizon forecasts are linearly related to short-horizon forecasts of some volatility measure and the expected stock return is quadratic in this measure.

Proposition 2. *The ARMA(2,2) representation of* $h_t^{\delta_i}$ *is*

$$(1-\beta_1 L)(1-\beta_2 L)h_t^{\delta_i} = \omega^2(1+\beta_1) + 2\omega a v_{1,t-1} + a(1-\beta_1 L)[a v_{2,t-1} + 2\beta v_{3,t-1}], \tag{2.12a}$$

$$v_{1,t-1} = f(\varepsilon_{t-1}) - k_1 h_{t-1}^{\frac{\delta_i}{2}}, \quad v_{2,t-1} = f(\varepsilon_{t-1})^2 - k_2 h_{t-1}^{\delta_i}, \tag{2.12b}$$

$$v_{3,t-1} = v_{1,t-1} h_{t-1}^{\frac{\delta_i}{2}}, \quad k_j = E[f(e_{t-1})^j], \tag{2.12c}$$

$$\beta_1 = a k_1 + \beta, \ \ \beta_2 = k_2 a^2 + \beta^2 + 2k_1 a\beta \tag{2.12d}$$

The stationarity condition for $h_t^{\delta_i}$ *is* $\beta_1, \beta_2 < 1$.
In addition, the ARMA(3,3) representation of the stock return is

$$\begin{aligned}(1-\phi L)(1-\beta_1 L)(1-\beta_2 L)r_t &= b^* + a\varsigma(1-\beta_1 L)[a v_{2,t-1} + 2\beta v_{3,t-1}] + 2\omega a\varsigma v_{1,t-1} + \\ &\quad +(1-\theta L)(1-\beta_1 L)(1-\beta_2 L)\varepsilon_t, \\ b^* &= b(1-\beta_1)(1-\beta_2) + \varsigma(1+\beta_1)\omega^2\end{aligned} \tag{2.13}$$

The stationarity condition for r_t *is* $\beta_1, \beta_2, |\phi| < 1$. *In what follows we examine only the case where* $\phi \neq \beta_1 \neq \beta_2$.

[9] Rather, the justification for including ςh_t is pragmatic: a number of researchers using GARCH models (e.g., Hentschel, 1995) have found a statistical significant relation between conditional variance and excess returns on stock market indices.

[10] In Mertons's (1973) intertemporal CAPM model, for example, the instantaneous expected excess return on the market portfolio is linear in its conditional variance if there is a representative agent with log utility. Merton's conditions (e.g., continious time, continuous trading, and a true "market portfolio") do not apply in our models.

Moreover, the auto covariances/correlations of the stock return are

$$
\begin{aligned}
cov_m(r_t) &= \varsigma^2 a^2[a^2 v(v_{2,t}) + 4\beta^2 v(v_{3,t}) + 4a\beta cov(v_{2,t}, v_{3,t})]\frac{1}{(1-\phi\beta_2)(\phi-\beta_2)}[\frac{\phi^{|m|+1}}{1-\phi^2} - \frac{\beta_2^{|m|+1}}{1-\beta_2^2}] \\
&+(a\varsigma)^2 2\omega[acov(v_{2,t}, v_{1t}) + 2\beta cov(v_{3t}, v_{1t})][\frac{\beta_2^{|m|+1}[2\beta_2 - \beta_1(1+\beta_2^2)]}{(1-\beta_2^2)(1-\phi\beta_2)(1-\beta_2\beta_1)(\beta_2-\phi)(\beta_2-\beta_1)} \\
&\frac{\phi^{|m|+1}[2\phi - \beta_1(1+\phi^2)]}{(1-\phi^2)(1-\phi\beta_2)(1-\phi\beta_1)(\phi-\beta_2)(\phi-\beta_1)} + \frac{\beta_1^{|m|+2}}{(\beta_1-\beta_2)(\beta_1-\phi)(1-\beta_1\phi)(1-\beta_1\beta_2)}] \\
&+\frac{\phi^{|m|}}{1-\phi^2}[1+\theta^2 - \theta(\phi+\phi^{-1})]v(\varepsilon_t) + \\
&+4\omega(a\varsigma)^2 v(v_{1t})[\frac{\phi^{|m|+2}}{(1-\phi^2)(1-\phi\beta_2)(1-\phi\beta_1)(\phi-\beta_2)(\phi-\beta_1)} \\
&+\frac{\beta_1^{|m|+2}}{(1-\beta_1^2)(1-\phi\beta_1)(1-\beta_1\beta_2)(\beta_1-\beta_2)(\beta_1-\phi)} \\
&+\frac{\beta_2^{|m|+2}}{(1-\beta_2^2)(1-\phi\beta_2)(1-\beta_1\beta_2)(\beta_2-\beta_1)(\beta_2-\phi)}] + \{\frac{\phi^{|m|}(\phi-\theta)}{(1-\phi^2)(1-\phi\beta_1)(1-\phi\beta_2)} + \\
&+\frac{\phi^{|m|+2}(\phi^{-1}-\theta)}{(1-\phi^2)(\phi-\beta_2)(\phi-\beta_1)} + \frac{\beta_2^{|m|+2}(\beta_2^{-1}-\theta)}{(1-\phi\beta_2)(\beta_2-\beta_1)(\beta_2-\phi)} \\
&+\frac{\beta_1^{|m|+2}(\beta_1^{-1}-\theta)}{(1-\phi\beta_1)(\beta_1-\beta_2)(\beta_1-\phi)}\}2\omega a\varsigma cov(v_{1t}, \varepsilon_t) + \{\frac{\phi^{|m|}(\phi-\theta)}{(1-\phi^2)(1-\phi\beta_2)} + \\
&+\frac{\phi^{|m|+1}(\phi^{-1}-\theta)}{(1-\phi^2)(\phi-\beta_2)} + \frac{\beta_2^{|m|+1}(\beta_2^{-1}-\theta)}{(1-\phi\beta_2)(\beta_2-\phi)}\}a\varsigma[acov(v_{2t}, \varepsilon_t) + 2\beta cov(v_{3t}, \varepsilon_t)], \quad m \neq 0 \qquad (2.14)
\end{aligned}
$$

$$
\begin{aligned}
var(r_t) &= \varsigma^2 a^2[a^2 v(v_{2,t}) + 4\beta^2 v(v_{3,t}) + 4a\beta cov(v_{2,t}, v_{3,t})]\frac{(1+\phi\beta_2)}{(1-\phi\beta_2)(1-\phi^2)(1-\beta_2^2)} + \\
&+(a\varsigma)^2 2\omega[acov(v_{2,t}, v_{1t}) + 2\beta cov(v_{3t}, v_{1t})]\{\frac{\beta_2[2\beta_2 - \beta_1(1+\beta_2^2)]}{(1-\beta_2^2)(1-\phi\beta_2)(1-\beta_2\beta_1)(\beta_2-\phi)(\beta_2-\beta_1)} \\
&+\frac{\phi[2\phi - \beta_1(1+\phi^2)]}{(1-\phi^2)(1-\phi\beta_2)(1-\phi\beta_1)(\phi-\beta_2)(\phi-\beta_1)} + \frac{\beta_1^2}{(\beta_1-\beta_2)(\beta_1-\phi)(1-\beta_1\phi)(1-\beta_1\beta_2)}\} \\
&+\frac{1}{1-\phi^2}[1+\theta^2 - 2\theta\phi]v(\varepsilon_t) + 4\omega(a\varsigma)^2 v(v_{1t})[\frac{\phi^2}{(1-\phi^2)(1-\phi\beta_2)(1-\phi\beta_1)(\phi-\beta_2)(\phi-\beta_1)} \\
&+\frac{\beta_1^2}{(1-\beta_1^2)(1-\phi\beta_1)(1-\beta_1\beta_2)(\beta_1-\beta_2)(\beta_1-\phi)} + \\
&+\frac{\beta_2^2}{(1-\beta_2^2)(1-\phi\beta_2)(1-\beta_1\beta_2)(\beta_2-\beta_1)(\beta_2-\phi)}] \\
&+\frac{4(\phi-\theta)\omega a\varsigma cov(v_{1t}, \varepsilon_t)}{(1-\phi^2)(1-\phi\beta_1)(1-\phi\beta_2)} + \frac{2(\phi-\theta)a\varsigma[acov(v_{2t}, \varepsilon_t) + 2\beta cov(v_{3t}, \varepsilon_t)]}{(1-\phi^2)(1-\phi\beta_2)}, \qquad (2.15a) \\
\rho_m(r_t) &= \frac{cov_m(r_t)}{var(r_t)} \qquad (2.15b)
\end{aligned}
$$

Furthermore, the cross covariances/correlations between the stock return and the $\frac{\delta_i}{2}$-th power of its conditional variance are

$$\rho(r_t, h_{t-m}^{\frac{\delta_i}{2}}) = \frac{cov(r_t, h_{t-m}^{\frac{\delta_i}{2}})}{\sqrt{var(r_t)var(h_t^{\frac{\delta_i}{2}})}} = \begin{cases} \frac{\gamma^+_{rh,m}}{\sqrt{\gamma_{r0}\times\gamma_{h0}}} & \text{if} \quad m \geq 0 \\ \frac{\gamma^-_{rh,m}}{\sqrt{\gamma_{r0}\times\gamma_{h0}}} & \text{if} \quad m < 0 \end{cases},$$

$$\gamma^+_{rh,m} = 2\omega a^2\varsigma v(v_{1t})[\frac{\beta_1^{m+2}}{(\beta_1-\phi)(\beta_1-\beta_2)(1-\beta_1^2)} + \frac{\beta_2^{m+2}}{(\beta_2-\phi)(\beta_2-\beta_1)(1-\beta_1\beta_2)} + \frac{\phi^{m+2}}{(\phi-\beta_2)(\phi-\beta_1)(1-\phi\beta_1)}] + a^2\varsigma[acov(v_{2t}, v_{1t}) + 2\beta cov(v_{3t}, v_{1t})] \times [\frac{\beta_2^{m+1}}{(\beta_2-\phi)(1-\beta_1\beta_2)} + \frac{\phi^{m+1}}{(\phi-\beta_2)(1-\phi\beta_1)}] + \frac{a\phi^m(\phi-\theta)cov(\varepsilon_t, v_{1t})}{1-\phi\beta_1}, m \geq 0, \quad (2.16a)$$

$$\gamma^-_{rh,m} = \frac{2\omega a^2\varsigma v(v_{1t})\beta_1^{|m|}}{(1-\beta_1^2)(1-\beta_1\phi)(1-\beta_1\beta_2)} + \frac{a^2\varsigma[acov(v_{2t}, v_{1t}) + 2\beta cov(v_{3t}, v_{1t})]\beta_1^{|m|}}{(1-\beta_1\phi)(1-\beta_1\beta_2)} + + \frac{a\beta_1^{|m|}(\beta_1^{-1}-\theta)cov(\varepsilon_t, v_{1t})}{1-\phi\beta_1}, \quad m < 0, \quad (2.16b)$$

where γ_{r0} and γ_{h0} are given by (2.15a) and (2.6, when m=0), respectively.
Finally, the auto covariances/correlations of the δ_i-th power of the conditional variance are

$$\rho_m(h_t^{\delta_i}) = \frac{cov(h_t^{\delta_i}, h_{t-m}^{\delta_i})}{var(h_t^{\delta_i})} = \frac{\gamma_{h2,m}}{\gamma_{h2,0}},$$

$$\gamma_{h2,m} = \frac{4\omega a^2 v(v_{1t})}{(1-\beta_1\beta_2)(\beta_1-\beta_2)}[\frac{\beta_1^{|m|+1}}{(1-\beta_1^2)} - \frac{\beta_2^{|m|+1}}{(1-\beta_2^2)}] + +a^2[a^2v(v_{2t}) + 4\beta^2 v(v_{2t}) + 4a\beta cov(v_{1t}, v_{2t})]\frac{\beta_2^{|m|}}{(1-\beta_2^2)} + +4\omega a^2[acov(v_{1t}, v_{2t}) + 2\beta cov(v_{1t}, v_{2t})] \times \times\{\frac{\beta_1^{|m|+1}}{(\beta_1-\beta_2)(1-\beta_1\beta_2)} + \frac{\beta_2^{|m|}[2\beta_2 - \beta_1(1+\beta_2^2)]}{(1-\beta_1\beta_2)(\beta_2-\beta_1)(1-\beta_2^2)}\} \quad (2.17)$$

The proof of Proposition 2 is given in Appendix A.

Note that all the formulae in Propositions 1 and 2 hold for the general Asymmetric Power GARCH[11]-in-mean model where 0< δ_i and -1<d_i < 1. For this model the E($h_t^{\frac{\delta_i}{2}+\frac{1}{2}}$) and E($h_t$) are fractional moments and their calculation is beyond the scope of this paper.

2.2 EGARCH-in-Mean-Models

One limitation of the GARCH models 1 and 2 results from the nonnegativity constraints on ω α and β in (2.2a) and (2.11a) which are imposed to ensure that $h_t^{\frac{\delta_i}{2}}$ remains nonnegative for all t. These constraints imply that increasing f(e_t) in any period increases $h_{t+m}^{\frac{\delta_i}{2}}$ for all m≥ 1 , ruling out oscillatory behavior in the $h_t^{\frac{\delta_i}{2}}$ process (Nelson, 1991). Alternatively, one can use a model in which the logarithm of the conditional variance of returns follows a GARCH-like process. This can be combined with the assumption that the expected return is linear in the logarithm of the conditional variance, as suggested by Engle et al (1987)[12].

- **Model 3**. Consider the ARMA(1,1)-EGARCH(1,1)-in-mean model

$$(1-\phi L)r_t = b + (1-\theta L)\varepsilon_t + \varsigma\ln(h_t), \ \varepsilon_t = e_t h_t^{\frac{1}{2}}, \ e_{t|t-1} \sim IID(0,1), \quad (2.18)$$

[11] The statistical properties of the simple Asymmetric Power GARCH(1,1) model have been examined in He and Terasvirta (2000).

[12] Engle, et al (1987) computed a series of LM tests for omitted variables [ϱ^{h_t}, $h_t^{\frac{1}{2}}$, and log(h_t)] to test the assumed linearity between the conditional variance and mean returns. Economic theory has little to say on the nature of this trade-off as it presumably depends on the risk preferences of the traders. Engle et al (1987) found that only the log variance was significant and they concluded that the final preffered model was the one with the log variance in the mean.

$$(1-\beta L)\ln(h_t)=\omega+cz_{t-1},\quad z_{t-1}=d\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}+\gamma[|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}|-E|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}|], \tag{2.19}$$

In this model the logarithm of the conditional variance follows an ARMA(1,1) process and it also enters in the specification of the mean. The stationarity condition for the $\ln(h_t)$ is $|\beta|<1$. Note that in (2.19) there are no inequality constraints whatever, and that cycling is permitted, since the β term can be negative or positive.

Proposition 3. *The univariate ARMA(2,2) representation of the stock return is*

$$(1-\phi L)(1-\beta L)r_t=[b(1-\beta)+\delta\omega]+(1-\theta L)(1-\beta L)\varepsilon_t+\varsigma cz_{t-1} \tag{2.20}$$

The stationarity condition for the stock return is $|\beta|,|\phi|<1$. In what follows we examine only the case where $\beta\neq\phi$.

In addition, the auto covariances/correlations of the stock return are

$$\rho_m(r_t)=\frac{cov(r_t,r_{t-m})}{var(r_t)}, \tag{2.21a}$$

$$\begin{aligned}cov(r_t,r_{t-m}) &= \gamma_{rm}=\frac{\phi^{|m|}}{1-\phi^2}[1+\theta^2-\theta(\phi+\phi^{-1})]E(h_t)+\\ &\frac{1}{(1-\phi\beta)(\phi-\beta)}[\frac{\phi^{|m|+1}}{1-\phi^2}-\frac{\beta^{|m|+1}}{1-\beta^2}]\varsigma^2c^2var\,(z_t)+\\ &\{\frac{\phi^{|m|}}{(1-\phi\beta)(1-\phi^2)}[-\theta+\phi]+\frac{\phi^{|m|+1}}{(\phi-\beta)(1-\phi^2)}[-\theta+\phi^{-1}]+\\ &\frac{\beta^{|m|+1}}{(\beta-\phi)(1-\beta\phi)}(-\theta+\beta^{-1})\}\varsigma cdE(h_t^{\frac{1}{2}}),\end{aligned} \tag{2.21b}$$

$$var(r_t)=\frac{(1+\theta^2-2\theta\phi)E(h_t)}{1-\phi^2}+\frac{(1+\phi\beta)\varsigma^2c^2var(z_t)}{(1-\phi\beta)(1-\phi^2)(1-\beta^2)}+\frac{(\phi-\theta)d2\varsigma cE(h_t^{\frac{1}{2}})}{(1-\beta\phi)(1-\phi^2)} \tag{2.21c}$$

where the first and second moments of h_t are given below and the variance of z_t is given by

$$var(z_t)=\begin{cases}d^2+\gamma^2(1-\frac{2}{\pi}) & e_t\sim \text{N}\\ d^2+\gamma^2\{1-\frac{[\Gamma(\frac{2}{v})]^2}{\Gamma(\frac{1}{v})\Gamma(\frac{3}{v})}\} & e_t\sim \text{GE}\\ d^2+\frac{\gamma^2}{2} & e_t\sim \text{DE}\end{cases} \tag{2.22}$$

where N, GE and DE denote the Normal, the Generalized Error and the Double Exponential distributions, respectively; v are the degrees of freedom of the generalized error distribution.

Moreover, 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

$$cov(h_t^{k_1},h_{t-m}^{k_2})=\varrho^{\frac{(k_1+k_2)[\omega-c\gamma f_D]}{1-\beta}}[\Pi^D_{m,k_1}\Pi^D_{m,k_1,k_2}-\Pi^D_{0,k_1,0}\Pi^D_{0,0,k_2}], \tag{2.23a}$$

$$E(h_t^{k_1+k_2})=\varrho^{\frac{(k_1+k_2)[\omega-c\gamma f_D]}{1-\beta}}\Pi^D_{0,k_1,k_2},\quad \rho(h_t^{k_1},h_{t-m}^{k_2})=\frac{[\Pi^D_{m,k_1}\Pi^D_{m,k_1,k_2}-\Pi^D_{0,k_1,0}\Pi^D_{0,0,k_2}]}{\sqrt{[\Pi^D_{0,k_1,k_1}-(\Pi^D_{0,k_1,0})^2][\Pi^D_{0,k_2,k_2}-(\Pi^D_{0,k_2,0})^2]}} \tag{2.23b}$$

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

$$\begin{aligned}
\Pi^N_{m,k_1} &= \prod_{i=0}^{|m|-1} \varrho^{\frac{(\gamma+d)^2(ck_1)^2\beta^{2i}}{2}}\{[\frac{1}{2}+\frac{ck_1(\gamma+d)\beta^i}{\sqrt{2\pi}}F_1(\frac{1}{2};\frac{3}{2};\frac{-(\gamma+d)^2(ck_1)^2\beta^{2i}}{2})]+ \\
&+\varrho^{-2\gamma d(ck_1)^2\beta^{2i}}[\frac{1}{2}+\frac{ck_1(\gamma-d)\beta^i}{\sqrt{2\pi}}F_1(\frac{1}{2};\frac{3}{2};\frac{-(\gamma-d)^2(ck_1)^2\beta^{2i}}{2})]\},\ m\neq 0 \qquad (2.24a) \\
\Pi^N_{m,k_1,k_2} &= \prod_{i=0}^{\infty}\varrho^{\frac{(\gamma+d)^2c^2\beta^{2i}(k_2+k_1\beta^{|m|})^2}{2}}\{[\frac{1}{2}+\frac{c(\gamma+d)\beta^i(k_2+k_1\beta^{|m|})}{\sqrt{2\pi}}\times \\
&\times F_1(\frac{1}{2};\frac{3}{2};\frac{-(\gamma+d)^2c^2\beta^{2i}(k_2+k_1\beta^{|m|})^2}{2})]+\varrho-2\gamma dc^2\beta^{2i}(k_2+k_1\beta^{|m|})^2\times \\
&\times[\frac{1}{2}+\frac{c(\gamma-d)\beta^i(k_2+k_1\beta^{|m|})}{\sqrt{2\pi}}F_1(\frac{1}{2};\frac{3}{2};\frac{-(\gamma-d)^2c^2\beta^{2i}(k_2+k_1\beta^{|m|})^2}{2})]\}, \qquad (2.24b) \\
f_N &= \sqrt{\frac{2}{\pi}} \qquad (2.24c)
\end{aligned}$$

where F *denotes the hypergeometric function.*
When the conditional distribution is the generalized error the $f_D, \prod^D_{m,k_1}$ *and* $\prod^D_{m,k_1,k_2}$ *terms are given by*

$$\begin{aligned}
\prod^g_{m,k_1} &= \prod_{i=0}^{|m|-1}\{\sum_{\tau=0}^{\infty}[2^{\frac{1}{v}}\lambda k_1c\beta^i]^\tau[(\gamma+d)^\tau+(\gamma-d)^\tau]\frac{\Gamma(\frac{1+\tau}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)}\}, m\neq 0 \qquad (2.25a) \\
\prod^g_{m,k_1,k_2} &= \prod_{i=0}^{\infty}\{\sum_{\tau=0}^{\infty}[2^{\frac{1}{v}}\lambda c\beta^i(k_2+k_1\beta^{|m|})]^\tau[(\gamma+d)^\tau+(\gamma-d)^\tau]\frac{\Gamma(\frac{1+\tau}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)}\}, \qquad (2.25b) \\
\lambda &= \{2^{\frac{-2}{v}}\Gamma(\frac{1}{v})[\Gamma(\frac{3}{v})]^{-1}\}^{\frac{1}{2}},\quad f_g=\frac{\Gamma(\frac{2}{v})\lambda 2^{\frac{1}{v}}}{\Gamma(\frac{1}{v})} \qquad (2.25c)
\end{aligned}$$

where v *are the degrees of freedom of the generalized error distribution*; when $v>1$ *the summations in (2.25a) and (2.25b) are finite whereas when* $v<1$ *the summations are finite if and only if* $k_1c\beta^i\gamma+|k_1c\beta^i d|\leq 0$ *and* $c\beta^i(k_2+k_1\beta^{|m|})\gamma+|c\beta^i(k_2+k_1\beta^{|m|})d|\leq 0$, *respectively (Nelson, 1991).*
When $v=1$ *(i.e. the conditional distribution is the double exponential), the above equations give*

$$\begin{aligned}
\prod^d_{m,k_1} &= \prod_{i=0}^{|m|-1}\frac{2-\sqrt{2}k_1c\beta^i\gamma}{2-2\sqrt{2}k_1c\beta^i\gamma+(k_1c)^2\beta^{2i}(\gamma^2-d^2)},\quad f_d=\frac{1}{\sqrt{2}},\ m\neq 0, \qquad (2.26a) \\
\prod^d_{m,k_1,k_2} &= \prod_{i=0}^{\infty}\frac{2-\sqrt{2}c\beta^i(k_2+k_1\beta^{|m|})\gamma}{2-2\sqrt{2}c\beta^i(k_2+k_1\beta^{|m|})\gamma+c^2\beta^{2i}(k_2+k_1\beta^{|m|})^2(\gamma^2-d^2)}, \qquad (2.26b)
\end{aligned}$$

The above equations hold if and only if $k_1c\beta^i\gamma+|k_1c\beta^i d|<\sqrt{2}$ and $c\beta^i(k_1\beta^{|m|}+k_2)\gamma+|c\beta^i(k_1\beta^{|m|}+k_2)d|<\sqrt{2}$. *(Nelson, 1991).*

Use of the logarithm of the conditional variance as a regressor in the mean equation induces a negative sign to the risk premium whenever h_t is less than one, and as $h_t\to 0$ the effect on r_t would be infinite (Pagan and Hong, 1991). Therefore, in what follows, we use the k-th power of the conditional variance as a regressor in the mean equation and we examine only a MA(1) mean specification. As a practical matter there is little difference between an AR(1) and a MA(1) model when the AR and 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).

- **Model 4.** Consider the MA(1)-EGARCH(1,1)-in-mean model

$$r_t=b+(1-\theta L)\varepsilon_t+\varsigma h_t^k,\ \varepsilon_t=e_th_t^{\frac{1}{2}},\ e_{t|t-1}\sim IID(0,1), \qquad (2.27)$$

$$(1-\beta L)\ln(h_t) = \omega + cz_{t-1}, \quad z_{t-1} = d\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} + \gamma[|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}| - E|\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}|] \tag{2.28}$$

Note that in this model the logarithm of the conditional variance follows an ARMA(1,1) process whereas the k-th power of the conditional variance enters the specification in the mean. Although for this model we do not have the ARMA representation of the h_t^k, and consequently of r_t, the covariance-stationarity condition for both processes is $|\beta| < 1$.

Proposition 4. *The autocorrelations of the stock return are*

$$\begin{aligned}
\rho_m(r_t) &= \frac{\gamma_{rm}}{\gamma_{r0}}, \quad \gamma_{rm} = \{\varsigma^2[\prod\nolimits_{m,k}^{D}\prod\nolimits_{m,k,k}^{D} - \prod\nolimits_{0,k,0}^{D}\prod\nolimits_{0,0,k}^{D}] + \varsigma\varrho^{\frac{(.5-k)(\omega-c\gamma f_D)}{1-\beta}}[\prod\nolimits_{|m|-1,k}^{D} D_{|m|-1,k}^{D}\prod\nolimits_{m,k,\frac{1}{2}}^{D} - \\
&\quad -\theta\prod\nolimits_{m,k}^{D} D_{mk}^{D}\prod\nolimits_{|m|+1,k,\frac{1}{2}}^{D}] - \theta\xi_m\varrho^{\frac{(1-2k)(\omega-c\gamma f_D)}{1-\beta}}\prod\nolimits_{0,\frac{1}{2},\frac{1}{2}}^{D}\}e^{\frac{2k(\omega-c\gamma f_D)}{1-\beta}}, &(2.29a)\\
\gamma_{r0} &= \{\varsigma^2[\prod\nolimits_{0k}^{D}\prod\nolimits_{0,k,k}^{D} - \prod\nolimits_{0,k,0}^{D}\prod\nolimits_{0,0,k}^{D}] - \\
&\quad -2\varsigma\theta\varrho^{\frac{(.5-k)(\omega-c\gamma f_D)}{1-\beta}} D_{0,k}^{D}\prod\nolimits_{1,k,\frac{1}{2}}^{D} + (1+\theta^2)\varrho^{\frac{(1-2k)(\omega-c\gamma f_D)}{1-\beta}}\prod\nolimits_{0,\frac{1}{2},\frac{1}{2}}^{D}\}e^{\frac{2k(\omega-c\gamma f_D)}{1-\beta}}, &(2.29b)\\
\xi_m &= \begin{cases} 1 & \text{if} \quad m=1 \\ 0 & \text{if} \quad \text{otherwise} \end{cases}, \quad \prod\nolimits_{0,k}^{D} = 1
\end{aligned}$$

where the $\Pi_{m,k}^D$ and $\prod_{m,k,k}^D$ *are defined by (2.24a)-(2.26b); when the conditional distribution is the normal, the double exponential and the generalized error one then the* D_{mk}^D *term is given by*

$$\begin{aligned}
D_{mk}^{N} &= \sqrt{\frac{1}{2\pi}}\varrho^{\frac{(ck)^2\beta^{2|m|}(\gamma+d)^2}{4}} \times \\
&\quad \{D_{-2}[-ck\beta^{|m|}(\gamma+d)] - \varrho^{-(ck)^2\beta^{2|m|}\gamma d}D_{-2}[-ck\beta^{|m|}(\gamma-d)]\}, &(2.30a)\\
D_{-2}[q] &= \varrho^{\frac{-1}{4}q^2}\int_0^\infty \varrho^{-qs-\frac{1}{2}s^2}s\,ds, \quad D_{-2}[0]=1, &(2.30b)\\
D_{mk}^{g} &= \sum_{\tau=0}^{\infty}(2^{\frac{1}{v}}ck\beta^{|m|}\lambda)^{\tau}[(\gamma+d)^{\tau} - (\gamma-d)^{\tau}]\frac{\Gamma(\frac{2+\tau}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)}, &(2.30c)\\
D_{mk}^{d} &= \frac{dck\beta^{|m|}[\sqrt{2}-ck\beta^{|m|}\gamma]}{\sqrt{2}[1+\frac{(ck\beta^{|m|})^2(\gamma^2-d^2)}{2} - \sqrt{2}ck\beta^{|m|}\gamma]^2}, \quad ck\beta^{|m|}\gamma + |ck\beta^{|m|}d| < \sqrt{2} &(2.30d)
\end{aligned}$$

where $D_q[\cdot]$ *denotes the parabolic cylinder function; when* $v > 1$ *the summation in (2.30c) is finite whereas when* $v < 1$ the *summation is finite if and only if* $ck\beta^{|m|}\gamma + |ck\beta^{|m|}d| \leq 0$.

Finally, the cross correlations between the stock return and the k_1*-th power of its conditional variance are*

$$\begin{aligned}
\rho_{m,k_1,rh} &= \frac{cov(r_t, h_{t-m}^{k_1})}{\sqrt{\gamma_{r0}\times\gamma_{h0,k_1}}} = &(2.31a)\\
&= \begin{cases} \dfrac{\varsigma[\prod_{m,k}^{D}\prod_{m,k,k_1}^{D} - \prod_{0,k,0}^{D}\prod_{0,0,k_1}^{D}]e^{\frac{(k+k_1)(\omega-c\gamma f_D)}{1-\beta}}}{\sqrt{\gamma_{r0}\times\gamma_{h0,k_1}}}, & m \geq 1 \\ e^{\frac{(k+k_1)(\omega-c\gamma f_D)}{1-\beta}}\{\dfrac{\varsigma[\prod_{m,k_1}^{D}\prod_{m,k_1,k}^{D} - \prod_{0,k_1,0}^{D}\prod_{0,0,k}^{D}]}{\sqrt{\gamma_{r0}\times\gamma_{h0},k_1}} + \\ \dfrac{\varrho^{\frac{(.5-k)(\omega-c\gamma f_D)}{1-\beta}}[\prod_{|m|-1,k_1}^{D} D_{|m|-1,k_1}^{D}\prod_{m,k_1,\frac{1}{2}}^{D} - \theta\prod_{m,k_1}^{D} D_{m,k_1}^{D}\prod_{|m|+1,k_1,\frac{1}{2}}^{D}]}{\sqrt{\gamma_{r0}\times\gamma_{h0,k_1}}}\}, & m \leq -1 \end{cases} &,(2.31b)\\
\gamma_{h0,k_1} &= \varrho^{\frac{2k_1[\omega-c\gamma f_D]}{1-\beta}}[\Pi_{0,k_1}^{D}\Pi_{0,k_1,k_1}^{D} - \Pi_{0,k_1,0}^{D}\Pi_{0,0,k_1}^{D}] &(2.31c)
\end{aligned}$$

where γ_{r0} *is given by (2.29b).*
The proof of Proposition 4 is given in Appendix C.

3 EMPIRICAL APPLICATION

Estimation and Inference

For our empirical analysis, we use the daily returns for the Korean stock price index (KOSPI) for July 1986 to December 1991. The reason that we choose this particular period is to avoid the structural changes that have characterized the Korean stock market during its evolution since 1980 (figure 1 plot the series for 1980-present). In particular prior to 1987, the authorities had imposed strict regulations on the pricing of stocks in the issuing market. The subsequent deregulation of the Korean stock market together with the late 80's economic boom increased the total value of the market from 6,600 billions at the end of 1985 to 73,000 billions in 1991. Moreover, until the end of 1991 only domestic investors were allowed to participate in the Korean stock exchange; in January 1992 it opened to foreign investors as well. Finally, the currency crisis of 1997 which affected all Asian economies created a turmoil in the Korean financial market as well.

An immediate problem in using this data is that we wish to model the excess returns but do not have access to any adequate daily riskless return series. However, Nelson (1991) fitted both models using excess returns and capital gains (ignoring both dividends and the riskless rate of return) and found actually no difference in either the estimated parameters or the fitted variances[13]. Therefore, we fit the models using the capital gain series.

We estimate our models using numerical maximum likelihood. This procedure is subject to the same caveat that applies to all empirical work with GARCH-in-mean models, namely that sufficient regularity conditions for consistency and asymptotic normality of the maximum likelihood estimator are not yet available. Below we treat our estimates as if they are indeed asymptotically normal.

Next, we examine the empirical issues raised earlier on in the paper. Table 1 presents various estimations of our models. We have estimated both Bollerslev and Taylor/Schwert GARCH(1,1) model using an AR(1) mean specification, two alternative conditional distributions (the gaussian and the student's t), and either the h_t or $h_t^{\frac{1}{2}}$ or $\ln(h_t)$ to capture the in-mean effect. Similarly, we have estimated Nelson's EGARCH(1,1) model using an AR(1) or a MA(1) mean specification, two alternative conditional distributions (the normal and the double exponential), and either the h_t or $h_t^{\frac{1}{2}}$ or $\ln(h_t)$ to capture the in-mean effect.

For all the autoregressive models the AR coefficient is statistically significant and very small[14]. In all the moving average EGARCH models the MA parameter is statistically significant and its value is very close to the AR parameter of the AR-EGARCH models. In addition, the estimated risk premium, as represented by ς, is positive and statistically significant[15]. This agrees with the significant positive relation between returns and conditional variance found by researchers using GARCH-M models (e.g., Chou, 1987, and French, Schwert and Stambough, 1987, but contrast with the findings of other researchers not using GARCH models (e.g., Pagan and Hong, 1991). Note that in Bollerslev's model a) when we include the leverage term the in-mean coefficient decreases and becomes less significant (for example, when the variance is used as a regressor and the conditional distribution is normal the in-mean coefficient drops from 7.58 to 6.27 and its p value increases from .036 to .061), and b) when we use the t distribution the in-mean coefficient increases and remains highly significant (for example when the variance is used as a regressor the in-mean coefficient increases from 7.58 to 11.79 and its p value drops from .036 to .028).

Furthermore, for all models the asymmetric relation between returns and changes in volatility, as represented by d, is negative and highly significant. Recall that a negative value of d indicates that volatility tends to rise(fall) when returns surprises are negative(positive). Note that in Bollerslev's model when we use the t distribution the leverage coefficient decreases (in absolute value) from .187 to .169 and remains highly significant. Moreover, for the asymmetric power GARCH model, the estimated power, as represented by δ is insignificantly different from two (for both the normal and the t-distributions). Therefore, it is of no surprise that the coefficients of this model are very similar to those of Bollerslev's model. This contrasts with the findings of Ding et al (1993) and Hentschel (1995). Finally, observe that in all cases where the t distribution is assumed the estimated degrees of freedom are greater than five which implies that the first two conditional moments exist.

Diagnostic Testing

Next we use a variety of diagnostic tests to determine whether various aspects of our different models are correctly specified. The diagnostics of the standardized residuals from the estimated models are given in

[13] The two series had a sample correlation of .9996. In other words, the series were practically identical, so ignoring dividends and interest payments appears likely to introduce no important errors in forecasting volatility of broad market indices.

[14] The autoregressive coefficient takes values from .037 to .079; its p value ranges between .1 and .003.

[15] When we include the variance, the standard deviation and the logarithm of the variance as a regressor in the mean equation ς takes values from 6.27 , .24 and .001 to 11.79, .32 and .001, respectively; the p values range from .06, .09, and .117 to .008, .008, and .01, respectively.

Table 2. In all cases the standardized residuals and squared standardized residuals are uncorrelated. Note that, although the Shapiro-Wilk test shows no normality it is very sensitive to outliers. When we add two dummies for the 26/06/86 and 27/04/90 data points we obtain a p value for the test which is above .05 for most of the models. Table 2 also gives a measure for the persistence of the conditional variance. For Nelson's model volatility persistence is approximately .82 and for Bollerslev's model it ranges from .82 to .88. Note that in the latter model persistence decreases to .72 when the t distribution is used.

In addition for all the estimated GARCH-in-mean models Table 3 reports those that are preferred to the ones without in-mean effect according to the Akaike information model selection criterion (AIC) and the likelihood ratio test. According to these two simplistic selection criteria Bollerslev's asymmetric GARCH(1,1) model with the variance as a regressor in the mean appears to be superior[16]. These tests leave many potential sources of misspecification unchecked. It therefore seems desirable to check the forecasting performance of the model. Table 4 presents four alternative measures of forecasting accuracy; Models 1a when the conditional distribution is the normal and 1a when the conditional distribution is the t have the minimum mean square error (MSE) and mean absolute error (MAE), respectively, whereas models 2b and 1b have the minimum logarithmic error (LE) and logarithmic absolute error (LAE), respectively.

Moreover, we use three diagnostic tests for volatility models suggested by Engle and Ng (1993): the Sign Bias Test, the Negative Size Bias Test, and the Positive size Bias Test. These tests examine whether we can predict the squared normalized residual by some variables observed in the past which are not included in the volatility model being used. If these variables can predict the squared normalized residual, then the variance model is misspecified. The sign and size bias tests do not indicate any asymmetric variance effects in the estimated models[17].

As we have already explained we have chosen the 1986-1991 estimation period because of its homogeneity in economic conditions. However, our results appear to be quite robust to the sample size. Table 5a presents some of our models which have been estimated for the 1980-1997 and 1992-1997 periods. Note the similarity of the estimated parameters to those given in Table 1. Also note that when we include dummy variables in our models the estimated parameters remain virtually unchanged (see Table 5b).

Estimated Theoretical Autocorrelations

The main objective of this study was not only to investigate which of the various GARCH-in-mean specifications model the conditional variance "best" but also to provide analytic expressions for existing moments of the various GARCH-in-mean processes. Figure 2 plots the estimated theoretical autocorrelations (eta) for models 1 to 4. Note that the eta of the stock return are positive for the AR models and negative for the MA ones (except for the first lag) (see figure 2a). In addition, for all models the eta between the stock return and lagged values of the conditional variance/standard deviation are positive (see figure 2b and 2d) whereas the eta between the conditional standard deviation and lagged values of the stock return are negative (see figure 2e). Finally, the eta of the conditional variance for models 2a and 3 are positive (see figure 2f).

4 CONCLUSIONS

This paper derived analytical expressions for existing moments of various GARCH-in-mean models. In particular we derived the autocorrelation function of the stock return and its conditional variance, and the cross correlations between the stock return and its conditional variance. We used an ARMA(1,1) form for the mean equation and we examined three alternative asymmetric GARCH specifications: Bollerslev's GARCH(1,1) model, Taylor/Schwert's GARCH(1,1) model and Nelson's EGARCH(1,1) model. For the later we assumed that the error term is drawn from either the Normal or the Double Exponential or the Generalized Error Distribution. We chosen the form in which the time varying variance enters the specification of the mean to determine the risk premium by assuming (i) linearity, i.e. we used h_t, $h_t^{\frac{1}{2}}$, and $\ln(h_t)$ as a regressor in Bollerslev's, Taylor/Schwert's and Nelson's models, respectively, and (ii) nonlinearity, i.e. we used h_t^2, h_t, and h_t^k in Bollerslev's, Taylor/Schwert's and Nelson's models, respectively. All our estimated models suggested a positive and statistically significant relation between risk and return. According to the selection criteria Bollerslev's asymmetric GARCH(1,1) model with the variance as a regressor in the mean appeared to be superior. We left the consideration of higher order ARMA-GARCH models for future research.

[16] Hentchel (1995) developed a family of asymmetric GARCH models that nests both the A-PGARCH model and the EGARCH model.

[17] Hagerud (1997a, b) proposed two new tests for asymmetric effects, based on a pair of LM statistics, and he showed that these two tests have superior power properties, compared to the standard asymmetric GARCH tests developed by Engle and Ng (1993).

References

Abadir, K. M., 1999, An introduction to hypergeometric functions for economists, Econometric Reviews, 18, 287-230.

Baillie, R. T. and T. Bollerslev, 1989, The message in daily exchange rates: A conditional variance tale, Journal of Business and Economic Statistics, 7, 297-305.

Baillie, R. T. and T. Bollerslev, 1992, Prediction in dynamic models with time-dependent conditional variance, Journal of Econometrics 52, 91-113.

Black, A. and P. Fraser, 1995, U.K. Stock Returns: Predictability and business conditions, Mancester School of Economics and Social Studies, 63(0), Supplement, 85-102.

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.

Bollerslev, T., 1987, A conditionally heteroskedastic time series model for speculative prices and rates of return, Review of Economics and Statistics, 69, 542-547.

Bollerslev, T., Engle, R.F. and J.M. Wooldridge, 1988, A capital asset pricing model with time varying covariances, Journal of Political Economy, 96, 116-131.

Breen, W., Glosten, L. R., and R. Jagannathan, 1989, Economic significance of predictable variations in stock index returns, Journal of Finance, 44, 1177-1189.

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.

Campbell, J. Y., 1987, Stock returns and the term structure, Journal of Financial Economics, 18, 373-399.

Campbell, J. Y. and L. Hentschel, 1992, No news is good news: An asymmetric model of changing volatility in stock returns, Journal of Financial Economics, 31, 281-318.

Caporale, T. and B. McKierman, 1996, The relationship between output variability and growth: evidence from post war UK data, Scottish Journal of Political Economy, 43, 229-236.

Chou, R. Y., 1987, Volatility persistence and stock returns-Some empirical evidence using GARCH, Journal of Applied Econometrics, 3, 279-294.

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.

Domowitz, I. and C. S. Hakkio, 1985, Conditional variance and the risk premium in the foreign exchange market, Journal of International Economics, 19, 47-66.

Ellul, A., 1999, As you like it: An investigation of trading behavior and price volatility on auction and dealership market architectures, Unpublished Mimeo, London School of Economics.

Elyasiani, E. and I. Mansur, 1998, Sensitivity of the bank stock returns distribution to changes in the level and volatility of interest rate: a GARCH-M model, Journal of Banking and Finance, 22, 535-563.

Engle, R. F., Lilien, D. M. and R. P. Robins, 1987, Estimating time varying risk premia in the term structure: the ARCH-M model, Econometrica 55(2), 391-407.

Engle, R. F. and C. Mustafa, 1992, Implied ARCH models from options prices, Journal of Econometrics, 52, 289-311.

Engle, R. F. and V. K. Ng, 1993, Measuring and testing the impact of news on volatility, Journal of Finance, 5, 1749-1778.

Fama, E. F. and G. W. Schwert, 1977, Asset returns and inflation, Journal of Financial Economics, 5, 115-146.

Fountas, S., Karanasos, M. and M. Karanassou, 2000, A GARCH model of inflation and inflation uncertainty with simultaneous feedback, Unpublished Mimeo, Dept of Economics, University of York.

Fraser, P., 1996, UK excess share returns: firm size and volatility, Scottish Journal of Political Economy, 43, 71-84.

French, K. R., Schwert, G. W. and R. F. Stambaugh, 1987, Expected stock returns and volatility, Journal of Financial Economics, 19, 3-29.

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.

Hagerud, G. E., 1997a, Specification tests for asymmetric GARCH, Unpublished Mimeo, Dept of Finance, Stocholm School of Economics.

Hagerud, G. E., 1997b, Modeling Nordic stock returns with asymmetric GARCH models, Unpublished Mimeo, Dept of Finance, Stockholm School of Economics.

Hall, S. G., 1991, An application of the stochastic GARCH-in-mean model to risk premia in the London metal exchange, Manchester School of Economics and Social Studies, 59(0), Supplement, 57-71.

Hansson, B. and P. Hordahl, 1997, Changing risk premia: evidence from a small open economy, Scandinavian Journal of Economics, 99(2), 335-350.

He, C. and T. Teräsvirta, 1999a, Fourth moment structure of the GARCH(p,q) model, Econometric Theory, 15(6), 824-846.

He, C. and T. Teräsvirta, 1999b, Statistical properties of the asymmetric power ARCH process, in Engle, R. F. and H. White, eds., Cointegration, Causality, and Forecasting. Festschrift in honour of Clive W. J. Granger, Oxford University Press, 462-474.

He, C., Teräsvirta, T. and H. Malmsten, 1999, Fourth moment structure of a family of first-order exponential GARCH models, Unpublished Mimeo, Department of Economic Statistics, Stockholm School of Economics

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

Hsieh, D.A., 1989, Modeling heteroskedasticity in daily foreign exchange rate Journal of Business and Economic Statistics, 7, 307-317.

Hurn, A. S., Mcdonald A. D. and T. Moody, 1995, In search of time-varying term premia in the London interbank market, Scottish Journal of Political Economy, 42, 152-164.

Jorion, P., 1988, On jump processes in the foreign exchange and stock markets, Review of Financial Studies, 1, 427-445.

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., 2000a, Modeling volatility persistence: some new results on the component GARCH model, Discussion Paper, Dept of Economics, University of York.

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

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, City University Business School.

Kroner, K. F. and W. D. Lastrapes, 1993, The impact of exchange rate volatility on international trade: reduced form estimates using the GARCH-in-mean model, Journal of International Money and Finance, 12, 298-318.

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.

Merton, R. C., 1973, An intertemporal capital asset pricing model, Econometrica, 41, 867-887.

Moosa, I. A. and N. E. Al-Loughani, 1994, Unbiasedness and time varying risk premia in the crude oil futures market, Energy economics, 16, 99-105.

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

Nelson, D. B.,and D. B. Foster, 1994, Asymptotic filtering theory for univariate ARCH models, Econometrica, 62, 1-41.

Paolella, M. S., 1999, Generalized GARCH models with nonnormal distributions, Unpublished Mimeo, Institute of Statistics and Econometrics, Kiel University.

Pagan, A. R. and Y.S. Hong, 1991, Nonparametric estimation and the risk premium, in: Barnett, W., Powell, J. and G. Tauchen, eds., Nonparametric and Semiparametric Methods in Economics and Statistics, Cambridge University Press, 51-75

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

Schwert, G. W., 1989, Why does stock market volatility change over time?, Journal of Finance, 44, 1115-1154.

Sentana, E., 1991, Quadratic ARCH models, Review of Economic Studies, 62(4), 639-661.

Taylor, S. J., 1986, Modelling financial time series, Wiley, New York, NY.

Turner, C. M. , Startz, R. and C. R. Nelson, 1989, A Markov model of heteroskedasticity, risk, and learning in the stock market, Journal of Financial Economics, 25, 3-22.

APPENDIX

A Proof of Proposition 1

In (2.2a) we add and subtract $ak_1 h_{t-1}^{\frac{\delta_i}{2}}$ and we get (2.3a).

We multiply (2.1) by $(1-\beta_1 L)$ and we use (2.3a) to get (2.4).

Equations (2.5b)-(2.7) follow from the canonical factorization of the auto/cross generating functions of the stock return and the $\frac{\delta_i}{2}$-th power of its conditional variance:

$$\begin{aligned} g_{rm}(z) &= \frac{(1-\theta z)(1-\theta z^{-1})}{(1-\phi z)(1-\phi z^{-1})} v(\varepsilon_t) + \frac{(\varsigma a)^2 v(v_{1t})}{(1-\phi z)(1-\phi z^{-1})(1-\beta_1 z)(1-\beta_1 z^{-1})} + \\ & \frac{2\varsigma a cov(\varepsilon_t, v_{1t})}{(1-\phi z)(1-\phi z^{-1})}[\frac{(1-\theta z)z^{-1}}{(1-\beta_1 z^{-1})} + \frac{(1-\theta z^{-1})z}{(1-\beta_1 z)}] = \sum_{m=-\infty}^{\infty} \gamma_{rm} z^m, \end{aligned} \tag{A.1a}$$

$$\begin{aligned} v(v_{1t}) &= [k_2 - k_1^2]E(h_t^{\delta_i}), \quad v(\varepsilon_t) = E(h_t), \quad cov(\varepsilon_t, v_{1t}) = E(h_t^{\frac{\delta_i}{2}+\frac{1}{2}})\lambda_{i1}, \\ g_{rh,m}^{+}(z) &= \frac{\varsigma a^2 v(v_{1t})}{(1-\phi z)(1-\beta_1 z)(1-\beta_1 z^{-1})} + \frac{a(1-\theta z)z^{-1} cov(\varepsilon_t, v_{1t})}{(1-\phi z)(1-\beta_1 z^{-1})} = \\ &= \sum_{m=0}^{\infty} \gamma_{rh,m}^{+} z^m + \sum_{m=-\infty}^{-1} \gamma_{rh,m}^{-} z^m, \quad g_{hm}(z) = \frac{a^2 v(v_{1t})}{(1-\beta_1 z)(1-\beta_1 z^{-1})} \end{aligned} \tag{A.1b}$$

>From equation (2.2a) we have

$$\begin{aligned} h_t^{\frac{\delta_i}{2}} &= \omega + \widehat{\beta} h_{t-1}^{\frac{\delta_i}{2}} \Rightarrow h_t^{\frac{\delta_i}{2}j} = (\omega + \widehat{\beta} h_{t-1}^{\frac{\delta_i}{2}})^j = \sum_{n=0}^{j} \binom{j}{n} \omega^{j-n} \widehat{\beta}^n h_{t-1}^{\frac{\delta_i}{2}n}, \\ \widehat{\beta}^n &= [af(e_{t-1}) + \beta]^n = \sum_{l=0}^{n} \binom{n}{l} \beta^{n-l} a^l f(e_{t-1})^l \end{aligned} \tag{A.2}$$

In the above equations taking expectations we get

$$\begin{aligned} E(h_t^{\frac{\delta_i}{2}j}) &= \frac{\sum_{n=0}^{j-1} \binom{j}{n} \omega^{j-n} \beta_n E(h_{t-1}^{\frac{\delta_i}{2}n})}{1-\beta_n}, \\ \beta_n &= E(\widehat{\beta}^n) = \sum_{l=0}^{n} \binom{n}{l} \beta^{n-l} a^l k_l \end{aligned} \tag{A.3}$$

where k_l denotes the l-th moment of $f(e_{t-1})$. For example, when the conditional distribution is the normal

$$k_l = \frac{1}{\sqrt{2\pi}}[(1+d_i)^{\delta_i l} + (1-d_i)^{\delta_i l}] 2^{\frac{\delta_i l - 1}{2}} \Gamma(\frac{\delta_i l + 1}{2}) \tag{A.4}$$

Equation (2.9) follows from (2.8a), when $j = 1, 2$.

■

B Proof of Proposition 2

Squaring (2.11a) we get

$$h_t^{\delta_i} = -\omega^2 + a^2 f(\varepsilon_{t-1})^2 + \beta^2 h_{t-1}^{\delta_i} + 2a\beta f(\varepsilon_{t-1}) h_{t-1}^{\frac{\delta_i}{2}} + 2\omega h_t^{\frac{\delta_i}{2}} \tag{B.1}$$

In the right hand side of the above equation we add and subtract $a^2 k_2 h_{t-1}^{\delta_i} + 2a\beta k_1 h_{t-1}^{\delta_i}$ to get

$$(1-\beta_2 L)h_t^{\delta_i} = -\omega^2 + a^2 v_{2,t-1} + 2a\beta v_{3,t-1} + 2\omega h_t^{\frac{\delta_i}{2}}, \tag{B.2a}$$

$$\beta_2 = \beta^2 + k_2 a^2 + 2ak_1\beta, \quad v_{2,t-1} = f(\varepsilon_{t-1})^2 - k_2 h_{t-1}^{\delta_i}, \tag{B.2b}$$

$$v_{3,t-1} = v_{1,t-1} h_{t-1}^{\frac{\delta_i}{2}} = f(\varepsilon_{t-1}) h_{t-1}^{\frac{\delta_i}{2}} - k_1 h_{t-1}^{\delta_i} \tag{B.2c}$$

Multiplying (B.2a) by $(1-\beta_1 L)$, where $\beta_1 = ak_1 + \beta$, and using

$$(1-\beta_1 L)h_t^{\frac{\delta_i}{2}} = \omega + av_{1,t-1} \tag{B.3}$$

we get (2.12a).
Moreover, the variances/covariances of the v_{1t}, v_{2t} and v_{3t} are

$$\begin{aligned} v(v_{1t}) &= (k_2-k_1^2)E(h_t^{\delta_i}), \quad v(v_{2t}) = (k_4-k_2^2)E(h_t^{2\delta_i}), & \text{(B.4a)}\\ v(v_{3t}) &= (k_2-k_1^2)E(h_t^{2\delta_i}), \quad cov(v_{2t},v_{3t}) = (k_3-k_1k_2)E(h_t^{2\delta_i}), & \text{(B.4b)}\\ cov(v_{1t},v_{2t}) &= (k_3-k_1k_2)E(h_t^{\frac{3\delta_i}{2}}), \quad cov(v_{1t},v_{3t}) = k_1(k_2-1)E(h_t^{\frac{3\delta_i}{2}}) & \text{(B.4c)} \end{aligned}$$

For example, the covariance between v_{2t} and v_{3t} can be derived by

$$\begin{aligned} cov(v_{2t},v_{3t}) &= E\{[f(\varepsilon_t)^2 - k_2 h_t^{\delta_i}][f(\varepsilon_t)h_t^{\frac{\delta_i}{2}} - k_1 h_t^{\delta_i}]\} = \\ &= E\{h_t^{2\delta_i}[f(e_t)^3 - k_1 f(e_t)^2 - k_2 f(e_t) + k_1k_2]\} = E(h_t^{2\delta_i})[k_3 - k_1k_2] \end{aligned} \tag{B.5}$$

Furthermore, the covariances between the ε_t term and the v_{1t}, v_{2t} and v_{3t} terms are

$$cov(\varepsilon_t, v_{1t}) = E(h_t^{\frac{\delta_i}{2}+\frac{1}{2}})\lambda_{i1}, \quad cov(\varepsilon_t, v_{2t}) = E(h_t^{\delta_i+\frac{1}{2}})\lambda_{i2}, \quad cov(\varepsilon_t, v_{3t}) = E(h_t^{\delta_i+\frac{1}{2}})\lambda_{i1}, \quad \lambda_{ij} = E[f(e_t)^j e_t]$$

Next, using equation (2.8a) we get the third and fourth moment of $h_t^{\frac{\delta_i}{2}}$

$$\begin{aligned} E(h_t^{\frac{3\delta_i}{2}}) &= \frac{\omega^3 - 3\omega^2 E(h_t^{\frac{\delta_i}{2}}) + 3\omega E(h_t^{\delta_i})}{1-\beta_3}, \\ E(h_t^{2\delta_i}) &= \frac{6\omega^4 + 4\omega^3 E(h_t^{\frac{\delta_i}{2}}) - 6\omega^2 E(h_t^{\delta_i}) + 4\omega E(h_t^{\frac{3\delta_i}{2}})}{1-\beta_4}, \\ \beta_3 &= k_3 a^3 + 3k_2 a_2^2 \beta + 3k_1 a\beta^2 + \beta^3, \\ \beta_4 &= k_4 a^4 + 4k_3 a^3 \beta + 6k_2 a^2 \beta^2 + 4k_1 a \beta^3 + \beta^4 \end{aligned} \tag{B.6}$$

The condition for the existence of $E(h_t^{\frac{3\delta_i}{2}})$, is $\beta_1, \beta_2, \beta_3 < 1$ and for the existence of $E(h_t^{2\delta_i})$ is $\beta_1, \beta_2, \beta_3, \beta_4 < 1$.

Moreover, to obtain the ARMA(3,3) representation of the stock return we multiply (2.10) by$(1-\beta_1 L)(1-\beta_2 L)$ and we use (2.11a).

Furthermore, the auto-covariances/correlations of the stock return can be derived using the canonical factorization of the autocovariance generating function:

$$\begin{aligned} g_{rm}(z) &= \frac{(1-\theta z)(1-\theta z^{-1})}{(1-\phi z)(1-\phi z^{-1})}v(\varepsilon_t) + \frac{a^2\varsigma^2[a^2 v(v_{2t}) + 4\beta^2 v(v_{3t}) + 4a\beta cov(v_{2t},v_{3t})]}{(1-\phi z)(1-\beta_2 z)(1-\phi z^{-1})(1-\beta_2 z^{-1})} + \\ &+ \frac{4\omega^2(a\varsigma)^2 v(v_{1t})}{(1-\phi z)(1-\phi z^{-1})(1-\beta_1 z)(1-\beta_1 z^{-1})(1-\beta_2 z)(1-\beta_2 z^{-1})} + \\ &+ \frac{(a\varsigma)^2 2\omega[acov(v_{2t},v_{1t}) + 2\beta cov(v_{3t},v_{1t})]}{(1-\beta_1 z)(1-\beta_1 z^{-1})}[\frac{1}{(1-\phi z)(1-\beta_2 z)} + \frac{1}{(1-\phi z^{-1})(1-\beta_2 z^{-1})}] \\ &+ \frac{a\varsigma[acov(v_{2t},\varepsilon_t) + 2\beta cov(v_{3t},\varepsilon_t)]}{(1-\phi z)(1-\phi z^{-1})}[\frac{(1-\theta z)z^{-1}}{(1-\beta_2 z^{-1})} + \frac{(1-\theta z^{-1})z}{(1-\beta_2 z)}] + \\ &\frac{2\omega a\varsigma cov(v_{1t},\varepsilon_t)}{(1-\phi z)(1-\phi z^{-1})}[\frac{(1-\theta z)z^{-1}}{(1-\beta_1 z^{-1})(1-\beta_2 z^{-1})} + \frac{(1-\theta z^{-1})z}{(1-\beta_1 z)(1-\beta_2 z)}] \\ &= \sum_{m=-\infty}^{\infty} \gamma_{rm} z^m \end{aligned} \tag{B.7}$$

In addition the cross-covariances/correlations) between the stock return and the $\frac{\delta_i}{2}$-power of its conditional variance can be derived using the canonical factorization of the cross-covariance generating function:

$$\begin{aligned} g^{+}_{rh,m}(z) &= \frac{2\omega\varsigma a^2 v(v_{1t})}{(1-\phi z)(1-\beta_1 z)(1-\beta_2 z)(1-\beta_1 z^{-1})} + \frac{a^2\varsigma[acov(v_{2t},v_{1t})+2\beta cov(v_{3t},v_{1t})]}{(1-\phi z)(1-\beta_2 z)(1-\beta_1 z^{-1})} + \\ &\quad + \frac{(1-\theta z)z^{-1}acov(v_t,\varepsilon_t)}{(1-\phi z)(1-\beta_1 z^{-1})} \\ &= \sum_{m=0}^{\infty}\gamma^{+}_{rh,m}z^m + \sum_{m=-\infty}^{-1}\gamma^{-}_{rh,m}z^m \end{aligned} \tag{B.8}$$

Finally, the auto-covariances/correlations of the δ_i-th power of the conditional variance can be derived using the canonical factorization of the autocovariance generating function:

$$\begin{aligned} g_{h2,m}(z) &= \frac{4\omega^2 a^2 v(v_{1t})}{(1-\beta_1 z)(1-\beta_2 z)(1-\beta_1 z^{-1})(1-\beta_2 z^{-1})} + \frac{a^2[a^2 v(v_{2t})+4\beta^2 v(v_{3t})+4a\beta cov(v_{2t},v_{3t})]}{(1-\beta_2 z)(1-\beta_2 z^{-1})} + \\ &\quad + \frac{2\omega a^2[acov(v_{1t},v_{2t})+2\beta cov(v_{1t},v_{3t})]}{(1-\beta_2 z)(1-\beta_1 z^{-1})}[\frac{1}{(1-\beta_1 z)}+\frac{1}{(1-\beta_2 z^{-1})}] \\ &= \sum_{m=-\infty}^{\infty}\gamma_{h2,m}z^m \end{aligned} \tag{B.9}$$

■

C Proof of Proposition 3

We multiply (2.18) by $(1-\beta L)$ and we use (2.19) to get the (2.20).The auto covariances/correlations of the stock return can be obtained from the cf of the agf of the stock return

$$\begin{aligned} g_{rm}(z) &= \frac{(1-\theta z)(1-\theta z^{-1})}{(1-\phi z)(1-\phi z^{-1})}v(\varepsilon_t) + \frac{(\varsigma c)^2 var(z_t)}{(1-\phi z)(1-\phi z^{-1})(1-\beta z)(1-\beta z^{-1})} + \\ &\quad \{\frac{(1-\theta z)\varsigma c z^{-1}}{(1-\phi z)(1-\phi z^{-1})(1-\beta z^{-1})} + \frac{(1-\theta z^{-1})\varsigma c z}{(1-\phi z)(1-\phi z^{-1})(1-\beta z)}\}cov(z_t,\varepsilon_t) \\ &= \sum_{m=-\infty}^{\infty}\gamma_{rm}z^m, \quad cov(z_t,\varepsilon_t) = dE(h_t^{\frac{1}{2}}), \quad v(\varepsilon_t) = E(h_t) \end{aligned} \tag{C.1}$$

In addition, the Wold representation of the conditional variance (eq. 2.19) is

$$\ln(h_t) = \frac{\omega}{1-\beta} + c\sum_{i=1}^{\infty}\beta^{i-1}z_{t-i} \Rightarrow \begin{cases} h_t^{k_1} = \varrho^{\frac{\omega k_1}{1-\beta}+ck_1\sum_{i=1}^{\infty}\beta^{i-1}z_{t-i}} \\ h_{t-m}^{k_2} = \varrho^{\frac{\omega k_2}{1-\beta}+ck_2\sum_{i=1}^{\infty}\beta^{i-1}z_{t-i-m}} \end{cases} \tag{C.2}$$

>From the above equation taking expectations we get

$$E(h_t^{k_1}h_{t-m}^{k_2}) = \varrho^{\frac{\omega(k_1+k_2)}{1-\beta}} \times E(\varrho^{ck_1\sum_{i=1}^{|m|}\beta^{i-1}z_{t-i}}) \times E(\varrho^{c(k_2+k_1\beta^{|m|})\sum_{i=1}^{\infty}\beta^{i-1}z_{t-i-|m|}}) \tag{C.3a}$$

$$E(h_t^{k_1}) = \varrho^{\frac{\omega k_1}{1-\beta}}E(\varrho^{ck_1\sum_{i=1}^{\infty}\beta^{i-1}z_{t-i}}) \tag{C.3b}$$

Next consider the expression A1.3 given in Theorem A1.1 in Nelson (1991)

$$E[e^{z_t b}] = e^{-b\gamma\sqrt{\frac{2}{\pi}}}[\Phi[(\gamma+d)b]\varrho^{\frac{b^2(\gamma+d)^2}{2}} + \Phi[(\gamma-d)b]\varrho^{\frac{b^2(\gamma-d)^2}{2}}] \tag{C.4}$$

where Φ denotes the cumulative normal distribution. The cumulative normal distribution can be expressed in terms of a hypergeometric function as (see Abadir, 1999)

$$\Phi(b) = \frac{1}{2} + \frac{b}{\sqrt{2\pi}}F(\frac{1}{2};\frac{3}{2};-\frac{b^2}{2}) \tag{C.5}$$

Using equations (C.3a)-(C.5) after some algebra we get equations (2.23a)-(2.24b).
Next, consider the expression A1.5 given in Theorem A1.2 in Nelson (1991)

$$E[\varrho^{z_t b}] = \varrho^{-b\gamma \frac{\Gamma(\frac{2}{v})\lambda 2^{\frac{1}{v}}}{\Gamma(\frac{1}{v})}} \sum_{\tau=0}^{\infty} (\lambda 2^{\frac{1}{v}} b)^{\tau} [(\gamma+d)^{\tau} + (\gamma-d)^{\tau}] \frac{\Gamma(\frac{\tau+1}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)} \tag{C.6}$$

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 (C.3a), (C.3b) and (C.6) after some algebra we get (2.23a), (2.23b), (2.25a) and (2.25b).
Finally, when $v = 1$, the above equation gives (if and only if $b\gamma + |bd| < \sqrt{2}$)

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

Using equations (C.3a), (C.3b) and (C.7) after some algebra we get (2.23a), (2.23b), (2.26a) and (2.26b). ■

D Proof of Proposition 4

From equation (2.27) we get

$$cov_m(r_t) = \varsigma^2 cov_m(h_t^k) + \varsigma E(h_t^k h_{t-m}^{\frac{1}{2}} e_{t-m}) - \varsigma\theta E(h_t^k h_{t-m-1}^{\frac{1}{2}} e_{t-m-1}) - \theta \xi_m E(h_{t-1}), \tag{D.1}$$

$$var(r_t) = \varsigma^2 var(h_t^k) + (1+\theta^2) E(h_t) - 2\varsigma\theta E(h_t^k h_{t-1}^{.5} e_{t-1}) \tag{D.2}$$

Using the Wold representation of the conditional variance (C.2) the above equation gives

$$\begin{aligned} cov_m(r_t) = {} & \varsigma^2 cov_m(h_t^k) + \varsigma \times \varrho^{\frac{\omega k.5}{1-\beta}} \{E(\varrho^{ck \sum_{i=1}^{|m|-1} \beta^{i-1} z_{t-i}}) \times \\ & E(e_{t-m} \varrho^{ck\beta^{|m|-1} z_{t-m}}) \times E(\varrho^{c(\frac{1}{2}+k\beta^{|m|}) \sum_{i=1}^{\infty} \beta^{i-1} z_{t-|m|-i}}) \\ & -\theta E(\varrho^{ck \sum_{i=1}^{|m|} \beta^{i-1} z_{t-i}}) \times E(e_{t-m-1} \varrho^{ck\beta^{|m|} z_{t-m-1}}) \times E(\varrho^{c(\frac{1}{2}+k\beta^{|m|+1}) \sum_{i=1}^{\infty} \beta^{i-1} z_{t-|m|-i}})\} - \\ & -\theta \xi_m \varrho^{\frac{\omega}{1-\beta}} E(\varrho^{c \sum_{i=1}^{\infty} \beta^{i-1} z_{t-i}}), \quad m \neq 0 \end{aligned} \tag{D.3}$$

$$\begin{aligned} var(r_t) = {} & \varsigma^2 var(h_t^k) + (1+\theta^2) \varrho^{\frac{\omega}{1-\beta}} \times E(\varrho^{ck \sum_{i=1}^{\infty} \beta^{i-1} z_{t-i}}) - \\ & -2\varsigma\theta \varrho^{\frac{\omega k.5}{1-\beta}} \times E(e_{t-1} \varrho^{ck z_{t-1}}) \times E(\varrho^{c(\frac{1}{2}+k\beta) \sum_{i=1}^{\infty} \beta^{i-1} z_{t-i-1}}) \end{aligned} \tag{D.4}$$

Next consider expression A1.4 given in Theorem A1.1 in Nelson (1991)

$$E(e_t \varrho^{z_t b}) = \varrho^{-b\gamma\sqrt{\frac{2}{\pi}}} \frac{1}{\sqrt{2\pi}} \{\varrho^{\frac{b^2(\gamma+d)^2}{4}} D_{-2}[-b(\gamma+d)] - \varrho^{\frac{b^2(\gamma-d)^2}{4}} D_{-2}[-b(\gamma-d)]\} \tag{D.5}$$

where $D_q[\cdot]$ denotes the parabolic cylinder function.
Using equations (2.23a), (2.23b), (C.4), (C.5), (D.3), (D.4) and (D.5) after some algebra we get (2.29b) and (2.30a).
Next consider expression A1.5 given in Theorem A1.2 in Nelson (1991)

$$E[e_t \varrho^{z_t b}] = \varrho^{-b\gamma \frac{\Gamma(\frac{2}{v})\lambda 2^{\frac{1}{v}}}{\Gamma(\frac{1}{v})}} \times 2^{\frac{1}{v}} \lambda \sum_{\tau=0}^{\infty} (\lambda 2^{\frac{1}{v}} b)^{\tau} [(\gamma+d)^{\tau} - (\gamma-d)^{\tau}] \frac{\Gamma(\frac{\tau+2}{v})}{2\Gamma(\frac{1}{v})\Gamma(1+\tau)} \tag{D.6}$$

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 (2.23a), (2.23b), (C.6), (D.3), (D.4) and (D.6) after some algebra we get (2.29b) and (2.30c). Finally, when $v = 1$, the above equation gives (if and only if $b\gamma + |bd| < \sqrt{2}$)

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

Using equations (2.23a), (2.23b), (C.7), (D.3), (D.4) and (D.7) after some algebra we get (2.29b) and (2.30d). >From equation (2.27) we get

$$cov(r_t, h_{t-m}^{k_1}) = \begin{cases} \varsigma cov(h_t^k, h_{t-m}^{k_1}) & \text{if} \quad m \geq 1 \\ \varsigma cov(h_t^k, h_t^{k_1}) - \theta E(h_t^{k_1} h_{t-1}^{.5} e_{t-1}) & \text{if} \quad m = 0 \\ \varsigma cov(h_t^{k_1}, h_{t-m}^k) + E(h_t^{k_1} h_{t-m}^{.5} e_{t-m}) - \theta E(h_t^{k_1} h_{t-m-1}^{.5} e_{t-m-1}) & \text{if} \quad m \leq -1 \end{cases} \tag{D.8}$$

Using a similar methodology to the one above, it can be seen that equation (D.8) gives (2.31a). ■