Sankhyā: The Indian Journal of Statistics
2000, Volume 62, Series B, Pt. 2, pp. 327–344

AN ARCH IN THE NONLINEAR MEAN (ARCH-NM) MODEL*

By SAMARJIT DAS**
AND
NITYANANDA SARKAR
Indian Statistical Institute, Calcutta

SUMMARY. This paper suggests a class of ARCH in the nonlinear mean (ARCH-NM) models. This class of models generalizes the usual ARCH-M model by considering the Box-Cox power transformation of the conditional variance for representing the risk premium. Thus, this generalization provides an approach by which a flexible specification of time-varying risk premium in the nonlinear form for ARCH is possible. Properties of this model are studied and the estimation procedure is described. The proposed model is then applied to the daily closing prices on the Bombay Stock Exchange Sensitive Index and its performance compared with the standard ARCH-M model using proper diagnostic checks.

1. Introduction

In a seminal paper in 1982, Engle introduced the autoregressive conditional heteroscedastic (ARCH) model. This model allows the conditional variance to change over time as a function of past errors keeping the unconditional variance constant. It has been observed that such models capture many temporal behaviours like thick tail distribution and volatility clustering of many economic and financial variables (see Bera and Higgins (1993), Bollerslev, Chou and Kroner (1992) and Bollerslev, Engle and Nelson (1994) for surveys on ARCH model and its generalizations).

The basic ARCH model has been generalized in different directions. One important generalization of ARCH model is what is known as ARCH in the mean (ARCH-M) model which was first introduced by Engle, Lilien and Robins (1987). The effects of risk and uncertainty on the returns on asset prices have increasingly attracted the economists and other researchers of capital markets and business finance in recent years. As the degree of uncertainty in asset return varies over

---

Paper received February 1998; revised December 1999.
AMS (1991) *subject classification.* 62J02, 62P20.
Key words ARCH, ARCH-M, Box-Cox power transformation, time-varying risk premium.

* An earlier version of the paper under a slightly different title 'A Nonlinear ARCH in the Mean (ARCH-M) Model' was presented at the 1996 India and South-East Asia Meeting of the Econometric Society held in New Delhi during December 28-30, 1996.

** Research of the first author was supported by the grant no. 9/93(37)95/EMR-I of C.S.I.R., India.

time, the compensation required by the risk averse economic agents for holding risky asset must also be varying. Existence of time varying risk premia i.e., increase in the expected rate of return due to an increase in the variance of the return, not only in asset pricing models but also in foreign exchange markets and term structure of interest rates, have been studied extensively (see, for instance, Amsler (1984), Domowitz and Hakkio (1985), Merton (1986) and Pesando (1983)). Considering a world of two assets - one is risky and the other riskless - Engle *et al.* (1987) showed that the mean and the variance of the return of the risky assets move in the same direction. They also showed that the conditional variance directly affects the expected return on a portfolio. In order to incorporate such aspects in the usual ARCH model framework, they suggested a new approach in which the ARCH model is extended in a direction so that it allows the conditional variance to influence the mean return. They also found that "variables which apparently were useful in forecasting excess return are correlated with the risk premia and lose their significance when a function of the conditional variance is included as a regressor" (p. 392). Engle *et al.* (1987) specified the class of ARCH-M models as follows :

$$y_t = x_t'\beta + \lambda\sqrt{h_t} + \epsilon_t \tag{1.1}$$

$$\epsilon_t \mid \Psi_{t-1} \sim N(0, h_t) \tag{1.2}$$

$$h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + ... + \alpha_p\epsilon_{t-p}^2 \tag{1.3}$$

where $\alpha_0 > 0$ and $\alpha_i \geq 0$ for all $i = 1, 2, \ldots, p$, y_t is the dependent variable, x_t is the $k \times 1$ vector of exogenous variables which may include lagged values of the dependent variable, $\Psi_{t-1} = \{y_{t-1}, x_{t-1}, \ldots\}$ is the information set at $t-1$, β is a $k \times 1$ vector of regression parameters and ϵ_t is the error term associated with the regression.

In financial literature $\lambda\sqrt{h_t}$ is known as risk premium. It may be noted that the ARCH-M specification requires the assumption of constant relative risk aversion utility function. However, this assumption is fairly strong and may not indeed be true in many practical situations. In most of the works on ARCH-M or close relatives of ARCH-M models the maintained hypothesis is that the risk premium can be expressed as an increasing function of the conditional variance of the asset return, say, $g(h_t)$. While in most applications $g(h_t) = \sqrt{h_t}$ has been used [see, for example, Bollerslev, Engle and Woolridge (1988), Domowitz and Hakkio (1985)], Engle *et al.* (1987) observed that $g(h_t) = \ln h_t$ worked better in estimating time varying risk premia in the term structure. In fact, they have discussed representation of risk premium as some function of the conditional variance. However, as pointed out by Pagan and Hong(1991), the use of $\ln h_t$ is somewhat restrictive in the sense that for $h_t < 1$, $\ln h_t$ will be negative, and for $h_t \longrightarrow 0$, the effect on y_t will be infinite. It may be pointed out that it is not enough that the risk premium is time varying. Since the expected rate of return will depend on the actual risk associated with decisions about y_t, it is imperative that proper functional forms of h_t are used to represent the risk premium. In fact, Backus and Gregory (1993), in a series of numerical examples, have shown that the relation between the risk premium and the conditional variance of the excess return can have virtually any shape - it can be increasing, decreasing, flat or even nonmonotonic depending on the parameters of the economy. Thus, although theory may lead to a monotonic relation between risk premium and conditional variance, it does not guarantee it. There are evidences also (e.g., Glosten, Jagannathan and Runkle (1989) and Harvey (1989, 1991)) that the monotonic relation between the risk premium and the conditional variance is not uniformly supported by the behaviour of actual prices. Given this somewhat unsatisfactory nature of parametric representation of risk premium which is basically an unobservable variable, some researchers like Pagan and Ullah (1988) and Pagan and Hong (1991) have suggested nonparametric methods. However, there are certain limitations in these methods, and hence as yet these cannot be recommended as standard tools of investigation where risk premium is involved (for details, see Pagan and Hong (1991)).

In the light of all these observations, it is clear that a more general and flexible specification of risk premium is called for. In this paper we propose a generalization of ARCH-M model in which $g(h_t)$ is assumed to have a general functional form as given by the Box-Cox (1964) family of power transformations. It is well known that this family of transformations encompasses all other standard functional forms as special cases. In this paper we study the proposed model in which h_t has the usual ARCH specification as given in (1.3). It may be noted that the usual ARCH-M specification is a special case of this generalization. Hence, for any given data the adequacy of the usual ARCH-M model against a class of generalized ARCH-M models as proposed by us [to be henceforth referred to as ARCH in the nonlinear mean (ARCH-NM) models] can now be studied.

The plan of this paper is as follows. In Section 2 we describe the proposed ARCH-NM model. The estimation of the model is described in Section 3. In Section 4 the advantage and appropriateness of this generalized approach in ARCH modelling is illustrated through an application of the proposed model to Bombay Stock Exchange Sensitive Index data. The paper concludes with some final comments in Section 5.

2. The Proposed Model

We propose a generalization of the usual ARCH-M specification by considering the risk premium in equation(1.1) as being $\lambda g(h_t)$ where $g(h_t)$ is defined as the Box-Cox (1964) transformation of h_t, i.e.,

$$\begin{aligned} g(h_t) &= \frac{h_t^{\xi} - 1}{\xi}, \text{ for } \xi \neq 0 \\ &= \ln h_t, \text{ for } \xi = 0. \end{aligned} \tag{2.1}$$

Thus, this generalization of the ARCH-M model allows for nonlinear representation in the mean of y_t through $g(h_t)$. The proposed ARCH-NM model is therefore represented as

$$y_t = x_t'\beta + \lambda g(h_t) + \epsilon_t, \ t = 1, 2, ..., T \tag{2.2}$$

where $g(h_t)$ is as defined in(2.1), $\epsilon_t|\Psi_{t-1} \sim N(0, h_t)$ and the conditional variance h_t has the usual ARCH specification as stated in (1.3), i.e.,

$$h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + ... + \alpha_p\epsilon_{t-p}^2$$

where $\alpha_0 > 0,\ \alpha_i \geq 0\ \forall i = 1, 2, ..., p,$

It may be noted at this stage that risk premium may also be negative, as discussed by Domowitz and Hakkio (1985) and Lintner (1965). However, unlike these models where this may be so due only to the negative sign of λ, in the proposed ARCH-NM model this may also due to $g(h_t)$ being negative. It is obvious that the model given by (2.1) and (2.2) reduces to the standard ARCH-M model when $\xi = 1$ or $\xi = 1/2$, the resulting constants in the transformation being adjusted through the appropriate regression parameters.

It is also evident from (2.2) that the $\{y_t\}$ series is autocorrelated. Hence, it implies that the proposed model crossbreeds, as in ARCH-M model, the random walk hypothesis which is assumed to hold for stock market returns. In fact, using a variance ratio test Lo and MacKinlay (1988) rejected the random walk hypothesis for weekly stock market returns. Obviously, the property of autocorrelatedness in y_t can be used to improve the accuracy of the forecasts. Empirical evidence towards this improvement in prediction in time varying risk premia has been provided by Shiller (1979), and Shiller, Campbell and Schoenholtz (1983) for term structure of interest rates, and by Domowitz and Hakkio (1985), Hodrick and Srivastava (1984) and Kendall (1989) for foreign exchange market. Unfortunately, the exact expressions for unconditional mean, variance and autocovariances are very difficult to obtain. However, some approximate expressions may be obtained for the ARCH-NM (1) model by considering Taylor series expansion of $g(h_t)$ upto squared term. For the sake of algebraic simplicity in deriving these expressions, the model for y_t in (2.2) is assumed to have no exogenous variable except the constant term. The approximate expressions for the unconditional mean, variance and autocovariances of y_t for such a model are stated in the following two results; the derivations are given in Appendix A.

RESULT 1. The approximate expressions for the unconditional mean and variance of y_t are given by

$$E(y_t) = \beta + \lambda\left[\frac{(\frac{\alpha_0}{1-\alpha_1})^\xi - 1}{\xi} + (\xi - 1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{\xi-2}\ \frac{(\alpha_0\alpha_1)^2}{(1-\alpha_1)^2(1-3\alpha_1^2)}\right] \quad (2.3)$$

and

$$\begin{aligned} V(y_t) \quad = \quad & \frac{\alpha_0}{1-\alpha_1} + \lambda^2\left[(2-\xi)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-1)}\right]\frac{2(\alpha_0\alpha_1)^2}{(1-\alpha_1)^2(1-3\alpha_1^2)} \\ & + \frac{\lambda^2}{4}(\xi-1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-1)} K_0 \\ & + \lambda^2\left[(\xi-1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-3} - 2\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-1}\right]\left\{\frac{6\alpha_0^3\alpha_1^2(1+2\alpha_1+2\alpha_1^2)}{(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)^2}\right\} \end{aligned} \quad (2.4)$$

where K_0 is appropriately defined, the expression being given in Appendix A.

RESULT 2. The first-order autocovariance of y_t can be approximated as

$$\begin{aligned} Cov\,(y_t, y_{t-1}) \quad = \quad & \lambda^2 \left[\left(\frac{\alpha_0}{1-\alpha_1} \right)^{2(\xi-1)} (4 - 4\xi + \xi^2) \right] K_1 \\ + \quad & \frac{\lambda^2}{2} \left[\left(\frac{\alpha_0}{1-\alpha_1} \right)^{2(\xi-1)} (\xi - 1) - \frac{(\xi-1)^2}{2} \left(\frac{\alpha_0}{1-\alpha_1} \right)^{2\xi-3} \right] K_2 \\ + \quad & \frac{\lambda^2}{2} \left[\left(\frac{\alpha_0}{1-\alpha_1} \right)^{2\xi-3} (\xi - 1) - \frac{(\xi-1)^2}{2} \left(\frac{\alpha_0}{1-\alpha_1} \right)^{2\xi-3} \right] K_3 \\ + \quad & \frac{\lambda^2}{4} \left[\left(\frac{\alpha_0}{1-\alpha_1} \right)^{2(\xi-2)} (\xi - 1)^2 \right] K_4, \end{aligned} \tag{2.5}$$

where the expressions of K_1, K_2, K_3 and K_4 are given in Appendix A.

The higher order autocorrelations of y_t are very cumbersome to evaluate when Taylor series expansion of $g(h_t)$ upto squared term is considered. However, if one considers only the first-order term of the Taylor series expansion of $g(h_t)$, then it is easy to find that $Corr\,(y_t, y_{t-k})(= \rho_k, \text{say}) = \sum_{i=1}^{p} \alpha_i \rho_{k-i}, k > p$ for an ARCH-NM (p) model. As for $0 < k \leq p$, no such recursive relation exists for $Corr\,(y_t, y_{t-k})$ and hence the expressions of these autocorrelations are to be separately obtained for each case.

It may be noted that the expressions of $E(y_t), V(y_t), Cov(y_t, y_{t-1})$ etc. for the standard ARCH and ARCH-M models may be obtained as special cases where $\lambda = 0$ and $\xi = 1$, respectively, in the expressions in (2.3), (2.4) and (2.5). Thus, by substituting $\xi = 1$ in (2.3) through (2.5), we find that

$$E(y_t) = \beta + \lambda \left(\frac{\alpha_0}{1-\alpha_1} - 1 \right) \tag{2.6}$$

$$V(y_t) = \frac{\alpha_0}{1-\alpha_1} + \frac{2(\lambda\alpha_0\alpha_1)^2}{(1-\alpha_1)^2(1-3\alpha_1^2)} \tag{2.7}$$

and

$$Cov(y_t, y_{t-1}) = \frac{2\lambda^2\alpha_0^2\alpha_1^3}{(1-\alpha_1)^2(1-3\alpha_1^2)}. \tag{2.8}$$

We can interpret $E(y_t)$ in the context of finance models as being the unconditional expected return for holding a risky asset. Since in the absence of a risk premium $V(y_t) = \alpha_0/(1-\alpha_1)$, the second component in (2.7) may be considered to be due to the presence of a risk premium which makes y_t more dispersed. Finally, as we know, ARCH-M effect makes y_t serially correlated and this serial correlation is given by (2.8). As for the other special case *viz.* ARCH model, which can be obtained by having $\lambda = 0$, it immediately follows from (2.3) through (2.5) that $E(y_t) = \beta, V(y_t) = \alpha_0/(1-\alpha_1)$ and $Cov(y_t, y_{t-1}) = 0$. Obviously, these correspond to the situation where risk premium is absent.

3. Estimation

The model is estimated by the method of maximum likelihood. The log-likelihood based on T observations and *conditional* on the initial values of all variables, is given by (omitting the constant)

$$L(\theta \mid \Psi_{t-1}) = \sum_{1}^{T} l_t(\theta \mid \Psi_{t-1}) \tag{3.1}$$

where

$$l_t(\theta \mid \Psi_{t-1}) = -(lnh_t)/2 - \epsilon_t^2/2h_t, \tag{3.2}$$

$\theta' = (\beta', \lambda, \alpha', \xi)$ is an $1 \times m(= k+p+3)$ vector of all parameters, and $\alpha' = (\alpha_0, \alpha_1, \ldots, \alpha_p)$ is the $1 \times (p+1)$ component vector of coefficients in the ARCH specification.

The first order condition of maximization of $l_t(\theta)$ (omitting henceforth from notation conditional on Ψ_{t-1}, for the sake of notational simplicity) yields

$$\frac{\partial l_t(\theta)}{\partial \theta} = \frac{1}{2h_t}\frac{\partial h_t}{\partial \theta}(\epsilon_t^2/h_t - 1) + \epsilon_t/h_t\frac{\partial \beta'}{\partial \theta}x_t + \lambda\frac{\partial g(h_t)}{\partial \theta} + g(h_t)\frac{\partial \lambda}{\partial \theta}. \tag{3.3}$$

It can be verified that $\frac{\partial \lambda}{\partial \theta}$ is an $m \times 1$ vector whose $(k+1)$th element is one and other elements are zeros; $\frac{\partial \beta'}{\partial \theta}$ is a $m \times k$ matrix of elements of zeros and ones with its first $k \times k$ submatrix being an identity matrix and the last $(m-k) \times k$ submatrix being a null matrix. All other derivatives *viz.*, $\frac{\partial l_t(\theta)}{\partial \theta}, \frac{\partial h_t}{\partial \theta}, \frac{\partial g(h_t)}{\partial \theta}$ are $m \times 1$ vector each. The evaluation of $\frac{\partial l_t(\theta)}{\partial \theta}$ requires the expressions of $\frac{\partial h_t}{\partial \theta}$ and $\frac{\partial g(h_t)}{\partial \theta}$ and the latter can again be expressed in terms of $\frac{\partial h_t}{\partial \theta}$. However, h_t as well as $g(h_t)$ are all functions of the previous innovations and all these derivatives will be of recursive structure. The required derivatives are thus obtained by the usual assumption that the initial values of $\frac{\partial h_t}{\partial \theta}$ and ϵ_t's do not depend upon the parameters. All the recursive relations are given in Appendix B.

Now let $S_{ti} = \frac{\partial l_t(\theta)}{\partial \theta_i}$, θ_i being the i-th element of the parameter vector θ. Then

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_t \frac{\partial l_t(\theta)}{\partial \theta} = \begin{pmatrix} S_{11} & S_{21} & \cdots & S_{T1} \\ \vdots & & & \\ S_{1m} & S_{2m} & \cdots & S_{Tm} \end{pmatrix} \times \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = S'e \text{ (say)},$$

where e is a $T \times 1$ vector of unity and S is a $T \times m$ matrix, the (t,i)th element of which is $\frac{\partial l_t(\theta)}{\partial \theta_i}$. The first order condition of maximization of the likelihood ensures that $\frac{\partial L(\theta)}{\partial \theta} = S'e = 0$. The information matrix corresponding to the $t-th$ observation is $I_t = E(\frac{\partial l_t(\theta)}{\partial \theta}\frac{\partial l_t(\theta)}{\partial \theta'}) = -E(\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'})$, and the same for the sample of T observations is given by $I = E(S'S/T)$. Furthermore, $S'S/T$ is also a consistent estimator of I under certain conditions. It may be noted that in this model conditional mean involves the parameters of the conditional variance, and hence the information matrix is not block diagonal. In order to solve the nonlinear equations for obtaining the ML estimate , we use the well known algorithm suggested by Berndt, Hall , Hall and Hausman (1974). This BHHH algorithm can be written as :

$$\theta^{(i+1)} = \theta^{(i)} + \eta(S^{(i)\prime}S^{(i)})^{-1}S^{(i)\prime}e$$

where $\theta^{(i)}$ is the estimate of θ at the $i-th$ step of iteration,$S^{(i)}$ the matrix of first order derivatives evaluated at $\theta^{(i)}$, and η the step length parameter. This algorithm ensures the existence of a consistent estimate of θ. If $\hat{\theta}_n$ be the maximum likelihood estimate of θ thus obtained, then applying Crowder's theorem (1976) we may easily conclude that $(S'S)^{-1/2}\ (\hat{\theta}_n - \theta_0) \overset{A}{\sim} N(0, I_m)$, where θ_0 is the true value of the parameter vector θ, and I_m, the identity matrix of order m.

As stated in the preceding section, the adequacy of the usual ARCH-M model for any given data can now be studied by considering the proposed ARCH-NM model as the alternative. In other words, null hypotheses like $\xi = 1$ and $\xi = 1/2$ may be tested against appropriate alternatives given by $\xi \neq 1$ and $\xi \neq 1/2$ respectively, in the ARCH-NM framework. Standard asymptotic tests may be used to carry out these hypotheses testing.

4. **An Illustration**

In this section we report the results of an application of the proposed model to daily closing prices on the Bombay Stock Exchange (BSE) as measured by the BSE Sensitive Index (SENSEX). The data cover the period October (4th week), 1989 to April (2nd week), 1996. The analysed series is the first differences of the logarithms of SENSEX. Hence, the data represent the continuously compounded rate of return for holding the (aggregate) securities for one day. It is evident from the plot of this return series which is given below, that the data exhibit episodes of both low and high volatility. Since dependence in "squared" data signifies nonlinearity and presence of conditional heteroscedasticity, we computed the Ljung-Box test statistic for the "squared" data, denoted as $Q^2(p)$, for lags upto 24. All these values were found to be highly significant (cf. Table I), suggesting thereby the presence of nonlinear dependence in the return data. We also obtained the skewness and kurtosis coefficients of y_t which are presented in Table I. The skewness coefficient conveys some evidence of asymmetry in the unconditional distribution. The kurtosis coefficient is significantly greater than 3, which indicates that the unconditional distribution of the data has heavier tail than a normal distribution.

Further, we carried out an ARCH test which yielded the value of the test statistic as 119.121. Obviously, this is highly significant both at 1 per cent and 5 per cent levels of significance. Since nonlinear dependence and a heavy-tailed unconditional distribution are typical characteristics of conditionally heteroscedastic data, we may thus conclude that the return data may be appropriately analysed by an ARCH model.

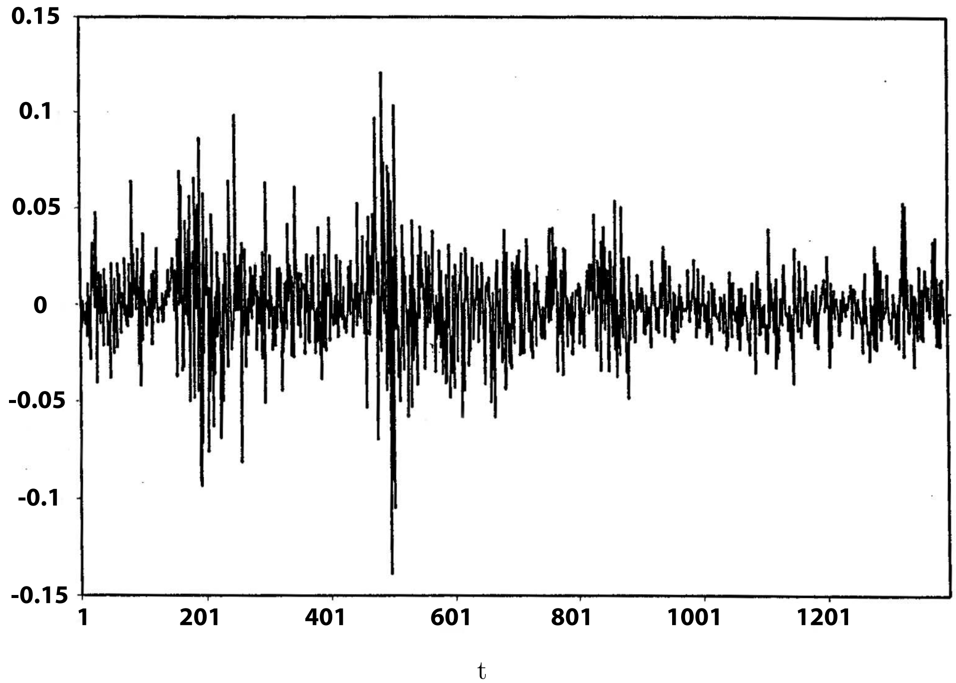

Figure 1. Time plot of the logarithmic growth rate of daily BSE SENSEX

Having recognized that ARCH effect is very strong in the data, we now discuss about the suitability of ARCH-M model or its generalization in the form of the proposed ARCH-NM model, for analysing the given return data. To this end it is quite clear from the plot that the changes in variance are similar to those hypothesized by ARCH-M/ARCH-NM model. Since both these models yield serial correlation in y_t, we computed Ljung-Box $Q(p)$ test statistic for all lags from 1 to 24 to find if the data support this property. These $Q(p)$ values are summarized in Table I. It is evident from the values of $Q(p)$ test statistic that the (linear) correlations are highly significant at 1 per cent level, indicating strongly that ARCH-M or its generalization in ARCH-NM would fit the data well. In what follows we report the empirical findings which, in fact, lend strong support to our approach in which a wider class of models given by ARCH-NM is suggested for proper representation of risk premium. Since it is well-known that GARCH is a better and generalized representation of the conditional variance than ARCH, the empirical exercise was carried out with GARCH representation for h_t. In other words, h_t was assumed to be given by

$$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \cdots + \alpha_p \epsilon_{t-p}^2 + \eta_1 h_{t-1} + \cdots + \eta_q h_{t-q} \tag{4.1}$$

where $\alpha_0 > 0, \alpha_i \geq 0$ for $i = 1, 2, \ldots p$ and $\eta_j \geq 0$ for $j = 1, 2, \ldots, q$.

Table 1. Diagnostic checks of models for BSE SENSEX data

Diagnostics	Observed Series	GARCH-M Standardized Residuals	GARCH-NM Standardized Residuals
Skewness Coefficient	0.0209	0.4726	0.3624
Kurtosis Coefficient	4.9399	4.1904	1.6875
ARCH	119.1211*	8.6736*	0.0934
Q (4)	16.7*	25.0*	40.9*
Q (8)	22.9*	26.0*	42.7*
Q (12)	32.5*	33.1*	45.6*
Q (16)	43.3*	47.3*	57.0*
Q (20)	56.6*	54.7*	62.4*
Q (24)	59.4*	55.9*	64.6*
Q^2 (4)	359.0*	34.9*	3.25
Q^2 (8)	582.0*	66.3*	4.90
Q^2 (12)	786.0*	134.0*	13.5
Q^2 (16)	1000.0*	189.0*	17.1
Q^2 (20)	1120.0*	225.0*	18.8
Q^2 (24)	1210.0*	252.0*	20.5

*indicates significance at 1 per cent level.

As reported below, GARCH (1,1) turned out to be the best model from consideration of maximization of the log-likelihood function. It may also be stated here that for the purpose of this application, it was assumed that excepting for an intercept term there is no other regressor in the model. By following the method of estimation outlined in the preceding section for the proposed ARCH-NM model, we found that the log-likelihood function given in (3.1) was maximized at $\xi = 0.05$. However, the value of the log-likelihood function corresponding to $\xi = 0$ being almost the same as the one at $\xi = 0.05$, (differing in the second decimal place only), we report the best fitted model as (with t-ratios given in parentheses)

$$\begin{array}{lllllll} y_t & = & 0.0152 & + & 0.0018\, ln h_t & + & \epsilon_t, \qquad\qquad L(\hat{\theta}) = 3628.77 \\ & & (2.5347) & & (2.4699) & & \\ h_t & = & 0.00001 & + & 0.12859\, \epsilon^2_{t-1} & + & 0.85151\, h_{t-1}. \\ & & (4.96431) & & (7.78673) & & (55.80221) \end{array} \tag{4.2}$$

Thus, the chosen model is GARCH-NM of order (1,1) with $\xi = 0$. It is evident that for this model all the parameters in h_t are significant at 1 per cent level of significance, and those in y_t are clearly significant at 5 per cent level and "almost" significant at 1 per cent level. It may be mentioned in this context that the significant positive value of λ is in conformity with the basic understanding in the risk premium literature that positive value of the risk aversion parameter is quite desirable. In order to compare this model with the standard GARCH-M model, we computed the maximum log-likelihood value corresponding $\xi = 1/2$ and found it to be only 3501.21, which is much smaller than the log-likelihood value of the chosen GARCH-NM model. As regards the significance of the estimated GARCH-M model, we find from (4.3) below that while all parameters in h_t are significant, the risk aversion parameter λ in y_t is highly insignificant, the t-statistic value being only

0.3031. Thus, not only that the standard GARCH-M model with $\xi = 1/2$ produces a much smaller log-likelihood value as compared to the chosen GARCH-NM model, but also that a straightforward GARCH-M fitting would choose an inappropriate model for the data. In order to formally test for the adequacy or otherwise of the standard GARCH-M model given the framework of a more generalized model as proposed in this paper, we carried out the likelihood ratio test with the null hypothesis specifying the standard GARCH-M model. The test obviously soundly rejected the null hypothesis in favour of the proposed GARCH-NM model.

$$\begin{array}{rclcccccl} y_t & = & \underset{(-0.1272)}{-0.0005} & + & \underset{(0.3031)}{0.0613\sqrt{h_t}} & + & \epsilon_t, & & L(\hat{\theta}) = 3501.21 \\ h_t & = & \underset{(8.3136)}{0.00002} & + & \underset{(8.1062)}{0.0991\,\epsilon_{t-1}^2} & + & \underset{(6.3815)}{0.3941\,h_{t-1}}. & & \end{array} \tag{4.3}$$

In order to judge the goodness of the fitted GARCH-NM model and compare it with the standard GARCH-M model, we now report in Table I the results of diagnostic checks for the standardized residuals given by $\widetilde{\epsilon}_t = \hat{\epsilon}_t/\sqrt{\hat{h}_t}$, where $\hat{\epsilon}_t$ and $\hat{h}_t$ are the ML residual and estimated conditional variance at t, respectively. We observe from this table that the autocorrelations of $\hat{\epsilon}_t$'s are significant; this suggests that there exist some linear correlations in the GARCH-NM standardized residuals. In other words, the autocorrelations of y_t have not been fully incorporated through the GARCH-NM framework. It may, however, be noted from a glance at Table I that the inference with regard to performance by $Q(p)$ test is the same for the usual GARCH-M (with $\xi = 1/2$) standardized residuals as well. Thus, irrespective of the framework being GARCH-M or GARCH-NM, we observe that the residuals exhibit serial correlation. In order to find if this correlation could be rectified by including an ARMA type component in the conditional mean part of the model, we considered an extension of the model in (2.2) by including y_{t-1} as an explanatory variable. While this led to some improvement in the serial correlation of these (standardized) residuals, the maximum value of the log-likelihood function reduced quite significantly to 3585.97, as compared to 3628.77 for the original GARCH-NM model of (2.2). Thus, we find, that, by ML criterion, the original GARCH-NM model turns out to be a more appropriate model for the given data set.

As far as the autocorrelation structure of $\widetilde{\epsilon}^2_t$ is concerned, $Q^2(p)$ based on standardized GARCH-NM residuals show that none of the 24 test statistic values corresponding to 24 lags is significant even at 5 per cent level. We may, therefore, conclude that the residuals contain no more nonlinearity. It is also evident from Table I that $Q^2(p)$ values are highly significant with standardized GARCH-M residuals. Thus, we observe that in terms of diagnostic checking with standardized residuals, the chosen GARCH-NM is a better model than GARCH-M.

Finally, we investigate the normality of $\widetilde{\epsilon}_t$. It is well-known that the family of ARCH models including its various generalizations are inherently non-normal, and hence it is expected that when appropriately estimated, the estimated models would account for part of the non-normality in the data. We find from Table I that the unconditional distribution of the observed series is skewed slightly to the right having skewness coefficient as 0.0209 and fat tailed with kurtosis coefficient being 4.9399. A glance at Table I suggests that for both GARCH-M and GARCH-NM standarized residuals, there are departures from normality as evidenced from the skewness and kurtosis coefficients computed from these residuals. Obviously, these findings were not entirely expected, at least for the GARCH-NM standardized residuals which otherwise performed better than GARCH-M model. In order to explain if these findings were due to any further conditional heteroscedasticity remaining in the residuals, ARCH test was carried out. The test statistic values for GARCH-NM and GARCH-M standardized residuals were found to be 0.0934 and 8.6736, respectively. The conclusions, therefore, are that while no ARCH effects were left with GARCH-NM residuals, significant ARCH effects still remained in the standard GARCH-M residuals.

Thus, we have established through this illustration that the suggested extension of the standard (G) ARCH-M model (in which the risk premium component is assumed to be a flexible function in the sense of Box-Cox transformation) can provide improvement over the standard (G) ARCH-M model. This notwithstanding, it is evident from this example that the fit is not entirely satisfactory from the point of view of some criteria of model selection. The explanations may lie in the fact that the proposed model is not nonlinear enough to model this data quite satisfactorily. It is now recognised that correct specification of the conditional variance function is important in several respects (see, for example, Granger and Newbold (1976) and Hopwood, Mckeown and Newbold (1984) for relevant details in ARCH literature). It is, in fact, clear from our discussions in Section 1 that the performance of ARCH-M and ARCH-NM models would be affected by the specification of the functional form of h_t. In the proposed ARCH-NM model a flexible functional form is all the more important since $g(h_t)$ appears in the conditional mean of y_t. Several alternative functional forms for h_t have been suggested in ARCH framework by Engle and Bollerslev (1986), Gweke (1986), Hentschel (1995), Higgins and Bera (1992), Pantula (1986) and several others. We are currently exploring possible extensions of our work along these directions.

5. **Conclusions**

In the literature on time-varying risk premium researchers have used conditional variance h_t or $\sqrt{h_t}$ or $\ln h_t$ to represent the risk premium in the model. It is, therefore, natural to argue that a flexible functional form of h_t for representing risk premium should be more useful and appropriate. Some of the existing empirical evidences lend support towards this direction. Keeping this in mind, we have proposed in this paper a generalization of the ARCH-M model by allowing for nonlinear representation in the mean of the dependent variable. This is done by considering the Box-Cox power transformation of the conditional variance representing the risk premium in the model. Obviously, this generalized model encompasses various other functional forms used in the risk premium literature as special cases. The estimation of this model by the method of maximum likelihood has been discussed, and an illustrative example in support of this generalized approach for representation of risk premium has also been given in this paper. Our illustration of BSE SENSEX data demonstorates that the suggested generalization provides improvement over the standard ARCH-M model. Before we conclude, we may state that keeping in mind the importance of functional form of h_t in the performance of ARCH-NM model, we are currently working on extending our work by considering alternative functional forms of h_t.

Appendix A

To prove Results 1 and 2, we need the unconditional expectation of $\omega_t' = (\epsilon_t^8, \epsilon_t^4, \epsilon_t^2)$. Engle (1982) proved that $E(\omega_t) = (I-A)^{-1}b$, where

$$A = \begin{bmatrix} 105\alpha_1^4 & 420\alpha_0\alpha_1^3 & 630\alpha_0^2\alpha_1^2 & 420\alpha_0^3\alpha_1 \\ 0 & 15\alpha_1^3 & 45\alpha_1^2 & 43\alpha_1^2\alpha_1 \\ 0 & 0 & 3\alpha_1^2 & 6\alpha_0\alpha_1 \\ 0 & 0 & 0 & \alpha_1 \end{bmatrix}$$

and $b' = (105\alpha_0^4 \quad 15\alpha_0^3 \quad 3\alpha_0^2 \quad \alpha_0)$.
It is quite clear that to obtain $E(\epsilon_t^8)$ we need only the first row of $(I-A)^{-1}$, i.e.,

$$\begin{bmatrix} \frac{1}{1-105\alpha^4} \\ -\frac{420\alpha_0\alpha_1^3}{(1-105\alpha_1^4)} \\ \frac{630\alpha_0^2\alpha_1^2(1+30\alpha_1^3)}{(1-105\alpha_1^4)`(1-15\alpha_1^3)(1-3\alpha_1^2)} \\ -\frac{420\alpha_0^3\alpha_1(45\alpha_1^5+30\alpha_1^3+6\alpha_1^2+1)}{(1-105\alpha_1^4)(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)} \end{bmatrix}'$$

and therefore

$$E(\epsilon_t^8) = \frac{105\alpha_0^4(1-5\alpha_1+15\alpha_1^2-114\alpha_1^3+495\alpha_1^4-315\alpha_1^5-405\alpha_1^6)}{(1-105\alpha_1^4)(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)}.$$

Similarly it may easily be checked that

$$\begin{aligned} E(\epsilon_t^6) &= \frac{105\alpha_0^3(1+2\alpha_1+6\alpha_1^2+3\alpha_1^3)}{(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)}, \\ E(\epsilon_t^4) &= \frac{3\alpha_o^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}, \end{aligned}$$

and

$$E(\epsilon_t^2) \quad = \quad \frac{\alpha_0}{1-\alpha_1} \ .$$

PROOF OF RESULT 1. As stated in the text, we take the model as

$$E(y_t) = \beta + \lambda E g(h_t).$$

Now, the Taylor series expansion of $g(h_t)$, around $E(h_t) = E(\epsilon_t^2)(= \bar{h}_t$, say) yields

$$g(h_t) = g(\bar{h}_t) + (h_t - \bar{h}_t) g'(h_t) \mid_{h_t=\bar{h}_t} + \frac{(h_t - \bar{h}t)^2}{2!} g''(h_t) \mid_{h_t=\bar{h}_t},$$

taking upto squared term only.

Therefore,

$$E(y_t) = g(\bar{h}_t) + \beta + \lambda \left\{ f(\bar{h}_t) + g''(\bar{h}_t) \frac{V(h_t)}{2} \right\}$$

Now, $V(h_t) = V(\alpha_0 + \alpha_1 \epsilon_{t-1}^2) = \frac{2(\alpha_0\alpha_1)^2}{(1-\alpha_1)^2(1-3\alpha_1^2)}$, and hence, we have

$$E(y_t) = \beta + \lambda \left[\frac{(\frac{\alpha_0}{1-\alpha_1})^\xi - 1}{\xi} + (\xi - 1) \left(\frac{\alpha_0}{1-\alpha_1} \right)^{\xi-2} \frac{(\alpha_0\alpha_1)^2}{(1-\alpha_1)^2(1-3\alpha_1^2)} \right].$$

As for the variance of y_t, it can easily be seen to be

$$\begin{aligned} V(y_t) \quad &= \quad \frac{\alpha_0}{1-\alpha_1} + \lambda^2 [g'(\bar{h}_t)^2 + \bar{h}_t^2 g''(\bar{h}_t)^2 - 2g'(\bar{h}_t) g''(\bar{h}_t) \bar{h}_t] V(h_t) + \\ & \frac{\lambda^2}{4} g''(\bar{h}_t)^2 V(h_t^2) + \lambda^2 [g'(\bar{h}_t) g''(\bar{h}_t) - 2\bar{h}_t g''(\bar{h}_t)^2] Cov(h_t, h_t^2). \end{aligned}$$

Since $V(h_t^2) = E(h_t^4) - \{E(h_t^2)\}^2$, we need the expressions for $E(h_t^2)$ and $E(h_t^4)$ for simplifying $V(y_t)$. These latter expressions are as follows.

$$\begin{aligned} E(h_t^2) \quad &= \quad \alpha_0^2(1+\alpha_1)/(1-\alpha_1)(1-3\alpha_1^2) \\ E(h_t^4) \quad &= \quad \frac{\alpha_0^4(1 + 3\alpha_1 + 15\alpha_1^2 - 6\alpha_1^3 - 50\alpha_1^3 - 1065\alpha_1^5 - 405\alpha_1^6 - 11335\alpha_1^7)}{(1-105\alpha_1^4)(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)} \\ & + \frac{\alpha_0^4(56700\alpha_4^8 - 9450\alpha_1^9 - 28350\alpha_1^{10})}{(1-105\alpha_1^4)(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)} \end{aligned}$$

Therefore,

$$\begin{aligned} V(h_t) \quad &= \quad \frac{\alpha_0^4(8\alpha_1^2 - 18\alpha_1^3 + 37\alpha_1^4 - 817\alpha_1^5 + 933\alpha_1^6 - 9160\alpha_1^7 + 6929\alpha_1^8)}{(1-105\alpha_1^4)(1-15\alpha_1^3)(1-3\alpha_1^2)^2(1-\alpha_1)^2} \\ & - \frac{\alpha_0^4(117555\alpha_1^9 + 154995\alpha_1^{10} + 113400\alpha_1^{11} - 113400\alpha_1^{12} - 85050\alpha_1^{13})}{(1-105\alpha_1^4)(1-15\alpha_1^3)(1-3\alpha_1^2)^2(1-\alpha_1)^2} \\ &= \quad K_0 \ (say). \end{aligned}$$

We also note that $Cov(h_t, h_t^2)$ involves $E(h_t^3)$ which is given by

$$\begin{aligned} E(h_t^3) &= E(\alpha_0^3 + 3\alpha_0^2\alpha_1\epsilon_{t-1}^2 + 3\alpha_0\alpha_1^2\epsilon_{t-1}^4 + \alpha_1^3\epsilon_{t-1}^6) \\ &= \frac{\alpha_0^3(1+2\alpha_1+6\alpha_1^2+3\alpha_1^3)}{(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)}. \end{aligned}$$

Hence,

$$Cov(h_t, h_t^2) = \frac{6\alpha_0^3\alpha_1^2(1+2\alpha_1+2\alpha_1^2)}{(1-15\alpha_1^3)(1-3\alpha_1^2)(1-\alpha_1)^2}.$$

Thus we obtain the expression for $V(y_t)$ given in (2.4).

Proof of Result 2. The first order autocovariance of y_t, denoted by ν_1, is

$$\begin{aligned} \nu_1 &= Cov[\beta + \lambda g(h_t), \quad \beta + \lambda g(h_{t-1})] \\ &= \lambda^2 Cov[g(h_t), g(h_{t-1})]. \end{aligned}$$

Again by using Taylor series expansion of $g(h_t)$, with respect to $\bar{h}_t$ and then taking upto the second term only, we have

$$\begin{aligned} \nu_1 = \; & \lambda^2 \left[(4-4\xi+\xi^2)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-1)}\right] Cov(h_t, h_{t-1}) \\ & + \frac{\lambda^2}{2}\left[(\xi-1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-3)} - \frac{(\xi-1)^2}{2}\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-3}\right] Cov(h_t, h_{t-1}^2) \\ & + \frac{\lambda^2}{2}\left[(\xi-1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-3} - \frac{(\xi-1)^2}{2}\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-3}\right] Cov(h_t^2, h_{t-1}) \\ & + \frac{\lambda^2}{4}\left[(\xi-1)^2\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-2)}\right] Cov(h_t^2, h_{t-1}^2). \end{aligned}$$

To evaluate ν_1, we note that it involves several covariances which in turn involve the expressions for $E(h_t\, h_{t-1}), E(h_t\, h_{t-1}^2), E(h_t^2\, h_{t-1})$ and $E(h_t^2\, h_{t-1}^2)$.

After algebraic simplifications, these expectation terms may be found to be as follows.

$$\begin{aligned} E(h_t h_{t-1}) &= \frac{\alpha_0^2(1+\alpha_1-\alpha_1^2)}{(1-\alpha_1)(1-3\alpha_1^2)} \\ E(h_t h_{t-1}^2) &= \frac{\alpha_1^3(1+2\alpha_1+2\alpha_1^2-9\alpha_1^3-12\alpha_1^4)}{(1-\alpha_1)(1-3\alpha_1^2)(1-15\alpha_1^3)} \\ E(h_t^2 h_{t-1}) &= \frac{\alpha_0^3(1+4\alpha_1+6\alpha_1^2-15\alpha_1^3-42\alpha_1^4-36\alpha_1^5+90\alpha_1^6)}{(1-\alpha_1)(1-3\alpha_1^2)(1-15\alpha_1^3)} \end{aligned}$$

and

$$\begin{aligned} E(h_t^2, h_{t-1}^2) = \; & \frac{\alpha_0^4(1+3\alpha_1+7\alpha_1^2+6\alpha_1^3-38\alpha_1^4-51\alpha_1^5+511\alpha_1^6+315\alpha_1^7)}{(1-\alpha_1)(1-3\alpha_1^2)(1-15\alpha_1^3)(1-105\alpha_1^4)} \\ & - \frac{\alpha_0^4(7050\alpha_1^8+27209\alpha_1^9+60930\alpha_1^{10}+99540\alpha_1^{11}+52380\alpha_1^{12})}{(1-\alpha_1)(1-3\alpha_1^2)(1-15\alpha_1^3)(1-105\alpha_1^4)}. \end{aligned}$$

Using the expressions for these expectation terms, the covariances in the expression for ν_1 can be reduced to

$$\begin{aligned} Cov(h_t, h_{t-1}) &= \frac{2\alpha_0^2\alpha_1^3}{(1-\alpha_1)^2(1-3\alpha_1^2)} = K_1(say), \\ Cov(h_t, h_{t-1}^2) &= \frac{3\alpha_0^3\alpha_1^2(1+4\alpha_1+4\alpha_1^2)}{(1-\alpha_1)^2(1-3\alpha_1^2)(1-15\alpha_1^3)} = K_2(say), \\ Cov(h_t^2, h_{t-1}) &= \frac{\alpha^3(2\alpha_1+2\alpha_1^2-6\alpha_1^3-12\alpha_1^4+6\alpha_1^5+12\alpha_1^6-90\alpha_1^7)}{(1-\alpha_1)^2(1-3\alpha_1^2)(1-15\alpha_1^3)} = K_3\ (say), \end{aligned}$$

and

$$\begin{aligned} Cov(h_t^2, h_{t-1}^2) &= \frac{\alpha_0^4(8\alpha_1^3+79\alpha_1^4+215\alpha_1^5+799\alpha_1^6-1504\alpha_1^7-12201\alpha_1^8)}{(1-\alpha_1)^2(1-3\alpha_1^2)(1-15\alpha_1^3)(1-105\alpha_1^4)} \\ &\quad -\frac{\alpha_0^4(21146\alpha_1^9+11626\alpha_1^{10}-21867\alpha_1^{11}-148323\alpha_1^{12})}{(1-\alpha_1)^2(1-3\alpha_1^2)^2(1-15\alpha_1^3)(1-105\alpha_1^4)} \\ &\quad +\frac{\alpha_0^4(168210\alpha_1^{13}-141480\alpha_1^{14}-157140\alpha_1^{15})}{(1-\alpha_1)^2(1-3\alpha_1^2)(1-15\alpha_1^3)(1-105\alpha_1^4)} \\ &= K_4(say). \end{aligned}$$

Thus, finally the expression for $Cov(y_t, y_{t-1})$ is given by

$$\begin{aligned} \nu_1 &= K_1\lambda^2\left[(4-4\xi+\xi^2)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-1)}\right] \\ &\quad + K_2\frac{\lambda^2}{2}\left[(\xi-1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-1)} - \frac{(\xi-1)^2}{2}\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-3}\right] \\ &\quad + K_3\frac{\lambda^2}{2}\left[(\xi-1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-3} - \frac{(\xi-1)^2}{2}\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2\xi-3}\right] \\ &\quad + K_4\frac{\lambda^2}{4}\left[(\xi-1)^2\left(\frac{\alpha_0}{1-\alpha_1}\right)^{2(\xi-2)}\right]. \end{aligned}$$

Appendix B

First order derivatives of $L(\theta \mid \Psi_{t-1})$ in (3.1) with respect to the parameter vector $\theta' = (\beta', \lambda, \alpha', \xi)$.

The argument of $l_t(\theta \mid \Psi_{t-1})$ as also the conditional sign are being dropped for convenience.

Case I : $\xi \neq 0$

$$\frac{\partial L}{\partial \theta} = \sum_{t=1}^{T}\frac{\partial l_t}{\partial \theta} = \sum_{t=1}^{T}\left(\frac{\partial l_t}{\partial \beta'}\ \frac{\partial l_t}{\partial \lambda}\ \frac{\partial l_t}{\partial \alpha'}\ \frac{\partial l_t}{\partial \xi}\right)'$$

Now, $\frac{\partial l_t}{\partial \theta} = \frac{1}{2h_t}\frac{\partial h_t}{\partial \theta}(\frac{\epsilon_t^2}{h_t} - 1) + \frac{\epsilon_t}{h_t}\frac{\partial \beta'}{\partial \theta}x_t + \lambda\frac{\epsilon_t}{h_t}\frac{\partial g(h_t)}{\partial \theta} + g(h_t)\frac{\epsilon_t}{h_t}\frac{\partial \lambda}{\partial \theta}$ and hence

$$\left.\begin{array}{rcl} \frac{\partial l_t}{\partial \beta} &=& \frac{1}{2h_t}\frac{\partial h_t}{\partial \beta}(\frac{\epsilon_t^2}{h_t} - 1) + \frac{\partial \beta'}{\partial \theta}x_t\frac{\epsilon_t}{h_t} + \lambda\frac{\partial g(h_t)}{\partial \beta}\frac{\epsilon_t}{h_t} \\ \frac{\partial l_t}{\partial \lambda} &=& \frac{1}{2h_t}\frac{\partial h_t}{\partial \lambda}(\frac{\epsilon_t^2}{h_t} - 1) + \lambda\frac{\partial g(h_t)}{\partial \lambda}\frac{\epsilon_t}{h_t} + g(h_t)\frac{\epsilon_t}{h_t} \\ \frac{\partial l_t}{\partial \alpha} &=& \frac{1}{2h_t}\frac{\partial h_t}{\partial \alpha}(\frac{\epsilon_t^2}{h_t} - 1) + \lambda\frac{\partial g(h_t)}{\partial \alpha}\frac{\epsilon_t}{h_t} \\ \frac{\partial l_t}{\partial \xi} &=& \frac{1}{2h_t}\frac{\partial h_t}{\partial \xi}(\frac{\epsilon_t^2}{h_t} - 1) + \lambda\frac{\partial g(h_t)}{\partial \xi}\frac{\epsilon_t}{h_t} \end{array}\right] \qquad (B.1)$$

All the above derivatives in (B.1) involve $\frac{\partial h_t}{\partial \theta}$ as well as $\frac{\partial g(h_t)}{\partial \theta}$, and these may be obtained as follows :

$$\begin{array}{rcl} \frac{\partial h_t}{\partial \beta} &=& -2\left\{\alpha_1\epsilon_{t-1}(x_{t-1} + \lambda h_{t-1}^{\xi-1}\frac{\partial h_{t-1}}{\partial \beta}) + ... + \alpha_p\epsilon_{t-p}\left(x_{t-p} + \lambda h_{t-p}^{\xi-1}\frac{\partial h_{t-p}}{\partial \beta}\right)\right\} \\ \frac{\partial h_t}{\partial \lambda} &=& -2\lambda\left(\alpha_1\epsilon_{t-1}h_{t-1}^{\xi-1}\frac{\partial h_{t-1}}{\partial \lambda} + ... + \alpha_p\epsilon_{t-p}h_{t-p}^{\xi-1}\frac{\partial h_{t-p}}{\partial \lambda}\right) \\ && -2\left(\alpha_1\epsilon_{t-1}g(h_{t-1}) + ... + \alpha_p\epsilon_{t-p}g(h_{t-p})\right) \\ \frac{\partial h_t}{\partial \alpha} &=& \eta_t - 2\lambda\left(\alpha_1 h_{t-1}^{\xi-1}\epsilon_{t-1}\frac{\partial h_{t-1}}{\partial \alpha} + ... + \alpha_p h_{t-p}^{\xi-1}\epsilon_{t-p}\frac{\partial h_{t-p}}{\partial \alpha}\right) \\ \frac{\partial h_t}{\partial \xi} &=& 2\lambda\alpha_1\epsilon_{t-1}\left[h_{t-1}^{\xi-1}\frac{\partial h_{t-1}}{\partial \xi} + \frac{h_{t-1}^{\xi}lnh_{t-1}}{\xi} - \frac{h_{t-1}^{\xi}-1}{\xi^2}\right] - \\ && \cdots - 2\lambda\alpha_p\epsilon_{t-p}\left[h_{t-p}^{\xi-1}\frac{\partial h_{t-p}}{\partial \xi} + \frac{h_{t-p}^{\xi}lnh_{t-p}}{\xi} - \frac{h_{t-p}^{\xi}-1}{\xi^2}\right] \end{array}$$

where $\eta_t' = \left(1, \epsilon_{t-1}^2, \ldots, \epsilon_{t-p}^2\right)$.

Since $\frac{\partial g(h_t)}{\partial \theta}$ can be easily expressed in terms of $\frac{\partial h_t}{\partial \theta}$, substituting all these above expressions of $\frac{\partial h_t}{\partial \theta}$ in (B.1) we have the final expression for $\frac{\partial l_t}{\partial \theta}$ as follows :

$$\begin{array}{rcl} \frac{\partial lt}{\partial \beta} &=& -\frac{1}{h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t h_t^{\xi-1}\right)\left\{(\alpha_1\epsilon_{t-1}\xi_{t-1} + \cdots + \alpha_p\epsilon_{t-p}\xi_{t-p})\right. \\ && \left. + \lambda\left(\alpha_1\epsilon_{t-i}h_{t-1}^{\xi-1}\frac{\partial h_{t-1}}{\partial \beta} + \cdots + \alpha_p\epsilon_{t-p}h_{t-p}^{\xi-1}\frac{\partial h_{t-p}}{\partial \beta}\right)\right\} + \frac{\epsilon_t}{h_t}x_t \\ \frac{\partial l_t}{\partial \lambda} &=& -\frac{1}{h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t h_t^{\xi-1}\right)\left\{\left(\alpha_i\epsilon_{t-1}h_{t-1}^{\xi-1}\frac{\partial h_{t-1}}{\partial \lambda} + \cdots + \alpha_p\epsilon_{t-p}h_{t-p}^{\xi-1}\frac{\partial h_{t-p}}{\partial \lambda}\right)\right. \\ && \left. + (\lambda_1\epsilon_{t-1}g(h_{t-1}) + \cdots + \alpha_{p'}g(h_{t-p}))\right\} + g(h_t)\frac{\epsilon_t}{h_t} \\ \frac{\partial l_t}{\partial \alpha} &=& \frac{1}{2h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t h_{t^{\xi-1}}\right)\left\{\eta_t - 2\lambda\left(\alpha_1 h_{t-1}^{\xi-1}\epsilon_{t-1}\frac{\partial h_{t-1}}{\partial \alpha}\right.\right. \\ && \left.\left. + \cdots + \ \alpha_p h_{t-p}^{\xi-1}\epsilon_{t-p}\frac{\partial h_{t-p}}{\partial \alpha}\right)\right\} \\ \frac{\partial l_t}{\partial \xi} &=& -\frac{\lambda}{h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t h_t^{\xi-1}\right)\left\{\alpha_1\epsilon_{t-1}\left(h_{t-1}^{\xi-1}\frac{\partial h_{t-1}}{\partial \xi} + \frac{h_{t-1}^{\xi}lnh_{t-1}}{\xi} - \frac{h_{t-1}^{\xi}-1}{\xi^2}\right)\right. \\ && \left. + \cdots + \alpha_p\epsilon_{t-p}\left(h_{t-p}^{\xi-1}\frac{\partial h_{t-p}}{\partial \xi} + \frac{h_{t-p}^{\xi-1}lnh_{t-p}}{\xi} - \frac{h_{t-p}^{\xi}-1}{\xi^2}\right)\right\} \\ && +\lambda\frac{h_t^{\xi-1}lnh_t\epsilon_t}{\xi} - \lambda\frac{h_t^{\xi}-1}{\xi^2}\frac{\epsilon_t}{h_t} \end{array}$$

Case II : $\xi = 0$. In this case $g(h_t) = ln\ h_t$, and hence we can easily find that the final expressions for the first order derivatives of l_t are as follows :

$$
\begin{aligned}
\frac{\partial l_t}{\partial \beta} &= -\frac{1}{h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t\right)\Big\{(\alpha_1\epsilon_{t-1}x_{t-1} + \cdots + \alpha_p\epsilon_{t-p}x_{t-p}) \\
&\quad + \lambda\left(\alpha_1\epsilon_{t-1}\frac{\partial h_{t-1}}{\partial \beta} + \cdots + \alpha_p\epsilon_{t-p}\frac{\partial h_{t-p}}{\partial \beta}\right)\Big\} + \frac{\epsilon_t}{h_t}x_t \\
\frac{\partial l_t}{\partial \lambda} &= -\frac{1}{h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t\right)(\alpha_1\epsilon_{t-1}h_{t-1} + \cdots + \alpha_p\epsilon_{t-p}h_{t-p}) \\
&\quad - \frac{\lambda}{h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t\right)\left(\alpha_1\epsilon_{t-1}\frac{\partial h_{t-1}}{\partial \lambda} + \cdots + \alpha_p\epsilon_{t-p}h_{t-p}\frac{\partial h_{t-p}}{\partial \lambda}\right) + \epsilon_t \\
\frac{\partial l_t}{\partial \alpha} &= \frac{1}{2h_t}\left(\frac{\epsilon_t^2}{h_t} - 1 + 2\lambda\epsilon_t\right)\left\{\eta_t - 1\lambda\left(\alpha_1\epsilon_{t-1}\frac{\partial h_{t-1}}{\partial \alpha} + \cdots + \alpha_p\epsilon_{t-p}\frac{\partial h_{t-p}}{\partial \lambda}\right)\right\}.
\end{aligned}
$$

Acknowledgement. The authors would like to thank Anil K. Bera for helpful comments on the paper. The authors are thankful also to an anonymous referee for very useful comments which have led to a significant improvement in the empirical content of the paper.

References

Amsler, Christine (1984) Term structure variance bounds and time varying liquidity premia. *Economics Letters*, **6**, 137 - 44.

Backus, D.V. and A.W. Gregory (1993) Theoretical relations between risk premiums and conditional variances. *Jour. Business and Economic Statistics*, **11**, 177 - 85.

Bera, A.K. and M.L. Higgins (1993) A survey of ARCH models : Properties, Estimation and Testing. *Jour. Economic Surveys* , **7** , 305 - 66.

Berndt, E.K., B.H. Hall, R.E. Hall and J.A. Hausman (1974) Estimation and inference in nonlinear structural models. *Ann. Economic and Social Measurement*, **4**, 653 - 65.

Bollerslev, T., R.Y. Chou and K.F. Kroner (1992) ARCH modelling in finance : A review of the theory and empirical evidences. *Jour. Econometrics*, **52**, 5 - 59.

Bollerslev, T., R.F. Engle and J.M. Wooldrige (1988) A capital asset pricing model with time varying covariances. *Jour. Political Economy*, **96**, 116 - 31.

Bollerslev, T., R.F. Engle and D.B. Nelson (1994) ARCH Models, in R.F. Engle and D.L. McFadden (eds.) *Handbook of Econometrics*, **Vol. IV**, Elsevier Science B.V.

Box, G.E.P. and D.R. Cox (1964) An analysis of transformations. *Jour. Royal Statist. Soc., Series B*, **26**, 211 - 43.

Crowder, M.J. (1976) Maximum likelihood estimation for dependent observations. *Jour. Royal Statist. Soc., Series B*, **38**, 45 - 53.

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

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) Modelling the persistence of conditional variances, *Econometric Reviews*, **5**, 1-87.

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

Glosten, L., R. Jagannathan and D. Runkle (1993) Relationship between the expected value and the volatility of the nominal excess return on stocks. *Jour. Finance*, **48**, 1779-1801.

Granger, C.W.J. and P. Newbold (1976) Forecasting transformed series. *Jour. Royal Statist. Soc., B*, **38**, 189 - 203.

Gweke, J. (1986) Modelling the persistence of conditional variances : Comment. *Econometric Reviews*, **5**, 57 - 61.

Harvey, C. (1989) Time varying conditional covariances in tests of asset pricing models. *Jour. Financial Economics*, **24**, 289 - 317.

– – –– (1991) The specification of conditional expectations. *Duke University, Department of Economics* (mimeo).

Hentschel, L. (1995) All in the Family : Nesting Symmetric and Asymmetric GARCH Models. *Jour. Financial Economics*, **39**, 71-104.

Higgins, M.L. and A.K. Bera (1992) A class of nonlinear ARCH models. *International Economic Review*, **33**, 137 - 58.

Hodrick, R. J. and S. Srivastava (1984) An investigation of risk and return in forward foreign exchange rates, *Jour. International Money and Finance*, **3**, 5 - 29.

Hong, E.P. (1991) The autocorrelation structure for GARCH-M process. *Economics Letters*, **37**, 129 - 32.

Hopwood, W.S., J.C. Mckeown and P. Newbold (1984) Time series forecasting models involving power transformations. *Jour. Forecasting*, **3**, 57 - 61.

Hsieh, D.A. (1989) Testing non-linear dependence in daily foreign exchanges rates. *Jour. Business*, **62**, 339 - 68.

Kendall, J. D. (1989) Role of exchange rate volatility in U.S. import price pass-through relationships, unpublished *Ph.D. dissertation*, Department of Economics, University of California, Davis, CA.

Lintner, J. (1965) The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, **47**, 13 - 37.

Lo, A. and C. Mackinlay (1988) Stock market prices do not follow random walks: Evidence from a simple specification test.*Review of Financial Studies, 1*, **1**, 41-46.

Merton, R.C. (1986) On estimating the expected return on the market : An exploratory investigation. *Jour. Financial Economics*, **8**, 323 - 61.

Pagan, A.R. and Y.S. Hong (1991) Nonparametric estimation and the risk premium, in W.A. Barnett, J. Powell and G. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Econometrics and Statistics, Cambridge University Press, Cambridge.*

Pagan, A.R. and A. Ullah (1988) The econometric analysis of models with risk terms. *Jour. Applied Econometrics*, **3**, 87 - 105.

Pantula, S.G. (1986) Modelling the persistence of conditional variances : Comment. *Econometric Review*, **5**, 79 - 97.

Pesando, J.E. (1983) On expectations, term premiums, and the volatility of long-term interest rates. *Jour. Monetary Economics*, **12**, 467 - 74.

Scheinkman, J.A. and B. LeBaron (1989) Non-linear dynamics and stock returns . *Jour. Business* , **62** , 311-37.

Shiller, R. J. (1979) The volatility of long-term interest rates and expectations models of the term structure, *Jour. Political Economy*, **87**, 1190-1219.

Shiller, R. J., J. Y. Campbell, and K. L. Schoenholtz (1983) Forward rates and future policy : Interpreting the term structure of interest rates, *Brookings Papers on Economic Activity*, 173-217.

Samarjit Das and Nityananda Sarkar
Economic Research Unit
Indian Statistical Institute
203 B.T. Road, Calcutta - 700 035
India
E-mail : tanya@isical.ac.in