



## Nonlinear Time-Series Analysis of Stock Volatilities

C. Q. Cao; R. S. Tsay

*Journal of Applied Econometrics*, Vol. 7, Supplement: Special Issue on Nonlinear Dynamics and Econometrics. (Dec., 1992), pp. S165-S185.

Stable URL:

<http://links.jstor.org/sici?sici=0883-7252%28199212%297%3CS165%3ANTAOSV%3E2.0.CO%3B2-6>

*Journal of Applied Econometrics* is currently published by John Wiley & Sons.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/jwiley.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

# NONLINEAR TIME-SERIES ANALYSIS OF STOCK VOLATILITIES

C. Q. CAO AND R. S. TSAY

*Graduate School of Business, University of Chicago, Chicago, IL 60637, USA*

## SUMMARY

The absolute value of the mean-corrected excess return is used in this paper to measure the volatility of stock returns. We apply various nonlinearity tests available in the literature to show that such volatility series are strongly nonlinear. We then explore the use of threshold autoregressive (TAR) models in describing monthly volatility series. The models built suggest that the volatility series exhibit significant lower-order serial correlations when the volatility is large, indicating certain volatility clustering in stock returns. Out-of-sample forecasts are used to compare the TAR models with linear ARMA models and nonlinear GARCH and EGARCH models. Based on mean squared error and average absolute deviation, the comparisons show that (a) the TAR models consistently outperform the linear ARMA models in multi-step ahead forecasts for large stocks, (b) the TAR models provide better forecasts than the GARCH and EGARCH models also for the volatilities of large stock returns, and (c) the EGARCH model gives the best long-horizon volatility forecasts for small stock returns.

## 1. INTRODUCTION

Recent years have seen an increasing interest in modelling the volatilities of stock returns. From a financial economist's point of view, understanding and modelling stock volatility is important. First, investment decisions, as characterized by asset pricing theories, depend heavily on the assessment of future returns and risks of various assets. The estimated volatility (or standard deviation) of a security return is widely used as a simple risk measure in many asset pricing models. Secondly, the expected volatility of a security return plays an important role in the option pricing theory. For example, the standard deviation of a return series enters directly into the well-known Black–Scholes model. Finally, an adequate specification of return volatilities may shed some light on the generating process of the returns.

In the empirical finance literature, many linear time-series models are used to describe the volatilities of stock returns. For example, Poterba and Summers (1986) specified a stationary AR(1) process for the volatility of monthly returns on the Standard & Poor's (S&P) composite index. French, Schwert, and Stambaugh (1987) used a nonstationary ARIMA(0,1,3) model to describe the logarithm of the volatility of S&P monthly returns and found that the model works reasonably well. Schwert (1990) and Schwert and Seguin (1990) used a linear autoregressive model of order 12, i.e. AR(12), as an approximation to monthly volatility process. The use of such linear models is not surprising, because (a) they provide good first-order approximations to many processes; (b) statistical theory is well developed for linear Gaussian models; and (c) computational tools for linear models, such as model specification, estimation and diagnostics, are readily available.

There are, however, certain features of a volatility series that cannot be described by linear

time series models. For instance, empirical evidence shows that stock returns tend to exhibit clusters of outliers, implying that the volatility series evolves over time in a nonlinear fashion. Such limitations of linear models have motivated many authors to consider nonlinear alternatives. See, for example, Gallant, Hsieh, and Tauchen (1991), Hsieh (1989), LeBaron (1992), Scheinkman and LeBaron (1989), Pagan and Schwert (1990), and Taylor (1986).

The most commonly used nonlinear time-series models in the finance literature are the autoregressive conditional heteroscedastic (ARCH) model of Engle (1982), the generalized ARCH (GARCH) model of Bollerslev (1986) and the exponential GARCH (EGARCH) model of Nelson (1991). These ARCH-family models have been found to be useful in capturing certain nonlinear features of financial time-series. In particular they are capable of producing heavy-tailed distributions and clusters of outliers. See Bollerslev, Chou, and Kroner (1992) for a nice review of these models and their applications. The success of the ARCH-type of models is understandable if we consider their relation to the random-coefficient autoregressive (RCA) models that, to a certain degree, are adaptive to the changing economical environment. As pointed out by Tsay (1987), a RCA model can produce, under some mild conditions, the first two conditional moments of an ARCH model. This is particularly so under the normality assumption.

The ARCH-family is only one of many time-series models that can capture the nonlinear characteristics of stock volatilities commonly observed in practice. The purpose of this paper is, therefore, to explore some other nonlinear models and to compare their forecasting accuracy with those of ARCH-family models. More specifically we have two goals. First, we consider directly a volatility process for a monthly stock return and test the nonlinearity of such a volatility process. Secondly, in the presence of nonlinearity, we address the issue of modelling the volatility process via a piecewise linear model and then compare the forecasts of such a model with those of GARCH and EGARCH models. The piecewise linear model employed is the threshold autoregressive model, (see Tong, 1990). The use of such models is motivated by their ability to capture the asymmetric pattern of a time-series and by their local linear property. We adopt the common belief that all statistical models are wrong, but a well-chosen one can provide accurate local approximations.

This paper is organized as follows. In section 2 we briefly introduce the three nonlinear models and the nonlinearity tests used in the paper. A modelling procedure for the TAR models is also given. Section 3 describes the data employed. For a monthly stock process the absolute value of its excess return minus the mean, pre-multiplied by  $\sqrt{(\pi/2)}$ , is used to obtain a volatility series. This definition is based on the fact that  $\sqrt{(\pi/2)} |R_t - \mu_t|$  is an unbiased estimator for the conditional standard deviation (volatility) of the monthly excess returns if the excess return  $R_t$  is normally distributed. We then apply in section 4 the nonlinearity tests of section 2 to the constructed volatility series. In section 5 we apply the TAR models to three volatility series of monthly stock returns. Finally, in section 6 we compare the TAR models with linear ARMA models and nonlinear GARCH and EGARCH models via out-of-sample forecasts.

## 2. NONLINEAR MODELS AND NONLINEARITY TESTS

### 2.1. Some Nonlinear Dynamic Models

Many nonlinear time-series models are available in the literature. Tong (1990) gives a nice summary of those models. Here we briefly discuss the three models employed in this paper.

A time series  $Y_t$  follows a GARCH( $p, q$ ) model if it satisfies:

$$Y_t = X_t b + \varepsilon_t \tag{1}$$

where  $X_t$  is a vector of explanatory variables including possibly lagged values of  $Y_t$ ,  $b$  is a constant vector and the error term  $\varepsilon_t$  is given by

$$\varepsilon_t = \sigma_t z_t, \quad z_t \sim \text{i.i.d. } N(0, 1), \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \tag{2}$$

where  $p$  and  $q$  are nonnegative integers,  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$  for  $i > 0$  and  $\beta_j \geq 0$  for all  $j$ . If  $(\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j) < 1$ , the innovational process  $\{\varepsilon_t\}$  of (1) is weakly stationary. For the most commonly used GARCH(1,1) model, the condition  $3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$  implies the existence of the fourth moment of  $\varepsilon_t$ , which is greater than that of a normal random variable. Consequently, the GARCH model is capable of producing outliers. The non-negative constraints on the coefficients  $\alpha_i$  and  $\beta_j$  of  $\sigma_t^2$  for  $i > 0$  and  $j \geq 0$  in (2) can be relaxed (see Tsay, 1987 and Nelson and Cao, 1992 for details). For further properties of the GARCH models see Engle (1982) and Bollerslev (1986).

The key feature of the GARCH model is that, by equation (2), the conditional variance of  $\varepsilon_t$  is predictable. Replacing  $\varepsilon_{t-i}^2$  of the variance equation in (2) by  $z_{t-i}^2 \sigma_{t-i}^2$ , we see that a large conditional variance tends to be followed by another large conditional variance. This feature implies that the GARCH model is capable of producing clusters of outliers. In financial time-series, clusters of outliers result in clusters of high volatilities.

A limitation of the GARCH model is that the conditional variance  $\sigma_t^2$  responds to positive and negative residuals  $\varepsilon_{t-i}$  in the same manner. On the contrary, empirical evidence in financial time-series shows that there is a negative correlation between the current returns and future returns volatility. To overcome such a limitation, Nelson (1991) proposed an exponential-GARCH model, which enables the conditional variance of the innovation to respond asymmetrically to positive and negative residuals. A time-series  $Y_t$  follows an EGARCH( $p, q$ ) model if it satisfies

$$Y_t = X_t b + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad z_t \sim \text{i.i.d. } N(0, 1), \tag{3}$$

$$\ln \sigma_t^2 = \alpha + \frac{1 + \Phi_1 B + \Phi_2 B^2 + \dots + \Phi_q B^q}{1 - \Delta_1 B - \Delta_2 B^2 - \dots - \Delta_p B^p} g(z_{t-1})$$

where  $g(z_t) = \theta z_t + \gamma [|z_t| - E|z_t|]$ ,  $B$  is the lag operator,  $\Phi_i, \Delta_j, \theta$ , and  $\gamma$  are parameters such that the two polynomials  $(1 + \sum_{i=1}^q \Phi_i B^i)$  and  $(1 - \sum_{j=1}^p \Delta_j B^j)$  have no common factors. Consider the  $g(z_t)$  function above. If  $z_t$  is positive then  $g(z_t)$  is a linear function of  $z_t$  with slope  $(\theta + \gamma)$ . If  $z_t$  is negative then the slope changes to  $(\theta - \gamma)$ . Consequently, the conditional variance  $\sigma_t^2$  responds asymmetrically to the sign of the innovation  $z_{t-i}$ .

A time-series  $Y_t$  is a TAR process if it satisfies the model

$$Y_t = \phi_0^{(i)} + \phi_1^{(i)} Y_{t-1} + \dots + \phi_p^{(i)} Y_{t-p} + \varepsilon_t^{(i)} \quad \text{if } Y_{t-d} \in L_i, \quad i = 1, 2, \dots, k \tag{4}$$

where the  $L_i$  form a non-overlapping partition of the real line i.e.  $\cup_{i=1}^k L_i = R$  and  $L_i \cap L_j = \emptyset$  if  $i \neq j$ ,  $k$  is the number of threshold regimes,  $d$  is the delay parameter (or threshold lag),  $p$  is the AR order and  $\{\varepsilon_t^{(i)}\}$  is a sequence of i.i.d. normal random variables with mean zero and variance  $\sigma_i^2$  such that  $\{\varepsilon_t^{(i)}\}$  and  $\{\varepsilon_t^{(j)}\}$  are independent if  $i \neq j$ . The TAR model in (4) is a piecewise linear model in the space of  $Y_{t-d}$  and is capable of providing accurate 'local approximations' in this space. It is, however, not a piecewise linear model in time. The key features of TAR models include time-irreversibility, asymmetric limit cycle and jump phenomenon. See Tong (1990) for other properties of TAR Models.

Alternatively, one can interpret a TAR model as a switching linear regression model for which many references are available in the literature. See, for example, Quandt and Ramsey (1978). The only difference is that for the TAR model in (4), the switching mechanism is controlled by the threshold variable  $Y_{t-d}$ , not by the time index  $t$ .

The idea of using a threshold variable  $Y_{t-d}$  to govern the dynamic pattern of a stock volatility series  $Y_t$  appears to be natural. As mentioned before, empirical evidence suggests that large returns tend to be followed by large returns and small returns by small returns. This evidence shows, among other possibilities, that the dynamic pattern of stock volatilities evolves with the past values of the returns. Such a state-dependent evolution is in good agreement with the concept of TAR models.

**2.2. A Modelling Procedure for TAR Models**

The modelling of GARCH and EGARCH processes has been discussed in the econometrics literature. For TAR models we shall use the modelling procedure of Tsay (1989, 1991). The main idea of the procedure is to transform a TAR model into a regular change-point problem in linear regression analysis. This is achieved by using the concepts of ‘arranged autoregression’ and ‘local estimation’. Here local estimation means that we perform a sequential estimation with a fixed-length window that controls the number of observations used. However, instead of using the time index, arranged autoregression uses the magnitude of the threshold variable  $Y_{t-d}$  to control the data flow of the window. By so doing we effectively transform the threshold problem into a switching regression problem for which statistics can be derived to test for model changes and to explore the dynamic structure of the process. For illustration, Part (a) of the Table I shows the usual set-up of fitting an AR(2) model to 10 consecutive observations of the TAR(2) model

$$Y_t = \begin{cases} 1.3Y_{t-1} - 0.4Y_{t-2} + \varepsilon_t^{(1)} & \text{if } Y_{t-2} < 0.0 \\ 0.3Y_{t-1} + 0.4Y_{t-2} + \varepsilon_t^{(2)} & \text{otherwise,} \end{cases}$$

where  $\varepsilon_t^{(i)}$  are independent sequences of i.i.d.  $N(0, 1)$  random variates,  $L_1 = (-\infty, 0)$  and  $L_2 = [0, \infty)$ . Obviously, such a set-up cannot yield consistent parameter estimates. On the other hand an arranged autoregression rearranges the set-up into that of Part (b) based on the magnitude of the threshold variable  $Y_{t-2}$ . It is then clear that in this particular instance the first four rows belong in the first regime of the TAR(2) model, whereas the last four rows belong in the second regime. A local estimation can thus provide consistent estimates of the

Table I

(a) Ordinary autoregression					(b) Arranged autoregression				
Time	$Y_t$	$Y_{t-1}$	$Y_{t-2}$	regime	Time	$Y_t$	$Y_{t-1}$	$Y_{t-2}$	regime
3	-0.41	1.21	1.31	$L_2$	8	0.12	-1.85	-3.08	$L_1$
4	0.21	-0.41	1.21	$L_2$	9	0.58	0.12	-1.85	$L_1$
5	-1.12	0.21	-0.41	$L_1$	7	-1.85	-3.08	-1.12	$L_1$
6	-3.08	-1.12	0.21	$L_2$	5	-1.12	0.21	-0.41	$L_1$
7	-1.85	-3.08	-1.12	$L_1$	10	1.28	0.58	0.12	$L_2$
8	0.12	-1.85	-3.08	$L_1$	6	-3.08	-1.12	0.21	$L_2$
9	0.58	0.12	-1.85	$L_1$	4	0.21	-0.41	1.21	$L_2$
10	1.28	0.58	0.12	$L_2$	3	-0.41	1.21	1.31	$L_2$

coefficients. In particular, the local estimates should be stable before the first change point enters the estimation window. In practice the local estimation can be done efficiently by a recursive least-squares algorithm or Kalman filter; the latter is particularly useful when there are missing values in the data.

We then consider the scatterplots of the local estimates of AR coefficients versus the threshold variable  $Y_{t-d}$ . Ideally, the scatterplots should look like a step function with jumps indicating the values of the threshold variable at which the regime changes. In practice, since the windows used overlap sequentially, the plots tend to show certain smooth transition from one regime to another. This is not a serious problem in model specification, because the main objective of arranged autoregression is to provide information on the possible partitions of the space of the threshold variable. Once the threshold values are located we can partition the space into several regimes and estimate an AR model with an appropriate order in each regime. Furthermore, we can adjust the threshold values to refine the estimated TAR model according to some information criterion, e.g. the AIC criterion.

In summary, the TAR modelling procedure used consists of the following steps:

1. Select a tentative AR order  $p$  and a set of possible threshold variables. In this paper the possible threshold lags entertained are  $\{1, 2, \dots, 12\}$ .
2. For each threshold variable considered, perform nonlinearity tests, especially a threshold nonlinear test.
3. Select the threshold variable  $Y_{t-d}$  based on the test results of step 2.
4. Perform an arranged autoregression to locate the possible threshold values.
5. Estimate the specified TAR model by conditional least-squares.
6. Check the estimated TAR model and refine it if necessary.

See Tsay (1989) for details and examples of such a procedure. It is worth mentioning that the computation of the above procedure is extremely simple. It merely requires a sorting routine and a recursive least-squares algorithm.

### 2.3. Some Nonlinearity Tests

To detect the nonlinearity of a stock volatility in section 4 we apply some nonlinearity tests available in the literature. The tests used include the  $F$ -test of Tsay (1986), the augmented  $F$ -test of Luukkonen, Saikkonen, and Teräsvirta (1988), the threshold test of Tsay (1989) and a general nonlinearity test of Tsay (1991). For model checking we use the BDS test of Brock, Dechert, and Scheinkman (1987) to test the i.i.d. assumption of the residuals. For completeness we briefly describe these tests in this subsection.

The  $F$ -test is a Lagrange multiplier test for detecting mainly second-order nonlinearity in an AR model. For a time series  $Y_t$  with observations  $(Y_1, Y_2, \dots, Y_T)$ , the test consists of three steps: (1) For a prespecified AR order  $p$ , regress  $Y_t$  on the vector  $W_t = (1, Y_{t-1}, \dots, Y_{t-p})$ , and compute the residual  $\hat{a}_t$ . (2) Regress the vector of squares and cross-products of  $Z_{t-i}$ , i.e.,  $Z_t = (Y_{t-1}^2, Y_{t-1}Y_{t-2}, \dots, Y_{t-1}Y_{t-p}, Y_{t-2}^2, Y_{t-2}Y_{t-3}, \dots, Y_{t-p}^2)$ , on  $W_t$  and save the residual vector  $\hat{V}_t$ . (3) Regress  $\hat{a}_t$  on  $\hat{V}_t$  and calculate the residual  $\hat{\varepsilon}_t$ . Form the  $F$ -statistic

$$\hat{F} = \frac{(\sum \hat{V}_t \hat{a}_t)(\sum \hat{V}_t \hat{V}_t)^{-1}(\sum \hat{V}_t \hat{a}_t) / m}{\sum \hat{\varepsilon}_t^2 / (T - p - m - 1)}$$

where  $m = \frac{1}{2}p(p+1)$  and the summation is summing over  $t$  from  $p+1$  to  $T$ . It can be shown that, under the null hypothesis that the  $\{Y_t\}$  is a stationary AR( $p$ ) process,  $\hat{F}$  is asymptotically an  $F$  distribution with degrees of freedom  $\frac{1}{2}p(p+1)$  and  $T - \frac{1}{2}p(p+3) - 1$ .

Luukkonen, Saikkonen, and Teräsvirta (1988) modified the above  $F$ -test by augmenting to the vector  $Z_t$  the cubic terms  $Y_{t-1}^3, Y_{t-2}^3, \dots, Y_{t-p}^3$ . These authors show via simulation that such an augmented- $F$ -test is more powerful than the original  $F$ -test in detecting certain types of nonlinear models, e.g. the logistic model.

To detect the threshold nonlinearity, Tsay (1989) proposed another  $F$ -test based on the arranged autoregression. The test consists of two steps. First, for a prespecified AR order  $p$  and a threshold lag  $d$ , fit recursively an arranged autoregression of order  $p$  to the series  $Y_t$ . Assuming that the recursion begins with the first  $b$  observations, calculate the standardized predictive residual  $\hat{e}_t$  for  $t > b$ . Secondly, regress the predictive residual  $\hat{e}_t$  on  $(1, Y_{t-1}, Y_{t-2}, \dots, Y_{t-p})$ , save the corresponding residuals  $\hat{e}_t$ , and form the  $F$ -statistic

$$\hat{F}(p, d) = \frac{(\sum \hat{e}_t^2 - \sum \hat{e}_t^2)/(p+1)}{\sum \hat{e}_t^2/(T-d-b-p-h)}$$

where  $h = \max(1, p+1-d)$  and the summation is summing over  $t$  from  $b+1$  to  $T-d-h+1$ . Under the null hypothesis that  $Y_t$  is an  $AR(p)$  process, the  $\hat{F}(p, d)$  statistic is asymptotically an  $F$  distribution with degrees of freedom  $p+1$  and  $T-d-b-p-h$ .

The general nonlinearity test of Tsay (1991) combines the ideas of the threshold test and  $F$ -tests. It is a test designed for detecting various types of nonlinearity. Similar to the threshold test, one computes first the predictive residual  $\hat{e}_t$  of an arranged autoregression of order  $p$  and threshold lag  $d$ . Next, one regresses  $\hat{e}_t$  on the regressors  $(1, Y_{t-1}, \dots, Y_{t-p}, (Y_{t-1}\hat{e}_{t-1}, \dots, Y_{t-p}\hat{e}_{t-p}), (\hat{e}_{t-1}\hat{e}_{t-2}, \dots, \hat{e}_{t-p}\hat{e}_{t-p-1}), (Y_{t-1} \exp(-Y_{t-1}^2/r), \Phi(Z_{t-d}), Y_{t-1}\Phi(Z_{t-d}))$ , where  $r = \max_t |Y_{t-1}|$  and  $Z_{t-d} = (\bar{Y}_{t-d} - \bar{Y}_d)/S_d$  with  $\bar{Y}_d$  and  $S_d$  the sample mean and standard deviation of  $Y_{t-d}$ ,  $\Phi(\cdot)$  is the cumulative distribution function of the standard normal random variable. The usual  $F$ -statistic of this second regression is the test statistic, which under the null hypothesis of an  $AR(p)$  model for  $Y_t$  follows asymptotically an  $F$ -distribution with degrees of freedom  $3(p+1)$  and  $T-b-3(p+1)$ , where  $b$  is defined as before.

The BDS test of Brock, Dechert, and Scheinkman (1987) is based on the idea of correlation dimension. It is a useful statistic for detecting i.i.d. random samples. Define an  $m$ -dimensional vector  $Y_t^m = (Y_t, Y_{t+1}, \dots, Y_{t+m-1})$ . Then the sample correlation dimension of this  $m$ -vector is:

$$C_m(\varepsilon) = \frac{2}{(T-m-1)(T-m)} \sum_{t < s} I_\varepsilon(Y_t^m, Y_s^m)$$

where

$$I_\varepsilon(Y_t^m, Y_s^m) = \begin{cases} 1 & \text{if } \|Y_t^m - Y_s^m\| < \varepsilon \\ 0 & \text{if } \|Y_t^m - Y_s^m\| \geq \varepsilon, \end{cases}$$

with  $\|\cdot\|$  the sup-norm. This correlation dimension measures the spatial correlation of the data. More precisely, the correlation dimension estimates the probability that any two  $m$ -dimensional vectors, say  $Y_t^m$  and  $Y_s^m$ , are within the distance  $\varepsilon$ . Under the null hypothesis that the  $\{Y_t\}$  are i.i.d. with a nondegenerated density function  $f(\cdot)$ , the BDS-U statistic

$$U_m^\varepsilon = \frac{\sqrt{T}[C_m(\varepsilon) - (C_1(\varepsilon))^m]}{\sigma_m(\varepsilon)}$$

is asymptotically  $N(0, 1)$ , where

$$\sigma_m^2(\varepsilon) = 4 \left[ K^m + 2 \sum_{j=1}^{m-1} K^{m-j} C^{2j} + (m-1)^2 C^{2m} - m^2 K C^{2m-2} \right]$$

$$C = \int (f(y+\varepsilon) - f(y-\varepsilon)) df(y), \quad K = \int (f(y+\varepsilon) - f(y-\varepsilon))^2 df(y).$$

In our implementation of the BDS-U statistic we use  $m = 2, \dots, 5$  and  $\varepsilon = 1, 1.5$  and 2 times the standard deviation of  $Y_t$ . These choices are based on suggestions of Hsieh and LeBaron (1988).

### 3. DATA DESCRIPTION

#### 3.1. Stock Return Series

The stock returns used in this paper are the monthly value-weighted, equal-weighted and S&P composite portfolio excess returns for NYSE & AMEX for the period from January 1928 to December 1989. The data are from the Center for Research in Security Prices (CRSP) at the University of Chicago and have 744 observations. We also use daily returns of these three stocks from 1985 to 1989 to construct benchmark series of monthly stock volatilities for model comparisons in section 6. A return series is defined as the logarithm of the price relative, which is the continuously compounded rate of return. Monthly excess return is simply the above stock return less the 1-month Treasury-bill return.

#### 3.2. Volatility Series of Stock Returns

Following Schwert (1990) and Schwert and Seguin (1990), we use in this paper the absolute value of monthly excess return minus its mean, pre-multiplied by  $\sqrt{(\pi/2)}$  to estimate the volatility of monthly stock returns. In other words the volatility series used is  $\sigma_t = \sqrt{(\pi/2)} |R_t - \mu_t|$  for the monthly excess return  $R_t$ . For value-weighted and equal-weighted portfolio excess returns  $R_t - \mu_t$  is taken as the residual of the AR(1) process for the excess returns. This is motivated by the fact that the AR(1) coefficients in the mean equation of the simple GARCH(1,1) and EGARCH(1,0) models are significant at the 5 per cent level for both excess return series. For S&P excess return, since the AR(1) specification in the mean equation of the GARCH(1,1) model is not statistically significant, the mean series  $\mu_t$  is treated as a constant and estimated by the sample mean of  $R_t$ .

#### 3.3. Power Transformation

Since the distribution of the volatility measure  $\sigma_t$  is skewed, we consider its Box-Cox transformation

$$Y_t = \frac{\sigma_t^\lambda - 1}{\lambda}, \quad (5)$$

to increase the efficiency of parameter estimation and to aid the model interpretation. It turns out that the distribution of  $Y_t$  is fairly symmetric when  $\lambda = \frac{1}{4}$ .

Alternatively, many authors in the literature modeled the logarithms of the volatility series  $\sigma_t$ , i.e. considering  $\ln(\sigma_t)$ . See, for example, Poterba and Summers (1987), French, Schwert, and Stambough (1987), Melino and Turnbull (1990), Scott (1987), Stein (1989), and Wiggins (1987). For the volatility series  $\sigma_t = \sqrt{(\pi/2)} |R_t - \mu_t|$ , the distribution of  $\ln(\sigma_t)$  is not symmetric and the normality assumption is questionable. Table II gives some summary statistics of  $\sigma_t$ ,  $\ln(\sigma_t)$  and the transformed series  $Y_t$  for the 3-monthly return series considered in the paper. From the table it is seen that the Box-Cox transformed series  $Y_t$  is much closer to normality than  $\sigma_t$  and  $\ln(\sigma_t)$  series are.

Table II. Summary statistics of volatility, logarithm of volatility and transformed volatility of S&amp;P, value-weighted and equal-weighted excess returns, 1928–1989

	S&P			VW			EW		
	$\sigma_t$	$\ln \sigma_t$	$Y_t$	$\sigma_t$	$\ln \sigma_t$	$Y_t$	$\sigma_t$	$\ln \sigma_t$	$Y_t$
Mean	0.0505	-3.5136	-2.2690	0.0495	-3.5111	-2.2719	0.0634	-3.3161	-2.1773
Variance	0.0030	1.4680	0.2108	0.0027	1.3582	0.2016	0.0050	1.5410	0.2478
Skewness	3.2692	-1.3245	0.0106	3.0303	-1.0863	-0.0006	3.0349	-1.2315	0.0657
Kurtosis	18.7543	7.0134	3.4924	16.8344	5.2033	3.3495	15.8402	6.2693	3.4188
Minimum	0.0000	-11.5129	-3.7751	0.0001	-9.6979	-3.6459	0.0000	-10.7534	-3.7280
Median	0.0371	-3.2938	-2.2444	0.0361	-3.3207	-2.2561	0.0419	-3.1723	-2.1902
Maximum	0.4476	-0.8037	-0.7281	0.4349	-0.8327	-0.7518	0.5488	-0.6001	-0.5573

#### 4. NONLINEARITY OF VOLATILITY SERIES

In this section we apply the nonlinearity tests of section 2 to the transformed volatility series  $Y_t$  of the 3-monthly stock returns considered. This exercise serves several purposes. First it explores the nonlinear nature, if any, of the transformed series. Secondly, since different nonlinearity tests have varying strength in detecting different types of nonlinearity, the exercise might shed some light on the nonlinear dynamic of the volatility series when the test results are contrasted. It is hoped that a better understanding of the stock return volatilities can result from the exercise.

Tables III (a)–(c) give the test results for the transformed volatility series  $Y_t$  of monthly

Table III(a). General nonlinearity test for the transformed volatility of monthly S&P excess returns (NYSE & AMEX), 1928–1989 (AR order:  $p = 3$ )

$d$	Ori-F ( $p$ -value): 4.2937 (0.0003)				Aug-F ( $p$ -value): 3.5281 (0.0003)			
	TAR-F	$p$ -value	$v_1$	$v_2$	NEW-F	$p$ -value	$v_1$	$v_2$
1	2.31066	0.05647	4	660	2.44416	0.00306	13	651
2	2.03273	0.08821	4	660	2.46895	0.00276	13	651
3	2.99801	0.01809	4	660	3.16262	0.00013	13	651
4	1.53873	0.18924	4	659	2.45258	0.00296	13	650
5	2.03750	0.08755	4	658	2.78141	0.00071	13	649
6	3.18851	0.01311	4	657	3.79524	0.00001	13	648
7	2.01772	0.09035	4	656	2.55017	0.00195	13	647
8	2.17950	0.06979	4	655	2.81681	0.00061	13	646
9	1.30206	0.26782	4	654	3.31052	0.00007	13	645
10	1.99718	0.09334	4	653	2.87587	0.00047	13	644
11	2.99421	0.01822	4	652	2.87265	0.00048	13	643
12	0.61106	0.65480	4	651	2.72927	0.00090	13	642

#### Notes

- (1)  $d$  is threshold lag used in the test.
- (2)  $v_1$  and  $v_2$  denote the degrees of freedom of  $F_{v_1, v_2}$ .
- (3) Ori-F refers to the  $F$ -test of Tsay (1986).
- (4) Aug-F refers to the augmented- $F$  test of Luukkonen, Saikkonen, and Terasvirta (1988).
- (5) TAR-F refers to the threshold test of Tsay (1989).
- (6) NEW-F refers to the general nonlinearity test of Tsay (1991).

Table III(b). General nonlinearity test for the transformed volatility of monthly value-weighted portfolio excess returns (NYSE & AMEX), 1928–1989 (AR order:  $p = 3$ )

$d$	Ori-F ( $p$ -value): 7.1646 (0.0000)				Aug-F ( $p$ -value): 5.4901 (0.0000)			
	TAR-F	$p$ -value	$v_1$	$v_2$	NEW-F	$p$ -value	$v_1$	$v_2$
1	4.47796	0.00141	4	660	2.76723	0.00076	13	651
2	1.94288	0.10167	4	660	2.80423	0.00064	13	651
3	4.73253	0.00090	4	660	3.44988	0.00003	13	651
4	0.52882	0.71460	4	659	2.36994	0.00419	13	650
5	3.07864	0.01579	4	658	3.47296	0.00003	13	649
6	0.54480	0.70289	4	657	2.75708	0.00079	13	648
7	5.34855	0.00030	4	656	3.42249	0.00004	13	647
8	0.60031	0.66254	4	655	2.50549	0.00237	13	646
9	1.78051	0.13099	4	654	3.87354	0.00000	13	645
10	3.07780	0.01582	4	653	2.97072	0.00031	13	644
11	2.39830	0.04897	4	652	3.52244	0.00002	13	643
12	0.85416	0.49124	4	651	3.13386	0.00015	13	642

Table III(c). General nonlinearity test for the transformed volatility of monthly equal-weighted portfolio excess returns (NYSE & AMEX), 1928–1989 (AR order:  $p = 4$ )

$d$	Ori-F ( $p$ -value): 2.3813 (0.0089)				Aug-F ( $p$ -value): 2.5232 (0.0016)			
	TAR-F	$p$ -value	$v_1$	$v_2$	NEW-F	$p$ -value	$v_1$	$v_2$
1	2.74321	0.01832	5	657	1.62865	0.05659	16	646
2	2.04864	0.07010	5	657	2.14205	0.00586	16	646
3	2.01140	0.07514	5	657	2.11350	0.00671	16	646
4	0.89265	0.48548	5	657	2.06468	0.00844	16	646
5	1.36919	0.23391	5	656	1.81744	0.02568	16	645
6	1.77481	0.11585	5	655	2.85989	0.00016	16	644
7	1.17578	0.31947	5	654	2.21981	0.00404	16	643
8	0.39319	0.85362	5	653	2.30309	0.00270	16	642
9	2.39319	0.03640	5	652	2.39850	0.00168	16	641
10	2.39532	0.03625	5	651	1.87675	0.01983	16	640
11	1.25899	0.28000	5	650	2.07094	0.00821	16	639
12	0.91282	0.47208	5	649	2.53350	0.00085	16	638

S&P, value-weighted and equal-weighted composite portfolio excess returns. For the S&P excess returns and value-weighted returns we use an AR order  $p = 3$  in the tests and the null hypothesis is that  $Y_t$  is a linear AR(3) process. For equal-weighted excess return  $p = 4$  is used. These AR orders are determined by the partial autocorrelations of the  $Y_t$  series. From the tables we observe that (a) the  $F$ -test and augmented- $F$ -test suggest that all three transformed volatility series are nonlinear at the 1 per cent level; (b) the threshold test and the general test indicate that the series exhibit certain threshold nonlinearity, especially when  $d = 1$  or 3; (c) the results show that the three volatility series do not have a unified nonlinear pattern. In summary the tests suggest that the transformed volatility series  $Y_t$  are nonlinear and the TAR model might be able to capture some nonlinear characteristics of the series.

5. MODELLING THE VOLATILITY OF STOCK RETURNS

5.1. TAR Models for the Volatility Series

Based on the results of nonlinearity tests we build in this section TAR models for the three transformed volatility series. Consider first the monthly S&P excess returns. Following the modelling procedure of Section 2 we begin with the determination of the threshold variable  $d$ . To this end Table III(a) suggests that  $Y_{t-1}, Y_{t-3}, Y_{t-6}$  and  $Y_{t-11}$  are possible threshold variables, because they correspond to the most significant testing results. Since  $d = 6$  or  $11$  is hard to interpret, we focus on the first two possible values. Further analysis of the estimation results led us to choose  $d = 1$ . In practice, if necessary, one can build a TAR model for each possible value of  $d$  and make a decision based on the models built. For instance, in this particular example one can build a TAR model for  $d = 1$  and another model for  $d = 3$ , and make a decision between these two TAR models by using a model selection criterion such as the AIC criterion.

With  $d = 1$  we turn to the specification of possible threshold values and, hence, the partition of the  $Y_{t-1}$  space. Figure 1 shows the scatterplots of local estimates of AR-1 and AR-3 coefficients versus the threshold variable  $Y_{t-1}$ . The local estimates are based on an arranged autoregression of order 3 with  $d = 1$  and a window size 120. By window size we mean the number of observations used in the local estimation, which is done recursively by adding a new observation and deleting an old one in each iteration. From the scatterplots, especially the one of AR-1 coefficient, the local estimates are relatively stable before and after  $Y_{t-1} = -2.2$ , suggesting that there is a change point around  $Y_{t-1} = -2.2$ . Consequently, we tentatively specify a two-regime TAR(3) model with a threshold  $Y_{t-1} = -2.2$  for the volatility series of the monthly S&P excess returns.

Finally, we refine the specification of the threshold value by searching through the neighbourhood of  $Y_{t-1} = -2.2$  and using AIC as the selection criterion. It turns out that  $Y_{t-1} = -2.16$  provides a two-regime TAR(3) model with the minimum AIC. With such a model specification, we obtain the model:

$$Y_t = \begin{cases} \phi_0^{(1)} + \phi_1^{(1)} Y_{t-1} + \phi_2^{(1)} Y_{t-2} + \phi_3^{(1)} Y_{t-3} + \varepsilon_t^{(1)} & \text{if } Y_{t-1} < -2.16 \\ \phi_0^{(2)} + \phi_1^{(2)} Y_{t-1} + \phi_2^{(2)} Y_{t-2} + \phi_3^{(2)} Y_{t-3} + \varepsilon_t^{(2)} & \text{if } Y_{t-1} \geq -2.16, \end{cases}$$

where  $Y_t = 4(\sigma_t^{1/4} - 1)$  is the transformed volatility series of the monthly S&P excess returns and the parameter estimates are:

$i$	$\phi_0^{(i)}$ (std)	$\phi_1^{(i)}$ (std)	$\phi_2^{(i)}$ (std)	$\phi_3^{(i)}$ (std)	$\sigma_i^2$	$N_i$
1	-1.75 (0.23)	-0.04 (0.07)	0.09 (0.05)	0.20 (0.05)	0.188	435
2	-0.77 (0.20)	0.44 (0.10)	0.18 (0.05)	0.12 (0.06)	0.193	306

where  $N_i$  denotes the number of observations in regime  $L_i$  and  $\sigma_i^2$  denotes the residual variance of  $\varepsilon_t^{(i)}$ , and std stands for standard error.

Model checking

To check the entertained TAR model we consider the standardized series  $\varepsilon_t^{(i)}/\sigma_i$ . The autocorrelation and partial autocorrelation functions of these standardized residuals are all within their (asymptotic) two standard-error limits, suggesting that there is no significant residual serial correlation. Further, we apply the BDS-U statistics of section 2 to check the

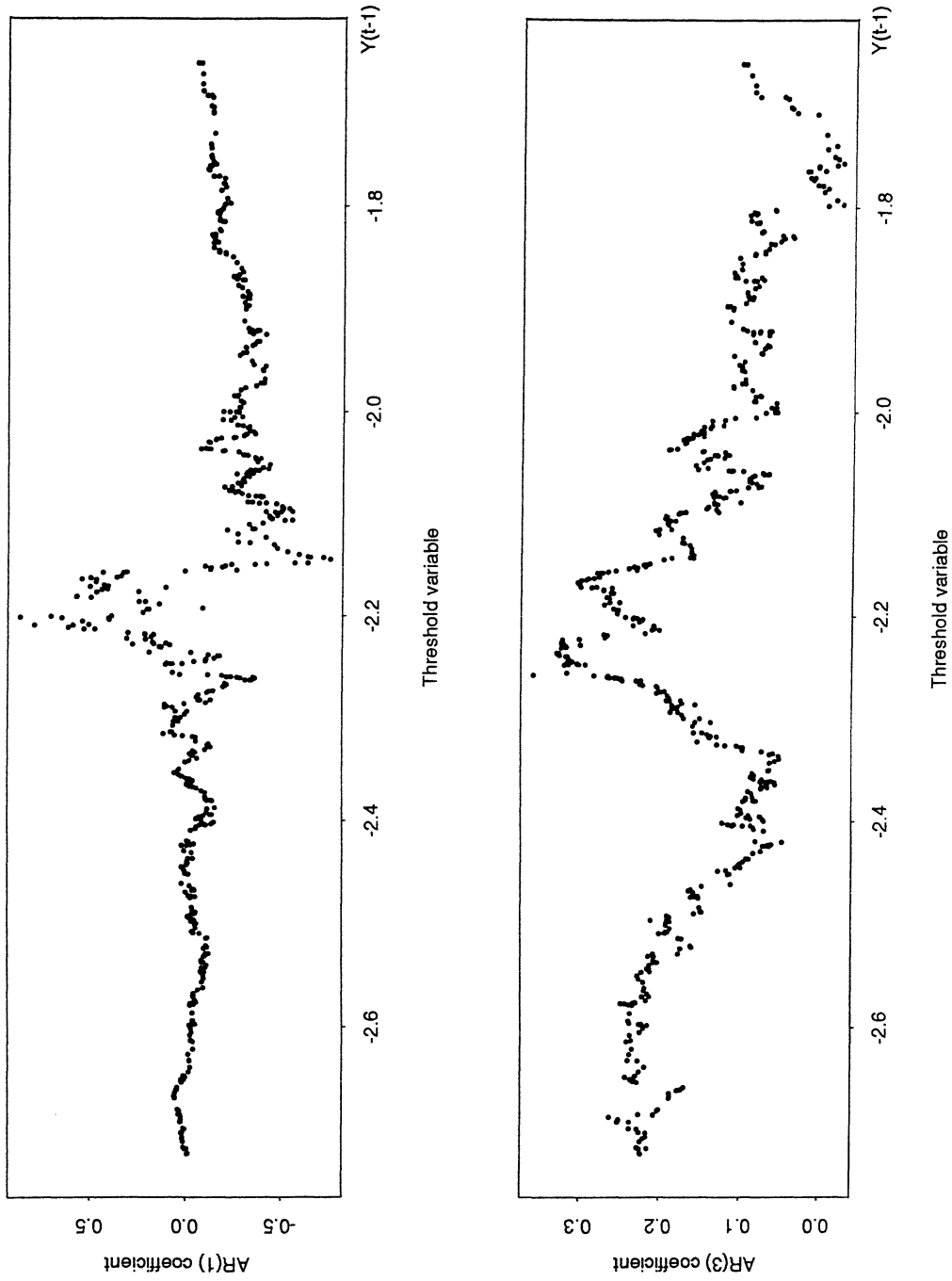


Figure 1. Plot of AR coefficients of an arranged autoregression vs. threshold variable for the monthly S&P volatility

i.i.d. assumption of the standardized residuals. The results all fail to reject the i.i.d. hypothesis.

*More interpretations*

Comparing with their standard errors the first two AR coefficients of Regime 1 are insignificant, implying that the model for this regime has weak serial correlations. Since a large negative value of  $Y_{t-1} = 4(\sigma_{t-1}^{1/4} - 1)$  corresponds to a small value in  $|R_{t-1} - \mu_{t-1}|$ , this weak AR(3) model says that the volatility of S&P monthly excess returns is not predictable when the past volatility is small. This is not surprising, since in the low-volatility period the local mean of  $\sigma_t$  is usually as good as any other predictors. On the other hand, the first two AR coefficients of the second-regime model are statistically significant, showing that certain predictability can be realized when the past volatility is sufficiently large. In other words the TAR model also supports the clustering of high volatilities.

The threshold  $Y_{t-1} = -2.16$  corresponds approximately to  $|R_t - \mu| = 3.6$  per cent, where  $\mu$  is the sample mean of the monthly S&P excess returns. Since the sample mean is essentially zero, the above TAR model says that, for the monthly S&P excess return, it is not fruitful to model the volatility when the previous excess return is between  $-3.6$  and  $3.6$  per cent. On the other hand, if the previous excess return is greater than  $3.6$  per cent in modulus, the volatility series shows some dynamic pattern, and hence is predictable.

Following the same procedure as above we built TAR models for the transformed volatility series of the monthly value-weighted and equal-weighted composite portfolio returns. For the value-weighted portfolio returns we chose a delay parameter  $d = 1$  and an AR order  $p = 3$  based on the results of nonlinearity tests in Table III(b). The scatterplots of local AR estimates suggest a two-regime TAR(3) model with a threshold around  $-2.4$ . Some refinement leads to the following two-regime TAR(3) model:

$$Y_t = \begin{cases} \phi_0^{(1)} + \phi_1^{(1)} Y_{t-1} + \phi_2^{(1)} Y_{t-2} + \phi_3^{(1)} Y_{t-3} + \varepsilon_t^{(1)} & \text{if } Y_{t-1} < -2.38 \\ \phi_0^{(2)} + \phi_1^{(2)} Y_{t-1} + \phi_2^{(2)} Y_{t-2} + \phi_3^{(2)} Y_{t-3} + \varepsilon_t^{(2)} & \text{if } Y_{t-1} \geq -2.38 \end{cases}$$

where, again,  $Y_t = 4(\sigma_t^{1/4} - 1)$  and the parameter estimates are:

$i$	$\sigma_0^{(i)}$ (std)	$\sigma_1^{(i)}$ (std)	$\sigma_2^{(i)}$ (std)	$\sigma_3^{(i)}$ (std)	$\sigma_i^2$	$N_i$
1	-1.92 (0.32)	-0.04 (0.09)	0.11 (0.06)	0.09 (0.06)	0.17	282
2	-1.07 (0.17)	0.34 (0.08)	0.11 (0.04)	0.13 (0.05)	0.20	459

where  $N_i$  and  $\sigma_i^2$  are defined as before. Similar checking procedures to those of S&P excess returns also fail to suggest any major discrepancy of the fitted model. Since the above model has similar structure to that of the monthly S&P volatility, it shares similar model interpretation.

For the transformed volatility of equal-weighted portfolio returns an AR order  $p = 4$  and a delay parameter  $d = 1$  are chosen. The modelling procedure of section 2 leads us to specify a three-regime TAR model with threshold value  $-2.49$  and  $-1.97$ . The estimated model is

$$Y_t = \begin{cases} \phi_0^{(1)} + \phi_1^{(1)} Y_{t-1} + \dots + \phi_4^{(1)} Y_{t-4} + \varepsilon_t^{(1)} & \text{if } Y_{t-1} < -2.49 \\ \phi_0^{(2)} + \phi_1^{(2)} Y_{t-1} + \dots + \phi_3^{(2)} Y_{t-3} + \varepsilon_t^{(2)} & \text{if } -2.49 \leq Y_{t-1} < -1.97 \\ \phi_0^{(3)} + \phi_1^{(3)} Y_{t-1} + \dots + \phi_3^{(3)} Y_{t-3} + \varepsilon_t^{(3)} & \text{if } Y_{t-1} \geq -1.97 \end{cases}$$

with parameter estimates:

<i>i</i>	$\phi_0^{(i)}$ (std)	$\phi_1^{(i)}$ (std)	$\phi_2^{(i)}$ (std)	$\phi_3^{(i)}$ (std)	$\phi_4^{(i)}$ (std)	$\sigma_i^2$	<i>N<sub>i</sub></i>
1	-0.76 (0.42)	0.06 (0.12)	0.20 (0.07)	0.16 (0.07)	0.21 (0.07)	0.19	184
2	-1.79 (0.45)	0.01 (0.19)	0.08 (0.06)	0.11 (0.05)		0.21	302
3	-0.59 (0.21)	0.52 (0.11)	0.13 (0.06)	0.18 (0.06)		0.24	255

The lag-4 coefficients of Regimes 2 and 3 were deleted as they are close to zero. Again, residual analysis and the BDS test fail to indicate any model discrepancy. This model can be interpreted along the same line as before, except that three regimes are needed.

**5.2. GARCH and EGARCH Models**

We build in this subsection the GARCH and EGARCH models for the 3-monthly stock returns considered in the paper. Since the ARCH-family models handle conditional variance directly we use the excess return *R<sub>t</sub>*, instead of the volatility series  $\sigma_t$ .

For a return series *R<sub>t</sub>* the GARCH model used is

$$R_t = \mu + b_1 R_{t-1} + \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{6}$$

This model was selected by using the Schwarz information criterion (Schwarz, 1978) among many GARCH models entertained. For the S&P excess return the AR(1) coefficient in the mean equation is not significant at the 5 per cent level, and is omitted. The parameter estimates of the GARCH(1,1) model are given in Table IV and they are all significant at the 5 per cent level, except for the mean  $\mu$  of the S&P excess returns.

The entertained EGARCH model is

$$R_t = \mu + b_1 R_{t-1} + \varepsilon_t, \quad \ln(\sigma_t^2) = \alpha + \frac{1}{1 - \Delta_1 B} g(z_{t-1}) \tag{7}$$

where  $g(z_t) = \theta z_t + \gamma [|z_t| - E|z_t|]$ . Some higher-order EGARCH models were also considered. However, the above EGARCH(1,0) model fits the data best according to the Schwarz information criterion. We reported the parameter estimates of the EGARCH(1,0) models in Table V. It is interesting to note that parameters  $\theta$  and  $\gamma$  are statistically significant,

Table IV. Estimates of the GARCH model (standard errors are in the parentheses)

Parameters	GARCH(1,1) S&P	GARCH(1,1) VW	GARCH(1,1) EW
$\mu \times 10$	0.268 (0.0168)	0.0585 (0.0170)	0.0670 (0.0207)
<i>b</i> <sub>1</sub>		0.0950 (0.0368)	0.1706 (0.0396)
$\alpha_0 \times 10^2$	0.0072 (0.0021)	0.0074 (0.0023)	0.0109 (0.0036)
$\alpha_1$	0.1168 (0.0178)	0.1184 (0.0189)	0.1111 (0.0211)
$\beta_1$	0.8632 (0.0159)	0.8613 (0.0177)	0.8714 (0.0235)
$\sigma_\varepsilon^2 \times 10^2$	0.3530	0.3259	0.5749

Table V. Estimates of the EGARCH model. Standard errors are in the parentheses

Parameters	EGARCH(1,0) S&P	EGARCH(1,0) VW	EGARCH(1,0) EW
$\mu \times 10$	0.0206 (0.0175)	0.0439 (0.0176)	0.0474 (0.0197)
$b_1$		0.1052 (0.0377)	0.1820 (0.0447)
$\alpha$	-0.5603 (0.3191)	-5.8102 (0.2721)	-5.1289 (0.2554)
$\theta$	-0.0504 (0.0172)	-0.0773 (0.0233)	-0.0818 (0.0175)
$\gamma$	0.2107 (0.0279)	0.1920 (0.0323)	0.1055 (0.0237)
$\Delta$	0.9777 (0.0065)	0.9710 (0.0078)	0.9873 (0.0078)
$\sigma_\varepsilon^2 \times 10^2$	0.3527	0.3252	0.5739

suggesting that the conditional variances of excess returns indeed respond differently to 'positive' and 'negative' past innovations.

Finally, we applied the BDS test to the standardized residuals of the fitted GARCH(1,1) and EGARCH(1,0) models for the three excess returns considered. The tests fail to reject the assumption that the standardized residuals are i.i.d. random variables.

## 6. FORECASTING COMPARISON

To evaluate the performance of the estimated TAR models in describing stock volatilities we compare their out-of-sample forecasts with that of (a) ARMA(1,1) models, (b) GARCH(1,1) models, and (c) EGARCH(1,0) models. The ARMA(1,1) models were used because they appear to fit the volatility series best among various ARMA models entertained. The comparison between TAR and ARMA(1,1) models is straightforward, because both models use the transformed volatility series. On the other hand, the comparison between TAR and the ARCH-family models is more complicated. The GARCH and EGARCH models use the excess return  $R_t$  directly whereas the TAR models use the transformed volatility. The main source of this difficulty is that the volatility is not observable in practice. To overcome this difficulty we use daily returns to construct a benchmark series for monthly stock volatilities. Ideally, daily returns are in a finer time-scale and should provide a reasonable measure of monthly volatilities. Details of the benchmark series are given later.

The post-sample forecast comparisons are carried out as follows. First, we reserve the last 60 observations (5 years from 1985 to 1989) for forecast comparison. Secondly, all the models used in forecasting are re-estimated using the first 684 observations. Thirdly, for a given model the time origins of forecast are  $t = 684, 685, \dots, 743$ . Fourthly, for a given model and a time origin  $t_0$ , we forecast the volatilities for time points  $t_0 + 1, t_0 + 2, \dots, 744$ . Therefore, for the time origin 684 we compute 1 to 60-step ahead forecasts; for the time origin 685 we compute 1 to 59-step ahead forecasts; and so on. Such a scheme provides 60 one-step ahead forecasts, 59 two-step ahead forecasts, etc. Finally, we summarize the forecast performance by considering mean squared error (MSE) and average absolute deviation (AAD) of the forecasts

by steps. Since fewer data points are available for longer-step ahead forecasts, we only report the comparison for 1 to 30-step.

### 6.1. Multi-step Ahead Forecasts

The multi-step ahead forecasts of TAR models are obtained by simulation. For a given TAR model for  $Y_t$ , suppose that we want to forecast  $Y_{t+1}, Y_{t+2}, \dots, Y_{t+h}$  at the time origin  $t$ . We first use the model and the data,  $Y_s$  for  $s \leq t$ , to generate 2000 realisations of  $\{Y_{t+l}\}_{l=1}^h$ . Then we average these realizations to obtain a mean series for time indices  $t+1, \dots, t+h$  and treat this mean series as our point forecasts. We ran some experiments with more than 2000 realizations and found that the forecasts are essentially the same.

### 6.2. TAR models versus ARMA models

Table VI shows the comparison between the ARMA and TAR models for the 3-monthly volatility series considered. It gives the MSE of TAR models for 1 to 30-step ahead forecasts

Table VI. Post-sample forecast comparison of MSE of TAR and ARMA models for the transformed volatility of monthly security excess returns, 1985–1989

Lead time	S&P		VW		EW	
	TAR	ARMA/TAR	TAR	ARMA/TAR	TAR	ARMA/TAR
1	0.161	1.065	0.214	1.021	0.281	0.907
2	0.168	1.028	0.217	1.004	0.261	0.963
3	0.159	1.046	0.204	1.019	0.216	0.979
4	0.161	1.060	0.206	1.026	0.230	0.972
5	0.162	1.061	0.207	1.026	0.239	0.946
6	0.166	1.048	0.211	1.029	0.242	0.949
7	0.162	1.052	0.201	1.028	0.234	0.969
8	0.159	1.036	0.201	0.992	0.225	0.956
9	0.158	1.046	0.198	1.009	0.228	0.969
10	0.163	1.009	0.204	0.979	0.232	0.958
11	0.167	1.025	0.208	0.993	0.240	0.952
12	0.166	1.067	0.209	1.026	0.250	0.963
13	0.171	1.039	0.215	1.010	0.252	0.950
14	0.166	1.067	0.215	1.014	0.245	0.960
15	0.165	1.093	0.215	1.031	0.250	0.955
16	0.168	1.089	0.219	1.033	0.254	0.947
17	0.175	1.088	0.230	1.018	0.265	0.955
18	0.172	1.113	0.227	1.048	0.268	0.967
19	0.175	1.086	0.198	1.028	0.271	0.943
20	0.172	1.102	0.194	1.042	0.272	0.941
21	0.174	1.093	0.197	1.027	0.278	0.939
22	0.169	1.069	0.191	1.018	0.283	0.930
23	0.173	1.051	0.195	0.999	0.290	0.927
24	0.178	1.038	0.192	1.000	0.291	0.930
25	0.180	1.056	0.193	1.022	0.299	0.955
26	0.169	1.063	0.184	1.026	0.302	0.939
27	0.177	1.023	0.190	0.995	0.310	0.931
28	0.183	1.027	0.194	1.007	0.315	0.959
29	0.187	1.028	0.199	1.007	0.327	0.949
30	0.158	1.079	0.164	1.033	0.288	0.959

and the ratio of MSE of ARMA models to TAR models. From the table it is seen that the TAR models outperform the ARMA models for the S&P and value-weighted monthly excess returns, but the ARMA models are slightly more accurate than the TAR models in forecasting the volatility of monthly equal-weighted excess returns. The table also shows that in terms of mean squared error the difference between ARMA and TAR models is usually less than 10 per cent. Similar results hold for the AAD criterion.

### 6.3. Comparison Between Nonlinear Models

We now compare the performance of TAR models with that of the ARCH-family models in (6) and (7). As mentioned above, the volatility of monthly excess returns for a stock is not observable, so that we construct a volatility series from the daily returns to serve as a benchmark series on which the forecasts can be compared. The constructed benchmark series for volatility consists of the monthly standard deviations computed from the daily portfolio returns. More specifically, consider the forecasting period from 1985 to 1989. Suppose that there are  $K_t$  trading days in month  $t$ , then the monthly standard deviation of a daily stock return is

$$s_t = \left[ \sum_{i=1}^{K_t} r_{it}^2 + 2 \sum_{i=1}^{K_t-1} r_{it}r_{i+1,t} \right]^{1/2}$$

where  $r_{i,t}$  denotes the  $i$ th trading day's return in month  $t$ . This measure of monthly stock volatility was used before in French, Schwert, and Stambaugh (1987). Since the daily returns of the three stocks considered have significant lag-1 autocorrelation, the cross-product  $r_{i,t}r_{i+1,t}$  is used in defining  $s_t$ .

The multi-step ahead forecasts of conditional variances for the GARCH(1,1) model in (6) were obtained by using an explicit functional form derived from the model in conjunction with the assumption that the error terms  $z_t$  are i.i.d.  $N(0, 1)$ . Details of the functional form is given in Appendix I. For the EGARCH(1,0) model in (7), we used the formula given in Appendix II to compute the forecasts of conditional variances. A square-root transformation is then taken to obtain monthly volatility forecasts. Finally, treating the data points of the benchmark series as observed monthly volatility, MSE and AAD of GARCH and EGARCH forecasts can be computed.

For the TAR models we modified the simulation procedure slightly in order to take care of the Box-Cox transformation used. The modification is that for each of the 2000 realizations of  $\{Y_{t+i}\}_{i=1}^h$  we take the inverse transformation  $\sigma_{t+i} = [(Y_{t+i}/4) + 1]^4$  to obtain a realization of  $\{\sigma_{t+i}\}$ . The mean series of these 2000 realizations of  $\{\sigma_{t+h}\}$  is then used to form the point forecasts of  $\sigma_{t+h}$ .

Tables VII and VIII report the results of 1 to 30-step ahead forecasts of TAR, GARCH(1,1) and EGARCH(1,0) models for the three monthly return series considered. The numbers shown in Table VII are the AAD of TAR models and the ratios of AAD between the ARCH-family models and the TAR models. From the table we observe that (a) the TAR models consistently have smaller AAD than the GARCH(1,1) and EGARCH(1,0) models for the monthly volatilities of S&P and value-weighted excess returns; (b) the gains of TAR models in both cases are more than 20 per cent for 8 to 25-step ahead forecasts; (c) for the equal-weighted excess returns the TAR model is slightly better at the short-term forecasts, but the EGARCH(1,0) model is superior for the long-term forecasts.

The numbers shown in Table VIII are MSE of the TAR models and the ratios of MSE

Table VII. Post-sample forecast comparison of AAD ( $\times 10^2$ ) of the TAR, GARCH and EGARCH models for the volatility of monthly security excess returns, 1985–1989

Lead time	S&P			VW			EW		
	GARCH		EGARCH	GARCH		EGARCH	GARCH		EGARCH
	TAR	TAR	TAR	TAR	TAR	TAR	TAR	TAR	TAR
1	1.66	1.09	1.04	1.76	1.07	1.03	2.65	1.17	1.11
2	1.77	1.10	1.06	1.80	1.11	1.09	2.71	1.19	1.13
3	1.81	1.15	1.12	1.80	1.18	1.15	2.79	1.20	1.14
4	1.95	1.17	1.13	1.94	1.18	1.14	3.01	1.16	1.10
5	1.96	1.19	1.17	1.96	1.20	1.19	3.13	1.12	1.06
6	1.97	1.18	1.17	1.96	1.19	1.18	3.18	1.11	1.04
7	1.98	1.20	1.20	1.96	1.21	1.19	3.25	1.10	1.02
8	1.98	1.20	1.22	1.96	1.21	1.22	3.32	1.07	1.00
9	1.91	1.25	1.28	1.90	1.25	1.28	3.30	1.08	0.99
10	1.98	1.23	1.24	1.94	1.25	1.26	3.36	1.08	0.98
11	1.96	1.24	1.25	1.93	1.27	1.27	3.47	1.05	0.95
12	1.99	1.23	1.22	1.96	1.26	1.23	3.48	1.06	0.94
13	1.99	1.26	1.26	1.96	1.27	1.26	3.60	1.04	0.92
14	2.04	1.25	1.26	1.99	1.27	1.27	3.66	1.03	0.91
15	2.05	1.26	1.28	2.03	1.28	1.28	3.73	1.02	0.89
16	2.10	1.28	1.30	2.05	1.30	1.30	3.71	1.04	0.90
17	2.10	1.30	1.34	2.08	1.32	1.34	3.88	1.02	0.89
18	2.06	1.33	1.39	2.03	1.35	1.38	3.88	1.03	0.90
19	2.08	1.33	1.40	2.05	1.34	1.38	3.88	1.03	0.91
20	2.14	1.29	1.38	2.10	1.32	1.38	4.00	1.00	0.88
21	2.18	1.27	1.37	2.14	1.31	1.35	4.07	0.98	0.86
22	2.16	1.24	1.33	2.15	1.27	1.31	4.09	0.98	0.86
23	2.19	1.20	1.29	2.16	1.24	1.26	4.09	0.97	0.84
24	2.20	1.18	1.24	2.17	1.23	1.22	4.11	0.95	0.81
25	2.19	1.18	1.25	2.17	1.21	1.20	4.13	0.92	0.78
26	2.31	1.08	1.12	2.26	1.13	1.10	4.30	0.89	0.75
27	2.41	1.01	1.03	2.33	1.07	1.00	4.34	0.86	0.72
28	2.44	1.01	1.04	2.36	1.08	1.00	4.37	0.87	0.73
29	2.42	1.04	1.05	2.36	1.09	1.01	4.45	0.88	0.74
30	2.47	1.03	1.03	2.41	1.09	0.99	4.47	0.89	0.74

between the ARCH-family models and TAR models. By and large the results are similar to that of the AAD criterion. The TAR models outperform the ARCH-family models for S&P and value-weighted monthly excess returns. However, the EGARCH(1,0) model fares well in long horizon forecasts of monthly volatility of equal-weighted excess returns.

Since the S&P composite portfolio and the value-weighted composite portfolio represent large stocks and the equal-weighted portfolio represents small stocks, the results of forecasting comparison indicate that the TAR models might be useful in forecasting monthly volatility of large stocks. On the other hand the EGARCH(1,0) model can yield accurate long-term forecasts of monthly volatility for small stocks. We note that the above out-of-sample forecasts of the TAR model are unbiased, whereas those of the GARCH and EGARCH models are biased. The unfavourable performance of the ARCH-type models might be due in part to the forecasting bias.

Table VIII. Post-sample forecast comparison of MSE ( $\times 10^3$ ) of the TAR, GARCH and EGARCH models for the volatility of monthly security excess returns, 1985–1989

Lead time	S&P			VW			EW		
	TAR	GARCH	EGARCH	TAR	GARCH	EGARCH	TAR	GARCH	EGARCH
		TAR	TAR		TAR	TAR		TAR	
1	1.21	1.09	1.07	1.07	1.06	1.04	1.54	1.17	1.08
2	1.26	1.11	1.11	1.06	1.12	1.13	1.55	1.22	1.12
3	1.30	1.13	1.14	1.08	1.18	1.20	1.54	1.30	1.18
4	1.41	1.12	1.13	1.13	1.20	1.22	1.67	1.30	1.15
5	1.39	1.16	1.18	1.13	1.23	1.26	1.77	1.25	1.10
6	1.38	1.17	1.19	1.14	1.23	1.26	1.71	1.30	1.14
7	1.39	1.20	1.23	1.15	1.25	1.32	1.76	1.29	1.14
8	1.39	1.21	1.25	1.15	1.26	1.33	1.78	1.28	1.12
9	1.38	1.21	1.26	1.13	1.27	1.36	1.76	1.30	1.14
10	1.44	1.25	1.27	1.17	1.32	1.35	1.84	1.30	1.10
11	1.48	1.23	1.25	1.21	1.30	1.35	1.94	1.25	1.06
12	1.44	1.25	1.28	1.20	1.29	1.33	1.88	1.29	1.08
13	1.54	1.21	1.23	1.26	1.25	1.28	1.98	1.23	1.02
14	1.54	1.26	1.32	1.25	1.34	1.41	2.01	1.25	1.05
15	1.60	1.25	1.31	1.30	1.32	1.39	2.13	1.19	0.99
16	1.58	1.31	1.41	1.30	1.38	1.48	2.11	1.26	1.07
17	1.66	1.27	1.37	1.35	1.35	1.46	2.29	1.18	1.02
18	1.68	1.27	1.37	1.37	1.36	1.47	2.31	1.18	1.02
19	1.72	1.26	1.38	1.40	1.33	1.46	2.32	1.17	1.02
20	1.75	1.26	1.39	1.42	1.33	1.47	2.40	1.13	0.99
21	1.80	1.26	1.39	1.45	1.33	1.47	2.51	1.09	0.95
22	1.84	1.21	1.34	1.49	1.28	1.41	2.46	1.10	0.96
23	1.85	1.21	1.33	1.51	1.26	1.37	2.48	1.06	0.92
24	1.88	1.19	1.29	1.54	1.22	1.31	2.51	1.01	0.87
25	1.90	1.19	1.28	1.55	1.23	1.30	2.56	0.99	0.85
26	2.04	1.09	1.18	1.65	1.13	1.19	2.75	0.91	0.79
27	2.12	1.04	1.12	1.71	1.06	1.10	2.82	0.85	0.74
28	2.14	1.05	1.14	1.73	1.07	1.11	2.84	0.86	0.74
29	2.20	1.04	1.12	1.77	1.07	1.10	2.92	0.86	0.74
30	2.18	1.09	1.16	1.78	1.09	1.12	2.85	0.90	0.76

## 7. CONCLUDING REMARKS

In this paper we presented empirical evidence that the volatilities of monthly stock returns are nonlinear. We also illustrated that simple threshold autoregressive models might be useful in modelling the volatility of monthly excess returns. This is achieved by constructing a volatility series from the excess returns and by modelling this constructed volatility series directly.

Out-of-sample forecasts suggest that certain gain in forecasting accuracy can be realized by using nonlinear TAR models instead of linear ARMA models. Further, using daily stock returns to derive a benchmark measure for the monthly volatility, we illustrated that TAR models could be useful in providing volatility forecasts as compared with the commonly used GARCH(1,1) and EGARCH(1,0) models, especially for large stock returns. We also found that the EGARCH(1,0) model can provide accurate long-horizon volatility forecasts for small stocks.

APPENDIX I

In this appendix we present the formula used to calculate the multi-step ahead forecasts of the conditional variance for the GARCH(1,1) model in (6). The variance equation of this model is  $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$ . Denote the forecast origin by  $n$  and the forecast horizon by  $j$ . Let  $I_n$  be the information set available at time  $n$ . For  $j=1$  the 1-step ahead forecast of the conditional variance is simply

$$E(\sigma_{n+1}^2 | I_n) = E(\alpha_0 + \alpha_1 \varepsilon_n^2 + \beta_1 \sigma_n^2 | I_n) = \alpha_0 + \alpha_1 \varepsilon_n^2 + \beta_1 \sigma_n^2.$$

For  $j \geq 2$ , by using the assumption that  $z_t$  are i.i.d.  $N(0,1)$  and the identity  $E(\varepsilon_{n+t}^2 | I_n) = E(\sigma_{n+t}^2 | I_n)$ , we have  $E(\sigma_{n+j}^2 | I_n) = \alpha_0 + (\alpha_1 + \beta_1) \cdot E(\sigma_{n+j-1}^2 | I_n)$ . Therefore, the forecasts of conditional variances of a GARCH(1, 1) model can be computed recursively.

APPENDIX II

Next we derive the formula for obtaining multi-step ahead forecasts of the conditional variance for the EGARCH model in (7). For simplicity we assume that the process is stationary so that it starts with  $t = -\infty$ . The conditional variance of the model is

$$\ln(\sigma_t^2) = \alpha + \frac{1}{1 - \Delta B} g(z_{t-1}) = \alpha + \sum_{k=1}^{\infty} \Delta^{k-1} g(z_{t-k})$$

where  $g(z_t) = \theta z_t + \gamma [ |z_t| - E|z_t| ]$ . Therefore,

$$\sigma_t^2 = \exp \left[ \alpha + \sum_{k=1}^{\infty} \Delta^{k-1} g(z_{t-k}) \right].$$

Let  $n, j$  and  $I_n$  be defined as in Appendix I. For  $j=1$  the 1-step ahead forecast of the conditional variance is simply

$$E(\sigma_{n+1}^2 | I_n) = \exp \left[ \alpha + \sum_{k=1}^{\infty} \Delta^{k-1} g(z_{n+1-k}) \right].$$

For  $j=2$ , the 2-step ahead forecast of the conditional variance is

$$\begin{aligned} E(\sigma_{n+2}^2 | I_n) &= E \left[ \exp \left( \alpha + \sum_{k=1}^{\infty} \Delta^{k-1} g(z_{n+2-k}) \right) | I_n \right] \\ &= \exp \left[ \alpha + \sum_{k=2}^{\infty} \Delta^{k-1} g(z_{n+2-k}) \right] E[\exp(g(z_{n+1})) | I_n] \\ &= \exp \left[ \alpha + \sum_{k=2}^{\infty} \Delta^{k-1} g(z_{n+2-k}) \right] \\ &\quad \times \exp \left( -\gamma \sqrt{\frac{2}{\pi}} \right) \left[ \exp \left( \frac{1}{2} (\gamma - \theta)^2 \right) \Phi(\gamma - \theta) + \exp \left( \frac{1}{2} (\gamma + \theta)^2 \right) \Phi(\gamma + \theta) \right], \end{aligned}$$

where  $\Phi(\cdot)$  is the cumulative distribution function of a standard normal variable. Similarly,

for  $j > 2$ , we have

$$\begin{aligned} E(\sigma_{n+j}^2 | I_n) &= \exp\left(\alpha + \sum_{k=j}^{\infty} \Delta^{k-1} g(z_{n+j-k})\right) E\left[\exp\left(\sum_{k=1}^{j-1} \Delta^{k-1} g(z_{n+j-k})\right) | I_n\right] \\ &= \exp\left[\alpha + \sum_{k=j}^{\infty} \Delta^{k-1} g(z_{n+j-k})\right] \prod_{k=1}^{j-1} \exp\left(-\Delta^{k-1} \gamma \sqrt{\frac{2}{\pi}}\right) \\ &\quad \times \left[ \exp\left(\frac{1}{2} (\gamma - \theta)^2 \Delta^{2(k-1)}\right) \Phi((\gamma - \theta) \Delta^{k-1}) \right. \\ &\quad \left. + \exp\left(\frac{1}{2} (\gamma + \theta)^2 \Delta^{2(k-1)}\right) \Phi((\gamma + \theta) \Delta^{k-1}) \right]. \end{aligned}$$

#### ACKNOWLEDGEMENTS

This research is supported in part by the National Science Foundation DMS-8902177 and DMS-9103250. The authors wish to thank George C. Tiao for useful discussions, Kenneth R. French, Simon M. Potter, Daniel B. Nelson and two anonymous referees for helpful comments and David Hsieh for providing the BDS-test program.

#### REFERENCES

- Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics*, **31**, 307–27.
- Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992), 'ARCH modeling in finance: a review of the theory and empirical evidence', *Journal of Econometrics*, **52**, 5–60.
- Brock, W. A., W. D. Dechert, and J. A. Scheinkman (1987), 'A test for independence based on the correlation dimension'. Working paper, University of Wisconsin-Madison.
- Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the variance of U.K. inflation', *Econometrica*, **50**, 987–1008.
- French, K. R., G. W. Schwert, and R. Stambaugh (1987), 'Expected stock returns and volatility', *Journal of Financial Economics*, **19**, 3–29.
- Gallant, A. R., D. Hsieh, and G. E. Tauchen (1991), 'On fitting a recalcitrant series: the pound/dollar exchange rate, 1974–83', in W. Barnett, J. Powell, and G. Tauchen (eds), *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, Cambridge University Press, Cambridge, pp. 199–240.
- Hsieh, D. (1989), 'Testing for nonlinear dependence in daily foreign exchange rates', *Journal of Business*, **62**, 339–368.
- Hsieh, D., and B. LeBaron (1988), 'Small sample properties of the BDS statistics, I, II and III'. Working paper, University of Chicago and University of Wisconsin-Madison.
- LeBaron, B. (1992), 'Some relations between volatility and serial correlations in stock market returns', *Journal of Business*, **65**, 199–220.
- Luukkonen, R., P. Saikkonen, and T. Teräsvirta, (1988), 'Testing linearity against smooth transition autoregressive models', *Biometrika*, **75**, 491–499.
- Melino, A., and S. M. Turnbull (1990), 'Pricing foreign currency options with stochastic volatility', *Journal of Econometrics*, **45**, 239–266.
- Nelson, D. B., and C. Q. Cao (1992), 'Inequality constraints in the univariate GARCH model', *Journal of Business & Economic Statistics*, **10**, 229–235.
- Nelson, D. B. (1991), 'Conditional heteroscedasticity in asset returns: a new approach', *Econometrica*, **59**, 347–370.
- Pagan A., and G. W. Schwert (1990), 'Alternative models for conditional stock volatility', *Journal of Econometrics*, **45**, 267–290.

- Poterba, J., and L. Summers (1986), 'The persistence of volatility and stock market fluctuations', *American Economic Review*, **76**, 1142–1151.
- Quandt, R. E., and J. Ramsey. (1978), 'Estimating mixtures of normal distributions and switching regressions', *Journal of the American Statistical Association*, **73**, 730–738.
- Schwarz, G. (1978), 'Estimating the dimension of a model', *Annual of Statistics*, **6**, 461–464.
- Scheinkman, J. A., and B. LeBaron (1989), 'Nonlinear dynamics and stock returns', *Journal of Business*, **3**, 311–337.
- Schwert, G. W. (1990), 'Stock volatility and the crash of '87', *Review of Financial Studies*, **3**, 77–102.
- Schwert, G. W., and P. J. Seguin (1990), 'Heteroscedasticity in stock returns', *Journal of Finance*, **4**, 1129–1155.
- Scott, L. O. (1987), 'Option pricing when the variance changes randomly: theory, estimation and an application', *Journal of Financial and Quantitative Analysis*, **22**, 419–438.
- Stein, J. (1989), 'Overreaction in the option market', *Journal of Finance*, **44**, 1011–1023.
- Taylor, S. (1986), *Modeling Financial Time Series*, John Wiley & Sons, Chichester.
- Tong, H. (1990), *Nonlinear Time Series: A Dynamical System Approach*, Oxford University Press, London.
- Tsay, R. S. (1986), 'Nonlinearity tests for time series', *Biometrika*, **73**, 461–466.
- Tsay, R. S. (1987), 'Conditional heteroscedastic time series models', *Journal of the American Statistical Association*, **82**, 590–604.
- Tsay, R. S. (1989), 'Testing and modeling threshold autoregressive processes', *Journal of the American Statistical Association*, **84**, 231–240.
- Tsay, R. S. (1991), 'Modelling nonlinearity in univariate time series analysis', *Statistical Sinica*, **1**, 431–452.
- Wiggins, J. B. (1987), 'Option values under stochastic volatility: theory and empirical estimations', *Journal of Financial Economics*, **19**, 351–372.