Statistics & Probability Letters 50 (2000) 365 374

Arch model with Box–Cox transformed dependent variable

Nityananda Sarkar

Indian Statistical Institute, Economic Research Unit, 203 Barrackpore Trunk Road, Calcutta 700 035, India

Abstract

Box-Cox power transformation has been used traditionally to linearise otherwise nonlinear models. In this paper, Engle's linear ARCH specification is considered for a regression model in which the dependent variable is Box Cox transformed. The consequent issues arising in both testing and estimation of the model are investigated. A Lagrange multiplier test is also developed to test Engle's linear ARCH model against this wider class of models. The usefulness of this generalisation is examined by applying it to the daily closing prices on the Bombay Stock Exchange Sensitive Index, and the findings strongly favour the proposed model.

Keywords: ARCH/GARCH model; Box–Cox ARCH model; Box–Cox (BC) transformation; Extended BCGARCH model; Extended BC transformation; Lagrange multiplier test

1. Introduction

Until the early 1980s the focus of most macroeconometric and financial time-series modelling centred on the conditional first moment. With the recognition of the increasingly important role played by risk and uncertainty considerations in modern economic theory, the development of new econometric time-series techniques that allow for modelling of time-varying variances and covariances began in right earnest. The most important contribution in this new development has been the autoregressive conditional heteroscedastic (ARCH) class of models introduced by Engle (1982). In time-series regressions, Engle's linear ARCH formulation allows the conditional variance of the present error to change over time as a function of past errors keeping the unconditional variance time invariant. It has been observed that such models can capture many important features like the thick-tail distribution, nonlinear dependence and volatility clustering of many economic and financial variables. Since Engle's paper there have been many extensions and generalisations of the ARCH model. Some of the well-known ones which have been also found to be extremely useful in applications are generalised ARCH (GARCH), nonlinear ARCH (NARCH), exponential GARCH (EGARCH), threshold

ARCH (TARCH) and ARCH-in-the-mean (ARCH-M) models (see Bera and Higgins (1993) for an excellent survey on ARCH models).

Although one of the most important features of the linear ARCH model is that it postulates a nonlinear relationship between the present and the past values of a time series, current evidence suggests that it is not nonlinear enough to model some financial time-series data. For example, Hsieh (1989, p. 336) found that the GARCH model cannot fit some exchange rates satisfactorily; Scheinkman and LeBaron (1989, p. 313) found evidence that volatility in stock market data cannot be captured completely by the linear ARCH model. To deal with this limitation of "inadequate" nonlinearity inherent in the linear ARCH specification, several researchers like Engle and Bollerslev (1986) and Pantula (1986) have suggested alternative functional forms for the conditional variance function. Recently, Higgins and Bera (1992) and Hentschel (1995) have further contributed to this specific area of research. While Higgins and Bera have proposed a generalisation, called the nonlinear ARCH (NARCH) model, which can be viewed as the Box and Cox (1964) power transformation on both sides of Engle's specification of conditional variance, Hentschel's suggestion is to consider a nested family of asymmetric GARCH models treating the variance equation as a law of motion for the Box–Cox transformation of the conditional standard deviation.

In this paper we propose to capture the strong nonlinearity in some financial time-series data by suggesting a different approach in which instead of assuming highly nonlinear and somewhat arbitrary specifications for the conditional variance, the time-series variable (dependent variable in the context of regression framework) is transformed by the Box–Cox (1964) family of power transformations. This transformation has traditionally been used to linearise otherwise nonlinear models (see Carroll and Ruppert, 1988, p. 118). This has also been used for reducing heterogeneity and achieving symmetric distribution of the transformed variable. In fact, Box–Cox proposed this family of power transformations of the dependent variable in a regression model so as to achieve all the three properties, viz., linearity, homoscedasticity and normality of the transformed dependent variable simultaneously. Although subsequent researchers, e.g., Draper and Cox (1969), Guerrero (1993), Poirier (1978) and Sarkar (1985), have shown that all the three desirable properties may not, in general, be achieved simultaneously, BC transformation is used in both time series and cross-section data analysis with the aim of achieving these properties as closely as possible. Insofar as the performance in terms of forecast accuracy of BC transformation is concerned, the available evidence is somewhat mixed. For instance, while Hopwood et al. (1984) have found that in time-series analysis incorporating this power transformation for ARIMA models is beneficial in terms of forecast accuracy, Granger and Newbold (1976) and Nelson and Granger (1979) observed while analysing a collection of macroeconomic time series, that there was little gain due to this transformation in forecast quality. Such mixed evidences have led Granger and Newbold (1986, p. 119) to conclude that "whether the additional effort is likely to lead to much reward in terms of superior forecast performance is not clear". This observation notwithstanding the use of BC transformation is often motivated by the possibility of obtaining improved forecasts. It may also be noted that BC transformation has been modified by Bickel and Doksum (1981) to accommodate negative values of the underlying variable. In this context it may be mentioned that in all the works on ARCH model and its various generalisations and extensions as well as numerous applications of these models, the underlying time-series variable has been assumed to be nontransformed.

While it is true that the analysis and the interpretations are made easier if the data are not transformed, we argue that the linear ARCH specification may not be able to capture adequately the strong nonlinearity of many financial time series, and that a suitable transformation of the time series may be able to do so. Following arguments of Carroll and Ruppert (1988) as also of Mills (1990, p. 41), we reason that for a given time series data having strong nonlinear dependence there would exist a value of λ, the transformation parameter under the BC family of transformations, such that the ARCH model with transformed variable appropriately models the strong nonlinearity in the data.

The proposed generalisation of the ARCH model might also be useful in another respect, viz., the transformation might affect the distributional shape of the variable favourably. It is well known that the unconditional distribution of data for which linear ARCH models are used is frequently skewed and leptokurtic. Since appropriate transformation may induce symmetry to the distribution, it is probable that the proposed generalisation would yield, in addition to capturing the presence of strong nonlinearity in the data, an approximate symmetric and mesokurtic distribution. In this context it may be pointed out that BC transformation does not, in general, permit large negative values for the transformed variable (cf. Poirier, 1978). Also, Yeo and Johnson (1998) have shown that the inducement of symmetric distribution may not be achieved always if the modified transformation of Bickel and Doksum (1981) is used. However, as stated by Davidson and MacKinnon (1985), Lahiri and Egy (1981) and others, it may safely be assumed that the probability for large negative values for the transformed variable is very small and hence the property of symmetry to the distributional shape may not be seriously affected. Thus, apart from capturing strong nonlinear dependence, the proposed model is likely to induce (near) symmetric and mesokurtic properties (i.e., normality) for the unconditional distribution. The paper is organised as follows. In Section 2, the proposed model is described. The estimation of the model is discussed in Section 3. A Lagrange multiplier test to find the adequacy of Engle's linear ARCH model against the generalisation proposed is developed in Section 4. The application of the generalised model to a given financial data is discussed in Section 5. The paper ends with some concluding remarks in Section 6.

2. The model

The generalisation of the ARCH model, to be henceforth called as the Box-Cox transformed ARCH (BCARCH) model, as proposed in this paper, is specified as follows:

$$z_t \mid \Psi_{t-1} \sim \mathrm{N}(x_t'\beta, h_t), \quad t = 1, 2, \ldots, T, \tag{2.1}$$

where z_t is the BC transformed value of the (original) dependent variable y_t i.e.,

$$z_t = \begin{cases} (y_t^\lambda - 1)/\lambda, & \lambda \neq 0, \\ \ln y_t, & \lambda = 0, \end{cases} \tag{2.2}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_p \varepsilon_{t-p}^2, \tag{2.3}$$

$\alpha_0 > 0, \alpha_i \geqslant 0$, $i = 1, 2, \ldots, p$, so that the conditional variance is strictly positive, $\varepsilon_t = z_t - x_t'\beta, x_t$ is the $k \times 1$ vector of fixed observations at time t on the k independent variables which may include the lagged values of the dependent variable, β is the $k \times 1$ vector of associated regression coefficients and Ψ_t is the information set at time t. It order that the transformation given in (2.2) is defined for all real values of λ, y_t must obviously be positive for all t. It may, however, be pointed out that the definition of the BC family of power transformations has been extended by Bickel and Doksum (1981) to accommodate negative values of y_t when $\lambda > 0$. The extended BC (EBC) transformation, as it may be called, is defined as

$$\tilde{z}_t = [\mathrm{sign}(y_t)|y_t|^\lambda - 1]/\lambda, \quad \lambda > 0. \tag{2.4}$$

Burbidge et al. (1988) have argued that λ is unlikely to approach 0 in the presence of negative y_t's because of the extreme effects, i.e., the transformed observations may have infinitely large negative values. We, however, include the case where $\lambda = 0$, by defining

$$z_t = \mathrm{sign}(y_t) \ln |y_t|, \quad \lambda = 0. \tag{2.5}$$

Without loss of generality, the first regressor is assumed to be unity so that the model allows for an intercept term. It may also be stated that all the independent variables can as well be transformed with the provision of different values of λ for different variables. However, this would not obviously give rise to any additional statistical problems than those arising in the model given in (2.1), nor would it change any of the essential conclusions – only the number of parameters in the model would increase.

It may be noted that λ is a parameter in this model, and this parameter indicates the degree of nonlinearity present in the data. Obviously, the model reduces to the linear ARCH model of Engle when $\lambda = 1$. Hence, this generalisation provides a framework for developing a specification test for the linear ARCH model.

3. Estimation

In this section we briefly outline the estimation procedure for the BCARCH model. As with the other generalisations of the ARCH model, we suggest the maximum likelihood (ML) method of estimation of BCARCH model. In this context it may be mentioned that Guerrero (1993) has suggested methods for estimating λ in the usual ARIMA model framework from consideration of variance-stabilizing transformation as well as bias reduction of the forecast in the original model. Obviously, these methods cannot be straightway applied to our model. The fact that these methods are essentially model independent holds only in the context "when using ARIMA, structural or unobservable components models for time series" (Guerrero, 1993, p. 46). Extension of these procedures for ARCH models like ours would have a basic problem in that ARCH process implies that while unconditional variance is homoscedastic, the conditional variance is time dependent in nonlinear relationship. Moreover, because of nonlinearity introduced due to power transformation and also due to the nonlinearity inherent in the ARCH specification, it would be extremely difficult to obtain the relevant analytical expressions so as to be able to use some kind of a generalised Guerrero's method.

Under the assumptions stated earlier, the log-likelihood function of the original deservations $y_1, y_2, \ldots, y_T$, is given, conditional upon Ψ_{t-1} (which is being dropped from notations for the sake of simplicity), as

$$l(\theta) = \sum_{t=1}^{T} l_t(\theta), \tag{3.1}$$

where

$$l_t(\theta) = \text{const.} - \frac{1}{2}\ln h_t - \frac{1}{2}\frac{(z_t - x_t'\beta)^2}{h_t} + (\lambda - 1)\ln y_t \tag{3.2}$$

is the log-density function for the tth observation, $\theta' = (\lambda, \beta', \alpha')$ is the $1 \times (k+p+2)$ complete parameter vector, $\alpha' = (\alpha_0, \alpha_1, \ldots, \alpha_p)$ is the $1 \times (p+1)$ component vector of coefficients in the ARCH specification, and the last term is the logarithm of the Jacobian of BC transformation. The Jacobian allows for the change of scale of the response due to the operation of the power transformation. It may be noted that to compare different values of λ, it is necessary to consider the likelihood in relation to that of the original observations, and hence the likelihood function has been accordingly written in (3.2). Mention may also be made of the fact that some values of y_t, especially of financial variables like the return data, are likely to be negative as well. In such situations we suggest, as already discussed in the preceding section, using the EBC transformation as given in (2.4) and (2.5). The log-density function in that case, say $\tilde{l}_t(\theta)$, would be given by

$$\tilde{l}_t(\theta) = \text{const.} - \frac{1}{2}\ln h_t - \frac{1}{2}\frac{(\tilde{z}_t - x_t'\beta)^2}{h_t} + (\lambda - 1)\ln |y_t|. \tag{3.3}$$

The ML estimates of the parameters are obtained by solving the normal equations which are obtained from (3.1) and (3.2). These equations are, however, highly nonlinear not only because of ARCH structure but also due to the transformation parameter λ. In order to solve the nonlinear equations for obtaining the ML estimates, we suggest using the well-known algorithm proposed by Berndt et al. (1974).

If $\hat{\theta}$ is the ML estimate of θ thus obtained, then it is well known that under standard regularity conditions (cf. Bollerslev, 1986; Crowder, 1976), $(\hat{S}'\hat{S})^{1/2}(\hat{\theta} - \theta_0) \underset{\sim}{A} \mathrm{N}(0, I_{k+p+2})$, where $\hat{S} = ((\partial l_t(\theta)/\partial\theta_i))_{\hat{\theta}}$ is the matrix of first-order derivatives evaluated at $\hat{\theta}$ and θ_0 is the true value of the parameter vector θ. The first-order derivatives required for the computations by BHHH algorithm can easily be obtained from (3.2) as follows:

$$\frac{\partial l(\theta)}{\partial \beta} = \sum_t \frac{\varepsilon_t x_t}{h_t} + \sum_t \frac{1}{2h_t}\frac{\partial h_t}{\partial \beta}\left(\frac{\varepsilon_t^2}{h_t} - 1\right), \tag{3.4}$$

$$\frac{\partial l(\theta)}{\partial \alpha} = \sum_t \frac{1}{2h_t}\frac{\partial h_t}{\partial \alpha}\left(\frac{\varepsilon_t^2}{h_t} - 1\right) \tag{3.5}$$

and

$$\frac{\partial l(\theta)}{\partial \lambda} = \sum_t \frac{1}{2h_t}\frac{\partial h_t}{\partial \lambda}\left(\frac{\varepsilon_t^2}{h_t} - 1\right) - \sum_t \frac{\varepsilon_t}{h_t}\frac{\partial \varepsilon_t}{\partial \lambda} + \sum_t \ln y_t. \tag{3.6}$$

If these derivatives were to be obtained for the likelihood function based on $\tilde{l}_t(\theta)$ in (3.3), $\partial l(\theta)/\partial \beta$ and $\partial l(\theta)/\partial \alpha$ would remain the same, only in $\partial l(\theta)/\partial \lambda$ the last term would change to $\sum_t \ln |y_t|$. Since all these derivatives in turn involve the derivatives of h_t with respect to the parameters α, β and λ and also $\partial \varepsilon_t/\partial \lambda$, the final expressions of these derivatives may be found by substituting the derivatives of $\partial h_t/\partial \alpha, \partial h_t/\partial \beta, \partial h_t/\partial \lambda$ and $\partial \varepsilon_t/\partial \lambda$ in (3.4), (3.5) and (3.6). The computational complexity introduced in the BCARCH model arises in evaluating these derivatives which are given in the appendix.

4. A Lagrange multiplier test

Since BCARCH model is a generalisation of the original ARCH model in which y_t has been transformed by the BC transformation so that strong nonlinear dependence present in many financial series may be adequately captured in the model, it is natural and also important to have a test to determine whether Engle's linear ARCH model provides an adequate description of the data. This is more so because once it is suspected that the conditional heteroscedasticity is present in the original data, one is usually inclined to consider the specification of the model with Engle's ARCH model. Since the Lagrange multiplier (LM) test is a very convenient test procedure in such situations, we derive an LM test for testing Engle's ARCH specification against the more general class of ARCH (i.e., BCARCH) models suggested here.

In deriving the LM test statistic, we first note that for the BCARCH model, the information matrix is no longer entirely block diagonal. This result is stated in the following theorem the proof of which is given in the appendix.

Theorem. *Like the ARCH regression model, the information matrix for the BCARCH model is block diagonal between the regression parameters β and the variance parameters α; however, the same does not hold for the information matrix between the regression parameters β and the transformation parameter λ as well as that between the variance parameters α and the transformation parameter λ.*

Since the information matrix is no longer entirely block diagonal, the testing for the mean, the variance and the transformation parameter has to be carried out jointly. Now, the LM test statistic for the null hypothesis H_0: $\lambda = 1$ (i.e., the linear ARCH regression model) against the alternative H_1: $\lambda \neq 1$ (i.e., the BCARCH regression model) is given by

$$\mathrm{LM} = d(\tilde{\theta})' I(\tilde{\theta})^{-1} d(\tilde{\theta}), \tag{4.1}$$

where the score vector $d(\theta) = \partial l(\theta)/\partial \theta$ and the information matrix $I(\theta) = -E(\partial^2 l(\theta)/\partial \theta \partial \theta')$ are evaluated at $\tilde{\theta}$, the ML estimate of θ under the null hypothesis $\lambda = 1$. Differentiating (3.4)–(3.6), we can easily obtain the second-order derivatives of the log-likelihood function.

In evaluating negative of the expectation of the Hessian matrix we note that this expectation operation may be simplified by taking iterated expectations on the information set Ψ_{t-1}. We thus have

$$E\left(-\frac{\partial^2 l(\theta)}{\partial\beta\partial\beta'}\right) = -E\left[E\left(\frac{\partial^2 l(\theta)}{\partial\beta\partial\beta'}\right) \middle| \Psi_{t-1}\right] = \sum_t E\left[\frac{x_t x_t'}{h_t} + \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\beta}\frac{\partial h_t}{\partial\beta'}\right] \tag{4.2}$$

and hence this can be consistently estimated by $I_\beta(\theta) = \sum_t [x_t x_t'/h_t + (1/2h_t^2)(\partial h_t/\partial\beta)(\partial h_t/\partial\beta')]$.

The other elements of the information matrix may be obtained in a similar fashion. To obtain the final form of the test statistic in (4.1), we evaluate the score vector $d(\theta)$ in (3.4)–(3.6) and the information matrix under the restriction $\lambda = 1$, and then substitute these in (4.1). The expressions for the derivatives in the LM test statistic involve those of $\partial h_t/\partial\theta$ and $\partial\varepsilon_t/\partial\lambda$ and it may easily be obtained from (7.5) to (7.9) in the appendix by substituting $\lambda = 1$.

One may alternatively construct the LM test from the matrix of scores, S. Suppose that S is calculated for the BCARCH model, but evaluated at the parameter estimates under the null hypothesis. The LM test can then be constructed as

$$\mathrm{LM} = i'\tilde{S}(\tilde{S}'\tilde{S})^{-1}\tilde{S}'i, \tag{4.3}$$

where i is a $T \times 1$ unit vector. Obviously, the LM test statistic can be expressed as

$$\mathrm{LM} = TR_0^2, \tag{4.4}$$

where R_0^2 is the unentered coefficient of determination (R^2) obtained by regressing the unit vector on the matrix of scores under the null hypothesis. This is easily computed from the R^2 of the first iteration starting from the estimates found by BHHH algorithm under the null hypothesis.

Since this LM test should be viewed as a diagnostic check of the adequacy of Engle's linear ARCH regression model after it has been estimated, in practical applications one should construct other diagnostic tests like the BDS test statistic of Brock et al. (1987), Ljung and Box (1978) test, etc., to convey more information on the validity of the chosen model.

5. An application

In this section we report the result of an application of the BCARCH model to daily closing prices on the Bombay Stock Exchange (BSE) as measured by the BSE Sensitive Index (SENSEX). The data cover the period January, 1984 to January, 1996. The analysed series is the first difference of the logarithms of this index. Hence the data represent the continuously compounded rate of return for holding the (aggregate) securities for one day. It may be mentioned in this context that some of the return figures were found to have small negative values (as expected from the definition of return itself), and thus in this example instead of the usual BC transformation (cf. (2.2)) the EBC as defined in (2.4) and (2.5) has been used. As it is well known that GARCH is a better representation of the conditional variance than ARCH., the former has been assumed to be the conditional variance function for this application i.e., for this example h_t has been assumed to be given by

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_p\varepsilon_{t-p}^2 + \delta_1 h_{t-1} + \cdots + \delta_q h_{t-q}, \tag{5.1}$$

where $\alpha_0 > 0, \alpha_i \geqslant 0$ for $i = 1,2,\ldots,p$ and $\delta_i \geqslant 0$ for $i = 1,2,\ldots,q$. We have thus fitted the EBCGARCH model to this data. For this purpose, we have assumed that apart from an intercept term there is a regressor in the form of z_{t-1} in the model. We had earlier fitted the model without the regressor only to find that there are high autocorrelations in the residuals, and this means that in such a case the conditional mean model is misspecified. In fact, many researchers have found that the absolute values of stock returns are highly autocorrelated.

Table 1
Estimated models for sensex data[a]

	$\hat{\lambda}$	$\hat{\beta}$	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\delta}_1$	Maximum log-likelihood
GARCH (1, 1)	1	0.1736 (0.0202)	0.3028×10^{-5} (0.5928×10^{-6})	0.0932 (0.0084)	0.9020 (0.0076)	7120.29
EBCGARCH (1, 1)	0.89	0.1778 (0.0201)	0.8769×10^{-5} (0.2260×10^{-5})	0.0911 (0.0102)	0.9022 (0.0101)	7153.13

[a]Figures in parentheses show the standard errors of the estimates.

Table 2
Ljung–Box test statistic values for the residuals

	$\hat{\lambda}$	$Q(1)$	$Q(3)$	$Q(6)$	$Q(7)$	$Q(9)$	$Q(10)$	$Q(11)$	$Q(12)$
GARCH (1, 1)	1	5.60[a]	12.55[b]	13.92[a]	14.64[a]	19.61[a]	19.97[a]	22.15[a]	23.93[a]
EBCGARCH (1, 1)	0.89	2.99	11.33[b]	11.93	11.95	16.84	17.58	18.87	19.96

[a]Indicates significance at 5% level.
[b]Indicates significance at 1% level.

On the basis of estimation of the EBCGARCH model with this data set, we have found that EBCGARCH (1, 1) fits the data best, and the log-likelihood function is maximum at $\lambda = 0.89$, the maximum value being 7153.13. Insofar as the performance by the usual linear GARCH model, i.e., where $\lambda = 1$, is concerned we have found that the maximum value is attained when the order is (1, 1) and the log-likelihood value, as given in Table 1, is 7120.29. Thus, it turns out that the estimated value of the transformation parameter λ is quite close to 1 for this data set. To find out if this closeness is statistically significant or not, we carried out an asymptotic test of H_0: $\lambda = 1$ against H_1: $\lambda \neq 1$, and found that the test rejected H_0 against H_1 at 1 per cent level of significance. It is then clear that the fitted EBCGARCH (1, 1) model is superior to the usual linear GARCH (1, 1) model. We have thus established through this example that an appropriate transformation of the variable in the GARCH framework is able to model the data better by capturing the nonlinear dependence in the time series more clearly than that explained by the linear GARCH model.

We have also studied the residuals of the chosen EBCGARCH (1, 1) model as well as those of the linear GARCH (1, 1) model to find if the residuals exhibit any dependence based on which the two models could be further compared. Towards this end we have used the diagnostic test given by the Ljung–Box $Q(m)$ test statistic where m indicates the lag value of the autocorrelation process. The values of this diagnostic test statistic are given in Table 2.

We observe from this table that all but one of the values of $Q(m)$ test statistic for the EBCGARCH model are insignificant at 5% level of significance. The same does not hold for the linear GARCH model. In the latter, $Q(m)$ values are significant for all lag values m at 5% level of significance; $Q(3)$ is significant even at 1% level of significance. We also carried out the Breusch–Godfrey LM test under ARMA alternative to check whether the residuals were nonautocorrelated. The value of this test statistic was found to be 5.60 for the linear GARCH (1, 1) model and 2.99 for the chosen EBCGARCH (1, 1) model. Thus, this diagnostic test suggests that while the residuals of the former model have significant autocorrelations at 2.5% level of significance, those of the latter are insignificant even at 5% level. This finding is quite in conformity with those obtained from the Ljung–Box test. Lastly, we compared the skewness and kurtosis coefficients of the residuals of the two models and found that EBCGARCH model was closer to normality than the linear GARCH model.

This is evident from the fact that the values of coefficient of skewness were 0.1112 and 0.0025 for the linear GARCH (1, 1) model and EBCGARCH (1, 1) model, respectively, and those of kurtosis were 5.8753 and 3.3387. We conclude therefore that for this example the data support the transformation approach proposed in this paper.

6. Conclusions

In time-series modelling the linear ARCH model is widely used for analysing data on financial and other similar economic variables. The original linear ARCH model of Engle has been generalised and extended in many ways so as to incorporate other features and complexities of such data.

In this paper we consider another generalisation in which the dependent variable is transformed by the well-known Box–Cox family of transformations; if some values are negative as the case may very well be with return data, the extended BC transformation of Bickel and Doksum is to be used. Such a generalisation is likely to capture the strong nonlinear dependence (in the time series) which otherwise remains unexplained by the linear ARCH model. We have discussed the maximum likelihood method of estimation of the BCARCH model, and also suggested a Lagrange multiplier test for testing Engle's linear ARCH specification against the wider class of alternative models given by the BCARCH specification. Finally, we have shown through the analysis of daily closing prices on the Bombay Stock Exchange Sensitive Index, that such a generalisation of the linear ARCH model may perform much better than the original linear ARCH model.

Acknowledgements

The author would like to thank Anil K. Bera for his valuable comments on the paper. Thanks are also due to an anonymous referee for his comments on an earlier draft of the paper. The computational assistance by Samarjit Das is duly acknowledged.

Appendix

Proof of the Theorem. We have already noted in Section 4 that the unconditional expectation is obtained through conditional expectation on the information set at time $t-1$. We thus have

$$E\left(-\frac{\partial^2 l(\theta)}{\partial\beta\partial\alpha'}\right) = -E\left(E\left[\frac{\partial^2 l(\theta)}{\partial\beta\partial\alpha'}\right]\middle| \Psi_{t-1}\right) \tag{A.1}$$

and this may easily be checked to simplify to

$$E\left(\frac{\partial^2 l(\theta)}{\partial\beta\partial\alpha'}\right) = \sum_t E\left(\frac{1}{2h_t^2}\frac{\partial h_t}{\partial\beta}\frac{\partial h_t}{\partial\alpha'}\right). \tag{A.2}$$

Similarly, we have

$$E\left(-\frac{\partial^2 l(\theta)}{\partial\beta\partial\lambda}\right) = \sum_t E\left\{-\frac{x_t}{h_t}E\left(\frac{\partial\varepsilon_t}{\partial\lambda}\right) + \frac{1}{2h_t^2}\frac{\partial h_t}{\partial\beta}\frac{\partial h_t}{\partial\lambda}\right\} \tag{A.3}$$

and

$$E\left(-\frac{\partial^2 l(\theta)}{\partial\alpha\partial\lambda}\right) = \sum_t E\left(\frac{1}{2h_t^2}\frac{\partial h_t}{\partial\alpha}\frac{\partial h_t}{\partial\lambda}\right). \tag{A.4}$$

Since the expressions in (A.1)–(A.4) involve $\partial h_t/\partial\alpha, \partial h_t/\partial\beta, \partial h_t/\partial\lambda$ and $\partial\varepsilon_t/\partial\lambda$, we differentiate h_t with respect to the parameters and ε_t with respect to λ, to obtain

$$\frac{\partial h_t}{\partial\beta_i} = -2\sum_{j=1}^{p}\alpha_j\varepsilon_{t-j}x_{t-j,i}, \quad i=1,2,\ldots,k, \tag{A.5}$$

$$\frac{\partial h_t}{\partial\alpha_0} = 1, \tag{A.6}$$

$$\frac{\partial h_t}{\partial\alpha_j} = \varepsilon_{t-j}^2, \quad j=1,2,\ldots,p, \tag{A.7}$$

$$\frac{\partial h_t}{\partial\lambda} = 2\sum_{j=1}^{p}\alpha_j\varepsilon_{t-j}(y_{t-j}^{\lambda}\ln y_{t-j}^{\lambda} - y_{t-j}^{\lambda}+1)/\lambda^2 \tag{A.8}$$

and

$$\frac{\partial\varepsilon_t}{\partial\lambda} = (y_t^{\lambda}\ln y_t^{\lambda} - y_t^{\lambda}+1)/\lambda^2. \tag{A.9}$$

Using Engle's definition of symmetric function for the ARCH model, we find (i) from (2.3) that h_t is a symmetric function of $\varepsilon_{t-1},\ldots,\varepsilon_{t-p}$, (ii) from (A.6) that $\partial h_t/\partial\alpha_0$ is a symmetric function, and (iii) from (A.7) that $\partial h_t/\partial\alpha_j$ $(j=1,2,\ldots,p)$ is a symmetric function of ε_{t-j}. However, it is evident from (A.5) that $\partial h_t/\partial\beta_i$ is an antisymmetric function of $\varepsilon_{t-1},\varepsilon_{t-2},\ldots,\varepsilon_{t-p}$. Since each of these ε_{t-j}, $j=1,2,\ldots,p$, is symmetrically distributed around zero, it is then obvious that $E(-\partial^2 l(\theta)/\partial\beta\partial\alpha')=0$. As for the other two expectations in (A.3) and (A.4), viz., $E(-\partial^2 l(\theta)/\partial\beta\partial\lambda)$ and $E(-\partial^2 l(\theta)/\partial\alpha\partial\lambda)$, it can easily be checked, after simplifying the expression for $\partial h_t/\partial\lambda$ in (A.8), that each of these is different from zero.

References

Bera, A.K., Higgins, M.L., 1993. A survey of ARCH models: properties, estimation and testing. J. Econom. Surveys 7, 305–366.

Berndt, E.K., Hall, B.H., Hall, R.E., Hausman, J.A., 1974. Estimation and inference in nonlinear structural models. Ann. Econom. Social Measurement 4, 653–665.

Bickel, P.J., Doksum, K.A., 1981. An analysis of transformations revisited. J. Amer. Statist. Assoc. 76, 296–311.

Bollerslev, T., 1986. A generalized autoregressive conditional heteroscedasticity. J. Econometrics 31, 307–327.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. Roy. Statist. Soc. Ser. B 26, 211–243.

Brock, W., Dechert, W.D., Scheinkman, J., 1987. A test for independence based on the correlation dimension. Mimeo, University of Wisconsin-Madison.

Burbidge, J.B., Magee, L., Robb, A.L., 1988. Alternative transformation to handle extreme values of dependent variable. J. Amer. Statist. Assoc. 83, 123–127.

Carroll, R.J, Ruppert, D., 1988. Transformation and Weighting in Regression. Chapman & Hall, New York.

Crowder, M.J., 1976. Maximum likelihood estimation of dependent observations. J. Roy. Statist. Soc. Ser. B 38, 45–53.

Davidson, R., MacKinnon, J.G., 1985. Testing linear and loglinear regressions against Box–Cox alternatives. Canadian J. Econom. XVIII, 499–517.

Draper, N.R., Cox, D.R., 1969. On distributions and their transformation to normality. J. Roy. Statist. Soc. Ser. B 31, 472–476.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987–1008.

Engle, R.F., Bollerslev, T., 1986. Modelling the persistence of conditional variances. Econometric Rev. 5, 1–50.

Granger, C.W.J., Newbold, P., 1976. Forecasting transformed series. J. Roy. Statist. Soc. Ser. B 38, 189–203.

Granger, C.W.J., Newbold, P., 1986. Forecasting Economic Time Series. Academic Press Inc., San Diego.

Guerrero, V.M., 1993. Time-series analysis supported by power transformations. J. Forecasting 12, 37–48.

Hentschel, L., 1995. All in the family: nesting symmetric and asymmetric GARCH models. J. Financial Econom. 39, 71–104.

Higgins, M.L., Bera, A.K., 1992. A class of nonlinear ARCH models. Int. Econom. Rev. 33, 137–158.

Hopwood, W.S., McKeown, J.C., Newbold, P, 1984. Time series forecasting models involving power transformations. J. Forecasting 3, 57–61.
Hsieh, D.A., 1989. Testing for nonlinear dependence in daily foreign exchange rate changes. J. Business 62, 339–368.
Lahiri, K., Egy, D., 1981. Joint estimation and testing for functional form and heteroscedasticity. J. Econometrics 15, 299–308.
Ljung, G.M., Box, G.E.P., 1978. On a measure of lack of fit in time series model. Biometrika 66, 67–72.
Mills, T.C., 1990. Time Series Techniques for Economists. Cambridge University Press, Cambridge.
Nelson H.L. Jr, Granger, C.W.J., 1979. Experience with using the Box-Cox transformation when forecasting economic time series. J. Econometrics 10, 57–69.
Pantula, S.G., 1986. Comment. Econometric Rev. 5, 71–73.
Poirier, J.J., 1978. Box-Cox transformation in limited dependent variable models. J. Amer. Statist. Assoc. 73, 284–289.
Sarkar, N., 1985. Box-Cox transformation and the problem of heteroscedasticity. Comm. Statist. Theory Methods 363–379.
Scheinkman, J.A., LeBaron, B., 1989. Nonlinear dynamics and stock returns. J. Business 62, 311–337.
Yeo, I.K., Johnson, R.A., 1998. A new family of power transformations to improve normality or symmetry. Technical Report No. 1002, Department of Statistics, University of Wisconsin-Madison.