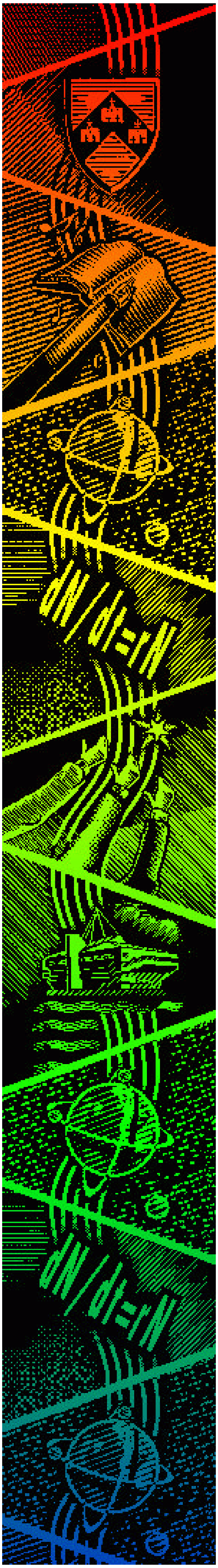

THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 1999/11

Prediction in ARMA models with GARCH in Mean Effects

by

Menelaos Karanasos

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

Prediction in ARMA models with GARCH in mean effects

Menelaos Karanasos ⋆

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

Abstract

This paper considers forecasting the conditional mean and variance from an ARMA model with GARCH in mean effects. Expressions for the optimal predictors and their conditional and unconditional MSE's are presented. We also derive the formula for the covariance structure of the process and its conditional variance.

Key Words: ARMA Model, GARCH in Mean Effects, Optimal Predictor, Autocovariances.
JEL Classification: C22

⋆ For correspodence: Email: mk16@york.ac.uk, Tel: 01904 433799, Fax: 01904 433759

1 Introduction

The autoregressive conditional heteroscedasticity (ARCH) model introduced by Engle (1982) and its generalisation, the GARCH model introduced by Bollerslev (1986) have become increasingly popular in modelling financial and economic variables (see for example, the surveys of Berra and Higgins (1992), Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), and for a more detailed description the book by Gourieroux (1997)). Following Engle's pathbreaking idea, several formulations of conditionally heteroscedastic models (e.g. Exponential GARCH, Fractional Integrated GARCH, Switching ARCH, Asymmetric Power ARCH, Component GARCH) have been introduced in the literature, forming an immense ARCH family.

Although the literature on GARCH type models is quite extensive relative fewer papers have examined the issue of forecasting in models where the conditional volatility is time-dependent. Engle and Kraft (1983) consider predictions from an ARMA process ARCH errors whereas Engle and Bollerslev (1986), hereafter EB, and Baillie and Bollerslev (1992), hereafter BB, consider predictions from an ARMA model with GARCH errors.

One important exclusion from this framework concerns the ARCH in mean model, introduced by Engle, Lilien and Robins (1987). This model was used to investigate the existence of time varying term premia in the term structure of interest rates. Such time varying risk premia have been strongly supported by a huge body of empirical research, in interest rates (Hurn, McDonald and Moody (1995)), in forward and future prices of commodities (Hall (1991), Moosa and Loughani (1994)), in GDP (Price (1994)), in industrial production (Caporale and McKierman (1996)) and especially in stock returns (see Campbell and Hentscel (1992), Glosten, Jagannathan and Runkle (1993), Black and Fraser (1995), Fraser (1996), Hansson and Hordahl (1997) , Elyasiani and Mansur (1998)).

In this paper we focus our attention on predictions from a general ARMA model with GARCH-in-mean effects. Many alternative expressions are available for the minimum mean square error (MMSE) predictor of the conditional mean from the univariate ARMA model (see Section 2 for a detailed discussion). In order to provide an analogy with subsequent material in Section 2 we present a new method for obtaining multiperiod predictions from the ARMA(r,s) model. In particular, we derive the optimal multistep predictor in terms of r past observations and s past errors. The coefficients in our formula are expressed in terms of the roots of the autoregressive polynomial (AR) and the parameters of the moving average (MA) one. We should mention that we only examine the case where the roots of the AR polynomial are distinct (the case of equal roots is left for future research). In Section 3 we use our method to derive the optimal predictor of future values for the conditional variance from the univariate GARCH model. Furthermore, in Section 4 of this paper we use our method to derive formulae for the multiperiod predictions of the ARMA model with GARCH in-mean effects. To point out the importance of our results we quote BB (1992): "Processes with feedback from the conditional variance to the conditional mean will considerably complicate the form of the predictor and its associated MSE. Analysis of such models is consequently left for future research".

The goal of our method is theoretical purity rather than the production of expressions intended for practical use. However, our method can be employed in the derivation of multistep predictions from a more complicated model with simultaneous feedback between the conditional mean and variance, namely the GARCH-M-X model [1] (see Christodoulakis, Hatgioannides and Karanasos,1999). Another value of our method is that it can easily be generalized to multivariate GARCH and GARCH in-mean models.

In addition to our method for obtaining a closed form expression for the optimal predictor (and its associated MSE) of the conditional mean from the ARMA(r,s) -GARCH(p,q) in-mean model, the following 3 substantive problems for which solutions do not exist in the GARCH literature are solved in this paper: (a) We give the infinite moving average (ima) representations of the conditional mean and variance (these formulae can be used to obtain alternative expressions for the MMSE predictors of the conditional mean and variance in terms of infinite past observations and errors); (b) We give the canonical factorization (cf) of the autocovariance generating function (agf) of the process and its conditional variance, the covariances between the squared errors[2] and the conditional variance, and finally the autocovariances and cross

[1] The GARCH-M-X model was introduced by Longstaff and Schwartz (1992). and includes the GARCH-X model of Brenner, Harjes and Kroner (1996) as a special case.

[2] The autocovariance function of the squared errors for the GARCH(p,q) models is given in the Karanasos (1999a) and He and Terasvirta (1997).

covariances for the process and its conditional variance; (iii) The MMSE predictor of future values of the squared conditional variance (which we subsequently use to obtain the conditional MSE associated with the MMSE predictor of the futures values of the conditional mean for the ARMA-GARCH in-mean model).

2 ARMA Model

2.1 *Forecasting with ARMA Models*

In this Section we consider multistep predictions from the ARMA(r,s) model:

$$\Phi(L)y_t = \phi + \Theta(L)\epsilon_t, \quad where \quad \Phi(L) = -\sum_{j=0}^{r} \phi_j L^j = \prod_{j=1}^{r}(1 - \lambda_j L), \qquad (2.\ 1)$$

$$and\ \Theta(L) = -\sum_{j=0}^{s} \theta_j L^j, \quad \phi_0 = \theta_0 = -1 \qquad (2.\ 1a)$$

Assumption 1: All the roots of the autoregressive polynomial $[\Phi(L)]$ and all the roots of the moving average polynomial $[\Theta(L)]$ lie outside the unit circle (Stationarity and Invertibility conditions).

Assumption 2: The polynomials $\Phi(L)$ and $\Theta(L)$ have no common left factors other than unimodular ones, i.e, if $\Phi(L) = U(L)\Phi_1(L)$ and $\Theta(L) = U(L)\Theta_1(L)$, then the common factor $U(L)$ must be unimodular (Irreducibility condition).

Many alternative expressions are available for the optimal predictor from the above ARMA model[3]. By expressing the ARMA(r,s) model in a companion representation form, BB (1992) derived the optimal multistep predictor in terms of r past observations and s past errors. In the following Proposition we obtain an expression for the optimal predictor, which is equivalent to the

[3]Yamamoto (1981) used the infinite order autoregressive representation of the multivariate ARMA model (which includes the univariate as a special case) to express the MMSE predictor as a function of an infinite number of past observations and he presented parametric expressions for the prediction weights. His prediction formula is particular convenient in obtaining the asymptotic prediction mean square error, when the prediction is formulated with estimated coefficients. Baillie (1980) obtained a formula for the MMSE predictor of an ARMAX model (which include the simple ARMA as a special case) in terms of an infinite number of past observations and he also gave parametric expressions for the prediction weights. Alternative prediction formulae, such as those based upon the Markovian representations of the ARMA model (Akaike, 1974) contain error terms in their formulae. The relative literature includes, among others, Yamamoto (1980, 1978, 1976), Baillie (1979), Schmidt (1974), Banshali (1974), Bloomfield (1972), and Davisson (1965).

BB one, by using a new method[4] for solving a linear homogeneous difference equation together with a technique for the manipulation of lag polynomials used in Sargent (1979). We believe that our method gives a useful insight into the treatment of stochastic difference equations. In what follows we only examine the case where the roots of the AR polynomial are distinct.

Proposition 2.1. Under assumptions 1-2, the minimum MSE predictor of y_t is

$$E_t(y_{t+i}) = \phi^{\star} + \sum_{n=0}^{s-1} z_{in}\epsilon_{t-n} + \sum_{n=0}^{r-1} x_{in}y_{t-n}, \quad where \tag{2. 2}$$

$$z_{in} = -\sum_{l=1}^{r}\sum_{j=n+1}^{min(i+n,s)} \zeta_{l0}\lambda_l^{i+n-j}\theta_j, \quad x_{in} = \sum_{j=1}^{r}\zeta_{ji}\gamma_{jn}, \quad \gamma_{j0}=1, \quad and \tag{2. 2a}$$

$$\gamma_{jn} = (-1)^n \prod_{l=1}^{n}[\sum_{\substack{k_l=k_{l-1}+1 \\ k_l\neq j}}^{r-(n-l)}]\prod_{l=1}^{n}(\lambda_{k_l}), \quad k_0=0, \quad \phi^{\star} = \phi[\frac{1}{\Phi(1)} - \sum_{l=1}^{r}\bar{\zeta}_{li}] \tag{2. 2b}$$

$$\zeta_{ji} = \frac{\lambda_j^{i+r-1}}{\prod_{\substack{l=1 \\ l\neq j}}^{r}(\lambda_j-\lambda_l)}, \quad \bar{\zeta}_{ji} = \frac{\zeta_{ji}}{(1-\lambda_j)} \tag{2. 2c}$$

The proof is given in the Appendix A. ■

Lemma 2.1. The i-period forecast error of the optimal predictor[5]

[4] We express the rth order difference equation as an AR(1) process with an error term which follows a r-1 difference equation. We obtain y_t as a function of r past values and the roots of the associated auxiliary equation. For an excellent discussion on solutions of linear difference equations see Wei (1989, pp. 27-30) or Brockwell and Davis (1983, Sect. 3.6).

[5] Pandit and Wu (1983, pp. 179-198) uses the ima representation of the ARMA(r,s) model in order to obtain the optimal predictor and its associated prediction error as a function of an infinite number pf past errors. The coefficients in their formula (which they called Green's function) are expressed in terms of the roots of the AR polynomial and the parameters of the MA. In the case of distinct roots these coefficients are given in Pandit and Wu (1983, pp. 105-106, when $s < r$) and in Pandit (1973, p. 100, when $s \gtrless r$). See also Pandit (1973, pp. 37-41) and Pandit and Wu (1983, pp. 177-179) for an excellent brief historical review of the prediction theory.

Wei (1989, pp. 23-27, 86-88) uses the ima representation to obtain MMSE forecasts, without giving a specific form for the prediction weights, and he also gives a recursive form for computing the optimal forecasts (pp. 91, 98). Brockwell and Davis (1987, Sect. 5.3, 5.5) present recursive methods for computing the best linear predictor and discuss ways to obtain the MMSE predictor based on the the infinite order ar and ma representations. Nerlove, Grether and Carvalho (1979, p. 93-94) provide a general scheme for computing least-squares forecasts which has been called unscrambling (see also Nerlove and Wage, 1964). Other conventional recursive expressions to obtain the predictor can also be found in Box and Jenkins (1971, pp. 126-132). Additional references on multistep predictions include the textbooks by Hamilton (1994, Ch 4), Granger and Newbold (1986, Ch 4), and Anderson (1976, Ch. 10).

$$FE_t(y_{t+i}) \;=\; \sum_{n=0}^{i-1} s_n \epsilon_{t+i-n}, \quad where \quad s_n = -\sum_{l=1}^{r} \sum_{j=0}^{min(n,s)} \zeta_{l0} \lambda_l^{n-j} \theta_j \qquad (2.\ 3)$$

With conditionally homoskedastic errors the conditional MSE for the optimal predictor is identical to the unconditional MSE of the same optimal predictor:

$$V_t[FE_t(y_{t+i})] = V_t(y_{t+i}) = V[FE_t(y_{t+i})] = E(\epsilon_t^2) \sum_{n=0}^{i-1} s_n^2 \qquad (2.\ 4)$$

The proof is given in the Appendix A. ∎

We now make a few remarks on the two prediction formulae. BB's method is not restricted to the case of distinct roots and is intended for practical computation whereas the goal of our method is theoretical purity rather than the production of expressions intended for practical use. The coefficients in BB's formula are expressed in terms of parameter matrices whereas in our formula they are expressed in terms of the roots of the AR polynomial and the parameters of the MA one. In the context of distinct roots the two expressions are equivalent[6]. While both methods can be used to derive the optimal predictor of future values for the conditional variance from the univariate GARCH model (see Section 3 of this paper, and Section 4 of BB, 1992), the advantage of our method is that it can be used to derive formulae for the multiperiod predictions from the ARMA model with GARCH-in-mean effects (see Section 4) and the GARCH-M-X model[7](see CHK, 1999). In addition our method can be applied to even more complicated GARCH models like the Component GARCH, the Asymmetric Power ARCH and the Switching ARCH[8]. Another value of our method is that it can easily be generalized to multivariate GARCH and GARCH-in-mean models[9].

[6]Algebraic manipulations of equations 16 and 17 in BB's paper should give the coefficients in our equation (2. 2*a*).

[7]The GARCH-M-X model was introduced by Longstaff and Schwartz (1992) as a discretization of their two-factor short-term interest rate model.

[8]The component GARCH model was introduced by Ding and Granger (1996),the Asymmetric Power ARCH was introduced by Ding, Engle and Granger (1992), and the Switching ARCH was introduced by Hamilton and Susmel (1996).

[9]The statistical properties of the multivariate GARCH (MGARCH) model have been recently examined by researchers. For example, Lin (1997) analyses the impulse response function for conditional volatility in various MGARCH models. For a theoretical and empirical analysis of the multivariate GARCH-in-mean models see Song (1996).

3 GARCH Model

3.1 *Forecasting the Volatility*

In this Section we consider multistep prediction from the GARCH(p,q) model:

$$B(L)h_t = \omega + A(L)\epsilon_t^2, \quad \text{where} \tag{3. 1}$$

$$B(L) = -\sum_{i=0}^{p} \beta_i L^i = \prod_{i=1}^{p}(1 - f_i L), \quad A(L) = \sum_{i=1}^{q} a_i L^i, \quad \beta_0 = -1 \tag{3. 1a}$$

Corollary 3.1. The ARMA representation of the conditional variance is given by

$$B^\star(L)h_t \quad = \quad \omega + A(L)v_t, \quad \text{where} \quad B^\star(L) = B(L) - A(L) = \prod_{i=1}^{p^\star}(1 - f_i^\star L) \tag{3. 2}$$

$p^\star = max(p, q)$ and $v_t = \epsilon_t^2 - h_t$. The v_t's are uncorrelated but they are not independent and have a very complicated distribution. It is not difficult to show that under conditional normality $var(v_t) = (2/3)E(\epsilon_t^4)$ (see K ,1999a). The fourth moment of the errors are given in Ling (1999), He and Terasvirta (1997) and Karanasos (1999a). A general condition for the existence of the 2mth moments is given by Ling (1999) (see also Ling and Li, 1997).

Proof. In (3. 1) we add and subtract $A(L)h_t$ and we get (3. 2).

Assumption 3: All the roots of the autoregressive polynomial $[B^\star(L)]$ *and all the roots of the MA polynomial* $[A(L)]$ *lie outside the unit circle (Stationarity and Invertibility condition).*

Assumption 4: The polynomials $B^\star(L)$ *and* $A(L)$ *are left coprime. In other words the representation* $\frac{B^\star(L)}{A(L)}$ *is irreducible.*

BB (1992) show that the squared errors (ϵ_t^2) correspond to an ARMA($p^\star$,p) process and express it in a first-order companion form in order to derive the optimal multistep predictor for the conditional variance from the GARCH model in terms of $p^\star$ values of past squared errors and p values of past conditional variances. In the following Proposition, in analogy to the derivation of (2. 2), we solve the linear stochastic difference equation (3. 2) in order to obtain the MMSE predictor for the conditional variance in terms of $p^\star$ values of past observations and q values of past squared errors. In what follows we only examine the case where the roots of the AR polynomial $[B^\star(L)]$ are distinct.

Proposition 3.1. Under assumptions 3 and 4 the MMSE i-step ahead forecast for the conditional variance from the GARCH(p,q) model is given by

$$E_t(h_{t+i}) = \omega^\star + \sum_{k=1}^{q-1} n_{ik} v_{t-k} + \sum_{n=0}^{p^\star-1} m_{in} h_{t-n}, \quad where \quad \omega^\star = \omega[\frac{1}{B^\star(1)} - \sum_{l=1}^{p^\star} \bar{k}^\star_{li}] \tag{3. 3}$$

$$n_{ik} = \sum_{l=1}^{p^\star} \sum_{j=k+1}^{min(i+k,q)} k^\star_{l0}(f^\star_l)^{i+k-j} a_j, \quad m_{in} = \sum_{j=1}^{p^\star} k^\star_{ji}\pi_{jn}, \quad \pi_{j0} = 1 \quad and \tag{3. 3a}$$

$$k^\star_{ji} = \frac{(f^\star_j)^{i+p^\star-1}}{\prod_{\substack{k=1\\k\neq j}}^{p^\star} (f^\star_j - f^\star_k)}, \quad \pi_{jn} = (-1)^n \prod_{l=1}^{n} [\sum_{\substack{k_l=k_{l-1}+1\\k_l\neq j}}^{p^\star-(n-l)}] \prod_{l=1}^{n} (f^\star_{k_l})$$

$$k_0 = 0, \quad \bar{k}^\star_{ji} = \frac{k^\star_{ji}}{1 - f^\star_j} \tag{3. 3b}$$

where the third term represents the definite solution of the deterministic (and homogeneous) part of the stochastic difference equation (3. 2), and $(f^\star_j)$, $j = 1, \cdots, max(p,q)$ are the inverse of the roots of the $B^\star(L)$ polynomial. The proof is similar to that of Proposition 2.1. ■

Note that the coefficients in BB's formula are expressed in terms of parameter matrices whereas in our formula they are expressed in terms of the roots of the AR polynomial and the parameters of the MA one.

Corollary 3.2. The ma representation of h_t as a function of infinite values of past v_t's is given by

$$h_{t+i} = \frac{\omega}{B^\star(1)} + \sum_{n=1}^{\infty} \delta_n v_{t+i-n}, \quad where \quad \delta_n = \sum_{l=1}^{p^\star} \sum_{j=1}^{min(n,q)} k^\star_{l0}(f^\star_l)^{n-j} a_j, \tag{3. 4}$$

The proof follows immediately from the univariate ARMA representations (3. 2) and the ima representation of an ARMA model which is given in Pandit (1973, p.100) and Pandit and Wu (1983, p. 105) and Pandit (1983).

The above representation is very useful because it can be used to obtain the forecast error of the optimal predictor for the conditional variance (see Lema 3.1) and its autocovariance function (see Proposition 3.2).

Moreover, the cf of the agf of h_t is given by

$$g_z(h) = \sum_{m=0}^{\infty} f_m \gamma_m (z^m + z^{-m}) = \frac{A(z)A(z^{-1})}{B^\star(z)B^\star(z^{-1})}\sigma_v^2 \tag{3. 5}$$

The γ_m are given in Proposition 3.2. The proof follows immediately from the univariate ARMA representation (3. 2) and the cf of the agf of an ARMA model discussed, for example, in NGC (1979, p. 39) and Sargent (1979, p. 228). ■

In many applications in financial economics the primary interest centres on the forecast for the future conditional variance. Such instances include option pricing as discussed by Day and Lewis (1992) and Lamourex and Lastrapes (1990), the efficient determination of the market rate of return as examined in Chou (1988), and the relationship between stock market volatility and the business cycle as analysed by Schwert (1989). In these situations it is therefore of interest to be able to characterise the uncertainty associated with the forecasts for the future conditional variances also. Some potentially useful results for this purpose are given by Lemma 3.1 and Theorem 3.1.

Lemma 3.1. The forecast error associated with the i-step-ahead predictor for the conditional variance from the GARCH(p,q) model is given by

$$FE_t(h_{t+i}) = h_{t+i} - E_t(h_{t+i}) = \sum_{n=1}^{i-1} \delta_n v_{t+i-n} \tag{3. 6}$$

In addition, the unconditional and conditional MSE are given by

$$V[FE_t(h_{t+i})] = var(v_t) \sum_{n=1}^{i-1} \delta_n^2 = (2/3) E(\epsilon_t^4) \sum_{n=1}^{i-1} \delta_n^2 \tag{3. 7}$$

$$V_t(h_{t+i}) = V_t[FE_t(h_{t+i})] = 2 \sum_{k=1}^{i-1} \delta_k^2 E_t(h_{t+i-k}^2) \tag{3. 8}$$

The proof follows directly from (3. 4). ■

Theorem 3.1 The i-period ahead forecast for the squared conditional variance h_t^2 is given by

$$E_t(h_{t+i-k}^2) = \bar{\omega} + \sum_{j=0}^{q-1} \psi_{j,i-k}^{\epsilon} \epsilon_{t-j}^2 + \sum_{j=0}^{p-1} \psi_{j,i-k}^{h} h_{t-j} + \sum_{j=0}^{q-1} \psi_{j,i-k}^{\epsilon^2} \epsilon_{t-j}^4$$

$$+ \sum_{k_1=1}^{q-1} \sum_{k_2=k_1+1}^{q} \epsilon_{t-k_1}^2 \epsilon_{t-k_2}^2 \psi_{k_1 k_2, i-k}^{\epsilon^2} + \sum_{j=0}^{p-1} \psi_{j,i-k}^{h^2} h_{t-j}^2$$

$$+ \sum_{l=0}^{q-1} \sum_{j=0}^{p-1} \epsilon_{t-l}^2 h_{t-j} \psi_{lj,i-k}^{\epsilon h} + \sum_{k_1=1}^{p-1} \sum_{k_2=k_1+1}^{p} h_{t-k_1} h_{t-k_2} \psi_{k_1 k_2, i-k}^{h^2} \qquad (3.\ 9)$$

where all the ψ's are functions of the GARCH parameters and are given, together with the proof, in the Appendix B. The above expression is very useful because the optimal forecasts of the squared conditional variance is needed in order to obtain the conditional variance of the forecast error associated with the optimal forecast for the conditional mean from the ARMA-GARCH in-mean model (see Section 4, Lemma 4.1). ■

In what follows we examine the covariance structure of the GARCH(p,q) model. The autocovariance function of the squared errors from the GARCH model is given in K (1999a) and He and Terasvirta (1997). In many cases it is useful to have the autocovariances of the conditional variance. These autocovariances can be used, for example, to obtain the autocovariances of the ARMA process with GARCH in-mean effects. (see Section 4, Theorem 4.1b).

Proposition 3.2. The autocovariance function (af) of the conditional variance h_t is given by

$$cov_j(h_t) = \gamma_j = \sum_{i=1}^{p^\star} e_{ij}\eta_{i,min(j,q)}\sigma_v^2, \quad where \tag{3. 10}$$

$$e_{ij} = \frac{(f_i^\star)^j (f_i^\star)^{p^\star-1}}{\prod_{l=1}^{p^\star}(1-f_l^\star f_i^\star)\prod_{\substack{k=1\\k\neq i}}^{p^\star}(f_i^\star - f_k)}, \quad and \tag{3. 10a}$$

$$\eta_{i,min(j,q)} = \sum_{k=1}^{q} a_k^2 + \sum_{l=1}^{j}\sum_{k=1}^{q-l} a_k a_{k+l}[(f_i^\star)^l + (f_i^\star)^{-l}] + \sum_{l=j+1}^{q}\sum_{k=1}^{q-l} a_k a_{k+l}[(f_i^\star)^l + (f_i^\star)^{l-2j}] \tag{3. 10b}$$

The covariances between the squared errors ($\epsilon_{t-l_1}^2$) and the conditional variances (h_{t-l_2}) can be derived by using the following equations:

$$E(\epsilon_t^2\epsilon_{t-m}^2) = E(h_t\epsilon_{t-m}^2), \;\; E(h_t h_{t-m}) = E(\epsilon_t^2 h_{t-m})$$

together with the af of the squared errors and the conditional variances.

The proof is similar to that of Theorem 1 in Karanasos (1999a). Alternatively, one can use either the ima representation (3. 4) or the cf of the agf (3. 5) to obtain the acf of the conditional variance (see K, 1999c). ■

3.2 *Forecasting with ARMA-GARCH Models*

In this subsection we consider the ARMA(r,s)-GARCH(p,q) model given by

$$\Phi(L)y_t = \phi + \Theta(L)\epsilon_t, \quad (\epsilon_t|\Omega_{t-1}) \sim N(0,h_t), \tag{3. 11}$$

$$B^\star(L)h_t = \omega + A(L)v_t \tag{3. 11a}$$

As BB (1992) note, in the absence of GARCH in mean effects, the actual form of the predictor of the future values of the conditional mean is the same as in the homoskedastic case, but the presence of GARCH changes the MSE of the predictor. The Proposition that follows gives the associated unconditional and conditional MSE.

Proposition 3.3. The conditional and the unconditional MSE associated with the optimal forecasts for the mean in the general ARMA(r,s)-GARCH(p,q) class of models are

$$V[FE_t(y_{t+i})] = E(\epsilon_t^2)\sum_{n=0}^{i-1} s_n^2 \tag{3. 12}$$

$$\star\, V_t[FE_t(y_{t+i})] = V_t(y_{t+i}) = \sum_{k=0}^{i-1} s_k^2 E_t(h_{t+i-k}) = \sum_{k=0}^{i-1} s_k^2 \{\sum_{l=0}^{q-1} n_{i-k,l} v_{t-l} + \sum_{l=0}^{p^\star-1} m_{i-k,l} h_{t-l}\} \tag{3. 13}$$

($\star$ apart from a constant).

The proof follows immediately from equations (2. 3) and (3. 3). ■

4 GARCH in-mean Model

To our knowledge, the analysis of the covariance structure and of the multistep predictions from a general ARMA model with GARCH errors and in-mean effects has not been considered yet. This Section attempts to fill this gap in the literature.

In what follows we will consider the ARMA(r,s)-GARCH(p,q)-M(1) process:

$$\Phi(L)y_t = \phi + \delta h_t + \Theta(L)\epsilon_t, \text{ and} \tag{4. 1}$$

$$B(L)h_t = \omega + A(L)\epsilon_t^2, \text{ or } B^\star(L)h_t = \omega + A(L)v_t \tag{4. 1a}$$

where $\Phi(L)$ and $\Theta(L)$ are given by (2. 1), (2. 1a), and B(L) and $B^\star(L)$ are given by (3. 1a) and (3. 2).

Corollary 4.1. The univariate ARMA representation of y_t is given by

$$B^{\star}(L)\Phi(L)y_t = \phi^{\circ} + \delta A(L)v_t + \Theta(L)B^{\star}(L)\epsilon_t, \;\; \phi^{\circ} = \phi B^{\star}(1) + \delta\omega \qquad (4.\ 2)$$

Proof. Multiplication of (4. 1) by $B^{\star}(L)$ and substitution of (4. 1a) into (4. 1) gives (4. 2).

■

Assumption 5. The polynomials $\Phi(L)$ and $A(L)$ are left coprime. In other words the representation $\frac{A(L)}{\Phi(L)}$ is irreducible.

In what follows we only examine the case where the roots of the AR polynomials $[\Phi(L), B^{\star}(L)]$ are distinct.

Corollary 4.2. Under assumptions 1-5, the canonical factorization (cf) of the autocovariance generating function (agf) for y_t is given by

$$g_z(y) = \frac{\delta A(z)A(z^{-1})\sigma_v^2}{\Phi(z)B^{\star}(z)\Phi(z^{-1})B^{\star}(z^{-1})} + \frac{\Theta(z)\Theta(z^{-1})\sigma_\epsilon^2}{\Phi(z)\Phi(z^{-1})} = \sum_{j=0}^{\infty} f_j\gamma_j(z^j + z^{-j}), \;\; f_j = \begin{cases} .5 & \text{if } j=0 \\ 0 & \text{otherwise} \end{cases} \qquad (4.\ 3)$$

where the γ_j's are given in Theorem 4.1b.

Proof. The proof follows immediately from the univariate ARMA representation (4. 2) and the cf of the agf of an ARMA model given in NCG (1979, pp. 70-78) and Sargent (1979, p. 228).

Under assumptions 1-5, the infinite-order ma representation of y_t is given by

$$y_t = \frac{\phi^{\circ}}{B^{\star}(1)\Phi(1)} + \sum_{n=0}^{\infty}[\delta g_n v_{t+i-n} + s_n\epsilon_{t+i-n}], \;\; \text{where } g_n = \sum_{l=1}^{r+p^{\star}} \sum_{j=1}^{min(n,q)} u_{l0} y_l^{n-j} a_j, \qquad (4.\ 4)$$

$$u_{lt} = \begin{cases} \dfrac{\lambda_l^{t+r+p^{\star}-1}}{\prod\limits_{\substack{j=1\\ j\neq l}}^{r}(\lambda_l-\lambda_j)\prod\limits_{j=1}^{p^{\star}}(\lambda_l-f_j^{\star})} & \text{if } l = 1,2\cdots,r \\[2ex] \dfrac{(f_m^{\star})^{t+r+p^{\star}-1}}{\prod\limits_{\substack{j=1\\ j\neq m}}^{p^{\star}}(f_m^{\star}-f_j^{\star})\prod\limits_{j=1}^{r}(f_m^{\star}-\lambda_j)} & \text{if } l = r+m, \;\; 1 \leq m \leq p^{\star} \end{cases} \qquad (4.\ 4a)$$

and $y_l = \lambda_l$, for $l = 1, \cdots, r$, $y_l = f^\star_m$, for $l = r+m,\ 1 \le m \le p^\star$ and s_n is given by (2. 3).

Proof. The proof follows directly from the univariate ARMA representation (4. 2) and the Wold representation of an ARMA model given in Pandit and Wu (1983, p. 105) and Pandit (1973). ■.

In the following Theorem we present closed form algebraic expressions for the optimal predictor (and its associated MSE) of future values for the conditional mean from the above model.

Theorem 4.1a. Under assumptions 1-5 the i-step-ahead predictor of y_t is readily seen to be

$$E_t(y_{t+i}) = \phi' + \delta \sum_{n=0}^{q-1} z^\circ_{in} v_{t-n} + \sum_{n=0}^{s-1} z_{in}\epsilon_{t-n} + \sum_{n=0}^{r+p^\star-1} x^\circ_{in} y_{t-n}, \text{ where} \tag{4. 5}$$

$$z^\circ_{in} = \sum_{l=1}^{r+p^\star} \sum_{j=n+1}^{min(i+n,q)} u_{l0}(\lambda^\circ_l)^{i+n-j} a_j,\ x^\circ_{in} = \sum_{j=1}^{r+p^\star} u_{ji}\gamma^\circ_{j'n},\ \gamma_{j'0} = 1 \tag{4. 5a}$$

$$\gamma^\circ_{j'n} = (-1)^n \prod_{l=1}^{n} [\sum_{\substack{k_l=k_{l-1}+1 \\ k_l \neq j'}}^{r+p^\star-(n-l)}] \prod_{l=1}^{n}(\lambda^\circ_{k_l}),\ k_0 = 0,\ \lambda^\circ_{k_l} = \begin{cases} \lambda_{k_l} & \text{if } k_l = 1, \cdots, r \\ f^\star_m & \text{if } k_l = r+m,\ 1 \le m \le p^\star \end{cases} \tag{4. 5b}$$

$$j' = \begin{cases} j & \text{if } j = 1, \cdots, r \\ m & \text{if } j = r+m,\ 1 \le m \le p^\star \end{cases},\ \phi' = \phi^\circ[\frac{1}{B^\star(1)\Phi(1)} - \sum_{l=1}^{r+p^\star} \bar{u}_{li}],\ \bar{u}_{li} = \frac{u_{li}}{1-\lambda^\circ_i} \tag{4. 5c}$$

where u_{lt} is given in (4. 4a) and the z_{in} and x_{in} are given in equation (2. 2a).

It is important to note that in the presence of GARCH in mean effects the optimal predictor for the conditional mean is a function of past values not only of the observations and the errors $(y_{t-n},\ \epsilon_{t-n})$ but of the conditional variances and the squared errors $(h_{t-n},\ \epsilon^2_{t-n})$ as well.

Proof. The proof follows directly from the univariate ARMA representation (4. 2) and the methodology used in Proposition 2.1. ■

Lemma 4.1. The forecast error for the above i-step-ahead predictor is given by

$$FE_t(y_{t+i}) = \sum_{n=0}^{i-1}[\delta g_n v_{t+i-n} + s_n \epsilon_{t+i-n}] \quad (4.\ 6)$$

with unconditional and conditional MSE given by

$$V[FE_t(y_{t+i})] = \delta^2(2/3)E(\epsilon_t^4)\sum_{n=1}^{i-1} g_n^2 + E(\epsilon_t^2)\sum_{n=0}^{i-1} s_n^2 \quad (4.\ 7)$$

$$V_t[FE_t(y_{t+i})] = 2\delta^2 \sum_{n=1}^{i-1} g_n^2 E_t(h_{t+i-n}^2) + \sum_{n=0}^{i-1} s_n^2 E_t(h_{t+i-n}) \quad (4.\ 8)$$

where s_n is given in (2. 3), and g_n is given in (4. 4).

Note that in the presence of GARCH in mean effects the conditional MSE is a function not only of the forecasts of the future values of the conditional variance (eq. 3. 3) but of the squared conditional variance (eq. 3. 9) as well. When we don't have GARCH in mean effects ($\delta = 0$) equations (4. 5), and (4. 6)-(4. 8) reduces to the equivalent expressions in Section 3. The proof follows immediately from the infinite-order ma representation (eq. 4. 4). ■

In the following Theorem we give a formula for the covariance structure of the ARMA-GARCH in-mean model which include several simpler models as special cases.

Theorem 4.1b. Under Assumptions 1-5 the autocovariance function of the above process is given by

$$\gamma_j = cov_j(y_t) = \sum_{i=1}^{r} e_{ij} z_{i,min(j,s)} var(\epsilon_t) + \sum_{i=1}^{r+p^\star} \pi_{ij} d_{i,min(j,q)} var(v_t), \quad where \quad (4.\ 9)$$

$$e_{ij} = \frac{\lambda_i^{j+r-1}}{\prod_{l=1}^{r}(1-\lambda_l\lambda_i)\prod_{\substack{k=1\\k\neq i}}^{r}(\lambda_i-\lambda_k)} \quad (4.\ 9a)$$

$$\pi_{ij} = \begin{cases} \dfrac{\lambda_i^{j+r+p^\star-1}}{\prod_{l=1}^{r}(1-\lambda_i\lambda_l)\prod_{\substack{k=1\\k\neq i}}^{r}(\lambda_i-\lambda_k)\prod_{l=1}^{p^\star}(1-\lambda_i f_l^\star)\prod_{k=1}^{p^\star}(\lambda_i-f_k^\star)} & \text{if } i = 1,\cdots,r \\ \dfrac{(f_n^\star)^{j+r+p^\star-1}}{\prod_{l=1}^{r}(1-f_n^\star\lambda_l)\prod_{l=1}^{p^\star}(1-f_n^\star f_l^\star)\prod_{k=1}^{r}(f_n^\star-\lambda_k)\prod_{\substack{k=1\\k\neq i}}^{p^\star}(f_n^\star-f_k^\star)} & \text{if } i = r+n,\ 1 \le n \le p^\star \end{cases} \quad (4.\ 9b)$$

$$z_{i,min(j,s)} = \sum_{k=0}^{s} \theta_k^2 + \sum_{l=1}^{j} \sum_{k=0}^{s-l} \theta_k \theta_{k+l} (\lambda_i^l + \lambda_i^{-l}) + \sum_{l=j+1}^{s} \sum_{k=0}^{s-l} \theta_k \theta_{k+l} (\lambda_i^l + \lambda_i^{l-2j}) \tag{4. 9c}$$

$$d_{i,min(j,q)} = \sum_{k=1}^{q} a_k^2 + \sum_{l=1}^{j} \sum_{k=1}^{q-l} a_k a_{k+l} [(\lambda_i^\circ)^l + (\lambda_i^\circ)^{-l}] + \sum_{l=j+1}^{q} \sum_{k=1}^{q-l} a_k a_{k+l} [(\lambda_i^\circ)^l + (\lambda_i^\circ)^{l-2j}] \tag{4. 9d}$$

and $\lambda_i^\circ = \lambda_i$, for $i = 1, \cdots, r$, $\lambda_i^\circ = f_n^\star$, for $i = r + n$, $1 \leq n \leq p^\star$

Observe that the above general formula incorporates the following results as special cases:

(a) the acf for the white noise process with GARCH(1,1) in mean effects given in Hong (1991), (b) the acf for the ARMA(r,s) model given in Zinde-Walsh (1988), and Karanasos (1998, 1999b), and (c) the acf of the conditional variance for the GARCH(p,q) model given in Section 3 (Proposition 3.2).

Proof. The covariance structure can be derived by using the following three alternative methods: (i) the one used in Karanasos (1999a), (ii) the one based on the cf of the agf (4. 3) and (iii) one based on the infinite-order ma representation[10] (4. 4).

■

Theorem 4.1c. The cf of the agf between y_t and h_t ($g_z(yh)$) is given by

$$g_z(yh) = \sum_{m=-\infty}^{\infty} \gamma_m z^m = \frac{A(z)A(z^{-1})}{\Phi(z)B^\star(z)B^\star(z^{-1})} \delta \sigma_v^2 \tag{4. 10}$$

The proof follows directly from the univariate ARMA representations (4. $1a$), (4. 2) and the cf of the agf of ARMA processes given in Sargent (1979, p. 228).

Moreover, the cross covariances (γ_m) are given by

[10] See Karanasos (1999c) for the use of methods (ii) and (iii) in the context of univariate and multivariate GARCH models.

$$\gamma_m = cov(y_t, h_{t-m}) = \begin{cases} \sum_{i=1}^{r+p^\star} e_{im}^{\lambda^\circ\star} z_{i,min(m,q)}^{\lambda^\circ} + \sum_{i=1}^{p^\star} e_{im}^{f\,\star} z_{i,m}^{f} & \text{if } m > 0 \\ \sum_{i=1}^{p^\star} e_{im}^{f\star} z_{i,min(m,q)}^{f} + \sum_{i=1}^{r+p^\star} e_{im}^{\lambda^\circ\star} z_{i,m}^{\lambda^\circ} & \text{if } m < 0 \end{cases}, \text{ and} \qquad (4.\ 11)$$

$$e_{im}^{\lambda^\circ\star} = \frac{e_{im}^{\lambda^\circ}}{\prod_{k=1}^{p^\star}(1-\lambda_i^\circ f_k^\star)}, \quad e_{im}^{\lambda^\circ} = \frac{(\lambda_i^\circ)^{r+p^\star-1+m}}{\prod_{\substack{k=1 \\ k\neq i}}^{r+p^\star}(\lambda_i^\circ - \lambda_k^\circ)}, \qquad (4.\ 11a)$$

$$e_{im}^{f\star} = \frac{e_{im}^{f}}{\prod_{k=1}^{r+p^\star}(1-f_i^\star \lambda_k^\circ)}, \quad e_{im}^{f} = \frac{(f_i^\star)^{p^\star-1+m}}{\prod_{\substack{k=1 \\ k\neq i}}^{p^\star}(f_i^\star - f_k^\star)}, \quad z_{i,m}^{f} = \sum_{l=m+1}^{q} \sum_{k=1}^{q-l} a_k a_{k+l} (f_i^\star)^{l-2m} \qquad (4.\ 11b)$$

$$z_{i,min(m,q)}^{f} = \sum_{k=1}^{q} a_k^2 + \sum_{l=1}^{q} \sum_{k=1}^{q-l} a_k a_{k+l} (\lambda_i^\circ)^l + \sum_{l=1}^{m} \sum_{k=1}^{q-l} a_k a_{k+l} (\lambda_i^\circ)^{-l} \qquad (4.\ 11c)$$

Proof. The cross covariances (γ_m) can be obtained by using either the cf of the agf (4. 10) or the infinite-order ma representations of the process and its conditional variance (4. 4), (3. 4) together with the techniques given in K(1999c). ■

Corollary 4.3. The bivariate ARMA representation of the GARCH-in-mean model is given by

$$\bar{\Phi}(L)\bar{y}_t = \bar{A}_0 + \bar{\Theta}(L)\bar{\epsilon}_t, \text{ where } \bar{\Phi}(L) = -\sum_{l=0}^{r^\star} \bar{\Phi}_l L^l, \ \bar{\Theta}(L) = \sum_{l=0}^{s^\star} \bar{\Theta}_l L^l, \ \begin{cases} r^\star = max(r, p^\star) \\ s^\star = max(s, q) \end{cases} \qquad (4.\ 12)$$

$$\bar{\Phi}_l = \begin{bmatrix} \phi_l' & 0 \\ 0 & \beta_l^{\star\prime} \end{bmatrix}, \ \bar{\Theta}_l = \begin{bmatrix} -\theta_l' & 0 \\ 0 & a_l' \end{bmatrix}, \ \bar{\Phi}_0 = -\begin{bmatrix} 1 & \delta \\ 0 & 1 \end{bmatrix}, \ \bar{\Theta}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad (4.\ 12a)$$

$$\phi_l' = \begin{cases} \phi_l & \text{if } l \le r \\ 0 & \text{if } l > r \end{cases}, \ \beta_l^{\star\prime} = \begin{cases} \beta_l^{\star\prime} & \text{if } l \le p^\star \\ 0 & \text{if } l > p^\star \end{cases}, \ \theta_l' = \begin{cases} \theta_l & \text{if } l \le s \\ 0 & \text{if } l > s \end{cases}, \ a_l' = \begin{cases} a_l & \text{if } l \le q \\ 0 & \text{if } l > q \end{cases} \qquad (4.\ 12b)$$

We can use the above bivariate ARMA representation and the techniques in Yamamoto

(1981) to obtain expressions for the optimal predictors and their MSE in computationally convenient algorithmic forms.

5 Concluding Remarks

Despite the extensive literature on GARCH and related models, relatively little attention has been given to the issue of forecasting in models where time-dependent conditional heteroscedasticity is present.

In this paper we focused on the prediction from an ARMA model with GARCH in mean effects. We showed that for processes with feedback from the conditional variance to the conditional mean the forms of the optimal predictor of the process and its MSE are considerably complicated. In addition, we gave the Wold representations of the conditional mean and variance of the process. These formulae can be used to obtain alternative expressions for the MMSE predictors of the process and its conditional variance in terms of an infinite number of past observations and errors. Moreover, we gave the cf of the agf for the process and its conditional variance which we subsequently used to obtain their autocovariances. We also obtained the covariances between the squared errors and the conditional variance, and the covariances between the process and its conditional variance. Furthermore, we gave expressions for the MMSE predictors of future values of both the conditional variance and the squared conditional variance. These optimal predictors were subsequently used to obtain the conditional MSE associated with the optimal predictor of the future values of the conditional mean. Finally, we gave the bivariate ARMA representation of the process and its conditional variance. This representation can be used in conjunction with the methodology in Yamamoto (1981) to obtain expressions for the MMSE predictors and their variances in computationally convenient algorithmic forms.

Note that this study only examined the case where the roots of the autoregressive polynomials of the processes are distinct. Thus one potentially important issue not addressed in this paper relates to the effect of equal roots. The potential generalisations of the simple ARMA-GARCH in mean model are numerous. To state a few: (a) The ARMA-Asymmetric Power GARCH in mean model, (b) The ARMA-GARCH-M-X model, (c) The ARMA-Component GARCH in mean model, (d) The Multivariate GARCH in mean model[11] (MGARCH-M).

References

Akaike, H., 1974, Markovian representation of stochastic process and its application to the analysis of autoregressive moving average processes, Ann. Inst. Statist. Math, 26, 363-387.

Anderson, O. D., 1976, *Time series analysis and forecasting: the Box and Jenkins approach*, Butterworths.

Baillie, R. T., 1979, Asymptotic prediction mean squared error for vector autoregressive models, Biometrika, 66, 675-678.

Baillie, R. T., 1980, Predictions from ARMAX models, Journal of Econometrics 12, 365-374.

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

Bhansali, R. J., 1974, Asymptotic mean square error of predicting more than one-step ahead using the regression method, Applied Statistics, 23, 35-42.

Berra, A. K. and M. L. Higgins, 1993, ARCH models: properties, estimation, and testing, Journal of Economic Surveys, 7, 305-362.

Black, A. and P. Fraser, 1995, U.K. Stock Returns: Predictability and business conditions, Supplement, 85-102.

Bloomfield, R. T., 1972, On the error prediction of a time series, Biometrika, 59, 501-507.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroscedasticity, Journal of Econometrics 31, 307-327.

Bollerlsev, T., Chou R. Y. and K. F. Kroner, 1992, ARCH modelling in finance: a review of the theory and empirical evidence, Journal of Econometrics, 52, 5-59.

Bollerslev, T., Engle R. F. and D. B. Nelson, 1994, ARCH models, in: R. F. Engle and D. L. McFadden eds., *Handbook of Econometrics*, Volume IV, 295-3038.

Brenner, R. J., Harjes, R. H. and K. F. Kroner, 1996, Another look at models of the short-

[11] The MGARCH in mean model as the univariate one has been widely used in the finance literature. (see for example, Kroner and Lastrapes, 1993, Lee and Koray, 1994, and Grier and Perry, 1996).

term interest rate, Journal of Financial and Quantitative Analysis, 31, 85-107.

Brockwell, P. J. and R.A. Davis, 1987, *Time series: theory and methods*, Springer Verlag.

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.

Chan, K. C., Karolyi A. and R. M. Stulz, 1992, Global financial markets and the risk premium on U. S. equity, 32, 137-167.

Chou, R. Y., 1988, Volatility persistence and stock valuations: Some empirical evidence using GARCH, Journal of Applied Econometrics 3, 279-294.

Christodoulakis G., Hatgioannides J. and M. Karanasos, 1999, A dynamic model for the short-term interest rate with simultaneous feedback between the conditional mean and the conditional variance, preprint.

Day, T. E.and C. M. Lewis, 1992, Stock market volatility and the information content of stock index options, Journal of Econometrics 52, 267-287.

Davisson, I. D., 1965, The prediction error of a stationary gaussian time series of unknown covariance, I.E.E.E trans. Info. Theory, IT-11, 527-532.

Ding, Z. and C. W. J. Granger, 1996, Modeling 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, Along memory property of stock market returns and a new model, Journal of Empirical Finance, 1, 83-106.

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., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of U.K inflation, Econometrica 50, 987-1008.

Engle, R. F.and T.Bollerslev, 1986, Modeling the persistence of conditional variances, Econometric Reviews 5(1), 1-50.

Engle, R. F.and D. F. Kraft,1983, Multiperiod forecast error variances of inflation estimated from ARCH models, Applied Time Series Analysis of Economic Data, (Bureau of the Census, Washington, DC).

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.

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

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.

Gourieroux, C., 1997, *ARCH models and financial applications*, Springer.

Grier, K. P. and M. J. Perry, 1996, Inflation, inflation uncertainty, and relative price dispersion: Evidence from bivariate GARCH-M models, Journal of Monetary Economics, 38, 391-405.

Granger, C. W. J.and P. Newbold, 1986, *Forecasting Economic Time Series*, Academic Press.

Hall, S. G., 1991, An application of the stochastic GARCH-in-mean model to risk premia in the london metal exchange, Manchester School, Supplement, 57-71.

Hamilton, J. D., 1994, *Time series analysis*, Princeton.

Hamilton, J. D. and R. Susmel, 1994, Autoregressive conditional heteroscedasticity and changes in regime, Journal of Econometrics, 16, 121-130.

Hansson, B. and P. Hordahl, 1997, Changing risk premia: evidence from a small open economy, Scandinavian Journal of Economics,

He, C. and T. Terasvirta, 1997, Fourth moment structure of the GARCH(p,q) process, Working Paper, Stockholm School of Economics, No 168, To appear in Econometric Theory.

Hong, E. P., 1991, The autocorrelation structure for the GARCH-M process, Economic Letters 37, 129-132.

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.

Karanasos, M., 1998, A new method for obtaining the autocovariance of an ARMA model:

an exact-form solution, Econometric Theory 14, 622-640.

Karanasos, M., 1999a, The second moment and the autocovariance function of the squared errors of the GARCH model, Journal of Econometrics 90, 63-76.

Karanasos, M., 1999b, The covariance structure of mixed ARMA and VARMA models, Unpublished Mimeo, University of York.

Karanasos, M., 1999c, The covariance structure of the S-GARCH and M-GARCH models, Unpublishe Mimeo, University of York.

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.

Lamoureux, C. G. and W. D. Lastrapes, 1990, Forecasting stock return variance: Toward an understanding of stochastic implied volatilities, Unpublished manuscript, (Department of Economics, University of Georgia, Atlanta, GA).

Lee, T. H. and F. Koray, 1994, Uncertainty in sales and inventory behaviour in the U.S trade sectors, Canadian Journal of Economics, 1.

Lin, W. L., 1997, Impulse response function for conditional volatility in GARCH models, Journal of Business and Economic Statistics, 1, 15-25.

Ling, S., 1999, On the probabilistic properties of a double threshold ARMA conditional heteroscedastic model, To appear in Journal of Applied Probability.

Ling, S. and W .K. Li, 1997, On fractional integrated autoregressive moving- average time series models with conditionally heteroskedasticity, Journal of American Statistical Association, 92, 1184-1194.

Longstaff, F. A. and E. S. Schwartz, 1992, Interest rate volatility and the term structure: a two-factor general equilibrium model, Journal of Finance, 4, 1259-1282.

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.

Nerlove, M., Grether, D. M. and J. L. Carvalho, 1979, *Analysis of economic time series: a synthesis*, New York: Academic Press.

Pandit, S. M., 1973, *Data dependent systems: modelling analysis and optimal control via time series*, Ph. D. Thesis, University of Wiskonsin-Madison.

Pandit, S. M. and S. M. Wu, 1983, *Time series and system analysis with applications*, New York: Wiley.

Price, S., 1994, Aggregate uncertainty, forward looking behaviour and the demand for manufacturing labour in the UK.

Sargent, T. J., 1979, Macroeconomic Theory, (Academic Press, Inc).

Schmidt, P., 1974, The asymptotic distribution of forecasts in the dynamic simulation of an econometric model, Econometrica, 42, 303-309.

Schwert, G. W., 1989, Business cycles, financial crises and stock volatility, Carnegie-Rochester Conference Series on Public Policy 31, 83-125.

Shiryaev, A. N., 1999, Essentials of Stochastic Finance- Facts, Models, Theory. Advanced Series on Statistical Science and Applied Probability, vol 3, World Scientific, Singapore.

Song, O. H., 1996, The effects of exchange rate volatilities on Korean trade: multivariate GARCH-in-mean model based on partial conditional variances, Ph. D., University of Hawaii.

Yamamoto, T., 1976, Asymptotic mean square prediction error for an autoregressive model with estimated coefficients, Applied Statistics, 25, 123-127.

Yamamoto, T., 1978, prediction error of parametric time-series models with estimated coefficients, Unpublished manuscript.

Yamamoto, T., 1981, Predictions of multivariate autoregressive moving average models, Biometrica 68, 485-492.

Zinde-Walsh, V., 1988, Some exact formulae for autoregressive moving average processes, Econometric Theory, 4, 384-402.

Wei, W. S., 1989, *Time series analysis: univariate and multivariate methods*, Addison Wesley.

Appendix

A *Proofs of Proposition 2.1, Lemma 2.1*

Let y_t follow an ARMA(r,s) process

$$y_t = \phi + \sum_{j=1}^{r} \phi_i y_{t-i} - \sum_{j=0}^{s} \theta_j \epsilon_{t-j} \quad \text{(A. 1)}$$

We will first give the definite solution of the homogeneous deterministic component (y_t^d) of the ARMA(r,s) process (y_t) and we will subsequently use a technique provided in Sargent (1987), together with the definite solution (y_t^d), in order to derive the optimal predictor and the associated MSE of y_t.

The definite solution of the r order deterministic difference equation $\Phi(L)y_{t+i} = 0$ is

$$y_{t+i} = \sum_{n=0}^{r-1} x_{in} y_{t-n} \quad \text{(A. 2)}$$

We will prove the above by induction. If we assume that (A. 2) holds for a $(r-1)$ order difference equation then it will be sufficient to prove that it holds for an r order difference equation.

y_{t+r} can be expressed as an AR(1) process with an error term which follows a $(r-1)$ order difference equation

$$y_{t+r} = \lambda_1 y_{t+r-1} + x_{t+r}, \quad where \quad \prod_{i=2}^{r}(1-\lambda_i L)x_{t+r} = 0 \quad \text{(A. 3)}$$

Using backward substitution in the above equation, we get

$$y_{t+r-1} = \sum_{i=1}^{t} \lambda_1^{i-1} x_{t+r-i} + \lambda_1^t y_{r-1} \quad \text{(A. 4)}$$

Since x follows a $(r-1)$ order difference equation we have

$$x_{t+r-1-l} = \sum_{n=0}^{r-2}\sum_{j=2}^{r} x_{r-1-n}\zeta_{jt-l}^{r-1}\gamma_{jn}^{r-1}, \quad where \quad \zeta_{jt-l}^{r-1} = \frac{\lambda_j^{t-l+r-2}}{\prod_{k=2,k\neq j}^{r}(\lambda_j-\lambda_k)} \tag{A. 5}$$

$$\gamma_{jn}^{r-1} = (-1)^n \prod_{l=1}^{n}[\sum_{k_l=k_{l-1}+1,k_l\neq j}^{(r-1)-(n-l)}]\prod_{l=1}^{n}(\lambda_{k_l}), \quad where \quad k_0=1, \quad \gamma_{j0}^{r-1}=1 \tag{A. 5a}$$

Substituting (A. 5) into (A. 4) and after some algebra, we get

$$y_{t+r-1} \;=\; \sum_{i=0}^{r-2}\sum_{j=2}^{r} x_{r-1-i}(\zeta_{jt}^{r}\gamma_{ji}^{r-1} - \lambda_1^t\zeta_{j0}^{r}\gamma_{ji}^{r-1}) + \lambda_1^t y_{r-1}, \quad where \quad \zeta_{jt}^{r} = \frac{\lambda_j^{t+r-1}}{\prod_{k=1,k\neq j}^{r}(\lambda_j-\lambda_k)} \tag{A. 6}$$

Finally, substituting sequentially in the above equation

$$x_{r-k} \;=\; y_{r-k} - \lambda_1 y_{r-k-1}, \quad k=1,\cdots,r-1 \tag{A. 7}$$

and using

$$1-\sum_{j=2}^{r}\zeta_{j0}^{r} \;=\; \zeta_{10}^{r}, \quad \sum_{j=2}^{r}\zeta_{j0}^{r}\gamma_{jk-1}^{r} = \zeta_{10}^{r}\gamma_{1k-1}^{r}, \quad for \quad k\geq 2 \tag{A. 8}$$

where γ_{jn}^{r} is given by (A. 5a) with $k_0=0$, we get equation (A. 2). ■

Using Sargent (1987) technique and (A. 2) we express y_t as

$$y_t = \phi + \sum_{i=1}^{r}\phi_i y_{t-i} - \sum_{j=0}^{s}\theta_j\epsilon_{t-j} = \frac{\phi}{\prod_{i=1}^{r}(1-\lambda_i)} - \frac{\sum_{j=0}^{s}\theta_j\epsilon_{t-j}}{\prod_{i=1}^{r}(1-\lambda_i L)} \tag{A. 9}$$

$$= \phi\sum_{i=1}^{r}\bar{\zeta}_{i0}^{r} - \sum_{j=0}^{s}\sum_{i=1}^{r}\theta_j\epsilon_{t-j}\beta_i\zeta_{i0}^{r} = \phi\sum_{i=1}^{r}\zeta_{i0}^{r}a_{i,t-1} - \sum_{j=0}^{s}\sum_{i=1}^{r}\theta_j\epsilon_{t-j}\beta_{i,t-1}\zeta_{i0}^{r} + y_d^r$$

where

$$\beta_i = \frac{1}{1-\lambda_i L}, \quad \zeta_{i0}^{r} = \frac{\lambda_i^{r-1}}{\prod_{j=1,j\neq i}^{r}(\lambda_i-\lambda_j)}, \quad \bar{\zeta}_{i0}^{r} = \frac{\zeta_{i0}^{r}}{1-\lambda_i}, \quad and \tag{A. 9a}$$

$$\beta_{i,t-1} = \frac{1}{1-\lambda_i L_{t-1}} = \sum_{j=0}^{t-1}(\lambda_i L)^j, \quad a_{i,t-1} = \frac{1}{1-\lambda_{i,t-1}} = \sum_{j=0}^{t-1}\lambda_i^j \tag{A. 9b}$$

and y_d^r is given by (A. 2).

From (A. 9) after some algebra we get

$$y_t = \phi[\frac{1}{\Phi_r(1)} - \sum_{l=1}^{r} \bar{\zeta}_{lt}^r] - \sum_{l=1}^{r}\sum_{i=0}^{t-1}\sum_{j=0}^{min(i,s)} \zeta_{l0}^r \lambda_l^{i-j} \theta_j \epsilon_{t-i} - \sum_{l=1}^{r}\sum_{i=0}^{s-1}\sum_{j=i+1}^{min(t+i,s)} \zeta_{l0}^r \lambda_l^{t+i-j} \theta_j \epsilon_{-i} + y_d^r \tag{A. 10}$$

Taking the conditional expectation of (A. 10), as of time 0, we get the t-period optimal predictor of y. In addition, using (A. 10), we get the t period forecast error. ■

B Proof of Theorem 3.1

Let h_t follow a GARCH(p,q) process (for simplicity we will assume that $p > q$.)

$$B(L)h_t = \omega + A(L)\epsilon_t^2 \tag{B.1}$$

From the above equation we get

$$\omega_{t+p-i} = \hat{\omega}\phi_{t+p-i} + \sum_{j=1}^{q^\star} a_j \lambda_{j,t+p-i} + \sum_{j=1}^{p^\star} \beta_j c_{j,t+p-i}, \quad where \tag{B.2}$$

$$\hat{\omega} = [\omega + \sum_{j=0}^{i} \beta_{p-i+j} h_{t-j} + \sum_{j=0}^{i-(p-q)} a_{q+j}\epsilon_{t-j}^2] \tag{B.2a}$$

$$\lambda_{j,t+i} = \hat{\omega}\phi_{t+p-i-j} + (3a_j + \beta_j)\omega_{t+p-i-j}$$

$$+ \sum_{k=1}^{j-1}(a_k + \beta_k)\lambda_{j-k,t+p-i-k} + \sum_{k=1}^{p^\star - j}(a_{k+j}\lambda_{k,t+p-i-j} + \beta_{k+j}c_{k,t+p-i-j}) \tag{B.2b}$$

$$c_{j,t+i} = \hat{\omega}\phi_{t+p-i-j} + (a_j + \beta_j)\omega_{t+p-i-j}$$

$$+\sum_{k=1}^{j-1}(a_k+\beta_k)c_{j-k,t+p-i-k}+\sum_{k=1}^{p^\star-j}(a_{k+j}\lambda_{k,t+p-i-j}+\beta_{k+j}c_{k,t+p-i-j}),\quad and \tag{B.2c}$$

$$\omega_{t+p-i}=E_t(h^2_{t+p-i}),\quad \phi_{t+p-i}=E_t(h_{t+p-i}),\quad \lambda_{j,t+p-i}=E_t(\epsilon^2_{t+p-i}\epsilon^2_{t+p-i-j}), \tag{B.2d}$$

$$c_{j,t+p-i}=E_t(h_{t+p-i},h_{t+p-i-j}),\quad and\quad a_k=0,\ for\ k>q^\star,\quad \beta_k=0,\quad for\quad k>p^\star, \tag{B.2e}$$

where, $q^\star=q-max[0,i-(p-q)+1]$, $p^\star=p-max(0,i+1)$ and

where i can be any negative number and a positive number less than $p-2$.

The expressions of $\lambda_{j,t+p-i}$ and $c_{j,t+p-i}$ (eq B.2b and B.2c) can be written in a VAR $(p^\star+q^\star,p^\star-1)$ form

$$\lambda^\star_{t+p-i}=\sum_{j=1}^{p^\star-1}A_jL\lambda^\star_{t+p-i-j}+\omega^\star_{t+p-i}\Rightarrow\bar{A}(L)\lambda^\star_{t+p-i}=\omega^\star_{t+p-i},\quad where \tag{B.3}$$

$$\bar{A}(L)=(I-\sum_{j=1}^{p^\star-1}A_jL) \tag{B.3a}$$

$\lambda^\star_{t+p-i}$ is a $((p^\star+q^\star)\times 1)$ vector matrix. It's j1th element is $\lambda^\star_{j1}=\lambda_{j,t+p-i}$ for $j\le q^\star$ and $\lambda^\star_{j1}=c_{k,t+p-i}$ for $j\ge q^\star+k$, $(1\le k\le p^\star)$.

$\omega^\star_{t+p-i}$ is a $((p^\star+q^\star)\times 1)$ vector matrix. It's j1th element is $\omega^\star_{j1}=\omega\phi_{t+p-i-j}+(3a_j+\beta_j)\omega_{t+p-i-j}$ for $j\le q^\star$ and $\omega^\star_{j1}=\omega\phi_{t+p-i-k}+(a_k+\beta_k)\omega_{t+p-i-k}$ for $j\ge q^\star+k$, $(1\le k\le p^\star)$.

A_δ is a $((p^\star+q^\star)\times(p^\star+q^\star))$ matrix. It consists of four submatrices

$$A_\delta = \begin{bmatrix} A^\delta_{\lambda\lambda} & A^\delta_{\lambda c} \\ A^\delta_{c\lambda} & A^\delta_{cc} \end{bmatrix} \tag{B.4}$$

$A^\delta_{\lambda\lambda}$ is a $(q^\star\times q^\star)$ matrix. It's ijth element is given by $a_{ij}=0$, for $i<\delta$, $a_{\delta j}=a_{j+\delta}$, for $i=\delta$, $(a_{j+\delta}=0$ for $j+\delta>q^\star)$, $a_{ij}=a_\delta+\beta_\delta$, for $j>\delta$, $j=i-\delta$, and $a_{ij}=0$, for $i>\delta$, $j\ne i-\delta$.

$A^{\delta}_{\lambda c}$ is a $(q^\star \times p^\star)$ matrix. It's ijth element is given by $\beta_{ij} = 0$, for $i \gtrless \delta$, $\beta_{\delta j} = \beta_{j+\delta}$, for $i = \delta$, $(\beta_{j+\delta} = 0$, for $j + \delta > p^\star)$.

$A^{\delta}_{c\lambda}$ is a $(p^\star \times q^\star)$ matrix. It's ijth element is given by $a_{ij} = 0$, for $i \gtrless \delta$, $a_{\delta j} = a_{j+\delta}$, for $i = \delta$, $(a_{j+\delta} = 0$, for $j + \delta > q^\star)$.

A^{δ}_{cc} is a $(p^\star \times p^\star)$ matrix. It's ijth element is given by $\beta_{ij} = 0$, for $i < \delta$, $\beta_{\delta j} = \beta_{j+\delta}$, for $i = \delta$, and $(\beta_{j+\delta} = 0$, for $j + \delta > p^\star)$, $\beta_{ij} = \beta_\delta + a_\delta$, for $i > \delta$, and $j = i - \delta$, and $\beta_{ij} = 0$, for $i > \delta$, and $j \neq i - \delta$.

After solving the above VAR$(p^\star + q^\star, p^\star - 1)$ model and substituting the solution into (B.2) we get

$$\mu(L)\omega_{t+p-i} = \xi(L)\phi_{t+p-i}, \quad where \tag{B.5}$$

$$\mu(L) = \sum_{j=0}^{2p^\star-1} \mu_j L^j = \prod_{j=1}^{2p^\star-1} (1 - \mu^\star_j L) = \gamma(L) - \sum_{j=1}^{p^\star}\{\sum_{k=1}^{q^\star}[a_j\gamma_{jk}(L) + \beta_j\gamma_{q^\star+j,k}(L)]$$

$$(3a_k + \beta_k) + \sum_{k=1}^{p^\star}[a_j\gamma_{j,q^\star+k}(L) + \beta_j\gamma_{q^\star+j,q^\star+k}(L)](a_k + \beta_k)\}L^k \tag{B.5a}$$

$$\xi(L) = \sum_{j=0}^{2p^\star-1} \xi_j L^j = \hat{\omega}\{\gamma(L) + \sum_{j=1}^{p^\star}\{\sum_{k=1}^{q^\star}[a_j\gamma_{jk}(L) + \beta_j\gamma_{q^\star+j,k}(L)]$$

$$+ \sum_{k=1}^{p^\star}[a_j\gamma_{j,q^\star+k}(L) + \beta_j\gamma_{q^\star+j,q^\star+k}(L)]\}L^k\}, \quad and \tag{B.5b}$$

$\gamma_{ij}(L)$ is the ijth element of , $(L) = [\bar{A}(L)]^{-1}$ and $\gamma(L)$ is the determinant of $\bar{A}(L)$.

The solution of the above system of difference equations is given by (3. 9). ■