Studies in Nonlinear Dynamics & Econometrics

Volume 15, *Issue* 2 2011 *Article* 1

Contemporaneous-Threshold Smooth Transition GARCH Models

Michael J. Dueker* Zacharias Psaradakis†

Martin Sola‡ Fabio Spagnolo**

*Russell Investments, mdueker@russell.com

†Birkbeck, University of London, zpsaradakis@econ.bbk.ac.uk

‡Birkbeck, University of London and Universidad Torcuato Di Tella, msola@econ.bbk.ac.uk

**Brunel University, fabio.spagnolo@brunel.ac.uk

Contemporaneous-Threshold Smooth Transition GARCH Models*

Michael J. Dueker, Zacharias Psaradakis, Martin Sola, and Fabio Spagnolo

Abstract

This paper proposes a contemporaneous-threshold smooth transition GARCH (or C-STGARCH) model for dynamic conditional heteroskedasticity. The C-STGARCH model is a generalization to second conditional moments of the contemporaneous smooth transition threshold autoregressive model of Dueker et al. (2007) in which the regime weights depend on the ex ante probability that a contemporaneous latent regime-specific variable exceeds a threshold value. A key feature of the C-STGARCH model is that its transition function depends on all the parameters of the model as well as on the data. The structural properties of the model are investigated, in addition to the finite-sample properties of the maximum likelihood estimator of its parameters. An application to U.S. stock returns illustrates the practical usefulness of the C-STGARCH model.

*We would like to thank Nicolas Caramp for research assistance.

1 Introduction

A general class of nonlinear models of dynamic heteroskedasticity is the class of generalized autoregressive conditionally heteroskedastic (GARCH) models with regime-dependent parameters subject to smooth changes. The defining characteristic of such so-called smooth transition GARCH (STGARCH) models is that transitions between regimes are modelled by using a continuous function (usually logistic or exponential) of some observable transition variable.

Dueker et al. (2007) introduced recently a new class of contemporaneous-threshold smooth transition autoregressive (C-STAR) models in which the mixing (or regime) weights depend on the ex ante probabilities that regime-specific latent variables exceed certain threshold values. A key feature of the C-STAR model is that its mixing (or transition) function depends on all the parameters of the model as well as on the data, a feature which allows the model to describe time series with a wide variety of conditional distributions.

This paper contributes to the literature on nonlinear GARCH models by proposing a contemporaneous-threshold smooth transition GARCH (C-STGARCH) model motivated by the approach of Dueker et al. (2007). Our model differs from other members of the STGARCH family in two important respects. Firstly, unlike models where the argument of the mixing function used in the definition of the conditional variance is either the lagged (squared) innovation (e.g., Hagerud, 1996; González-Rivera, 1998; Lundbergh and Teräsvirta, 1998; Anderson et al., 1999; Lubrano, 2001; Medeiros and Veiga, 2009) or the lagged conditional variance (e.g., Lanne and Saikkonen, 2005), the mixing weights for C-STGARCH models are a function of both. Secondly, the mixing weights of a C-STGARCH model depend on all the parameters of the model. This implies that, unlike other STGARCH models, there is no need to choose an appropriate transition variable using a selection criterion since, by construction, all the variables that enter the model's information set also enter the transition function.

The paper is organized as follows. Section 2 introduces the model discusses its properties. Section 3 considers maximum likelihood (ML) estimation of the parameters of the model. Section 4 presents an illustrative empirical application to U.S. stock returns. Section 5 summarizes and concludes.

2 The C-STGARCH Model

The C-STGARCH model in this paper is a member of the family of STGARCH models. We say that a real-valued time series $\{y_t\}$ follows a STGARCH$(1,1)$ model if it satisfies the following equations:

$$y_t - \mu = \sigma_t u_t, \qquad t = 1, 2, \ldots, \tag{1}$$

$$\sigma_t^2 = G(\mathbf{z}_{t-1})\sigma_{0t}^2 + \{1 - G(\mathbf{z}_{t-1})\}\sigma_{1t}^2, \tag{2}$$

$$\sigma_{it}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{t-1}^2, \qquad i = 0, 1. \tag{3}$$

In (1)–(3), $\{u_t\}$ are independent and identically distributed (i.i.d.) random variables such that u_t is independent of y_{t-j} for $j \geq 1$ and $\mathsf{E}(u_t) = \mathsf{E}(u_t^2 - 1) = 0$, $G(\mathbf{z}_{t-1})$ is a continuous function of a vector of exogenous and/or predetermined variables $\mathbf{z}_{t-1}$ such that $0 \leq G(\mathbf{z}_{t-1}) \leq 1$, $\varepsilon_t := y_t - \mu$, and $\omega_i > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$ $(i = 0, 1)$ and μ are constants.

A popular choice for the mixing (or transition) function G in (2) is the logistic formulation

$$G(s_{t-1}) = \left[1 + \exp(-\gamma\{s_{t-1} - k\})\right]^{-1}, \qquad \gamma > 0, \tag{4}$$

where s_{t-1} is a so-called transition variable. The location parameter k in (4) may be interpreted as the threshold between the two regimes associated with the limiting values of $G(s_{t-1})$ (as s_{t-1} diverges to positive and negative infinity), while the slope parameter γ determines the smoothness of the transitions between the two regimes. Existing STGARCH models set s_{t-1} equal to ε_{t-1}, ε_{t-1}^2 or σ_{t-1}^2.

To define the contemporaneous-threshold STGARCH model, let F be the cumulative distribution function of u_t^2 (which is assumed to be non degenerate). The C-STGARCH$(1,1)$ model is formulated by specifying the mixing function G in (2) as

$$G(\mathbf{z}_{t-1}) = \frac{F(\dfrac{k}{\omega_0 + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2})}{F(\dfrac{k}{\omega_0 + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2}) + 1 - F(\dfrac{k}{\omega_1 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2})}, \tag{5}$$

where $\mathbf{z}_{t-1} = (\varepsilon_{t-1}^2, \sigma_{t-1}^2)^\intercal$ and k is a non-negative threshold parameter. It can be easily seen that

$$G(\mathbf{z}_{t-1}) = \frac{\mathsf{P}(\varepsilon_{0t}^2 < k | \mathbf{z}_{t-1})}{\mathsf{P}(\varepsilon_{0t}^2 < k | \mathbf{z}_{t-1}) + \mathsf{P}(\varepsilon_{1t}^2 \geq k | \mathbf{z}_{t-1})}, \tag{6}$$

where $\varepsilon_{it}^2 = \sigma_{it}^2 u_t^2$ $(i = 0, 1)$. Hence, under the C-STGARCH specification, (1) may be re-written as

$$\varepsilon_t^2 = \left\{ \frac{\mathsf{P}(\varepsilon_{0t}^2 < k|\mathbf{z}_{t-1})\sigma_{0t}^2 + \mathsf{P}(\varepsilon_{1t}^2 \geq k|\mathbf{z}_{t-1})\sigma_{1t}^2}{\mathsf{P}(\varepsilon_{0t}^2 < k|\mathbf{z}_{t-1}) + \mathsf{P}(\varepsilon_{1t}^2 \geq k|\mathbf{z}_{t-1})} \right\} u_t^2.$$

Since the mixing weights are determined by the probability that the contemporaneous latent variable ε_{0t}^2 (ε_{1t}^2) is below (above) the threshold level k, we call this a contemporaneous-threshold STGARCH model.

The first-order C-STGARCH model can be straightforwardly generalized to allow for higher order dynamics by replacing the specification in (3) with

$$\sigma_{it}^2 = \omega_i + \sum_{j=1}^{q} \alpha_{ij} \varepsilon_{t-j}^2 + \sum_{r=1}^{p} \beta_{ir} \sigma_{t-r}^2, \qquad i = 0, 1,$$

for some $p \geq 1$ and $q \geq 1$. The mixing function of the resulting $C - STGARCH(p, q)$ model is defined in a way analogous to (5)–(6) with $\mathbf{z}_{t-1} = (\varepsilon_{t-1}^2, ., \varepsilon_{t-q}^2, \sigma_{t-1}^2, ., \sigma_{t-p}^2)^\intercal$. For the sake of simplicity and clarity of exposition, and since the GARCH$(1, 1)$ specification is by far the most popular in applications, we shall focus hereafter on the C-STGARCH$(1, 1)$ model.

The C-STGARCH model differs from other models that belong to the STGARCH family in two notable respects. First, unlike models where the argument of the mixing function used in the definition of the conditional variance is ε_{t-1}, ε_{t-1}^2 or σ_{t-1}^2, the mixing weight in (5) is a function of both ε_{t-1}^2 and σ_{t-1}^2. Second, the mixing weights depend on all of the model parameters. This means that for a C-STGARCH there is no need to use any selection criteria to choose the appropriate threshold variables since, by construction, all the variables that enter the information set of the model are also present in the mixing function.

3 Properties of the C-STGARCH Model

In this section, we investigate some of the key characteristics of the C-STARCH model. In particular, we consider: (i) the stability of the model; (ii) the response of the mixing function to changes in the parameters of the model; (iii) the empirical distribution of data generated by the model; (iv) the news impact curve implied by the model; (v) implications of the model regarding persistence. In the discussion that follows, it is assumed that the parameters of the C-STGARCH$(1, 1)$ model satisfy the identification condition $\omega_0/(1 - \alpha_0 - \beta_0) < \omega_1/(1 - \alpha_1 - \beta_1)$, which is sufficient (but not necessary) for ensuring

that $F(k/\{\omega_0 + \alpha_0\varepsilon_{t-1}^2 + \beta_0\sigma_{t-1}^2\})$ and $1 - F(k/\{\omega_1 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2\})$ do not both tend to zero simultaneously. It is also assumed that F is strictly monotone and differentiable.

3.1 Stability of the Skeleton

As discussed in Tong (1990), the stability properties of a nonlinear model may be analyzed by considering the noiseless part, or skeleton, of the model. The skeleton of the C-STGARCH$(1,1)$ model is defined as

$$\bar{\varepsilon}_t^2 = S(\bar{\varepsilon}_{t-1}^2, \boldsymbol{\theta}),$$

where

$$S(\bar{\varepsilon}_{t-1}^2, \boldsymbol{\theta}) = \bar{G}(\bar{\varepsilon}_{t-1}^2)\{\omega_0 + (\alpha_0 + \beta_0)\bar{\varepsilon}_{t-1}^2\} + \{1 - \bar{G}(\bar{\varepsilon}_{t-1}^2)\}\{\omega_1 + (\alpha_1 + \beta_1)\bar{\varepsilon}_{t-1}^2\},$$

$$\bar{G}(\bar{\varepsilon}_{t-1}^2) = \frac{F(\dfrac{k}{\omega_0 + (\alpha_0 + \beta_0)\bar{\varepsilon}_{t-1}^2})}{F(\dfrac{k}{\omega_0 + (\alpha_0 + \beta_0)\bar{\varepsilon}_{t-1}^2}) + 1 - F(\dfrac{k}{\omega_1 + (\alpha_1 + \beta_1)\bar{\varepsilon}_{t-1}^2})},$$

and $\boldsymbol{\theta}$ denotes the vector of all the parameters of the model. A fixed point of the skeleton is any value $\bar{\varepsilon}_e^2$ which satisfies the equation

$$S(\bar{\varepsilon}_e^2, \boldsymbol{\theta}) = \bar{\varepsilon}_e^2, \tag{7}$$

and $\bar{\varepsilon}_e^2$ is said to be an equilibrium point of the model. Since the C-STGARCH model is nonlinear, there may be one, several or no equilibrium points satisfying (7).

An examination of the local stability of each equilibrium point may be carried out by considering the following first-order Taylor expansion of $S(\bar{\varepsilon}_{t-1}^2, \boldsymbol{\theta})$ about $\bar{\varepsilon}_e^2$:

$$\begin{aligned} \bar{\varepsilon}_t^2 - \bar{\varepsilon}_e^2 &= S(\bar{\varepsilon}_{t-1}^2, \boldsymbol{\theta}) - S(\bar{\varepsilon}_e^2, \boldsymbol{\theta}) \\ &\approx \left(\frac{\partial S(\bar{\varepsilon}_{t-1}^2, \boldsymbol{\theta})}{\partial \bar{\varepsilon}_{t-1}^2}\bigg|_{\bar{\varepsilon}_{t-1}^2 = \bar{\varepsilon}_e^2} \right) (\bar{\varepsilon}_{t-1}^2 - \bar{\varepsilon}_e^2). \end{aligned} \tag{8}$$

If the absolute value of the partial derivative in (8) is less than unity, then the equilibrium is locally stable and $\bar{\varepsilon}_t^2$ is a contraction in the neighborhood of $\bar{\varepsilon}_e^2$.

It is straightforward to verify that

$$\frac{\partial S(\bar{\varepsilon}^2_{t-1},\boldsymbol{\theta})}{\partial \bar{\varepsilon}^2_{t-1}} = (\alpha_1+\beta_1)+\{(\alpha_0+\beta_0)-(\alpha_1+\beta_1)\}\bar{G}(\bar{\varepsilon}^2_{t-1}) + \{(\omega_0-\omega_1)+[(\alpha_0+\beta_0)-(\alpha_1+\beta_1)]\bar{\varepsilon}^2_{t-1}\}\frac{\partial \bar{G}(\bar{\varepsilon}^2_{t-1})}{\partial \bar{\varepsilon}^2_{t-1}} \quad (9)$$

and

$$\frac{\partial \bar{G}(\bar{\varepsilon}^2_{t-1})}{\partial \bar{\varepsilon}^2_{t-1}} = -\frac{\{1-F(\tau_1)\}F'(\tau_0)(\alpha_0+\beta_0)\tau_0^2+F(\tau_0)F'(\tau_1)(\alpha_1+\beta_1)\tau_1^2}{k\{F(\tau_0)+1-F(\tau_1)\}^2}, \quad (10)$$

where $\tau_0 = k/\{\omega_0+(\alpha_0+\beta_0)\bar{\varepsilon}^2_{t-1}\}$ and $\tau_1 = k/\{\omega_1+(\alpha_1+\beta_1)\bar{\varepsilon}^2_{t-1}\}$.

As a numerical illustration, consider an C-STGARCH$(1,1)$ model with $u_t \sim \mathsf{N}(0,1)$ and the following parameter configuration:

$$\mu = 0.3, \quad (\omega_0,\omega_1) = (0.01,0.02), \quad (\alpha_0,\alpha_1) = (0.51,0.1), \quad (11)$$

$$(\beta_0,\beta_1) = (0.40,0.75), \quad k \in \{0.3,1\}. \quad (12)$$

We use a grid of starting values to solve equation (7) numerically and find the number of equilibrium points; the local stability of each equilibrium point is then assessed by considering the expansion in (8)–(10). A single equilibrium point $\bar{\varepsilon}^2_e = 0.118$ and $\bar{\varepsilon}^2_e = 0.111$ is found for $k = 0.3$ and $k = 1$, respectively; the associated partial derivative in (8) is respectively equal to 0.897 and 0.952, suggesting that the model is locally for both values of k. More generally, an examination of (9) and (10) reveals that the partial derivative in (8) increases (decreases) with the value of the threshold k when regime 0 is more (less) persistent than regime 1 in the sense that $\alpha_0+\beta_0$ is larger (smaller) than $\alpha_1+\beta_1$.

3.2 Properties of the Mixing Function

As mentioned before, a key feature of the C-STGARCH model is that its mixing function G depends on all the parameters of the model as well as on both ε^2_{t-1} and σ^2_{t-1}. The signs of $\partial G/\partial\alpha_0$ and $\partial G/\partial\beta_0$ are both negative. An increase in α_0 and/or β_0 raises $\omega_0+\alpha_0\varepsilon^2_{t-1}+\beta_0\sigma^2_{t-1}$ and reduces the probability $\mathsf{P}(\varepsilon^2_{0t} < k|\mathbf{z}_{t-1})$ and thus $G(\mathbf{z}_{t-1})$. A similar argument applies for a change in α_1 and/or β_1, with the signs of both $\partial G/\partial\alpha_1$ and $\partial G/\partial\beta_1$ being negative. A change in α_1 and/or β_1 raises $\omega_1+\alpha_1\varepsilon^2_{t-1}+\beta_1\sigma^2_{t-1}$, increases the probability $\mathsf{P}(\varepsilon^2_{1t} > k|\mathbf{z}_{t-1})$, thus reducing $G(\mathbf{z}_{t-1})$.

The sign of $\partial G/\partial k$ is always positive since the higher the threshold is the bigger is the area of the conditional density of ε_{0t}^2 which is below the threshold and the smaller is the area of the conditional density of ε_{1t}^2 which is above the threshold. In other words, an increase in k results in an increase in $F(k/\{\omega_0+\alpha_0\varepsilon_{t-1}^2+\beta_0\sigma_{t-1}^2\})$ and a decrease in $1-F(k/\{\omega_1+\alpha_1\varepsilon_{t-1}^2+\beta_1\sigma_{t-1}^2\})$. The sign of $\partial G/\partial\omega_1$ is always negative since $1-F(k/\{\omega_1+\alpha_1\varepsilon_{t-1}^2+\beta_1\sigma_{t-1}^2\})$ is higher the larger ω_1 is. Analogously, the sign of $\partial G/\partial\omega_0$ is always negative. Note also that the signs of $\partial G/\partial\varepsilon_{t-1}^2$ and $\partial G/\partial\sigma_{t-1}^2$ are negative.

3.3 Empirical Distribution of the Data

There is a large variety of empirical distributions and time series that can be generated by the C-STGARCH model. In Figures 1 and 2, we show the conditional state-dependent distributions (for two different conditioning values), the threshold, the histogram of ε_t^2, and the time series of ε_t^2 and $G(\mathbf{z}_{t-1})$ generated by a C-STGARCH$(1,1)$ model. We used 400 realizations for the time-series evolution of ε_t and $G(\mathbf{z}_{t-1})$, with the parameter values given in (11)–(12) and u_t having Student's t-distribution with 3 degrees of freedom (rescaled to have unit variance).

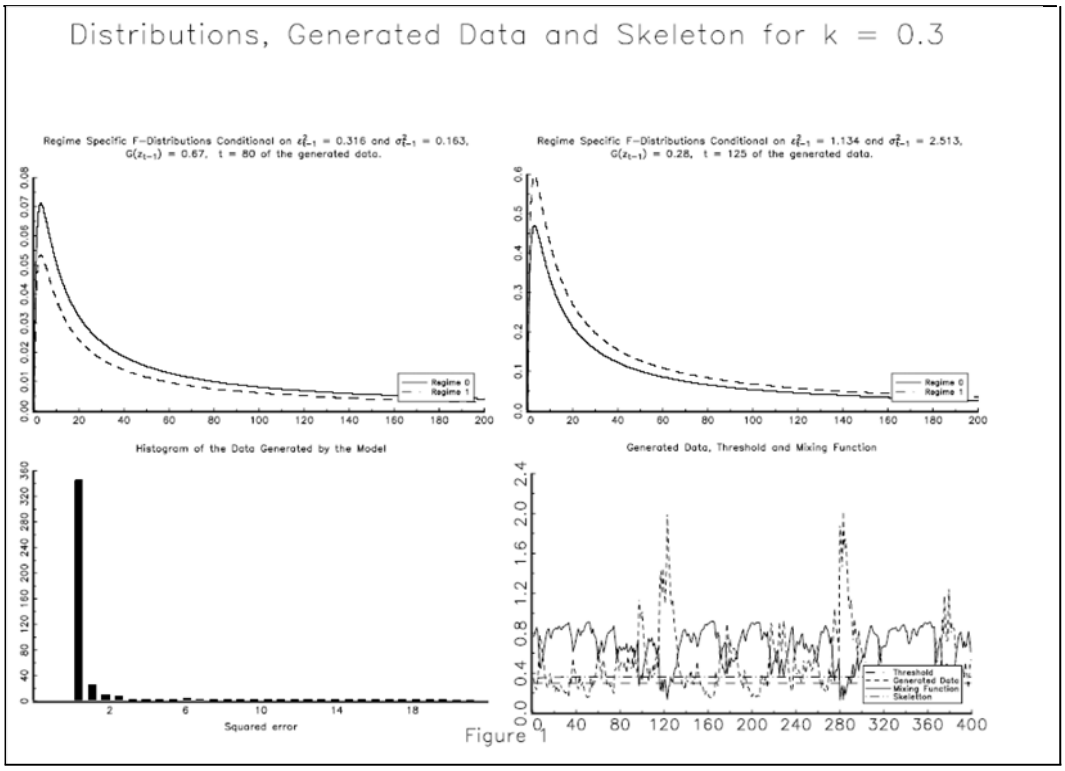

Figure 1

We present results for $k = 0.3$ and $k = 1$. The results are qualitatively similar for the two threshold values. However, the separation of the regimes and the time spent in each regime are quite different in the two cases. When the threshold is high ($k = 1$) most of the observations are below the threshold and are mostly associated with regime 0, with a few occasional shifts to regime 1. For relatively small values of the threshold ($k = 0.3$) the mixing weights are such that the process remains in each regime for a considerable amount of time.

The upper left panel of Figure 1 shows the probability density functions of the latent regime-specific random variables ε_{1t}^2 and ε_{2t}^2 conditional on $\mathbf{z}_{t-1} = (0.316, 0.163)^\intercal$ (observation 80 of the simulated data) when $k = 0.3$. For this value of $\mathbf{z}_{t-1}$, the value of the mixing function is 0.67. It can be seen in the lower right panel that it coincides with relatively moderate values of the conditional variance and therefore it assigns two thirds of the weight to the regime which is associated with values smaller than the threshold. The upper right panel of Figure 1 shows the regime-dependent probability density functions of ε_{1t}^2 and ε_{2t}^2 conditional on $\mathbf{z}_{t-1} = (1.134, 2.513)^\intercal$ (observation 125 of the simulated data). For this value of $\mathbf{z}_{t-1}$, the mixing function takes the value 0.28. It can be seen in the lower right panel that it coincides with relatively high values of the conditional variance and therefore it assigns a high weight to the regime which is associated with values smaller than the threshold.

The lower left panel shows the histogram of the generated data, which reveals that many of the observations which come from regime 0 with high probability have values smaller than 0.3, while those observations which come from regime 1 with high probability have values greater than 0.3. It can be seen from the histogram that approximately two thirds of the observations in the sample are associated with high probability with regime 1.

Finally in the lower right panel we can see that $G(\mathbf{z}_{t-1})$ changes considerably over the sample, taking low values for periods of very high volatility. The skeleton of the model converges extremely fast to the equilibrium point $\bar{\varepsilon}_e^2 = 0.118$. We found that the value of the partial derivative in (8) is 0.9055, suggesting that this equilibrium is locally stable.

The corresponding results when $k = 1$ are shown in Figure 2, and are qualitatively similar to those obtained with $k = 0.3$.

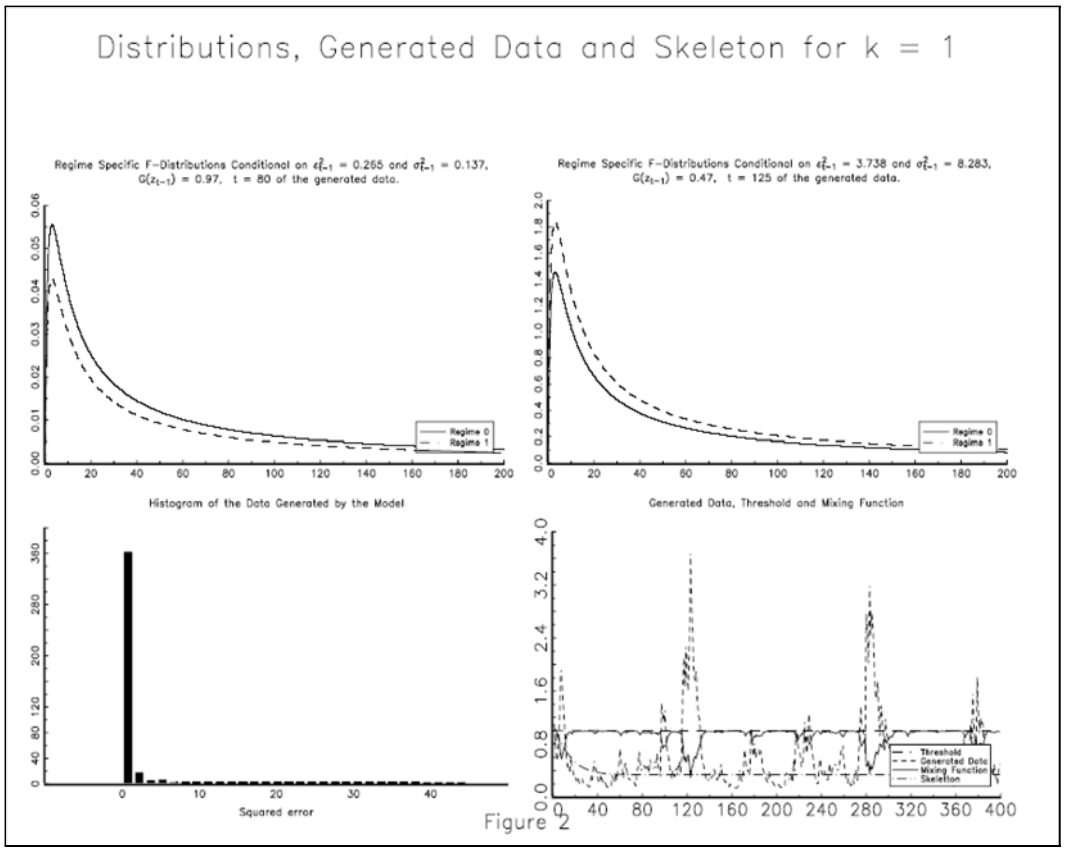

Figure 2

The skeleton of the model converges again very rapidly to the equilibrium point $\bar{\varepsilon}_e^2 = 0.111$, which is locally stable since the partial derivative in (8) is 0.9524. The most significant difference between the $k = 1$ and $k = 0.3$ cases is that when $k = 1$ the mixing function $G(\mathbf{z}_{t-1})$ takes values close to unity most of the time, with values close to zero reached only occasionally in periods of very high volatility. Given that these plots are constructed using the same random numbers, it is reasonable to conclude that the higher the threshold is the smaller is the number of observations which are likely to exceed the threshold.

3.4 News Impact Curve

A useful way of comparing GARCH-type models is by using the so-called 'news impact curve' (cf. Pagan and Schwert, 1990; Engle and Ng, 1993), which represents the relationship between the current shock and next period's conditional volatility, assuming that all other past and current information is held constant. Although the news impact curve is typically computed by

evaluating all past conditional variances at the unconditional variance implied by the model, given the complex structure of our model, we will evaluate it at the fixed point of the skeleton.

The news impact curve for C-STGARCH$(1,1)$ model is thus defined as

$$\text{NIC}(\varepsilon_t|\sigma_t^2=\bar{\varepsilon}_e^2)=G(\varepsilon_t^2|\sigma_t^2=\bar{\varepsilon}_e^2)\tilde{\sigma}_{0,t+1}^2+\{1-G(\varepsilon_t^2|\sigma_t^2=\bar{\varepsilon}_e^2)\}\tilde{\sigma}_{1,t+1}^2,$$

where

$$\tilde{\sigma}_{i,t+1}^2=\delta_i+\alpha_i\varepsilon_t^2,\qquad \delta_i=\omega_i+\beta_i\bar{\varepsilon}_e^2,\qquad i=0,1,$$

and

$$G(\varepsilon_t^2|\sigma_t^2=\bar{\varepsilon}_e^2)=\frac{F(\frac{k}{\delta_0+\alpha_0\varepsilon_t^2})}{F(\frac{k}{\delta_0+\alpha_0\varepsilon_t^2})+1-F(\frac{k}{\delta_1+\alpha_1\varepsilon_t^2})}.$$

The news impact curve will typically separate large and small shocks rather than positive and negative shocks, as is the case with some alternative models. In our model, $\text{NIC}(\varepsilon_t|\sigma_t^2=\bar{\varepsilon}_e^2)$ is always a weighted average of $\delta_0+\alpha_0\varepsilon_t^2$ and $\delta_1+\alpha_1\varepsilon_t^2$. For extreme shocks, we have

$$\text{NIC}(\varepsilon_t|\sigma_t^2=\bar{\varepsilon}_e^2)\approx\delta_0+\alpha_0\varepsilon_t^2,\qquad \text{as } G(\varepsilon_t^2|\sigma_t^2=\bar{\varepsilon}_e^2)\to 1,$$

$$\text{NIC}(\varepsilon_t|\sigma_t^2=\bar{\varepsilon}_e^2)\approx\delta_1+\alpha_1\varepsilon_t^2,\qquad \text{as } G(\varepsilon_t^2|\sigma_t^2=\bar{\varepsilon}_e^2)\to 0.$$

The top panel of Figure 3 shows plots of the news impact curve for different values of the threshold parameter k. We also plot the news impact curve of the regime-specific GARCH models given by

$$\text{NIC}(\varepsilon_{it}|\sigma_{it}^2=\bar{\varepsilon}_e^2)=\delta_i+\alpha_i\varepsilon_{it}^2,\qquad i=0,1.$$

The C-STGARCH news impact curves lie between the two regime-specific curves. The bottom panel plots the mixing functions evaluated at the fixed point of the skeleton. It is clear that the higher the threshold is the more likely it is that the data would come from regime 0, which is associated with a lower unconditional variance but with higher persistence than regime 1. This in turn implies that the corresponding news impact curve is steeper.

These observations highlight the fact that models of the C-STGARCH type have rather complicated dynamics and that the evaluation of the news impact curve at skeleton values is not a trivial issue because it has significant implications for the probability of the process being in a specific regime and thus for the steepness of the news impact curve.

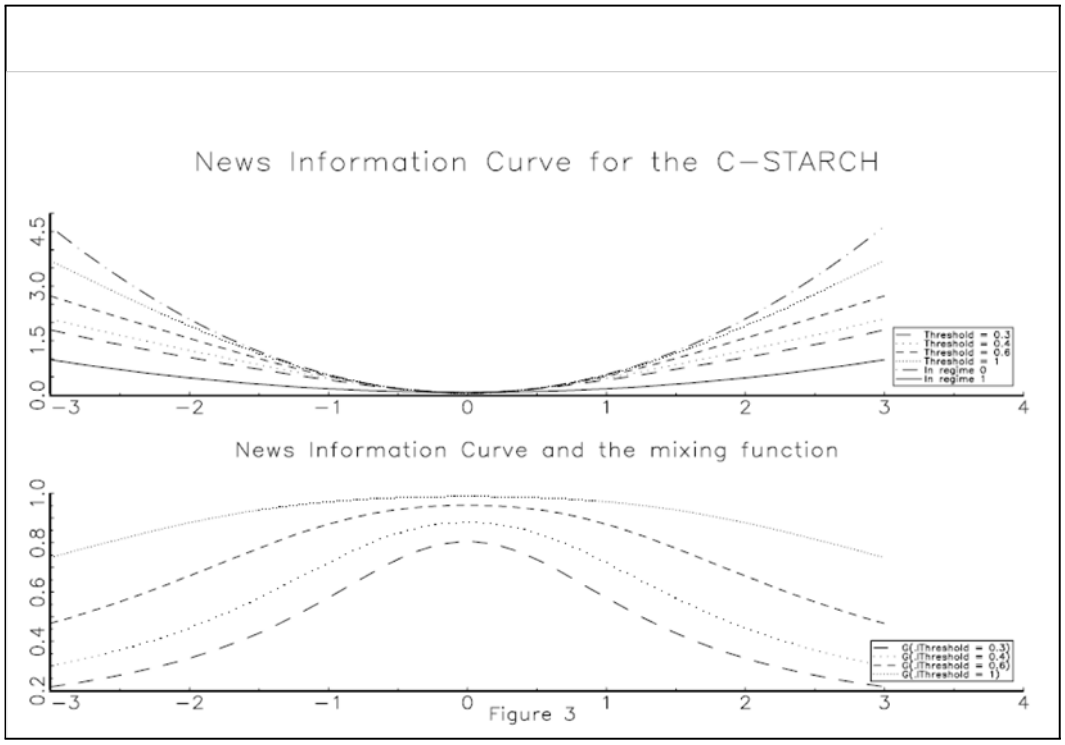

Figure 3

3.5 Persistence

Since Diebold (1986) and Lamourex and Lastrapes (1990), it has been recognized that strong persistence in the conditional variance of economic and financial time series may be a spurious feature due to unaccounted structural breaks and regime shifts. It is, therefore, reasonable to expect a linear GARCH specification to exhibit strong persistence in the presence of neglected nonlinear dynamics of the C-STGARCH type. Perhaps unsurprisingly, in simulation experiments based on different parameter configurations and sample sizes, we found that a Wald test almost never rejected the hypothesis of integrated GARCH(1, 1) volatility when the data came from a C-STGARCH(1, 1) model.

An insight into this result may be gained by observing that the C-STGARCH(1, 1) model may be written as

$$\begin{aligned}\varepsilon_t^2 &= G\left(\varepsilon_{t-1}^2, v_{t-1}\right)\left\{\omega_0 + \left(\alpha_0 + \beta_0\right)\varepsilon_{t-1}^2 - \beta_0 v_{t-1} + v_t\right\} \\ &\quad + \left\{1 - G\left(\varepsilon_{t-1}^2, v_{t-1}\right)\right\}\left\{\omega_1 + \left(\alpha_1 + \beta_1\right)\varepsilon_{t-1}^2 - \beta_1 v_{t-1} + v_t\right\},\end{aligned}$$

where $v_t := \varepsilon_t^2 - \sigma_t^2$ and

$$G(\varepsilon_{t-1}^2, v_{t-1}) = \frac{F(\frac{k}{\omega_0+(\alpha_0+\beta_0)\varepsilon_{t-1}^2-\beta_0 v_{t-1}})}{F(\frac{k}{\omega_0+(\alpha_0+\beta_0)\varepsilon_{t-1}^2-\beta_0 v_{t-1}}) + 1 - F(\frac{k}{\omega_1+(\alpha_1+\beta_1)\varepsilon_{t-1}^2-\beta_1 v_{t-1}})}.$$

By a first-order Taylor expansion of ε_t^2 about $(\varepsilon_{t-1}^2 = \bar{\varepsilon}_e^2, v_{t-1} = 0, v_t = 0)$, we deduce after some straightforward algebra that

$$\varepsilon_t^2 \approx \zeta_0 + \zeta_1(\varepsilon_{t-1}^2 - \bar{\varepsilon}_e^2) + \zeta_2 v_{t-1} + v_t, \tag{13}$$

where

$$\zeta_0 = G(\bar{\varepsilon}_e^2, 0)\{\omega_0 + (\alpha_0+\beta_0)\,\bar{\varepsilon}_e^2\} + \{1 - G\left(\bar{\varepsilon}_e^2, 0\right)\}\{\omega_1 + (\alpha_1+\beta_1)\,\bar{\varepsilon}_e^2\},$$

$$\begin{aligned}\zeta_1 \;=\;& G_1'\left(\bar{\varepsilon}_e^2, 0\right)\{\omega_0 + (\alpha_0+\beta_0)\,\bar{\varepsilon}_e^2\} + G\left(\bar{\varepsilon}_e^2, 0\right)(\alpha_0+\beta_0)\\ & -G_1'\left(\bar{\varepsilon}_e^2, 0\right)\{\omega_1 + (\alpha_1+\beta_1)\,\bar{\varepsilon}_e^2\} + \{1 - G\left(\bar{\varepsilon}_e^2, 0\right)\}(\alpha_1+\beta_1),\end{aligned}$$

$$\begin{aligned}\zeta_2 \;=\;& G_2'\left(\bar{\varepsilon}_e^2, 0\right)\{\omega_0 + (\alpha_0+\beta_0)\,\bar{\varepsilon}_e^2\} - G\left(\bar{\varepsilon}_e^2, 0\right)\beta_0\\ & -G_2'\left(\bar{\varepsilon}_e^2, 0\right)\{\omega_1 + (\alpha_1+\beta_1)\,\bar{\varepsilon}_e^2\} - \{1 - G\left(\bar{\varepsilon}_e^2, 0\right)\}\beta_1.\end{aligned}$$

Here, $G_1'(\varepsilon_{t-1}^2, v_{t-1}) = \dfrac{\partial G(\varepsilon_{t-1}^2, v_{t-1})}{\partial \varepsilon_{t-1}^2}$ and $G_2'(\varepsilon_{t-1}^2, v_{t-1}) = \dfrac{\partial G(\varepsilon_{t-1}^2, v_{t-1})}{\partial v_{t-1}}$, with

$$G_1'\left(\bar{\varepsilon}_e^2, 0\right) = -\frac{\{1-F(w_1)\}F'(w_0)(\alpha_0+\beta_0)w_0^2 + F(w_0)F'(w_1)(\alpha_1+\beta_1)w_1^2}{k\{F(w_0)+1-F(w_1)\}^2},$$

$$G_2'\left(\bar{\varepsilon}_e^2, 0\right) = \frac{\{1-F(w_1)\}F'(w_0)\beta_0 w_0^2 + F(w_0)F'(w_1)\beta_1 w_1^2}{k\{F(w_0)+1-F(w_1)\}^2},$$

$w_0 = k/\{\omega_0 + (\alpha_0+\beta_0)\bar{\varepsilon}_e^2\}$ and $w_1 = k/\{\omega_1 + (\alpha_1+\beta_1)\bar{\varepsilon}_e^2\}$.

Rearranging terms, the approximation in (13) may be written as

$$\varepsilon_t^2 \approx \omega + \varphi\varepsilon_{t-1}^2 - \vartheta v_{t-1} + v_t, \tag{14}$$

where

$$\begin{aligned}\omega \;=\;& G\left(\bar{\varepsilon}_e^2, 0\right)\omega_0 + \{1 - G\left(\bar{\varepsilon}_e^2, 0\right)\}\omega_1\\ & -G_1'\left(\bar{\varepsilon}_e^2, 0\right)\{\omega_0 - \omega_1 + (\alpha_0+\beta_0-\alpha_1-\beta_1)\,\bar{\varepsilon}_e^2\}\bar{\varepsilon}_e^2,\end{aligned}$$

$$\begin{aligned}\varphi &= G_1'\left(\bar{\varepsilon}_e^2, 0\right)\left\{\omega_0 - \omega_1 + \left(\alpha_0 + \beta_0 - \alpha_1 - \beta_1\right)\bar{\varepsilon}_e^2\right\} \\ &\quad + G\left(\bar{\varepsilon}_e^2, 0\right)\left(\alpha_0 + \beta_0 - \alpha_1 - \beta_1\right) + \alpha_1 + \beta_1, \\ \vartheta &= -G_2'\left(\bar{\varepsilon}_e^2, 0\right)\left\{\omega_0 - \omega_1 + \left(\alpha_0 + \beta_0 - \alpha_1 - \beta_1\right)\bar{\varepsilon}_e^2\right\} \\ &\quad + G\left(\bar{\varepsilon}_e^2, 0\right)\left(\beta_0 - \beta_1\right) + \beta_1.\end{aligned}$$

It can be seen that persistence, as measured by the coefficient on ε_{t-1}^2 in (14), is a complicated function of all the parameters of the model rather than a simple weighted average of the regime-specific persistence measures $\alpha_0 + \beta_0$ and $\alpha_1 + \beta_1$.

For the baseline parameter configuration (11)–(12) with $k = 0.3$, we have $\varphi = 0.9055$. When ω_1 is increased to $\omega_1 = 0.025$ we have $\varphi = 0.9124$, and when $\omega_1 = 0.03$ we have $\varphi = 0.9145$. The corresponding values of φ when $k = 1$ are 0.9524, 1.028 and 1.099, respectively. These numerical results demonstrate that, for some parameter configurations, the persistence parameter φ can be close to unity, or even larger than unity, even though $\alpha_0 + \beta_0 < 1$ and $\alpha_1 + \beta_1 < 1$.

4 Parameter Estimation

Once the probability distribution of u_t in (1) is specified, the parameters of a C-STGARCH model can be estimated by the ML method. Letting f_u denote the probability density function of u_t, the log-likelihood function for a sample $(y_1, \ldots, y_T)$ from the C-STGARCH(1, 1) model (ignoring initial conditions) is

$$\mathcal{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T}\left\{-\ln\sigma_t + \ln f_u(\varepsilon_t/\sigma_t)\right\},$$

where

$$\sigma_t^2 = G(\mathbf{z}_{t-1})(\omega_0 + \alpha_0\varepsilon_{t-1}^2 + \beta_0\sigma_{t-1}^2) + \{1 - G(\mathbf{z}_{t-1})\}(\omega_1 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2),$$

$G(\mathbf{z}_{t-1})$ is given by (5), $\boldsymbol{\theta} = (\mu, \omega_0, \alpha_0, \beta_0, \omega_1, \alpha_1, \beta_1, k, \boldsymbol{\lambda}^\intercal)^\intercal$, and $\boldsymbol{\lambda}$ is a vector of (unknown) shape parameters specifying f_u.

In the simulations and empirical application that follow, f_u is specified to be the probability density function of Student's t-distribution with m degrees of freedom (rescaled to have unit variance), so that

$$f_u(x) = \frac{\Gamma\left(\{m+1\}/2\right)}{\Gamma(m/2)\sqrt{\pi(m-2)}}\left(1 + \frac{x^2}{m-2}\right)^{-(1+m)/2}, \quad -\infty < x < \infty, \quad m > 2, \tag{15}$$

where $\Gamma(\cdot)$ stands for the gamma function. ML estimation based on the t-distribution is arguably more appropriate than estimation based on a Gaussian likelihood since many financial time series exhibit substantial leptokurtosis which may not be adequately accounted for by conditional heteroskedasticity alone (see, e.g., Bollerslev, 1987). Note that, under the maintained assumption in (15), the values of F in (5) can be computed as

$$F(x) = H\left(\frac{mx}{m-2}\right), \qquad x > 0,$$

where H is the cumulative distribution function of the central F-distribution with 1 and m degrees of freedom, i.e.,

$$H(x) = \frac{\Gamma\left(\{m+1\}/2\right)}{\Gamma(m/2)\sqrt{\pi m}} \int_0^x \frac{1}{\sqrt{z}} \left(1 + \frac{z}{m}\right)^{-(1+m)/2} \mathrm{d}z, \qquad x > 0.$$

The asymptotic properties of the ML estimator have not been formally established for our model. However, if a C-STGARCH model satisfies suitable stationarity, ergodicity and identifiability conditions, it is reasonable to expect that standard asymptotic results for statistical inference (e.g., Crowder, 1976) apply.

To throw some light on the sampling properties of the maximum likelihood estimator of the parameters of a C-STGARCH$(1,1)$, we now discuss the results of a simulation study. The data-generating process in the sampling experiments is the model defined by (1)–(5), with $\{u_t\}$ being i.i.d. random variables having the probability density function (15) with $m = 3$. The experiments are a factorial design of:

$$\begin{gathered}\mu = 0.3, \quad (\omega_0, \omega_1) = (0.01, 0.02), \quad (\alpha_0, \alpha_1) = (0.51, 0.1), \\ (\beta_0, \beta_1) = (0.40, 0.75), \\ k \in \{0.3, 0.4, 0.6, 1\}, \quad T \in \{400, 800, 1600, 3200\}.\end{gathered}$$

In each Monte Carlo replication, $50 + T$ data points for y_t are generated with $y_0^2 = \sigma_0^2 = k$, but only the last T of these are used in order to attenuate the effect of the starting values. The ML estimate $\widehat{\boldsymbol{\theta}} = (\widehat{\mu}, \widehat{\omega}_0, \widehat{\alpha}_0, \widehat{\beta}_0, \widehat{\omega}_1, \widehat{\alpha}_1, \widehat{\beta}_1, \widehat{k}, \widehat{m})^\intercal$ is obtained by means of a quasi-Newton algorithm that approximates the Hessian according to the Broyden–Fletcher–Goldfarb–Shanno (BFGS) update computed from numerical derivatives. For each design point, a grid of 7 values for each parameter (including the true value) are used as starting values for the BFGS iterations; the starting values that result in the largest likelihood are then selected. Finally, since the

computation of ML estimates is particularly time consuming (given the large number of design points and the grid for initial values), the number of Monte Carlo replications per experiment is 2000.

Table 1 records the finite-sample bias of the ML estimator of $\boldsymbol{\theta}$. The results show that the bias of the ML estimator is a decreasing function of the sample size. As a measure of the accuracy of estimated asymptotic standard errors as approximations to the sampling standard deviation of the ML estimator, Table 2 shows the ratio of the exact standard deviation of the ML estimates to the estimated standard errors averaged across replications for each design point. The standard errors are calculated in the familiar manner from the inverse of the Hessian of the log-likelihood function evaluated at the ML estimates. For the vast majority of cases, the estimated asymptotic standard errors are downward biased. These biases are not, however, substantial and should not have significant adverse effects on inference.

Table 1. Finite-Sample Bias of the ML Estimator

k	T	$\widehat{\mu}$	$\widehat{\omega}_0$	$\widehat{\omega}_1$	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{k}$	$\widehat{m}$
0.3	400	0.102	0.117	0.440	0.127	0.114	0.109	0.099	0.023	0.126
	800	0.085	0.093	0.354	0.176	0.044	0.098	0.068	0.014	0.077
	1600	0.045	0.020	0.234	0.034	0.012	0.045	0.056	0.011	0.056
	3200	0.011	0.008	0.047	0.019	0.002	0.020	0.013	0.005	0.012
0.4	400	0.124	0.120	0.012	0.021	0.076	0.090	0.087	0.045	0.084
	800	0.073	0.078	0.033	0.097	0.058	0.087	0.090	0.020	0.027
	1600	0.015	0.054	0.015	0.003	0.014	0.015	0.033	0.012	0.007
	3200	0.008	0.020	0.004	0.003	0.009	0.002	0.015	0.004	0.000
0.6	400	0.002	0.124	0.109	0.339	0.082	0.288	0.082	0.094	0.088
	800	0.025	0.091	0.099	0.253	0.087	0.139	0.087	0.062	0.073
	1600	0.012	0.039	0.031	0.109	0.036	0.096	0.020	0.014	0.018
	3200	0.007	0.010	0.011	0.023	0.010	0.021	0.009	0.005	0.006
1.0	400	0.055	0.430	0.117	0.523	0.048	0.111	0.045	0.012	0.160
	800	0.020	0.328	0.086	0.325	0.024	0.097	0.018	0.004	0.097
	1600	0.007	0.218	0.022	0.209	0.001	0.023	0.016	0.001	0.001
	3200	0.001	0.055	0.003	0.053	0.000	0.009	0.003	0.000	0.000

Table 2. Ratio of Sampling Standard Deviation to Estimated Standard Error

k	T	$\widehat{\mu}$	$\widehat{\omega}_0$	$\widehat{\omega}_1$	$\widehat{\alpha}_0$	$\widehat{\alpha}_1$	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{k}$	$\widehat{m}$
0.3	400	1.103	1.167	1.235	1.118	1.099	1.224	1.213	1.096	1.063
	800	1.009	0.895	1.220	1.120	0.974	1.187	1.123	1.054	1.076
	1600	1.012	1.085	1.263	1.091	1.003	0.989	1.165	1.037	1.022
	3200	1.005	1.044	1.097	1.025	1.002	1.000	1.032	1.012	0.999
0.4	400	1.092	1.113	1.103	1.244	0.860	1.209	1.154	1.096	1.068
	800	1.016	1.090	0.901	1.109	1.095	1.149	0.872	1.008	1.013
	1600	1.009	1.002	0.995	1.147	0.998	1.160	1.035	1.002	1.002
	3200	1.000	1.000	1.002	1.031	1.000	1.037	1.016	1.000	1.000
0.6	400	1.093	1.371	1.194	1.184	1.129	1.186	1.172	1.043	1.156
	800	1.076	1.225	1.093	1.067	1.174	1.109	1.176	1.054	1.092
	1600	1.008	1.187	1.087	1.015	1.140	1.076	1.099	1.023	1.011
	3200	1.002	1.014	1.006	0.998	1.031	1.002	1.008	1.000	1.000
1.0	400	1.100	1.406	0.983	0.912	0.874	1.295	1.106	1.094	1.143
	800	1.090	1.216	1.093	0.946	1.030	0.956	1.097	1.039	1.054
	1600	1.083	1.101	1.007	1.022	0.992	0.988	0.999	1.015	1.026
	3200	1.005	1.012	1.001	1.000	1.000	1.003	1.000	1.002	1.005

5 Empirical Application

In this section, we illustrate the practical use of the proposed C-STGARCH model using a time series of U.S. daily stock returns. Our data set consists of continuously compounded daily returns of the S&P 500 index over the period January 1, 1964 to March 12, 2007. The returns are pre-filtered by means of a first-order autoregressive model to remove serial correlation. Several financial crises took place during our sample period, which are associated with periods of very high volatility. We are interested, therefore, in models which are capable of discriminating between periods of high and low volatility rather than between periods associated with positive and negative volatility shocks (as, for example, Medeiros and Veiga (2009)).

The C-STGARCH$(1,1)$ model defined by (1), (2), (3), and (5) is compared with the linear GARCH and some alternative models that are intended to capture the same type of nonlinear volatility dynamics. The models considered are:

(i) GARCH:

$$\sigma_t^2 = \omega + \alpha_0 \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2,$$

(ii) STGARCH(a):

$$\sigma_t^2 = \omega + [\alpha_0 G(\sigma_{t-1}^2) + \alpha_1 \{1 - G(\sigma_{t-1}^2)\}] \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2,$$

(iii) STGARCH(b):

$$\sigma_t^2 = \omega + \alpha_0 \varepsilon_{t-1}^2 + [\beta_0 G(\sigma_{t-1}^2) + \beta_1 \{1 - G(\sigma_{t-1}^2)\}] \sigma_{t-1}^2,$$

(iv) STGARCH(c):

$$\sigma_t^2 = \omega + (\alpha_0 \varepsilon_{t-1}^2 + \beta_0 \sigma_{t-1}^2) G(\sigma_{t-1}^2) + (\alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2)\{1 - G(\sigma_{t-1}^2)\},$$

where

$$G(\sigma_{t-1}^2) = \left[1 + \exp(-\gamma\{\sigma_{t-1}^2 - k\})\right]^{-1}, \qquad \gamma > 0.$$

Note that, in all nonlinear models, the constant term in the volatility equation (which determines the level of the conditional variance) is assumed to be time-invariant. This restriction is imposed in order to avoid associating one of the regimes with the outlier observations that are present in our sample[1].

The ML estimates of the parameters of the five volatility models (assuming t-distributed innovations) are shown in Table 3, together with corresponding asymptotic standard errors. We also report the value of the Ljung and Box (1978) portmanteau statistic (Q_b) based on the first $b = 35$ and $b = 45$ sample autocorrelations of the squared standardized residuals, the value of the maximized log-likelihood ($\mathcal{L}_{\max}$), and the value of the Akaike information criterion (AIC).

[1] We also considered specifications in which the constant term in the volatility equation was regime-dependent. Such models were found to identify a few outliers as one regime (with no significant GARCH effects). Since this is not the type of dynamics that STGARCH models are intended to capture, we restricted the constant term to be constant. Needless to say, such a restriction would be unnecessary for samples with no extreme events.

Table 3. ML Estimates

	GARCH	STGARCH(a)	STGARCH(b)	STGARCH(c)	C-STGARCH
μ	0.0152 (0.0057)	0.0157 (0.0062)	0.0157 (0.0013)	0.0153 (0.0062)	0.0153 (0.0079)
ω	0.0044 (0.0008)	0.0020 (0.0011)	0.0021 (0.0001)	0.0026 (0.0013)	0.0006 (0.0007)
α_0	0.0571 (0.0046)	0.0180 (0.0074)	0.0593 (0.0009)	0.1078 (0.0379)	0.0254 (0.0111)
α_1	–	0.0095 (0.0722)	–	0.0417 (0.0537)	0.1762 (0.0414)
β_0	0.9365 (0.0049)	0.9398 (0.0044)	0.9329 (0.0009)	0.8918 (0.0338)	0.9630 (0.0106)
β_1	–	–	0.9404 (0.0012)	0.9579 (0.0569)	0.8131 (0.0485)
k	–	1.8726 (0.4513)	1.3344 (0.0290)	1.4625 (0.6329)	2.3793 (1.3162)
γ	–	4.0400 (3.7244)	6.6149 (0.7498)	1.2935 (0.5540)	–
m	6.9663 (0.3988)	6.6507 (0.3889)	6.5706 (0.0643)	6.8889 (0.4019)	6.7463 (0.5300)
Q_{35}	22.2401	23.1115	116.4828	33.8614)	25.0754
	[0.9536]	[0.9384]	[0.0000]	[0.5230]	[0.8925]
Q_{45}	29.2586	30.0474	127.8812	40.8075	32.8151
	[0.9666]	[0.9575]	[0.0000]	[0.6501]	[0.9116]
$\mathcal{L}_{\max}$	–13057.4	–13053.3	–13054.1	–13051.7	–13043.7
AIC	26124.8	26122.2	26124.2	26121.4	26103.4
MSE	4.6082	4.7541	5.1725	4.6096	4.6045
CR	0.3009	0.3001	0.3294	0.3037	0.2812
R^2	0.0786	0.0792	0.0510	0.0794	0.0800

Figures in parentheses are asymptotic standard errors. Figures in square brackets are asymptotic p-values. MSE is the mean square error and CR is the confusion rate.

The parameter estimates of the linear GARCH model are consistent with the literature in that the fitted model implies strong persistence in the variance as measured by the estimate of $\alpha_0 + \beta_0$. For the three logistic STGARCH models, the estimated threshold parameter (k) varies between 1.34

and 1.87. The STGARCH(a) model has an estimated adjustment parameter (γ) with a large standard error and the estimated mixing weights plotted in Figure 4 do not show marked movement. The STGARCH(b) model exhibits signs of misspecification and has a large adjustment parameter ($\widehat{\gamma} = 6.61$). This is probably due to the fact that the smooth-transition mechanism of the model essentially captures outliers in the squared returns, as can be seen from the estimated mixing weights shown in Figure 5. The STGARCH(c) model, shown in Figure 6 appears to be the most successful of the three logistic STGARCH models.

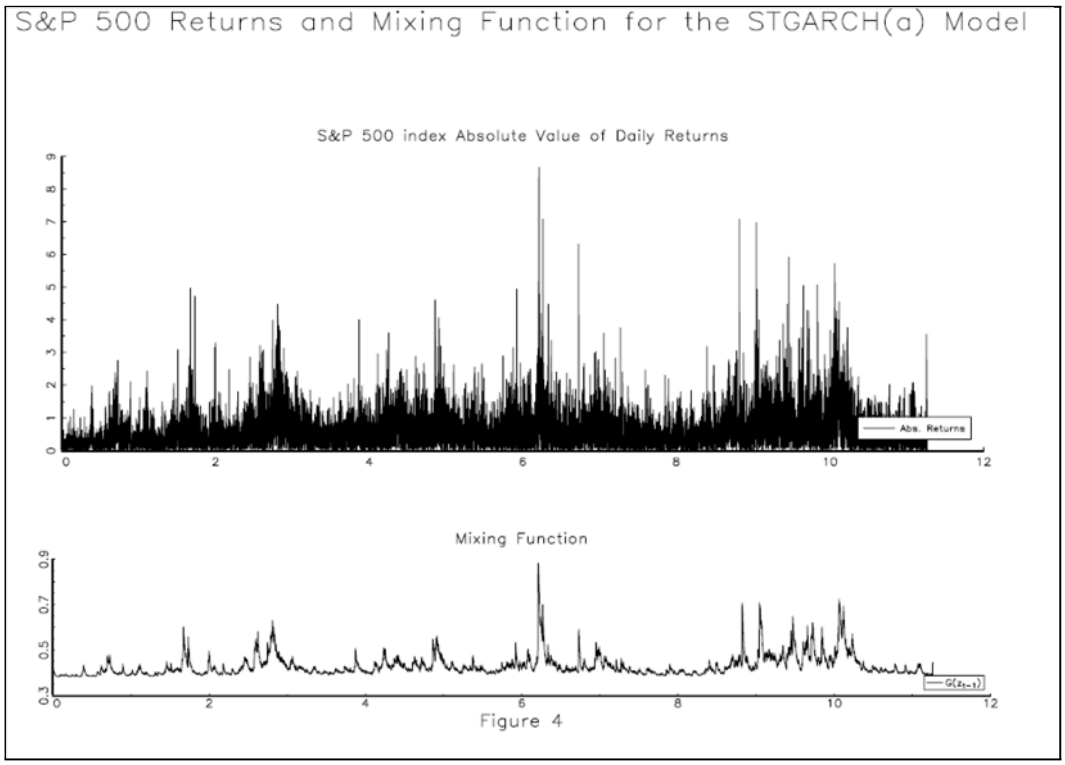

Figure 4

The estimated parameters of the C-STGARCH model reveal remarkably different behaviour in the two regimes. The response to the lagged squared shock is much more substantial in regime 1 ($\widehat{\alpha}_1 = 0.18$) than in regime 0 ($\widehat{\alpha}_0 = 0.03$). This in turn implies that big shocks are amplified in regime 1, which is therefore a regime associated with periods of high conditional volatility.

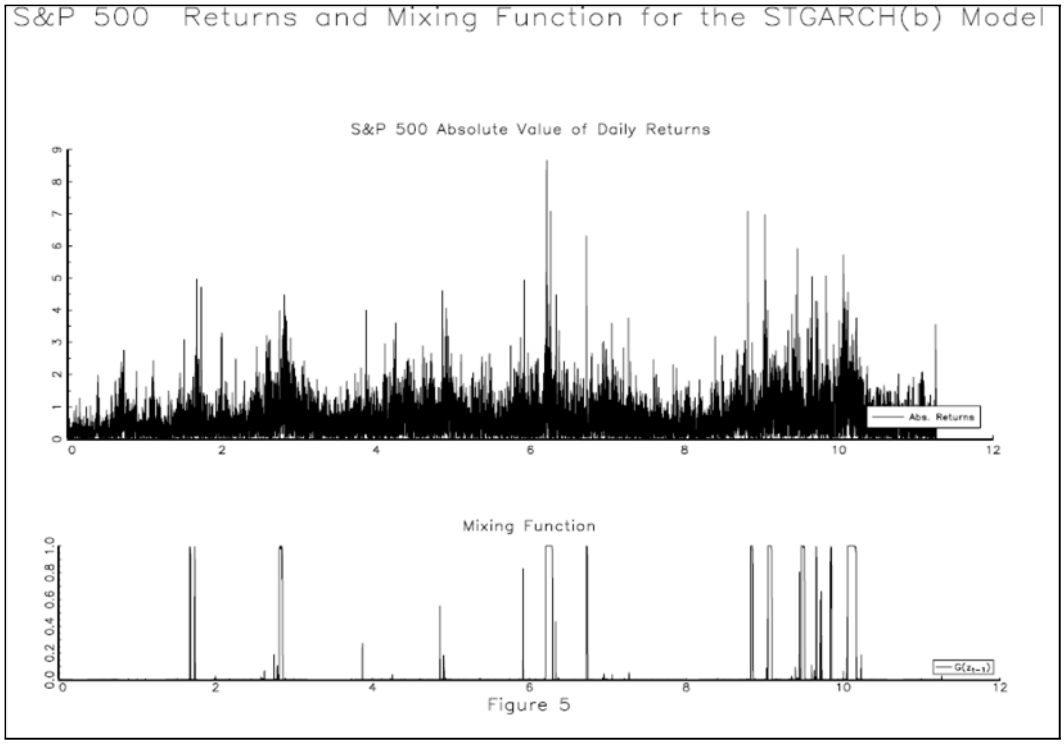

Figure 5

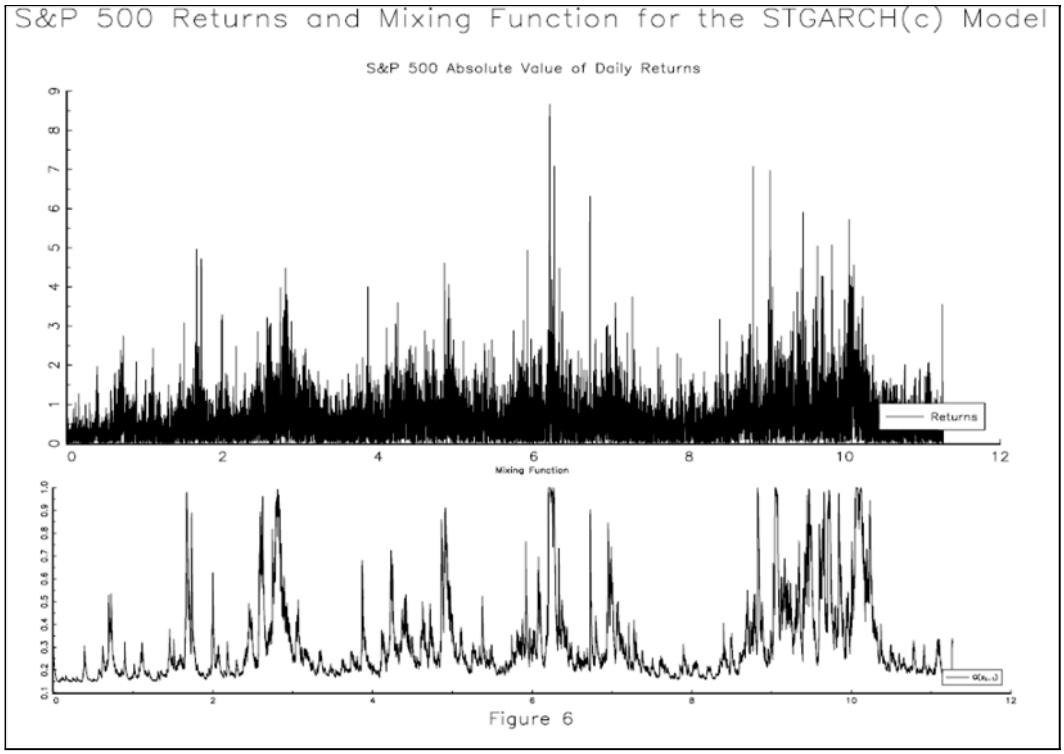

Figure 6

Nevertheless, the estimated regime-specific persistence parameter is greater for regime 0 than for regime 1 ($\widehat{\beta}_0 = 0.96$ and $\widehat{\beta}_1 = 0.81$, respectively). The stability of the empirical C-STGARCH model is assessed by numerical simulation.

The skeleton of the model is found to have a single fixed point $\bar{\varepsilon}_e^2 = 0.345$. The derivative $\partial S(\bar{\varepsilon}_{t-1}^2, \boldsymbol{\vartheta})/\partial \bar{\varepsilon}_{t-1}^2$ in (9) is 0.9856 when evaluated at $\bar{\varepsilon}_{t-1}^2 = \bar{\varepsilon}_e^2$, suggesting that the empirical model is locally stable.

The values of the mixing function shown in Figure 7 suggest that the regimes are highly persistent and that the separation mostly associates periods of high conditional volatility with regime 1. The separation of regimes is quite similar to that implied by the STGARCH(c) model, as can be seen in Figure 6.[2]

The models in Table 3 are compared in terms of three additional criteria based on the residuals $\widehat{\varepsilon}_t$ and the estimated conditional variance $\widehat{\sigma}_t^2$ from each model: the mean square error loss MSE $= (1/T)\sum_{t=1}^{T}(\widehat{\varepsilon}_t^2 - \widehat{\sigma}_t^2)^2$; the coefficient of determination (R^2) in the regression of $\ln \widehat{\varepsilon}_t^2$ on $\ln \widehat{\sigma}_t^2$ and a constant (cf. Pagan and Schwert, 1990); and the confusion rate (CR), i.e., the percentage of times the direction of the change in the returns is incorrectly predicted. The MSE, R^2 and CR criteria all favour the C-STGARCH model, as does the AIC.

Finally, the news impact curves and mixing functions of the estimated models, evaluated at the fixed point of their skeleton, are shown in Figure 8. Note that the mixing functions of the logistic STGARCH models are constant when evaluated at the skeleton. The bottom panel of Figure 8 shows that, at the fixed point of the skeleton, the mixing weight associated with regime 0 is close to zero for the STGARCH(b) model and close to unity for the STGARCH(c). Equal weights are assigned to each regime in the case of the STGARCH(a) model. With respect to the news impact curve, we find that the STGARCH(c) and the GARCH predict a much higher response of volatility to past return shocks than the other models. The news impact curve for the C-STGARCH is steeper for values higher than the threshold.

[2] Note that the labeling of regimes as 0 and 1 for the C-STGARCH model is the reverse of that for the logistic STGARCH models.

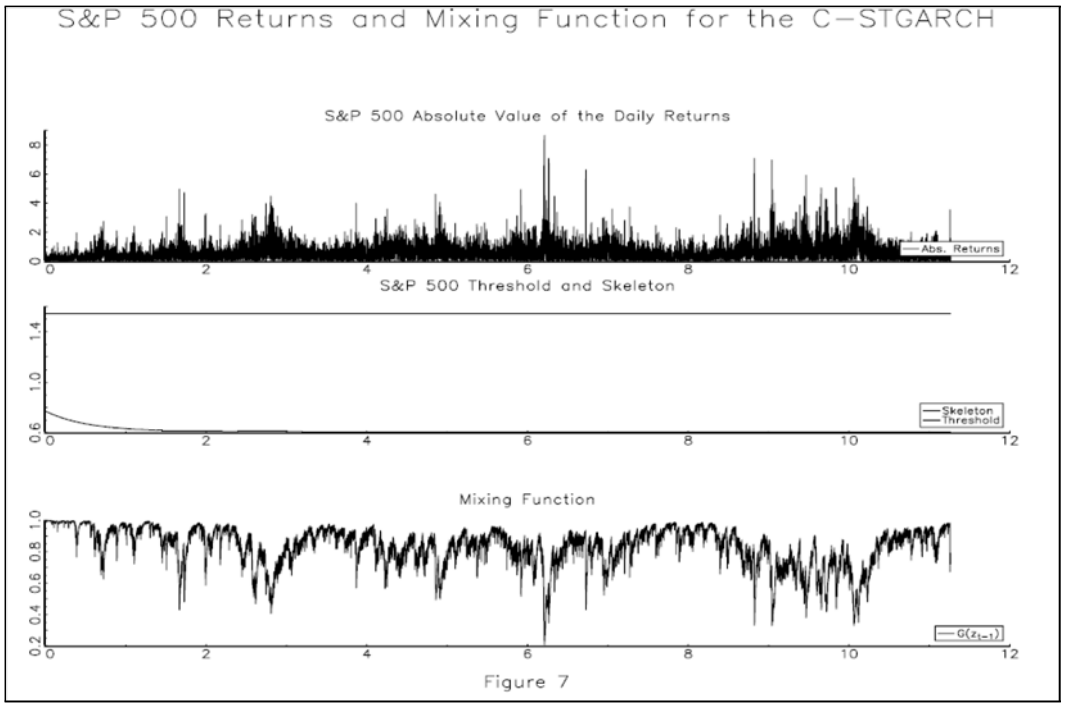

Figure 7

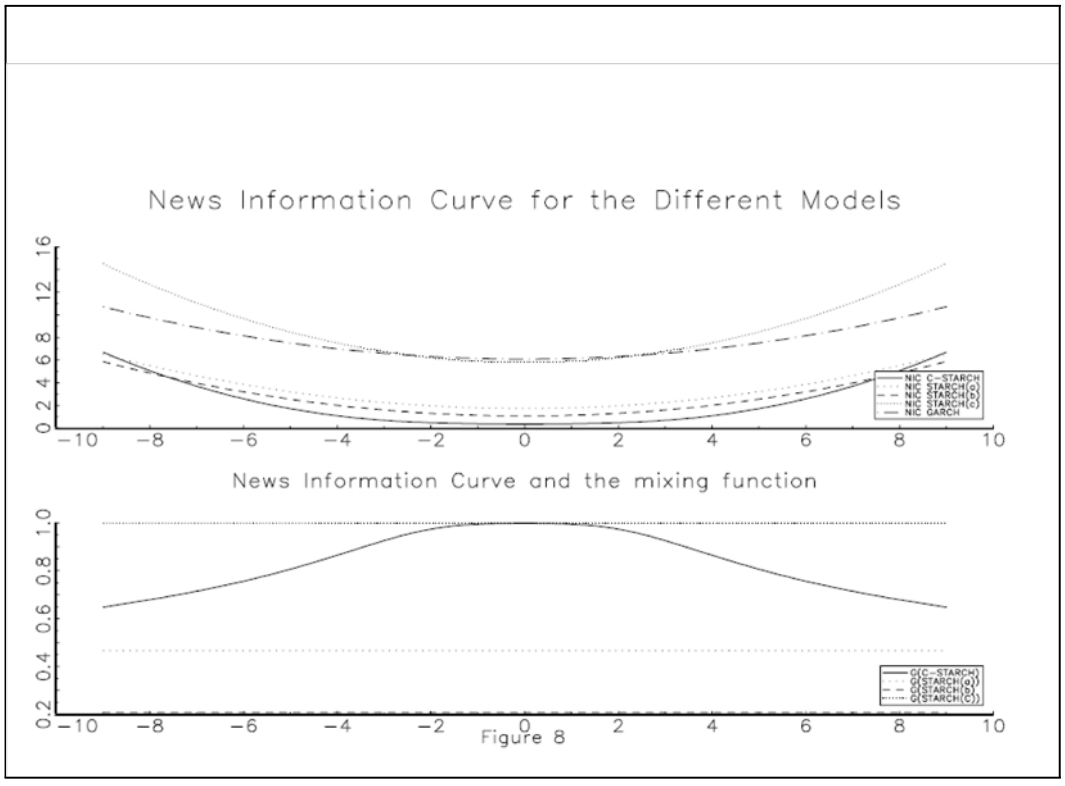

Figure 8

6 Summary

In this paper, we have proposed a contemporaneous-threshold autoregressive conditionally heteroskedastic model which belongs to the family of STGARCH models. A key feature of the C-STGARCH model is that its transition function depends on all the parameters of the model as well as on the data. We have discussed the structural properties of the model, including its stability characteristics, news impact curve, and implications about persistence, and evaluated the finite-sample properties of the ML estimator of its parameters. We have also presented an empirical application to the daily returns on the S&P 500 index, which has shown that the proposed C-STGARCH model is capable of outperforming some competing nonlinear GARCH models.

References

[1] Anderson, H.M., Nam, K. and Vahid, F. (1999), "Asymmetric Nonlinear Smooth Transition GARCH Models", in P. Rothman (ed.), *Nonlinear Time Series Analysis of Economic and Financial Data*, Boston: Kluwer, 191–207.

[2] Bollerslev, T. (1987), "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", *Review of Economics and Statistics* 69, 542–547.

[3] Crowder, M.J. (1976), "Maximum Likelihood Estimation for Dependent Observations", *Journal of the Royal Statistical Society* Ser. B 38, 45–53.

[4] Diebold, F.X. (1986), "Modelling the Persistence of Conditional Variances: A Comment," *Econometric Reviews* 5, 51–56.

[5] Dueker, M.J., Sola, M. and Spagnolo, F. (2007), "Contemporaneous Threshold Autoregressive Models: Estimation, Testing and Forecasting", *Journal of Econometrics* 141, 517–547.

[6] Engle, R.F. and Ng, V.K. (1993), "Measuring and Testing the Impact of News on Volatility", *Journal of Finance* 48, 1749–1778.

[7] González-Rivera, G. (1998), "Smooth-Transition GARCH Models", *Studies in Nonlinear Dynamics and Econometrics* 3, 61–78.

[8] Hagerud, G.E. (1996), "A Smooth Transition ARCH Model for Asset Returns", Working Paper Series in Economics and Finance No. 162, Stockholm School of Economics.

[9] Lamoureux, C.G. and Lastrapes, W.D. (1990), "Persistence in Variance, Structural Change, and the GARCH Model," *Journal of Business and Economic Statistics* 8, 225–234.

[10] Lanne, M. and Saikkonen, P. (2005), "Non-linear GARCH Models for Highly Persistent Volatility", *Econometrics Journal* 8, 251–276.

[11] Ljung, G.M. and Box, G.E.P. (1978), "On a Measure of Lack of Fit in Time Series Models", *Biometrika* 65, 297–303.

[12] Lubrano, M. (2001), "Smooth Transition GARCH Models: A Bayesian Perspective", *Recherches Economiques de Louvain* 67, 257–287.

[13] Lundbergh, S. and Teräsvirta, T. (1998), "Modelling Economic High-Frequency Time Series with STAR-STGARCH Models", Working Paper Series in Economics and Finance No. 291, Stockholm School of Economics.

[14] Medeiros, M.C. and Veiga, A. (2009), "Modeling Multiple Regimes in Financial Volatility with a Flexible Coefficient GARCH(1,1) Model", *Econometric Theory* 25, 117–161.

[15] Pagan, A.R. and Schwert, G.W. (1990), "Alternative Models for Conditional Stock Volatility", *Journal of Econometrics* 45, 267–290.

[16] Tong, H. (1990), *Non-linear Time Series: A Dynamical System Approach*, Oxford: Oxford University Press.