# An Empirical Evaluation of STAR-GARCH Models in the Presence of Aberrant Observations

**Felix Chan and Michael McAleer**

*Department of Economics, University of Western Australia*

**Abstract:** This paper presents an empirical evaluation of quasi-maximum likelihood estimation of Smooth Transition Autoregressive (STAR) models with GARCH errors (STAR-GARCH). The empirical evidence shows that different algorithms can produce substantially different estimates, so that the interpretation of the model is sensitive to the choice of algorithm. Convergence, the choice of different algorithms for maximising the likelihood function, and the sensitivity of the estimates to aberrant observations (that is, extreme observations and outliers), are examined using daily data for S&P 500, Hang Sang and Nikkei 225 for the period January 1986 to April 2000.

***Keywords:*** STAR, GARCH, smooth transition, extreme observations, outliers.

## 1 Introduction

Regime switching time series models have become popular in the class of non-linear models in recent years. Given the substantial research activity in analysing time-varying volatility, it is of interest to investigate regime switching models with GARCH errors (see Engle (1982) and Bollerslev (1986)). One of the most popular specifications in this class is the Smooth Transition Autoregressive (STAR) - Generalised Autoregressive Conditional Heteroscedasticity (STAR-GARCH) model.

Chan and McAleer (2002) have examined the structural and statistical properties of the STAR-GARCH model, and the finite sample properties of the Quasi-Maximum Likelihood Estimator (QMLE). Nevertheless, difficulties in selecting the most efficient optimization algorithm remain. As noted by Lundbergh and Teräsvirta (1999), van Dijk, Teräsvirta and Franses (2002), and Brooks, Burke and Persand (2001), the convergence of the QMLE is sensitive to the choice of initial values.

This paper provides empirical evidence to show that different algorithms can produce substantially different estimates and interpretations for the same model. This is contrary to the common belief that different algorithms will produce similar estimates (Lundbergh and Teräsvirta (2000)). The paper also examines the effects of aberrant observations (that is, extreme observations and outliers) on the QMLE for the STAR-GARCH model. This is of interest because the STAR-GARCH model was not designed specifically to accommodate aberrant observations. However, such models are often used to model financial time series, which frequently exhibit excessive kurtosis. Empirical evidence presented in the paper shows that the effects of aberrant observations on the QMLE for the STAR component in a STAR-GARCH model depend on the choice of transition functions.

The plan of the paper is as follows. Section 2 gives a brief outline of recent theoretical developments for the GARCH, STAR and STAR-GARCH models. Section 3 discusses the data. Section 4 presents a detailed discussion of various optimisation algorithms and their effects on the QMLE for STAR-GARCH models. Section 5 investigates the effects of aberrant observations on the estimates for STAR-GARCH models. Section 6 gives some concluding remarks.

## 2 Recent Theoretical Developments for GARCH, STAR and STAR-GARCH Models

### 2.1 GARCH

Consider the ARMA$(r, s)$ model:

$$y_t = \sum_{i=1}^{r} \phi_i y_{t-i} + \sum_{i=1}^{s} \theta_i \varepsilon_{t-i} + \varepsilon_t \qquad (2.1)$$

in which $\varepsilon_t$ follows Engle's (1982) Autoregressive Conditional Heteroscedasticity, ARCH$(p)$, process if

$$\begin{aligned} \varepsilon_t &= \eta_t \sqrt{h_t}, \qquad \eta_t \sim i.i.d.(0,1) \\ h_t &= \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 \end{aligned} \qquad (2.2)$$

where $\omega > 0$ and $\alpha_i \geq 0$ for all $i = 1, .., p$, are sufficient conditions for $h_t > 0$ for all $t = 1, .., T$.

Defining $\delta = (\omega_0, \alpha_1, ..., \alpha_p)'$, the conditional log-likelihood function of (2.1) given observations $\varepsilon_t$, $t = 1, .., T$, can be written as:

$$l(\delta) = \frac{-1}{2T} \sum_{t=1}^{T} \left( \ln h_t + \frac{\varepsilon_t^2}{h_t} \right) \qquad (2.3)$$

so that the MLE is given as

$$\hat{\delta} = \text{argmax}_{\delta \in \Theta} l(\delta)$$

assuming that $\delta \in \Theta$, a compact subset of $\mathbb{R}^{p+1}$. Engle (1982) showed that the information matrix of this function is block diagonal, which implies the parameters in the conditional mean and the conditional variance can be estimated separately without loss of asymptotic efficiency. The estimated errors given by the estimated conditional mean equation can be used to estimate the equation of the conditional variance.

The MLE is referred to as the Quasi-MLE (QMLE) when $\eta_t$ is not normal. Weiss (1986) and Pantula (1989) showed that the QMLE is consistent and asymptotic normal if $E(\varepsilon_t^4) < \infty$. This result was extended by Ling and McAleer (2002c), who showed that the QMLE is consistent and asymptotic normal if $E(\varepsilon_t^2) < \infty$.

Bollerslev (1986) extended ARCH by including the lags of the conditional variance in (2.2) to yield the GARCH$(p,q)$ model, namely

$$h_t = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i h_{t-i} \tag{2.4}$$

where $\omega > 0$, $\alpha_i \geq 0$ for all $i = 1,..,p$, and $\beta_i \geq 0$ for all $i = 1,..,q$, are sufficient conditions for $h_t > 0$ for all $t = 1,..,T$. The necessary and sufficient condition for the second-order stationarity of (2.4) was established by Bollerslev (1986) as $\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i < 1$. He and Teräsvirta (1999a) obtained the moment conditions of a family of GARCH(1,1) models using a similar method as in Bollerslev (1986). Ling and McAleer (2002b) derived the sufficient condition for the existence of the stationary solution of this family of GARCH(1,1) models, showed that He and Terasvirta's (1999a) condition was necessary but not sufficient, and provided the sufficient condition. He and Teräsvirta (1999b) examined the fourth moment structure of the GARCH$(p,q)$ process. Ling (1999) obtained a sufficient condition for the existence of the $2m$th moment for the GARCH$(p,q)$ model, based on Theorem 2.1 in Ling and Li (1997) and Theorem 2 in Tweedie (1988). This condition is also necessary for the existence of the $2m$th moment, as demonstrated by Ling and McAleer (2002a). Thus, the moment structure of a general GARCH$(p,q)$ has been fully established.

The parameters in a GARCH$(p,q)$ model are often estimated by MLE, or by QMLE when the normality of $\eta_t$ is not assumed. Ling and Li (1997) showed that the *local* QMLE for GARCH$(p,q)$ is consistent and asymptotic normal if $E(\varepsilon_t^4) < \infty$. For the *global* QMLE, Ling and McAleer (2002a) showed that $E(\varepsilon_t^2) < \infty$ is sufficient for consistency and $E(\varepsilon_t^6) < \infty$ is sufficient for asymptotic normality.

### 2.2 STAR and STAR-GARCH

Non-linear time series models such as regime switching models have become very popular in recent years, so it is of interest to investigate regime switching models with GARCH errors. Smooth Transition Autoregressive (STAR) models will be dscussed as a special type of regime switching model.

Tong (1978) and Tong and Lim (1980) proposed the Threshold Autoregressive (TAR) model. The TAR model assumes that the regimes switch from one to another, as determined by the *threshold variables*, $s_t$, relative to the *threshold value*, $c$. In the two regime case:

$$\begin{aligned} y_t =& (\phi_{11} + \sum_{i=2}^{r} \phi_{1i} y_{t-i+1})(1 - I(s_t - c)) \\ &+ (\phi_{21} + \sum_{i=2}^{r} \phi_{2i} y_{t-i+1}) I(s_t - c) + \varepsilon_t \end{aligned} \tag{2.5}$$

where

$$I(s_t - c) = \begin{cases} 0, & s_t < c \\ 1, & s_t \geq c. \end{cases}$$

Model (2.5) can also be written as:

$$y_t = \begin{cases} \phi_{11} + \sum_{i=2}^{r} \phi_{1i} y_{t-i+1} + \varepsilon_t, & s_t < c \\ \phi_{21} + \sum_{i=2}^{r} \phi_{2i} y_{t-i+1} + \varepsilon_t, & s_t \geq c. \end{cases}$$

The threshold variable, $s_t$, is usually (but not always) defined as a linear combination of the lagged values of $y_t$, that is, $s_t = \sum_{i=1}^{k} \pi_i y_{t-i}$, which is often referred to as a Self Exciting TAR (SETAR).

Equation (2.5) is similar to a standard model of structural change, apart from the definition of threshold variable and threshold value, which assumes that the regimes switch from one to another instantly. To allow for a smooth transition, Terašvirta (1994) proposed the Smooth Transition Autoregressive (STAR) model:

$$\begin{aligned} y_t =& (\phi_{11} + \sum_{i=2}^{r} \phi_{1i} y_{t-i+1})(1 - G(s_t;\gamma,c)) + \\ & (\phi_{21} + \sum_{i=2}^{r} \phi_{2i} y_{t-i+1}) G(s_t;\gamma,c) + \varepsilon_t \end{aligned} \tag{2.6}$$

in which $G(s_t;\gamma,c)$ is the transition function, assumed to be twice differentiable, ranging from 0 to 1, and $\gamma$ is the *transition rate*.

There are many choices of transition function, with the most popular being the first-order logistic transition function:

$$G(s_t;\gamma,c) = \frac{1}{1 + \exp(-\gamma(s_t - c))}.$$

Although the Logistic STAR (LSTAR) is widely used, a useful alternative is the Exponential STAR (ESTAR) model:

$$G(s_t;\gamma,c) = 1 - \exp(-\gamma(s_t - c)^2), \qquad \gamma > 0.$$

In order to use this model effectively, it is important to choose the appropriate transition function and threshold variable. When the transition function and the threshold variable have been determined, the parameters in STAR can be estimated by Non-linear Least Squares (NLS). If

$$y_t = F(x_t; \phi) + \varepsilon_t$$

the NLS estimator is given by

$$\hat{\phi} = argmin_\phi \sum_{t=1}^{T} (y_t - F(x_t; \phi))^2 = argmin_\phi \sum_{t=1}^{T} \varepsilon_t^2.$$

If $\varepsilon_t$ is normal, NLS is equivalent to MLE, otherwise NLS can be interpreted as QMLE. A STAR-GARCH model allows $\varepsilon_t$ in equation (2.6) to follow a GARCH process (see Lundbergh and Teräsvirta (1999) for further details). Chan and McAleer (2002) provide the regularity conditions for the existence of the moments and for stationarity of the STAR-GARCH model, and also the statistical properties of the QMLE. As the information matrix of the log-likelihood function of STAR-GARCH is block diagonal, the parameters in the conditional mean and conditional variance equations can be estimated separately, as in the case of ARMA-GARCH.

## 3 Data

Following Lundbergh and Teräsvirta (1999), both regimes for each model estimated below are assumed to follow an AR(1) process, $\varepsilon_t$ is assumed to be GARCH(1,1), $s_t = y_{t-1}$, and $r_t = \varepsilon_{t-1}$.

All models are estimated using three different stock indices, namely Standard and Poor's 500 Composite Index (S&P), Hang Sang Index (HSI) and Nikkei 225 Index. The data were obtained through the DataStream database service and the sample is from 1/1/1986 to 11/4/2000, giving a total of 3725 data points for each index. S&P is possibly the most widely used financial data series for empirical purposes. As the S&P series do not contain a significant conditional mean component, the empirical results reported below display a similar pattern to the experimental results in Chan and McAleer (2002), who concentrated on the conditional variance and not the conditional mean.

Of primary concern are stock returns, $R_t$, namely

$$R_t = \frac{Y_t - Y_{t-1}}{Y_{t-1}}$$

where $Y_t$ denotes the index at time $t$. S&P appears to be less volatile than either Nikkei or HSI, especially during the early to late 1990's (observations 1500 to 3000) before the Asian economic and financial crises. Furthermore, S&P seems to have fewer aberrant observations than Nikkei and HSI, while Nikkei seems to have more positive shocks than S&P and HSI. Although Nikkei seems more volatile, the volatility is relatively low compared with HSI, which seems to be the most volatile. There are some obvious aberrant observations for all three indices. HSI also appears to have the highest number of outliers.

An obvious similarity among the three indices is the enormous decrease in returns at observation 474, which corresponds with the share market crash in October 1987. This is also the most significant outlier in the three indices. The second largest decrease in returns is observation 894 for HSI, which corresponds with the Tianenman Square incident in Beijing on 4 June 1989.

## 4 Optimisation Algorithms

Estimation of STAR-GARCH is problematic as existing econometric software packages do not yet include the appropriate algorithms. STAR-GARCH can be estimated by a two-stage procedure, which involves estimating STAR by NLS, then using the residuals to estimate GARCH by QMLE.

In practice, estimation of STAR and STAR-GARCH can be problematic. An attempt was made in EViews to estimate LSTAR by NLS with $s_t = Y_{t-1}$ for S&P, HSI and Nikkei, but was unsuccessful because of computational problems in calculating the near singular Hessian matrix.

A GAUSS version 3 program was used to estimate STAR-GARCH by optimising the likelihood function. Several attempts were made to estimate LSTAR and ESTAR for S&P, HSI and Nikkei, for which the estimates of the variance did not converge. This suggests three possibilities: (i) the variance is not constant, so that STAR-GARCH should be used; (ii) the use of alternative optimisation algorithms; and (iii) the use of alternative initial values.

In order to investigate the effects of different algorithms on the estimates of STAR-GARCH, two were chosen to maximise the log-likelihood function of an LSTAR-GARCH model using S&P, HSI and Nikkei data. The two algorithms used are the Newton method and Broyden-Fletcher-Goldfarb-Shanno method (BFGS), a quasi-Newton method, with the same initial values, yielding the estimates in Table 1 (in which * denotes significance at the 1% level).

As shown in Table 1, the two algorithms produce substantially different estimates for the conditional mean but not for the conditional variance. In fact, a similar set of estimates for the conditional variance can be obtained by estimating a simple AR(1)-GARCH(1,1) model using the Berndt-Hall-Hall-Hausman (1974) (BHHH) algorithm. This suggests that only the estimates of the conditional mean would seem to be sensitive to the choice of algorithm, which is contrary to the findings in Lundbergh and Teräsvirta (2000).

The high threshold values, $c$, for HSI imply that the first regime dominates the second. However, since BFGS gives substantially different estimates from Newton, it is difficult to determine which set of estimates will produce better forecasts. Moreover, the estimates from the two algorithms are highly significant, which makes interpretation problematic.

| | S&P | | HSI | | Nikkei | |
|---|---|---|---|---|---|---|
| | Newton | BFGS | Newton | BFGS | Newton | BFGS |
| $\hat{\phi}_{11}$ | 1.7833* | -3.0150* | -0.2238* | 0.0142 | -0.2796* | -1.5166* |
| $\hat{\phi}_{12}$ | 1.1204* | -4.1005* | 2.0523* | -4.9601* | -0.1930* | 0.1396* |
| $\hat{\phi}_{21}$ | -0.2862* | 3.0088* | 0.2260* | 0.0007 | 0.0013* | 1.5131* |
| $\hat{\phi}_{22}$ | -1.3476* | 3.8090* | -1.8170* | 0.3180* | 0.0126 | -0.2892* |
| $\hat{\gamma}$ | 2.4515* | 1.4818* | 4.3080* | 3.9604* | 1.6278* | 1.0020* |
| $\hat{c}$ | 0.8634* | -0.0402* | -0.0162* | -0.8311* | -2.0104* | -0.0586* |
| $\hat{\omega}$ | 1.17e-6 | 1.17e-6* | 7.08e-6* | 7.10e-6* | 3.2e-6* | 3.21e-6* |
| $\hat{\alpha}$ | 0.0753* | 0.0754* | 0.1488* | 0.1486* | 0.1475* | 0.1465* |
| $\hat{\beta}$ | 0.9166* | 0.9166* | 0.8394* | 0.8394* | 0.8519* | 0.8529* |

Table 1: LSTAR-GARCH Estimates

A possible explanation of these differences is that the Hessian matrix of the log-likelihood function of the LSTAR-GARCH model is not accurately approximated by BFGS, in which case the covariance matrix will also be unreliable. Another possibility is that there exists more than one optimum for the log-likelihood function, or that it is flat. This is supported by the similar mean likelihood scores of the two algorithms, as shown in Table 2.

| Index | Newton | BFGS |
|---|---|---|
| S&P | 4.24018 | 4.24018 |
| HSI | 3.75686 | 3.75664 |
| Nikkei | 3.92471 | 3.92487 |

Table 2: Maximum Likelihood Score

## 5 Aberrant Observations

Extreme observations are often referred to as being 2 to 3 standard deviations from the mean, while outliers are often defined as being more than 3 standard deviations from the mean. The difference between extreme observations and outliers is that outliers can also be defined as observations that were not generated from the same population as the other observations in the sample. Stock returns often contain more aberrant observations as compared with a normal distribution. Consequently, the distribution seems to have fatter tails, or excessive kurtosis, than a normal distribution.

In order to examine the effects of aberrant observations on the estimates for LSTAR-GARCH and ESTAR-GARCH, each of the data sets is adjusted using the following trimming algorithm:

1. Calculate the standard deviation for the sample;
2. If an observation is 4, between 3 and 4, between 2.5 and 3 times larger than the standard deviation, it is reduced to 4, 3, 2.5 times the standard deviation, respectively;
3. Repeat steps 1 and 2 above for every observation in the sample.

An LSTAR-GARCH model is estimated by the Newton algorithm using both adjusted and unadjusted S&P, HSI and Nikkei data. The estimates can be found in Table 3.

As shown in Table 3, $\phi_{12}$ exceeds 1 using both adjusted and unadjusted S&P data, which suggests that the first regime follows a non-stationary process. The same is also true for HSI using unadjusted data. In all three cases, the estimates of $\hat{\alpha}$ decreased and those of $\hat{\beta}$ increased when the data were adjusted. However, the effects of aberrant observations on the estimates of STAR are not entirely clear, though it appears that, if such data have a positive (negative) effect on $\hat{c}$, they will have a negative (positive) effect on $\hat{\gamma}$. This is an unusual result as there is no obvious reason why the threshold value should be related to the transition rate. However, for S&P and HSI, the threshold values are closer to 0 for the adjusted data, suggesting that the adjusted data exhibit asymmetric behaviour. Moreover, it appears that if $\hat{\phi}_{12}$ increases (decreases) after the data are adjusted, then $\hat{\phi}_{22}$ will also increase (decrease), which suggests that aberrant observations have the same effects on the coefficients of $y_{t-1}$ in both regimes.

A similar analysis is conducted to examine the effects of aberrant observations for ESTAR-GARCH. The estimates are given in Table 4.

The estimates of ESTAR-GARCH using the unadjusted data seem more plausible than those for LSTAR-GARCH in Table 3. All the regimes follow a stationary AR(1) process, but the low threshold values suggest that the second regime dominates the first for all three indices. Furthermore, the estimates for the GARCH component are very similar to those for LSTAR-GARCH. For Nikkei, the GARCH estimates are identical, which suggests that the choice of transition function for the conditional mean does not affect the GARCH estimates at all.

| | S&P | | HSI | | Nikkei | |
|---|---|---|---|---|---|---|
| | Unadjusted | Adjusted | Unadjusted | Adjusted | Unadjusted | Adjusted |
| $\hat{\phi}_{11}$ | 1.7833* | -1.7092* | -0.2238* | -0.5797* | -0.2796* | -0.6837* |
| $\hat{\phi}_{12}$ | 1.1204* | 2.1950* | 2.0523* | -4.4924* | -0.1930* | 0.8175* |
| $\hat{\phi}_{21}$ | -0.2862* | 1.5478* | 0.2260* | 0.6832* | 0.0013* | 0.3890* |
| $\hat{\phi}_{22}$ | -1.3476* | -0.2446* | -1.8170* | -4.9560* | 0.0126 | 0.1554* |
| $\hat{\gamma}$ | 2.4515* | 2.9175* | 4.3080* | 15.6068* | 1.6278* | 1.0431* |
| $\hat{c}$ | 0.8634* | 0.1849* | -0.0162 | -0.1252* | -2.0104* | 0.7372* |
| $\hat{\omega}$ | 1.17e-6 | 3.87e-7* | 7.08e-6* | 5.44e-6* | 3.22e-6* | 2.05e-6* |
| $\hat{\alpha}$ | 0.0753* | 0.0378* | 0.1488* | 0.1022* | 0.1475* | 0.1017* |
| $\hat{\beta}$ | 0.9166* | 0.9579* | 0.8394* | 0.8750* | 0.8519* | 0.8909* |

Table 3: LSTAR-GARCH Estimates for Adjusted and Unadjusted Data

| | S&P | | HSI | | Nikkei | |
|---|---|---|---|---|---|---|
| | Unadjusted | Adjusted | Unadjusted | Adjusted | Unadjusted | Adjusted |
| $\hat{\phi}_{11}$ | -0.7497* | -0.5490* | -0.9955* | -0.4233* | -0.4259* | -0.2144* |
| $\hat{\phi}_{12}$ | 0.0790 | -2.3958* | -0.9307* | 0.0004 | 0.1168* | 0.6474* |
| $\hat{\phi}_{21}$ | 0.0008* | 1.8336* | 0.0013* | 0.0013* | 0.0011* | 0.3901* |
| $\hat{\phi}_{22}$ | 0.0360 | 0.9974* | 0.1269* | 0.1126* | 0.0137 | 0.1911* |
| $\hat{\gamma}$ | 0.9474* | 0.7736* | 0.7997* | 1.2412* | 1.6688* | 0.8317* |
| $\hat{c}$ | -3.0813* | -0.5821* | -4.3532* | -2.5382* | -2.1846* | 0.7273* |
| $\hat{\omega}$ | 1.22e-6* | 3.87e-7* | 7.27e-7* | 6.07e-6* | 3.22e-6* | 2.05e-6* |
| $\hat{\alpha}$ | 0.0767* | 0.0378* | 0.1506* | 0.1076* | 0.1475* | 0.1017* |
| $\hat{\beta}$ | 0.9149* | 0.9579* | 0.8370* | 0.8669* | 0.8519* | 0.8909* |

Table 4: ESTAR-GARCH Estimates for Adjusted and Unadjusted Data

The effects of aberrant observations on the transition rate, $\hat{\gamma}$, and the threshold value, $\hat{c}$, for ESTAR-GARCH seem to be different from LSTAR-GARCH. In particular, the inverse relationship between $\hat{\gamma}$ and $\hat{c}$ no longer holds. However, the estimated threshold values increased when the adjusted data were used. For all three data sets, the estimated threshold values using the adjusted data were closer to 0 than for the unadjusted data. Furthermore, the sign of the estimated threshold value changed from negative to positive for Nikkei.

However, the effects of aberrant observations on the estimates of the transition rates are unclear. The effects of such data on $\hat{\phi}_{12}$ and $\hat{\phi}_{22}$ for ESTAR-GARCH are different from LSTAR-GARCH. In this case, if aberrant observations have positive (negative) effects on $\hat{\phi}_{12}$, they have negative (positive) effects on $\hat{\phi}_{22}$ for both S&P and HSI. The aberrant observations have negative effects on both $\hat{\phi}_{12}$ and $\hat{\phi}_{22}$ for Nikkei.

There does not seem to be a clear pattern between the estimates using adjusted and unadjusted data. However, due to the increase in the threshold value, the second regime no longer dominates the first, so that the adjusted data exhibit regime switching behaviour. Empirical evidence suggests that the estimate of the threshold value is sensitive to the sign and magnitude of outliers. If the magnitude of positive outliers is greater (smaller) than their negative counterparts, the threshold estimate will be greater (smaller) than 0. This result explains the low threshold estimates, because the magnitude of the negative outliers is often larger than the positive outliers for all three indices, and also the increase in the threshold estimates when the magnitude of the outliers is reduced. These results also show that the effects of aberrant observations on the estimates of STAR-GARCH are sensitive to the choice of transition function.

## 6 Conclusion

This paper provided a brief survey of recent developments for analysing the GARCH, STAR and STAR-GARCH models. Empirical evidence using the S&P, Hang Sang and Nikkei indexes showed that the QMLE for STAR-GARCH models were sensitive to the choice of optimisation algorithm. It was also shown that the estimates for STAR-GARCH are highly sensitive to aberrant observations. Furthermore, the effects of aberrant observations on the estimates of STAR-GARCH depend on the choice of transition function. The convergence of the estimates can be highly sensitive to the choice of algorithm.

## 7 Acknowledgments

The authors are grateful for helpful comments and suggestions from seminar participants at Bocconi University, Carlos III University of Madrid, Complutense University, Erasmus University Rotterdam, National University of Singapore, University of Amsterdam, University of the Basque Country, and University of Venice, and to Rob Engle, Clive Granger and Timo Teräsvirta for helpful discussions. The first author wishes to acknowledge an Australian Postgraduate Award and an Individual Research Grant at the University of Western Australia. The second author gratefully acknowledges the financial support of the Australian Research Council.

## 8 References

E.K. Berndt, B.H. Hall, R.E. Hall, and J.A. Hausman. Estimation and Inference in Non-linear Structural Models. *Annals of Economic and Social Measurement*, 3:653–665, 1974.

T. Bollerslev. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327, 1986.

C. Brooks, S.P. Burke, and G. Persand. Benchmarks and the Accuracy of GARCH Model Estimation. *International Journal of Forecasting*, 17:45–56, 2001.

F. Chan and M. McAleer. Maximum Likelihood Estimation of STAR and STAR-GARCH Models: A Monte Carlo Analysis. To appear in *Journal of Applied Econometrics*, 2002.

C.H. Chen, editor. *Pattern Recognition and Signal Processing.* Amsterdam: Sijhoff and Noordoff, 1978.

D. van Dijk, T. Teräsvirta, and P.H. Franses. Smooth Transition Autoregressive Models - A Survey of Recent Developments. To appear in *Econometric Reviews*, 2002.

R.F. Engle. Autoregressive Conditional Heteroscedasticity, with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50:987–1007, 1982.

C. He and T. Terašvirta. Properties of Moments of a Family of GARCH Processes. *Journal of Econometrics*, 92:173–192, 1999a.

C. He and T. Teräsvirta. Fourth Moment Structure of the GARCH$(p, q)$ Process. *Econometric Theory*, 15:824–846, 1999b.

S. Ling. On the Probabilitistic Properties of a Double Threshold ARMA Conditional Heteroscedasticity Model. *Journal of Applied Probability*, 36:1–18, 1999.

S. Ling and W.K. Li. On Fractional Integrated Autoregressive Moving Average Time Series Models with Conditional Heteroscedasticity. *Journal of the American Statistical Association*, 92:1184–1194, 1997.

S. Ling and M. McAleer. Necessary and Sufficient Moment Conditions for the GARCH$(r, s)$ and Asymmetric Power GARCH$(r, s)$ Models. *Econometric Theory*, 18:722–729, 2002a.

S. Ling and M. McAleer. Stationarity and the Existence of Moments of a Family of GARCH Processes. *Journal of Econometrics*, 106:109–117, 2002b.

S. Ling and M. McAleer. Asymptotic Theory for a Vector ARMA-GARCH Model. To appear in *Econometric Theory.* 2002c.

S. Lundbergh and T. Teräsvirta. Modelling Economic High Frequency Time Series with STAR-STGARCH Models. *SSE/EFI Working Paper Series in Economics and Finance*, No. 291, 1999.

S. Lundbergh and T. Teräsvirta. Forecasting with Smooth Transition Autoregressive Models. *SSE/EFI Working Paper Series in Economics and Finance*, No. 390, 2000.

S.G. Pantula. Estimation of Autoregressive Models with ARCH Errors. *Sankhya B*, 50:119–138, 1989.

T. Terašvirta. Specification, Estimation and Evaluation of Smooth Transition Autoregressive Models. *Journal of the American Statistical Association*, 89:208–218, 1994.

H. Tong. On a Threshold Model. In Chen (1978), pages 101-141.

H. Tong and K.S. Lim. Threshold Autoregressive, Limit Cycles and Data. *Journal of the Royal Statistical Society B*, 42:245–292, 1980.

R.L. Tweedie. Invariant Measure for Markov Chains with No Irreducibility Assumptions. *Journal of Applied Probability*, 25A:275–285, 1988.

A.A. Weiss. Asymptotic Theory for ARCH Models: Estimation and Testing. *Econometric Theory*, 2:107–131, 1986.