Journal of the Korean Statistical Society (2002), **31**: *2, pp 199–212*

Statistical Analysis of Transfer Function Models with Conditional Heteroscedasticity

J. S. Baek[1], K. T. Sohn[2] and S. Y. Hwang[3]

Abstract

This article introduces transfer function model (TFM) with conditional heteroscedasticity where ARCH concept is built into the traditional TFM of Box and Jenkins (1976). Model building strategies such as identification, estimation and diagnostics of the model are discussed and are illustrated via empirical study including simulated data and real data as well. Comparisons with the classical TFM are also made.

Keywords. Conditional heteroscedasticity, cross correlation, noise, transfer function model.
AMS 2000 subject classifications. Primary 62M10.

1. Introduction

Transfer function model (TFM, hereafter), especially designed for analyzing relationships between output series and input series, has proved useful in the context of time series. Comprehensive statistical treatments for TFM are given, for instance, in Wei (1990, ch.13) and Box and Jenkins (1976, ch.10). TFM is usually referred to as ARMAX model, X here standing for "exogenous" and can also be viewed as a special case of stochastic regression model (*cf*. Granger and Teräsvirta, 1993). Let zero mean stationary processes $\{x_t\}$ and $\{y_t\}$ be input series and output series, respectively. Typically $\{x_t\}$ influences $\{y_t\}$, but not vice versa. The classical TFM model is defined by the equation

$$y_t = v(B)x_t + n_t, \tag{1.1}$$

Received August 2001; accepted February 2002.
[1]Statistical Research Center for Complex System, Seoul National University, Seoul 151-747, Korea
[2]Department of Statistics, Pusan National University, Pusan 609-735, Korea
[3]Department of Statistics, Sookmyung Women's University, Seoul 140-742, Korea

where $v(B) = \sum_{j=0}^{\infty} v_j B^j$ is called a (transfer function) filter, $\{n_t\}$ is a noise series which is assumed to be independent of the input series $\{x_t\}$, and $\{x_t\}$ and $\{n_t\}$ are modeled to follow stationary ARMA processes of appropriate order. The main objectives of TFM are to figure out the filter $v(B)$ and to identify ARMA model for the noise $\{n_t\}$.

Since the seminal work of Engle (1982), a lot of research has been directed to the conditional heteroscedastic autoregressive (ARCH, for short) models in an attempt to explain volatility in the time series data, especially in the field of econometrics. The process $\{\epsilon_t\}$ is said to follow ARCH(m) if

$$\epsilon_t = \sqrt{h_t} \cdot u_t,$$
$$h_t = \alpha_0 + \sum_{k=1}^{m} \alpha_k \epsilon_{t-k}^2, \tag{1.2}$$

where u_t is a sequence of *iid* random variables with mean zero and variance unity. In order to incorporate serial correlations in (1.2), Weiss (1984) introduced the ARMA model with ARCH(m) errors. We refer to the process $\{n_t\}$ as ARMA(p,q)-ARCH(m) provided

$$\phi_p(B) n_t = \theta_q(B) \epsilon_t, \tag{1.3}$$

where $\{\epsilon_t\}$ is an ARCH(m) process defined in (1.2).

Contrary to the usual practice in TFM that $\{n_t\}$ just follows ARMA process, the present paper deals with TFM where the noise $\{n_t\}$ in (1.1) forms a ARMA(p,q)-ARCH(m) in (1.2) and (1.3). The main advantage of the research model defined (1.1) through (1.3) over the classical TFM appears to be in that our model is capable of capturing the conditional heteroscedastic patterns (volatility, risk) of the data. This article is organized as follows. Model description and estimation problem for the model are addressed in Section 2. Model building strategies are discussed in Section 3. Empirical study is presented in Section 4 with an application of the model to the financial time series data illustrating that the suggested model beats the usual TFM.

2. The Model and Parameter Estimation

Consider the following TFM with conditional heteroscedasticity (TF-ARCH, hereafter).

$$y_t = \frac{\omega_s(B)}{\delta_r(B)} B^b x_t + n_t, \tag{2.1}$$

$$\phi_p(B)n_t = \theta_q(B)\epsilon_t,\ \epsilon_t = h_t^{1/2}u_t,\ h_t = \alpha_0 + \sum_{k=1}^{m}\alpha_k\epsilon_{t-k}^2,$$

where $\omega_s(B) = \omega_0 - \omega_1 B - \cdots - \omega_s B^s$, $\delta_r(B) = 1 - \delta_1 B - \cdots - \delta_r B^r$, and b is a delay parameter, $\theta_q(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$, $\phi_p(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$. $\{u_t\}$ is a sequence of iid $N(0,1)$ variables, α_0 is positive and α_k $(k = 1, 2, \cdots, m)$ are non-negative. Here the filter $v(B)$ is parsimoniously represented by

$$v(B) = \frac{\omega_s(B)}{\delta_r(B)}B^b \tag{2.2}$$

and $\{u_t\}$ is assumed Gaussian. Let $\{(x_1, y_1), \cdots, (x_N, y_N)\}$ be given data and Ψ_j denote the information set available at time j, *i.e.*, Ψ_j is a σ-field generated by the data up to time j, $j = 1, \cdots, N$.

It is to be noted that the model (2.1) can be written as

$$\begin{aligned}\epsilon_t \ &= \ y_t + d_1 y_{t-1} + \cdots + d_{p+r} y_{t-p-r} + c_0 x_{t-b} + c_1 x_{t-b-1} \\ &\quad + \cdots + c_{p+s} x_{t-b-p-s} + b_1 \epsilon_{t-1} + b_2 \epsilon_{t-2} + \cdots + b_{r+q}\epsilon_{t-r-q},\end{aligned} \tag{2.3}$$

where the coefficients $d_1, \cdots, d_{p+r}$ are obtained from δ and ϕ, $c_0, c_1 \cdots, c_{p+s}$ from ω and ϕ, and $b_1, \cdots, b_{r+q}$ are a function of δ and θ.

Let l be the log-likelihood and l_j denote the log-likelihood of the jth observation. Then the log-likelihood function apart from constants is seen to be

$$l = \sum_{j=1}^{N} l_j(\beta, \alpha|\Psi_{j-1}), \tag{2.4}$$

$$l_j(\beta, \alpha) = l_j(\beta, \alpha|\Psi_{j-1}) = -\frac{1}{2}\log h_j - \frac{1}{2h_j}\epsilon_j^2,$$

where $\beta = (\omega_0, \omega_1, \cdots, \omega_s; \delta_1, \cdots, \delta_r; \phi_1, \cdots, \phi_p; \theta_1, \cdots, \theta_q)', \alpha = (\alpha_0, \alpha_1, \cdots, \alpha_m)'$ are parameters to be estimated.

Below we collect conditions needed in the sequel.

(C.1) For the $r \times 1$ vector of parameters $\delta = (\delta_1, \cdots, \delta_r)'$, all the roots of $1 - \sum_{j=1}^{r}\delta_j z^j = 0$ lie outside the unit circle of z.

(C.2) For the $p \times 1$ vector of parameters $\phi = (\phi_1, \cdots, \phi_p)'$, all the roots of $1 - \sum_{j=1}^{p}\phi_j z^j = 0$ lie outside the unit circle of z.

(C.3) For the $q \times 1$ vector of parameters $\theta = (\theta_1, \cdots, \theta_q)'$, all the roots of $1 - \sum_{j=1}^{q}\theta_j z^j = 0$ lie outside the unit circle of z. Furthermore, the polynomials $\phi_p(B)$ and $\theta_q(B)$ have no common zeros for identifiability.

(C.4) $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_m)'$ satisfies $\alpha_0 > 0$, $\sum_{j=1}^{m} \alpha_j < 1$.

(C.5) The fourth order moment $E(\epsilon_t^4)$ is finite.

Remark. (C.2) and (C.3) imply that $\{n_t\}$ is stationary and invertible and (C.4) guarantees the stationarity of the ARCH process $\{\epsilon_t\}$. Regarding the finite fourth order moment condition (C.5), An, Chen and Huang (1997) discussed the sufficient conditions for existence of the fourth (and higher) order moments of $\{\epsilon_t\}$. Let $m \times m$ matrix

$$A_t = \begin{pmatrix} \alpha_1 u_t^2 & \cdots & \alpha_{m-1} u_t^2 & \alpha_m u_t^2 \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix}$$

and $\Sigma_m = E(A_t^{\otimes m})$, where $A^{\otimes m} = A \otimes A \otimes \cdots \otimes A$ and $\otimes$ is the Kronecker product operator. They showed that if the spectral radius $\rho(\Sigma_2) < 1$, then $|\epsilon_t|^2 \in L^2$ where L^2 stands for the collection of all random variables with finite second moment and $\rho(A) = \max|\lambda_i|$, λ_i's are eigenvalues of $n \times n$ matrix A. For ARCH(1) process, it is easily seen that $A_t = \alpha_1 u_t^2$ and $\Sigma_2 = E(\alpha_1^2 u_t^4) = 3\alpha_1^2$. Thus $\rho(\Sigma_2) < 1$ is equivalent to $3\alpha_1^2 < 1$, as is Engle (1982). For higher order models, ARCH(2), say, it can be verified that A_t and Σ_2 are

$$A_t = \begin{pmatrix} \alpha_1 u_t^2 & \alpha_2 u_t^2 \\ 1 & 0 \end{pmatrix},$$

$$\Sigma_2 = E\begin{pmatrix} \alpha_1^2 u_t^4 & \alpha_1\alpha_2 u_t^4 & \alpha_1\alpha_2 u_t^4 & \alpha_2^2 u_t^4 \\ \alpha_1 u_t^2 & 0 & \alpha_2 u_t^2 & 0 \\ \alpha_1 u_t^2 & \alpha_2 u_t^2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3\alpha_1^2 & 3\alpha_1\alpha_2 & 3\alpha_1\alpha_2 & 3\alpha_2^2 \\ \alpha_1 & 0 & \alpha_2 & 0 \\ \alpha_1 & \alpha_2 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Since it is a hard task to analytically identify (α_1, α_2) for $\rho(\Sigma_2) < 1$, we calculate $\rho(\Sigma_2)$ for each values of (α_1, α_2) from (0.05,0.05) to (0.95,0.95) by 0.05 and obtain the following table.

TABLE 2.1 *Spectral radius $\rho(\Sigma_2)$ for ARCH(2)*

α_1	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
α_2	.55	.50	.50	.45	.40	.35	.30	.25	.15	.10
$\rho(\Sigma_2)$	.966	.918	.977	.962	.958	.961	.973	.992	.918	.956

For the first pair of $(\alpha_1, \alpha_2) = (0.05, 0.55)$, $\rho(\Sigma_2)$ turns out to be 0.966 which implies that for $\alpha_1 = 0.05$, any values $0 \le \alpha_2 \le 0.55$ ensure $\rho(\Sigma_2) < 1$, *i.e.*, the finite fourth order moment of ϵ_t. Similarly the table provides the approximated region of (α_1, α_2) for $E\epsilon_t^4 < \infty$. These results for ARCH(2) have not yet been properly spelled out in the literature.

We are now in a position to discuss the maximum likelihood estimators (MLE) of β and α. Let $\gamma = (\beta', \alpha')'$: $k \times 1$ vector with $k = s + r + p + q + m + 2$. The derivatives with respect to α and β are given by

$$\frac{\partial l_j}{\partial \alpha} = -\frac{1}{2h_j}\frac{\partial h_j}{\partial \alpha} + \frac{1}{2h_j^2}\frac{\partial h_j}{\partial \alpha}\epsilon_j^2 = \frac{1}{2h_j}\left(\frac{\epsilon_j^2}{h_j} - 1\right)\frac{\partial h_j}{\partial \alpha},$$

$$\begin{aligned}\frac{\partial l_j}{\partial \beta} &= -\frac{1}{2h_j}\frac{\partial h_j}{\partial \beta} - \frac{1}{2}\left(-\frac{\epsilon_j^2}{h_j^2}\frac{\partial h_j}{\partial \beta} + \frac{2\epsilon_j}{h_j}\frac{\partial \epsilon_j}{\partial \beta}\right)\\ &= \frac{1}{2h_j}\left(\frac{\partial h_j}{\partial \beta}\right)\left(\frac{\epsilon_j^2}{h_j} - 1\right) - \frac{\epsilon_j}{h_j}\left(\frac{\partial \epsilon_j}{\partial \beta}\right)\\ &= \frac{1}{h_j}\left(\frac{\epsilon_j^2}{h_j} - 1\right)\sum_{k=1}^{m}\alpha_k\epsilon_{j-k}\left(\frac{\partial \epsilon_{j-k}}{\partial \beta}\right) - \frac{\epsilon_j}{h_j}\left(\frac{\partial \epsilon_j}{\partial \beta}\right),\end{aligned}$$

and the Hessian matrix is obtained as

$$\frac{\partial^2 l_j}{\partial \alpha \partial \alpha'} = -\frac{1}{2h_j^2}\left(\frac{\epsilon_j^2}{h_j}\right)\left(\frac{\partial h_j}{\partial \alpha}\right)\left(\frac{\partial h_j}{\partial \alpha'}\right) + \left(\frac{\epsilon_j^2}{h_j} - 1\right)\frac{\partial}{\partial \alpha'}\left(\frac{1}{2h_j}\frac{\partial h_j}{\partial \alpha}\right),$$

$$\begin{aligned}\frac{\partial^2 l_j}{\partial \beta \partial \beta'} &= -\frac{1}{h_j}\left(\frac{\partial \epsilon_j}{\partial \beta}\right)\left(\frac{\partial \epsilon_j}{\partial \beta'}\right) + \frac{2\epsilon_j}{h_j^2}\left(\frac{\partial \epsilon_j}{\partial \beta}\right)\left(\frac{\partial h_j}{\partial \beta'}\right) - \frac{1}{2h_j^2}\left(\frac{\epsilon_j^2}{h_j}\right)\left(\frac{\partial h_j}{\partial \beta}\right)\left(\frac{\partial h_j}{\partial \beta'}\right)\\ &\quad - \frac{\epsilon_j}{h_j}\left(\frac{\partial^2 \epsilon_j}{\partial \beta \partial \beta'}\right) + \left(\frac{\epsilon_j^2}{h_j} - 1\right)\frac{\partial}{\partial \beta'}\left(\frac{1}{2h_j}\frac{\partial h_j}{\partial \beta}\right),\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 l_j}{\partial \alpha \partial \beta'} &= \frac{\epsilon_j}{h_j^2}\left(\frac{\partial h_j}{\partial \alpha}\right)\left(\frac{\partial h_j}{\partial \beta'}\right) - \frac{1}{2h_j^2}\left(\frac{\epsilon_j^2}{h_j}\right)\left(\frac{\partial h_j}{\partial \alpha}\right)\left(\frac{\partial h_j}{\partial \beta'}\right)\\ &\quad + \left(\frac{\epsilon_j^2}{h_j} - 1\right)\frac{\partial}{\partial \beta'}\left(\frac{1}{2h_j}\frac{\partial h_j}{\partial \alpha}\right).\end{aligned}$$

Thus, after some calculations, the information matrix, defined as the expected Hessian matrix. turns out to be

$$F_{\alpha\alpha} = \sum_{j=1}^{N} E\left(\frac{1}{2h_j^2}\frac{\partial h_j}{\partial \alpha}\frac{\partial h_j}{\partial \alpha'}\right), \tag{2.5}$$
$$F_{\beta\beta} = \sum_{j=1}^{N} E\left\{\frac{1}{h_j}\left(\frac{\partial \epsilon_j}{\partial \beta}\right)\left(\frac{\partial \epsilon_j}{\partial \beta'}\right) + \frac{1}{2h_j^2}\left(\frac{\partial h_j}{\partial \beta}\right)\left(\frac{\partial h_j}{\partial \beta'}\right)\right\},$$
$$F_{\alpha\beta} = 0.$$

It is to be emphasized that the off-diagonal block $F_{\alpha\beta} = 0$, which is a direct consequence of the symmetric and regular ARCH (*cf*. Engle, 1982). This fact implies that the MLE of α and β are asymptotically independent and thus α and β can be estimated separately.

Define the score vector $S_N(\gamma) = l'(\gamma) = \partial l(\gamma)/\partial\gamma$: $k \times 1$ vector and the sample Fisher information matrix $F_N(\gamma) = -\partial^2 l(\gamma)/\partial\gamma\partial\gamma'$: $k \times k$ matrix. Denote also the information matrix $F(\gamma)$ as

$$F(\gamma) = E(F_N(\gamma)).$$

Lemma 2.1. *Under conditions (C.1) to (C.5), as N tends to infinity,*

(i) $F^{-1/2}(\gamma)S_N(\gamma) \xrightarrow{d} N(0, I_k)$,

(ii) *uniform smoothness of information matrix :*

$$\sup ||F^{-1/2}(\gamma)F_N(\gamma^*)F^{-1/2}(\gamma) - I_k|| = o_p(1),$$

where the "sup" is taken over $\gamma^* \in D(\gamma)$ with, for some constant M,

$$D(\gamma) = \{\gamma^*; ||F^{1/2}(\gamma)(\gamma^* - \gamma)|| < M\}.$$

Here $||B||$ denotes the norm of the matrix B, $||B||^2 = tr(B'B)$, and I_k is used for the $k \times k$ identity matrix. Also $F^{1/2}(\gamma)$ stands for the square-root matrix from the spectral decomposition of $F(\gamma)$. Notice that $D(\gamma)$ converges to a single point γ as $N \to \infty$.

Proof. Let $U_j = \partial l_j(\gamma)/\partial\gamma$ to get $S_N(\gamma) = \sum_{j=1}^{N} U_j$. From (C.5), it can be shown that

$$E(U_j|\Psi_{j-1}) = 0 \quad \text{and} \quad E(U_j^2) < \infty,$$

which implies $S_N(\gamma)$ is a martingale and thus (i) follows by using martingale central limit theorem (Billingsley, 1961).

Notice also that for verifying the assertion (ii) it suffices to show that

$$\sup_{\gamma^* \in D(\gamma)} \left|\lambda'_N[F_N(\gamma^*) - F(\gamma)]\lambda_N\right| \Big/ \lambda'_N F(\gamma)\lambda_N = o_p(1)$$

for every sequence of vectors $\{\lambda_N\}$ with unit length. From (C.2) and (C.4) it follows that the noise process $\{n_t\}$ is ergodic, which implies together with (C.1) that the observation process is also ergodic. One can obtain by the ergodic theorem that as $N \to \infty$,

$$\lambda'_N F_N(\gamma)\lambda_N / \lambda'_N F(\gamma)\lambda_N \xrightarrow{p} 1. \tag{2.6}$$

Split

$$\lambda'_N[F_N(\gamma^*) - F(\gamma)]\lambda_N = \lambda'_N[F_N(\gamma^*) - F_N(\gamma)]\lambda_N + \lambda'_N[F_N(\gamma) - F(\gamma)]\lambda_N. \tag{2.7}$$

The second term in the R.H.S. of (2.7) is clearly $o_p(\lambda'_N F(\gamma)\lambda_N)$ due to (2.6). Also it follows from the special structure of the second order derivatives of the log-likelihood previously mentioned that for all sufficiently large N,

$$\left|\lambda'_N[F_N(\gamma^*) - F_N(\gamma)]\lambda_N\right| \le C \cdot \lambda'_N F_N(\gamma)\lambda_N \cdot ||\gamma^* - \gamma||,$$

where C is a constant which is free from N and may depend on γ.

Consequently

$$\sup_{\gamma^* \in D(\gamma)} \left|\lambda'_N[F_N(\gamma^*) - F_N(\gamma)]\lambda_N\right| = o_p\Big(\lambda'_N F(\gamma)\lambda_N\Big),$$

which readily yields the assertion (ii). □

The theorem below addresses the asymptotic distribution of the MLE $\hat{\gamma}$ of γ.

Theorem 2.1. *Under the same conditions as in Lemma 2.1, the maximum likelihood estimator $\hat{\gamma} = (\hat{\beta}', \hat{\alpha}')'$, which is obtained by solving $S_N(\hat{\gamma}) = 0$, is asymptotically normal, viz.,*

$$F^{1/2}(\gamma)(\hat{\gamma} - \gamma) \xrightarrow{d} N(0, I_k).$$

Proof. By Taylor's expansion about γ, it can be written as

$$S_N(\hat{\gamma}) = S_N(\gamma) - F_N(\gamma^*)(\hat{\gamma} - \gamma)$$

where $||\gamma^* - \gamma|| \le ||\hat{\gamma} - \gamma||$. This readily yields

$$F^{1/2}(\gamma)(\hat{\gamma} - \gamma) = \{-F^{1/2}(\gamma)F_N^{-1}(\gamma^*)F^{1/2}(\gamma)\} \cdot \{F^{-1/2}(\gamma)S_N(\gamma)\}. \tag{2.8}$$

It then follows from (i) and (ii) in Lemma 2.1 that the first term in the R.H.S. of (2.8) converges to identity matrix and thus get the theorem. □

Theorem 2.2. *Let t_N be $F^{1/2}(\gamma)$-consistent estimator of γ, i.e., $F^{1/2}(\gamma)(t_N-\gamma)$ is bounded in probability. Then the one-step (maximum likelihood) estimator $\hat{\gamma}_N$ defined as*

$$\hat{\gamma}_N = t_N + F_N^{-1}(t_N)S_N(t_N)$$

is asymptotically normal, viz.,

$$F^{1/2}(\gamma)(\hat{\gamma}_N - \gamma) \xrightarrow{d} N(0, I_k).$$

Proof. Write

$$\begin{aligned} F^{1/2}(\gamma)(\hat{\gamma}_N - \gamma) &= F^{1/2}(\gamma)\left[t_N + F_N^{-1}(t_N)S_N(t_N) - \gamma\right] \\ &= F^{1/2}(\gamma)\left[t_N + F_N^{-1}(t_N)\{S_N(\gamma) - F_N(\gamma^*)(t_N - \gamma)\} - \gamma\right] \\ &= F^{1/2}(\gamma)F_N^{-1}(t_N)F^{1/2}(\gamma)F^{-1/2}(\gamma)S_N(\gamma) + A, \end{aligned} \tag{2.9}$$

where $A = F^{1/2}(\gamma)\left[I_k - F_N^{-1}(t_N)F_N(\gamma^*)\right](t_N - \gamma)$.

Using the uniform smoothness of the information matrix addressed (ii) in Lemma 2.1, the term inside square bracket in A converges to zero. Therefore $F^{1/2}(\gamma)$-consistency of t_N implies $A = o_p(1)$, which in turn yields the theorem via (2.9). □

Note that above theorem justifies the iterative scoring algorithm (see, for instance, Engle, 1982). Each iteration produces $\hat{\gamma}_{i+1}$ based on $\hat{\gamma}_i$ according to

$$\hat{\gamma}_{i+1} = \hat{\gamma}_i + F_N^{-1}(\hat{\gamma}_i)S_N(\hat{\gamma}_i). \tag{2.10}$$

3. Simulation Example

In this section, we suggest procedures of modeling for TF-ARCH models. The main model building block certainly consists of identification, estimation and diagnostics. With slight modifications for the ARCH components we will recommend the stages of analysis for traditional TFM suggested, for instance, by Wei (1990, ch.13) and Box and Jenkins (1976, ch.10). Thus omitting some details only outlines for modeling strategies will be provided. More detailed explanations will be given via simulated data.

3.1. Identification

Identification of models refers to specifications of the transfer function, noise and ARCH processes.

Step 1. Prewhiten the input series,

$$\eta_t = \frac{\phi_x(B)}{\theta_x(B)} x_t.$$

Step 2. Transform the output series using the same filter and generate the filtered output series,

$$\zeta_t = \frac{\phi_x(B)}{\theta_x(B)} y_t.$$

Step 3. Calculate the sample cross correlation function (CCF) between η_t and ζ_t, and identify the transfer function, *i.e.*, figure out the orders (b, r, s) using CCF. Thus a preliminary estimate of the transfer function $v(B)$ is obtained by

$$\hat{v}(B) = \frac{\hat{\omega}_s(B)}{\hat{\delta}_r(B)} B^b$$

and the noise n_t can be "estimated" by

$$\hat{n}_t = y_t - \hat{v}(B) x_t.$$

Step 4. Identify an appropriate ARMA model for noise series by examining sample ACF and PACF from $\hat{n}_t$, which in turn yields the following preliminary model

$$y_t = \frac{\hat{\omega}_s(B)}{\hat{\delta}_r(B)} x_{t-b} + \frac{\hat{\theta}_q(B)}{\hat{\phi}_p(B)} \epsilon_t. \tag{3.1}$$

Recall the equation (2.3) to calculate $\hat{\epsilon}_t$ using the preliminary estimates $\hat{\omega}_j$, $\hat{\delta}_j$, $\hat{\phi}_j$ and $\hat{\theta}_j$ in equation (3.1), via,

$$\begin{aligned} \hat{\epsilon}_t &= y_t + \hat{d}_1 y_{t-1} + \cdots + \hat{d}_{p+r} y_{t-p-r} + \hat{c}_0 x_{t-b} + \hat{c}_1 x_{t-b-1} \\ &\quad + \cdots + \hat{c}_{p+s} x_{t-b-p-s} + \hat{b}_1 \hat{\epsilon}_{t-1} + \hat{b}_2 \hat{\epsilon}_{t-2} + \cdots + \hat{b}_{r+q} \hat{\epsilon}_{t-r-q}. \end{aligned}$$

Step 5. Model an appropriate ARCH model by examining sample ACF and PACF of $\hat{\epsilon}_t^2$ using LM-statistic and portmanteau-statistic. Consequently, we reach at the following transfer function model with ARCH noise process

$$\begin{aligned} y_t &= \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\phi_p(B)} \epsilon_t, \qquad (3.2) \\ h_t &= \alpha_0 + \sum_{k=1}^{m} \alpha_k \epsilon_{t-k}^2. \end{aligned}$$

3.2. Estimation of parameters

After identifying the TF-ARCH model in (3.2), the parameters $\beta = (\omega_0, \omega_1, \cdots, \omega_s; \delta_1, \cdots, \delta_r; \phi_1, \cdots, \phi_p; \theta_1, \cdots, \theta_q)'$ and $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_m)'$ can be estimated as discussed in Theorem 2.1 and (2.10).

3.3. Diagnostic checking

In terms of model inadequacy, three situations may arise.

Case 1. The transfer function $v(B)$ is incorrect.

Case 2. The transfer function $v(B)$ is correct and only the noise model is incorrect. These two situations can be diagnosed and then may be resolved via standard methods as in Wei (1990).

Case 3. The transfer function $v(B)$ and the noise model are correct, and only the ARCH process is not correct. In this case the ACF based on squared residuals $\{\hat{\epsilon}_t^2\}$ does not vanish for some time lag $k \geq 1$. Accordingly, the portmanteau test proposed by Li and Mak (1994) can be adopted :

$$Q = N \sum_{j=m+1}^{K} \hat{\rho}_{\hat{\epsilon}^2}^2(j), \tag{3.3}$$

where

$$\hat{\rho}_{\hat{\epsilon}^2}(k) = \frac{\sum_{t=k+1}^{N} \left(\hat{\epsilon}_t^2/\hat{h}_t - 1\right)\left(\hat{\epsilon}_{t-k}^2/\hat{h}_{t-k} - 1\right)}{\sum_{t=1}^{N} \left(\hat{\epsilon}_t^2/\hat{h}_t - 1\right)^2}$$

is the squared residual autocorrelation of lag k. Under the null hypothesis of model adequacy, the Q statistic approximately follows χ^2 distribution with $K-m$ degrees of freedom where m specifies the order of ARCH model.

A simulation experiment is now conducted to illustrate the model building steps discussed above. Random variates in simulation are generated by using IMSL (International Mathematical and Statistical Libraries).

We generate a set of time series (x_t, y_t) with length 230 from the following TF-ARCH model :

$$y_t = \frac{2.5}{1 - 0.5B}\, x_t + \epsilon_t,$$

$$(1 + 0.7B)x_t \sim N(0, 50),$$

where ϵ_t has a zero mean and conditional variance h_t given by

$$h_t = 100 + 0.5\epsilon_{t-1}^2$$

so that $\sigma_\epsilon^2 = 200$.

We discard the first 30 values to minimize the effect of the starting values and use the rest 200 values for model building.

1. Prewhitening of the input series. The first ten lags of the sample ACF and PACF of input series x_t indicate AR(1) model. The fitted model is found as

$$\eta_t = (1 + 0.662B)x_t,$$

where η_t denotes the white noise with a zero mean and $\hat{\sigma}_\eta^2 = 48.29$. Diagnostic checking does not indicate any inadequacies in the model. Thus, the prewhitening filter for the input series x_t is used to filter the output series y_t.

2. Identification of the transfer function. The pattern of the sample CCF between η_t and the filtered output series $\zeta_t = (1 + 0.662B)y_t$ with $\hat{\sigma}_\zeta^2 = 659.99$ suggests a transfer function with $b = 0$, $r = 1$ and $s = 0$. That is,

$$v(B)\; x_t = \frac{\omega_0}{(1 - \delta_1 B)}\; x_t.$$

3. Identification of the noise model. The estimated noise series is then obtained by

$$\hat{n}_t = y_t - \frac{1.978}{(1 - 0.143B)}\; x_t.$$

The pattern of the sample ACF and PACF based on the estimated noise series $\hat{n}_t$ suggests a white noise, *i.e.*,

$$n_t = \epsilon_t.$$

4. Identification of the ARCH process. The Lagrange Multiplier (LM) test for heteroscedasticity is summarized in Table 3.1. Note that the differences between adjacent LM values are not significant and therefore ARCH(1) is selected.

TABLE 3.1 *LM test statistic for ARCH*

Lag	1	2	3	4	5	6	7	8	9	10
LM	68.78	68.81	69.73	71.01	71.06	71.85	72.43	74.28	74.28	74.87
p-value	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

5. Estimation of the transfer function model with ARCH noise process. Combining the above results, we obtain the tentative transfer function model with ARCH noise process

$$y_t = \frac{\omega_0}{(1-\delta_1 B)}\, x_t + \epsilon_t, \tag{3.4}$$

$$\epsilon_t|\Psi_{t-1} \sim N(0, h_t), \quad h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2.$$

Relevant estimates of the parameters and standard errors (S.E.) are listed below.

Parameter	ω_0	δ_1	α_0	α_1	σ_ϵ^2
True	2.50	.50	100.00	.50	200.00
Estimate	2.37	.44	112.35	.53	239.54
S.E.	.109	.046	18.461	.137	

6. Diagnostic checking. The CCF between prewhitened input series η_t and residual $\hat{\epsilon}_t$ implies the adequacy of the transfer function. The residual autocorrelation of $\hat{\epsilon}_t$ is checked to pass the noise model adequacy. Also based on ten lags ($k = 10, m = 1$) the test Q in (3.3) for ARCH modeling is found to be 15.327 which indicates that the ARCH(1) in (3.4) is correct.

4. Empirical Study

In this section, the proposed model is applied to a Korean monthly stock price index (stock end of period) and a monthly yields on corporate bonds. The data set is obtained from National Statistical Office and the time period is from February 1992 to September 2000. We are interested in modeling the dynamic relationship between yields on corporate bonds and a stock price index. In order to stabilize the time series logarithmic transformation and differences are taken. Let $x_t = \ln(X_t/X_{t-1})$ and $y_t = \ln(Y_t/Y_{t-1})$ where X_t is for corporate bonds and Y_t is a stock price index.

Applying the Box-Jenkins methodology, we obtain the following classical transfer function model :

$$y_t = \frac{\omega_0}{(1-\delta_1 B - \delta_2 B^2)} x_t + \frac{1}{(1-\phi_1 B)} \epsilon_t. \tag{4.1}$$

Estimates of the parameters are listed below.

Parameter	ω_0	δ_1	δ_2	ϕ_1
Estimate	−.411	−.914	−.551	.331
S.E.	.095	.104	.122	.097

The LM test for ARCH is seen to be significant at 5% significance level and hence we are led to consider the transfer function model with ARCH noise process.

A careful inspection of the steps discussed in Section 3 brings out the following tentative model with estimates.

$$y_t = \frac{\omega_0}{1 - \delta_1 B - \delta_2 B^2}\, x_t + \frac{1}{1 - \phi_1 B}\, \epsilon_t, \tag{4.2}$$

$$\epsilon_t | \Psi_{t-1} \sim N(0, h_t), \quad h_t = \alpha_0 + \alpha_2 \epsilon_{t-2}^2.$$

Parameter	ω_0	δ_1	δ_2	ϕ_1	α_0	α_2
Estimate	−.750	−.629	−.281	.262	.004	.510
S.E.	.107	.112	.110	.039	.001	.226

Also, diagnostic checking is safely passed to suggest that the proposed model in equation (4.2) be adequate. For comparision purpose, recall the classical model (TF) in (4.1) and the proposed model (TF-ARCH) in (4.2). RMSE is computed using the last 30 output series $\{y_t\}$ and the corresponding one-step-ahead forecast. Estimates and RMSE's for each model are summarized below, indicating that TF-ARCH beats TF in the sense of the smaller RMSE.

Parameter	ω_0	δ_1	δ_2	ϕ_1	α_0	α_2	RMSE
TF-ARCH	−.750	−.629	−.281	.262	.004	.510	.123
TF	−.411	−.914	−.551	.331			.129

Acknowledgments

We thank the two referees for constructive comments on an earlier version. This research was supported (in part) by KOSEF through Statistical Research Center for Complex Systems at Seoul National University. S. Y. Hwang's work was supported by a grant from KOSEF (1999-1-104-002-3).

REFERENCES

An, H. Z., Chen, M. and Huang, F. C. (1997). "The geometric ergodicity and existence of moments for a class of non-linear time series model", *Statistics & Probability Letters*, **31**, 213–224.

Billingsley, P. (1961). "The Lindeberg-Levy theorem for martingale", *Proceeding of American Mathematical Society*, **12**, 788–792.

Box, G. E. P. and Jenkins, J. M. (1976). *Time Series Analysis : Forecasting and Control*, 2nd ed., Holden-Day, San Francisco, CA.

Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation", *Econometrica*, **50**, 987–1007.

Granger, C. W. J. and Teraäsvirta, T. (1993). *Modeling Nonlinear Economic Relationships*, Oxford University Press, London.

Li, W. K. and Mak, T. K. (1994). "On the squared residual autocorrelation in non-linear time series with conditional heteroscedasticity", *Journal of Time Series Analysis*, **15**, 627–636.

Wei, W. W. S. (1990). *Time Series Analysis : Univariate and Multivariate Methods*, Addison Wesley.

Weiss, A. A. (1984). "ARMA models with ARCH errors", *Journal of Time Series Analysis*, **5**, 129–143.