Evaluating the Forecast Performance of Smooth Transition Volatility Models

F. Chan and M. McAleer

Department of Economics, University of Western Australia (fchan@ecel.uwa.edu.au)

Abstract: Interest in non-linear time series models has increased dramatically in recent years. Regime switching models such as the Smooth Transition Autoregressive (STAR) Model and the Markov Switching Model have been the most popular in the class of non-linear models. The lack of established structural and statistical properties of these models has prevented a formal evaluation of their adequacy and validity. An alternative empirical strategy is to evaluate different models based on their forecast performance. This paper discusses several issues regarding the forecast performance of the STAR-Generalised Autoregressive Conditional Heteroscedasticity (GARCH), or STAR-GARCH model, and the STAR-Smooth Transition GARCH (or STAR-STGARCH) model. The forecast performance of each model is evaluated for two important stock indexes, namely Standard and Poor's Composite 500 Index and the Hang Seng Index, using two different forecasting criteria. An emphasis is placed on the importance of obtaining structural and statistical properties of the models, which have generally been ignored in the literature. Moreover, issues regarding the choice of an appropriate algorithm for maximising the likelihood function are also discussed. It is shown that different algorithms can produce different parameter estimates for similar likelihood values. Moreover, different parameter estimates lead to different forecasts and different forecasting performance. A simple trimming method designed to reduce the effects of extreme observations and outliers is used to evaluate the effects of these observations on forecast performance.

Keywords: Volatility; Smooth Transition; GARCH; Forecasting

1 INTRODUCTION

Interest in non-linear time series models has increased dramatically over the past decade, with the Smooth Transition Autoregressive (STAR) Model and the Markov Switching (MSW) Model being the most popular. Combining various STAR-type models with Bollerslev's [1986] Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model has now become established in the financial volatility literature. This combination has resulted in two highly flexible non-linear models of financial volatility, namely STAR-GARCH and STAR-Smooth Transition GARCH (STAR-STGARCH) models (see van Dijk et al. [2000] and Lundburgh and Teräsvirta [2000] for recent developments of STAR models).

Despite their flexibility, the structural and statistical properties of these models are essentially unknown, which prevents a reliable evaluation of their adequacy and validity. An alternative empirical approach is to evaluate their forecast performance. This paper evaluates the forecast performance of six different STAR-type models using the returns of two important stock indexes, namely Standard and Poor's 500 Composite Index and the Hang Seng Index. The out-of-sample forecast is evaluated according to two different criteria, namely Mean Square Error (MSE) and Mean Absolute Error (MAE).

The paper addresses important empirical issues regarding the Maximum Likelihood Estimators (MLE) of STAR-type models. Specifically, the paper presents empirical evidence which shows that the MLE is sensitive to the choice of algorithm for maximising the likelihood function, but that the likelihood values may remain essentially unchanged. More importantly, model selection based on forecast performance may be sensitive to the algorithm chosen to estimate the models, which raises issues regarding the validity and reliability of MLE for such models.

Furthermore, reducing the magnitudes of extreme observations and outliers within the sample may be ineffective for improving the out-of-sample forecast. Indeed, in some cases, such transformations can worsen the forecast performance of individual models. The paper presents empirical evidence to support these claims.

The structure of the paper is as follows. Section 2 discusses some issues relating to STAR-type models. Information on the data and the methodology adopted in the paper are given in section 3. Empirical results are presented in section 4, and section 5 provides some concluding remarks.

2 MODELS

A simple first-order STAR model with two regimes is defined as follows:

$$\begin{aligned} y_t =& (\phi_{11} + \phi_{12} y_{t-1})(1 - G(s_t; \gamma, c)) + \\ & (\phi_{21} + \phi_{22} y_{t-1}) G(s_t; \gamma, c) + \varepsilon_t, \end{aligned}$$

where $G(s_t; \gamma, c)$ is the transition function, assumed to be twice differentiable and bounded between 0 and 1, γ is the transition rate, and c is the threshold value. Although there are few theoretical results regarding the stationarity of the STAR model, a sufficient condition is $\phi_{ij} < 1 \ \forall i, j$ (see van Dijk et al. [2001] for further discussion). The transition variable, s_t, is usually (but not always) defined as a linear combination of the lagged values of y_t, namely

$$s_t = \sum_{i=1}^{k} \theta_i y_{t-i}.$$

The STAR model was proposed by Teräsvirta [1994] as an extension of Tong [1978]. Lundbergh and Teräsvirta [2000] extended the STAR model by specifying the error term to follow a GARCH(p, q) process, as defined in Bollerslev [1986]:

$$\begin{aligned} \varepsilon_t &= \eta_t \sqrt{h_t}, \\ h_t &= \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j h_{t-j}. \end{aligned}$$

Sufficient conditions for $h_t > 0$ are $\omega > 0$, $\alpha_i \geq 0$ $\forall i = 1, .., p$, and $\beta_j \geq 0 \ \forall j = 1, .., q$.

Lundbergh and Teräsvirta [2000] extended the model further by incorporating the concept of smooth transition into the GARCH component, namely the STAR-STGARCH model:

$$\begin{aligned} h_t =& (\omega_1 + \sum_{i=1}^{p} \alpha_{1i} \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_{1i} h_{t-i}) \times \\ & (1 - H(p_t; \xi, d)) + \\ & (\omega_2 + \sum_{i=1}^{p} \alpha_{2i} \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_{2i} h_{t-i}) \times \\ & H(p_t; \xi, d), \end{aligned}$$

where $H(p_t; \xi, d)$ satisfies the same conditions as $G(s_t; \gamma, c)$, ξ is the transition rate, and d is the threshold value. Regarding the choice of transition function, the two most widely used in the literature are the first-order logistic function

$$L(s_t; \gamma, c) = \frac{1}{1 + \exp(-\gamma(s_t - c))}, \qquad \gamma \geq 0,$$

and the first-order exponential function

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

STAR models with logistic transition functions are denoted Logistic STAR (or LSTAR) models, while STAR models with Exponential transition functions are denoted Exponential STAR (or ESTAR) models.

The models considered in this paper are as follows:

1. Autoregressive (AR(1))-GARCH (AR)
2. Logistic STAR - GARCH (LSTAR)
3. Exponential STAR - GARCH (ESTAR)
4. Logistic STAR - Logistic Smooth Transition GARCH (LSTARLSTG)
5. Exponential STAR - Exponential Smooth Transition GARCH (ESTARESTG)
6. Logistic STAR - Exponential Smooth Transition GARCH (LSTARESTG)
7. Exponential STAR - Logistic Smooth Transition GARCH (ESTARLSTG).

Unless otherwise stated, all the estimated STAR-type models have two regimes, and each follows an AR(1) process with $s_t = y_{t-1}$. Moreover, ε_t is assumed to follow a GARCH(1,1) process and, in the case of Smooth Transition GARCH-type models, $p_t = \varepsilon_{t-1}$.

3 DATA AND METHODOLOGY

3.1 Data

The returns of two stock indexes are used to estimate all the models given above, namely Standard and Poor's Composite 500 Index (S&P) and the Hang Seng Index (HS). Returns on these indexes are calculated as follows:

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

Data are obtained from the DataStream database service, with the sample period 1/1/1986 to 1/4/2000, giving a total of 3726 data points for each index. The first 3226 data points are used for estimation, leaving the last 500 data points for evaluating the one-day ahead out-of-sample forecasts.

3.2 Methodology

Each of the seven models is estimated twice by applying two different algorithms to maximise their log-likelihood functions. The effects of algorithmic choice on model forecast performance can be investigated by comparing the out-of-sample forecast performance based on the two sets of estimates.

Two forecasting criteria are calculated for the purpose of comparison, namely Mean Square Error (MSE) and Mean Absolute Error (MAE). Forecast performance between models estimated by the same algorithm are compared, then a comparison is made between the best models from each algorithm.

The algorithms used in this paper are the Newton and Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods. There are two main reasons for choosing these algorithms. First, they belong to two different classes of optimisation algorithms, namely Newton and Quasi-Newton. The difference between these two classes is that Newton requires the use of the analytical Hessian matrix of the objective function, which is often costly and time consuming to compute. Quasi-Newton requires only the numerical approximation of the Hessian matrix, which can be obtained efficiently by methods such as finite-differences. The results can be unreliable if the approximation of the Hessian matrix is inaccurate. Moreover, the covariance matrix, which is typically estimated as the inverse of the Hessian matrix, can also be affected, so that inferences may be similarly affected.

Second, these two algorithms are the most popular algorithms among other optimisation algorithms in their respective classes. Most computer software packages, such as GAUSS and MatLab, contain options for both algorithms.

As noted in van Dijk et al. [2000] and Brooks, Burke and Persands [2001], the estimation of the MLE for GARCH and STAR-type models appears to be sensitive to the choice of initial values. For this reason, all the models estimated in this paper used the same initial values.

A simple trimming algorithm was implemented to examine the effects of extreme observations and outliers on the forecast performance of each model.

The algorithm can be summarised as follows:

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

4 EMPIRICAL RESULTS

4.1 Estimates

The two algorithms seemed to produce the same estimates for the simple AR(1)-GARCH(1,1) model, but the estimates of STAR-type models differed substantially. As shown in Table 1, Newton estimates of the STAR-GARCH models differed substantially from BFGS estimates in the conditional mean, but not in the conditional variance. This reflects the block diagonal nature of the information matrix.

As the information matrices of STAR-STGARCH models are generally not block diagonal, it is not surprising to find that the BFGS estimates differed substantially from Newton estimates in both the conditional mean and the conditional variance, as given in Table 2.

Although the estimates obtained from using different algorithms can be substantially different, their log-likelihood values remain similar, and in some cases are identical. This suggests that the likelihood functions of these models are flat, as suggested in van Dijk et al. [2000].

The following implications follow from the estimates:

1. If the MLE is sensitive to the choice of algorithm, it is possible to obtain two different interpretations for the same model using the same data set. This is worrying because the interpretation of the empirical models should be independent of the method of estimation.
2. Differences in estimates can lead to a situation whereby BFGS estimates satisfy various regularity conditions (such as stationarity, ergodicity, consistency and asymptotic normality), while Newton estimates satisfy only some or none of these conditions, and vice-versa.
3. Stationarity may be of concern because the estimated coefficients of the lagged dependent variable often exceed one. However, there are presently no theoretical results regarding the stationarity of STAR-type models, so it is unclear whether the estimated models are stationary (see van Dijk et al. [2000] for further discussion regarding the stationarity of STAR-type models).
4. There are two plausible explanations for the differences in estimates obtained by the two algorithms, namely that (Q)MLE is not an appropriate estimator for these models, or that the optimisation algorithms are inappropriate. It is difficult to determine the specific cause

without understanding the structural and statistical properties of these models. Thus, a deeper understanding of these properties is crucial for applying these models sensibly.

4.2 Forecast Performance

Similar likelihood values might suggest that these models are likely to produce similar forecast performance. Although this is true to some extent, model selection based on forecast performance seems to be sensitive to the choice of optimisation algorithms. Table 3 provides the list of models that produced the best forecasts based on two different criteria, namely Mean Square Error (MSE) and Mean Absolute Error (MAE).

Denote MSE(BFGS) as MSE from models estimated by BFGS, and MAE(Newton) as MAE from models estimated by Newton. According to MSE(BFGS), the best model to forecast S&P returns is LSTAR. However, the forecast performance based on models estimated by Newton suggests that ESTARESTG is optimal. Similar observations are obtained for adjusted S&P and adjusted Hang Seng. ESTARLSTG appears to be superior for the adjusted S&P according to MSE(BFGS), but MSE(Newton) suggests that LSTARLSTG produces the smallest forecast error. A similar conclusion holds for MAE. LSTAR produces the minimum MAE(BFGS) for S&P, but ESTARESTG is marginally superior according to MAE(Newton). In fact, given the same algorithm, both forecast criteria seem to select the same model for the mean.

This outcome does not hold for the variance. According to MSE(BFGS), a simple GARCH model seems to be best for predicting volatility for S&P, adjusted S&P and Hang Seng. However, this is not supported by MAE(BFGS). It is well known that MSE penalises large errors, so there is a temptation to conclude that GARCH(1,1) captures outliers and extreme observations better than do STAR-type models. This is not necessarily the case, as it is not clear whether these models are correctly specified. Since there is no theoretical result regarding the effects of misspecification on the MLE of STAR-type models, their ability to accommodate extreme observations and outliers is unclear. It is also worth noting that the existing specification tests for STAR-type models rely on several rather restrictive assumptions, such as normality, stationarity and the existence of the necessary moments (see Lundbergh and Teräsvirta [1999] and van Dijk et al. [2000] for informal derivations of these tests). As the regularity conditions necessary for a formal derviation of these tests are still unknown, the validity and reliability of these tests remain questionable.

An interesting observation from table 3 is that the best model for the mean (return) is not necessarily the best model for the variance (volatility). Indeed, in most cases, the best model for the mean is, in fact, different from that for the variance. This interesting pattern may reflect the fact that:

1. models are inadequate or misspecified;
2. MLE is not robust for these models;
3. there is a trade-off between power in predicting the mean and variance.

Table 4 contains the differences in forecast errors between models estimated using BFGS and Newton. It is interesting to note that Newton estimates for LSTARESTG seem to perform worse in forecasting volatility than BFGS estimates. It is also interesting and comforting to note that GARCH appears to be robust and invariant to the choice of optimisation algorithm, though it is sensitive to initial conditions, as noted above.

It is not clear which algorithm is likely to produce estimates that minimise the forecast errors. Interestingly, Newton seems to produce superior estimates for predicting the mean according to Table 3, but it is unclear which algorithm produces superior estimates for predicting volatility.

4.3 Effects of Extreme Observations and Outliers

Another interesting observation from the results is that symmetric trimming of outliers and extreme observations does not seem to improve forecast performance. In most cases, forecast performance appears to be worse after such adjustment. However, this conclusion is based on the assumption that these models are correctly specified. As the effects of misspecification of non-linear time series models are generally unknown, it is difficult to draw strong conclusions about the effects of outliers and extreme observations on the MLE of STAR-type models. It is clear, however, that outliers and extreme observations should be handled with caution. Simple symmetric trimming algorithms may not reduce the effects of extreme observations and outliers in nonlinear models, and they do not necessarily improve out-of-sample forecast performance. If the underlying distribution of the process is asymmetric, then symmetric trimming may induce biased estimates. Moreover, if the sample used for forecasting contains an excessive number of extreme observations and outliers, it would not be surprising for the estimated models based on the adjusted data to perform poorly.

Table 1: STAR-GARCH estimates

S&P(LSTAR)	$\hat{\phi}_{11}$	$\hat{\phi}_{12}$	$\hat{\phi}_{21}$	$\hat{\phi}_{22}$	$\hat{\gamma}$	$\hat{c}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
Newton	0.0431	-2.1097	-0.0420	2.2661	1.6996	0.0016	1.3793e-6	0.0782	0.9095
BFGS	-0.0341	-2.7766	0.0389	3.0998	1.2579	0.0788	1.3807e-6	0.0782	0.9095
HS(ESTAR)	$\hat{\phi}_{11}$	$\hat{\phi}_{12}$	$\hat{\phi}_{21}$	$\hat{\phi}_{22}$	$\hat{\gamma}$	$\hat{c}$	$\hat{\omega}$	$\hat{\alpha}$	$\hat{\beta}$
Newton	-0.3415	-1.3078	0.2538	-0.3164	3.4741	-0.4967	8.136e-6	0.1626	0.8201
BFGS	-0.1516	-2.8288	0.0089	-0.0701	217.7126	-0.1183	8.7847e-6	0.1511	0.8205

Table 2: STAR-STGARCH estimates

S&P(LSTARESTG)									
Newton	$\hat{\phi}_{11}$	$\hat{\phi}_{12}$	$\hat{\phi}_{21}$	$\hat{\phi}_{22}$	$\hat{\gamma}$	$\hat{c}$	$\hat{\omega}_1$	$\hat{\alpha}_{11}$	$\hat{\beta}_{11}$
	-0.2826	-0.2860	0.0021	0.0338	2.2733	-2.3266	1.4137e-6	0.0799	0.9069
	$\hat{\omega}_2$	$\hat{\alpha}_{21}$	$\hat{\beta}_{21}$	$\hat{\xi}$	$\hat{d}$				
	0.006196	0.1566	0.1098	3.2164	3.0245				
BFGS	$\hat{\phi}_{11}$	$\hat{\phi}_{12}$	$\hat{\phi}_{21}$	$\hat{\phi}_{22}$	$\hat{\gamma}$	$\hat{c}$	$\hat{\omega}_1$	$\hat{\alpha}_{11}$	$\hat{\beta}_{11}$
	52.8016	57.1852	-2.1135	3.8145	2.8685	-1.1218	8.0038e-6	0.1627	0.8197
	$\hat{\omega}_2$	$\hat{\alpha}_{21}$	$\hat{\beta}_{21}$	$\hat{\xi}$	$\hat{d}$				
	1.3800e-6	0.07817	0.9095	3.0930	2.9143				
HS(LSTARLSTG)									
Newton	$\hat{\phi}_{11}$	$\hat{\phi}_{12}$	$\hat{\phi}_{21}$	$\hat{\phi}_{22}$	$\hat{\gamma}$	$\hat{c}$	$\hat{\omega}_1$	$\hat{\alpha}_{11}$	$\hat{\beta}_{11}$
	0.8093	-3.2471	-0.0034	0.1961	8.6617	-0.5962	8.2113e-6	0.1635	0.8190
	$\hat{\omega}_2$	$\hat{\alpha}_{21}$	$\hat{\beta}_{21}$	$\hat{\xi}$	$\hat{d}$				
	8.2302e-6	0.1190	0.8312	3.0935	2.9147				
BFGS	$\hat{\phi}_{11}$	$\hat{\phi}_{12}$	$\hat{\phi}_{21}$	$\hat{\phi}_{22}$	$\hat{\gamma}$	$\hat{c}$	$\hat{\omega}_1$	$\hat{\alpha}_{11}$	$\hat{\beta}_{11}$
	-0.0390	-0.9386	0.0304	-0.5178	55.0001	-0.0056	2.676e-5	0.5298	0.8268
	$\hat{\omega}_2$	$\hat{\alpha}_{21}$	$\hat{\beta}_{21}$	$\hat{\xi}$	$\hat{d}$				
	9.7314e-9	0.02056	0	0.0274	-32.9896				

Table 3: Forecasting under different criteria and error measurements[1]

Mean	MSE		MAE	
	BFGS	NEWTON	BFGS	NEWTON
S&P	LSTAR	**ESTARESTG**	**LSTAR**	ESTARESTG
Adjusted S&P	ESTARLSTG	**LSTARLSTG**	ESTARLSTG	**LSTARLSTG**
HS	AR	AR	AR	AR
Adjusted HS	ESTARESTG	**LSTAR**	ESTARESTG	**AR**
Variance	MSE		MAE	
	BFGS	NEWTON	BFGS	NEWTON
S&P	**AR**	LSTARLSTG	AR	AR
Adjusted S&P	**AR**	ESTARLSTG	**LSTAR**	LSTARLSTG
HS	AR	AR	ESTARLSTG	**AR**
Adjusted HS	ESTARESTG	**LSTAR**	ESTARLSTG	**LSTAR**

[1] The better model from using the two algorithms is given in bold letters.

Table 4: Differences in forecast errors between algorithms

Mean	MSE AR	LSTAR	ESTAR	LSTARLSTG	LSTARESTG	ESTARESTG	ESTARLSTG
S&P	0.000E+00	-1.800E-09	-1.657E-04	-1.150E-06	6.480E-08	1.770E-07	-1.240E-06
Adjusted S&P	0.000E+00	-2.310E-08	1.100E-09	5.116E-07	1.777E-04	9.087E-05	-3.963E-06
HS	0.000E+00	8.038E-06	1.861E-05	1.989E-05	-1.748E-07	-1.996E-07	2.040E-05
Adjusted HS	0.000E+00	1.492E-05	7.704E-07	-2.161E-05	-2.154E-05	-3.565E-05	8.908E-07
Mean	**MAE AR**	**LSTAR**	**ESTAR**	**LSTARLSTG**	**LSTARESTG**	**ESTARESTG**	**ESTARLSTG**
S&P	0.0000E+00	-6.7200E-07	-9.5685E-03	-1.7665E-05	2.1680E-06	8.4470E-06	-2.5399E-05
Adjusted S&P	0.0000E+00	-5.6700E-07	-1.7010E-06	1.1326E-05	5.3831E-04	3.7000E-04	-9.1308E-05
HS	0.0000E+00	9.0270E-05	1.5117E-04	2.2513E-04	-9.2000E-06	-1.4410E-05	1.6277E-04
Adjusted HS	0.0000E+00	8.6460E-05	6.9200E-06	-3.2906E-04	-3.2798E-04	-4.7782E-04	6.3400E-06
Variance	**MSE AR**	**LSTAR**	**ESTAR**	**LSTARLSTG**	**LSTARESTG**	**ESTARESTG**	**ESTARLSTG**
S&P	0.0000E+00	8.7100E-12	-9.6724E-08	2.0260E-08	-2.1730E-11	6.5327E-09	-1.1445E-08
Adjusted S&P	0.0000E+00	5.4000E-12	1.9600E-11	3.2613E-08	-9.2282E-04	-9.2426E-04	1.7432E-08
HS	0.0000E+00	9.7587E-09	1.2381E-08	9.6747E-08	-5.4980E-10	-2.3140E-09	6.1286E-08
Adjusted HS	0.0000E+00	6.5456E-09	5.0020E-10	-1.4726E-09	-1.4972E-09	-1.4131E-08	-1.1816E-08
Variance	**MAE AR**	**LSTAR**	**ESTAR**	**LSTARLSTG**	**LSTARESTG**	**ESTARESTG**	**ESTARLSTG**
S&P	0.0000E+00	1.6100E-08	-1.6149E-04	6.7180E-05	-2.9300E-08	4.3186E-05	-1.8195E-05
Adjusted S&P	0.0000E+00	-7.3000E-09	2.4700E-08	1.0135E-04	-2.7014E-02	-2.7099E-02	6.9709E-05
HS	0.0000E+00	-4.3238E-06	6.2320E-07	-6.9701E-05	-2.4160E-07	-1.0795E-06	-8.6589E-05
Adjusted HS	0.0000E+00	6.7106E-06	3.7370E-07	-7.7246E-06	-7.7509E-06	-1.5697E-05	-2.0927E-05

5 CONCLUSION

Forecast performance based on the MLE of a simple GARCH model and six different STAR-type models has been evaluated. This paper has presented empirical evidence to show that MLE is sensitive to the choice of optimisation algorithm, so that model selection based on forecast performance may be affected. Ideally, it would be desirable to estimate these models using several different algorithms and to evaluate their forecast performance. It would appear that some algorithms are more appropriate than others, and it is difficult to select the most appropriate and efficient algorithm without an adequate understanding of the structure of the likelihood functions.

Another important empirical finding is that the best model for the mean is not necessarily that for the variance. These findings raise serious issues regarding the robustness and reliability of MLE for STAR-type models.

The problems raised above emphasize the importance of obtaining structural and statistical properties of STAR-type models and their likelihood functions, and would seem to be crucial in order to apply the models correctly. Other important issues, such as the effects of misspecification on MLE, should be investigated further for purposes of determining the validity and reliability of such models for practical purposes.

ACKNOWLEDGEMENTS

The first author would like to acknowledge an Australian Postgraduate Award and an Individual Research Grant, Faculties of Economics and Commerce, Education and Law, University of Western Australia, for financial support. The second author wishes to acknowledge the financial support of the Australian Research Council.

REFERENCES

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

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

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

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

Lundbergh, S. 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.

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

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

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