Stat Methods Appl (2010) 19:399–430
DOI 10.1007/s10260-010-0131-2

ORIGINAL ARTICLE

Fractionally integrated time varying GARCH model

Adnen Ben Nasr · Mohamed Boutahar · Abdelwahed Trabelsi

Published online: 20 March 2010

Abstract This paper introduces the new FITVGARCH model to describe both long memory and structural change behaviour in the volatility process by allowing for time varying dynamic structure in the conditional variance. The parameters of the conditional variance in the FIGARCH model are allowed to change smoothly over time. We derive an LM-type test for parameter constancy of the FIGARCH model against the alternative of time dependent parameters. Simulation analysis shows that both empirical size and power of the constancy test are quite good. An empirical application to the stock market volatility indicates that this new class of model seems to outperform the FIGARCH model in the description of the daily NASDAQ composite index returns.

Keywords Modeling volatility · Long memory · Structural changes · Model specification

1 Introduction

A large body of research suggests that there is significant evidence of long memory in the conditional volatility of various financial and economic time series; see Ding

A. Ben Nasr (✉) · A. Trabelsi
BESTMOD, Institut Supérieur de Gestion de Tunis, 41 rue de la liberté-Cité Bouchoucha, Le Bardo, 2000 Tunis, Tunisia
e-mail: adnen.bennasr@isg.rnu.tn

A. Trabelsi
e-mail: Abdel.Trabelsi@isg.rnu.tn

M. Boutahar
GREQAM, Université de la Méditerranée, 2 Rue de la Charité, 13236 Marseille cedex 02, France
e-mail: mohammed.boutahar@univmed.fr

et al. (1993), Baillie et al. (1996), Andersen and Bollerslev (1997), Bollerslev and Mikkelsen (1996), Lobato and Savin (1998) and Davidson (2004) for evidence that persistence in the volatility can be characterized as a long memory process. The fractionally integrated GARCH (FIGARCH) model, introduced by Baillie et al. (1996), proved to be successful in modeling the observed persistence in the volatility of many time series such as stock market returns and option prices (Bollerslev and Mikkelsen 1996), exchange rates (Tse 1998) and inflation rates (Baillie et al. 2002). Breidt et al. (1998) propose the long memory stochastic volatility model (LMSV) as a time series representation of persistence in conditional volatility.

Another related discussion on financial time series suggested that there is strong evidence for the occurrence of structural changes in the volatility process. See for example, Bos et al. (1999) and Andreou and Ghysels (2002). Recent Econometric models allow for stochastic time variation in the parameters of a GARCH specification. Hamilton and Susmel (1994) and Cai (1994) have introduced independently the Markov switching ARCH-model while Dueker (1997) has extended the approach to GARCH models. In this approach, the conditional variances are allowed to switch between a finite numbers of regimes with the transition between regimes governed by an unobserved Markov chain. So et al. (1998) generalized the stochastic volatility model by incorporating the Markov regime switching properties. A nonlinear version of the GARCH process based on smooth transition approach has been proposed by Hagerud (1997) and Gonzalez-Rivera (1998). They introduced the Smooth Transition GARCH (STGARCH) process to model the asymmetric behaviour of the conditional variance. Anderson et al. (1999) proposed the Asymmetric Nonlinear Smooth Transition GARCH (ANSTGARCH) model. The STGARCH model has been extended to model the structural change in the conditional variance in Amado and Teräsvirta (2008). They proposed the Time varying GARCH (TVGARCH) model to allow for time dependent parameters in the GARCH process. Engle and Rangel (2008) introduced the Spline-GARCH model that allows for time-variation in the unconditional volatility level.

A related line of research on long memory and structural changes in the volatility discusses the connection between these phenomena. In fact, the volatility persistence may be due to structural breaks in the volatility process. This approach has been originally suggested by Diebold (1986) and Lamoureux and Lastrapes (1990). Hamilton and Susmel (1994), and Cai (1994), among Others, suggest that regime switching may be the main reason for the persistence of the volatility. Articles by Beine and Laurent (2001); Breidt and Hsu (2002) and Granger and Hyung (2004) show that presence of occasional structural breaks in the data can produce slowly decaying autocorrelations which corresponds to long-memory behavior generally observed in the conditional volatility of exchange rates and stock returns. This literature concludes that it is very difficult to distinguish between true and spurious long memory processes.

However, recent contributions to this literature have attempted to discriminate between long memory and structural changes in the volatility process. Stărică and Granger (2005) concluded that log-absolute returns of the S&P 500 index are best described as an *iid* series affected by occasional shifts in the unconditional variance. Mikosch and Stărică (2004) have analyzed the properties of the autocorrelation function of the S&P 500 absolute returns over the period 1953–1977. They found that the autocorrelation function has the properties of long memory process when taking the whole sample but resembles to that of short memory process when taking only the period 1953–1973. They explained this finding by the fact that the volatility has increased over the period 1973–1977. Perron and Qu (2007) have shown that the behavior of the log-periodogram estimate of the fractional integration parameter for the processes of short memory with breaks is not equivalent to that of the short memory process. They note that the estimates of the fractional integration parameter will vary with the number of frequencies m for a short-memory series with breaks, but it seems to be independent of m for truly long-memory process. Perron and Qu (2009) have shown that stock market volatility may be better characterized by a short-memory process affected by occasional level shifts. They also present a test designed to distinguish between long memory and short-memory process with level shifts.

Another body of research has suggested that both long memory and structural change characterize the structure of financial returns volatility. Relevant references on this issue include Lobato and Savin (1998), Beine and Laurent (2001), Morana and Beltratti (2004) and Martens et al. (2004). More recently, Baillie and Morana (2009) introduced a new long-memory volatility model, denoted by Adaptive FIGARCH, which allows for jointly modeling long-memory and structural change behaviors in the conditional variance process. The structural change is modeled in this Adaptive FIGARCH model by allowing the intercept to follow a slowly varying function.

Motivated from the above summary of literature, in particular the line of research that suggests the co-existence of both long memory and structural change in the volatility process of financial markets data, we present a new model that allows the volatility to have such behaviors. The idea is to allow the parameters in the conditional variance equation of the FIGARCH model to be time dependent. More precisely, the change of the parameters is assumed to be smooth over time using logistic smooth transition function.

The paper is organized as follows. Section 2 presents the classical GARCH and FIGARCH models. In Sect. 3 we introduce the new Fractionally Integrated Time Varying GARCH (FITVGARCH) model. Testing for parameter constancy is derived in Sect. 4 while the empirical size and power properties are evaluated by means of Monte Carlo experiments in Sect. 5. The estimation procedure for the FITVGARCH parameters and the covariance matrix are discussed in Sect. 6. This section also includes simulation results of estimating the model by QMLE method. In Sect. 7 we discuss the test of serial dependence in the squared standardized errors as a misspecification test. Section 8 contains empirical evidence on financial market volatility. Section 9 concludes.

2 The FIGARCH model

As a starting point, we present the generalized autoregressive conditional heteroskedasticity (GARCH) model, developed by Engle (1982), Bollerslev (1986), Nelson (1991) and others, to model the time varying volatility. The GARCH(p, m) model is defined as

$$y_t = \mu + \varepsilon_t \tag{1}$$

$$\begin{aligned} \varepsilon_t &= \eta_t h_t^{1/2}, \quad \eta_t \sim N(0,1) \\ h_t &= \omega + \alpha(L)\varepsilon_t^2 + \beta(L)h_t \end{aligned} \tag{2}$$

where μ is the mean of the process, h_t is the conditional variance of $y_t, \alpha(L) = \alpha_1 L + \cdots + \alpha_m L^m, \beta(L) = \beta_1 L + \cdots + \beta_p L^p$ where L denotes the lag operator. The GARCH process is covariance stationary if the following restriction is satisfied: $\alpha(1) + \beta(1) < 1$. The conditions $\omega \geq 0, \alpha_i \geq 0$ for $i = 1, \ldots, m$, and $\beta_j \geq 0$ for $j = 1, \ldots, p$, are assumed, in Bollerslev (1986), to ensure that the conditional variance h_t is positive. However, Nelson and Cao (1992) showed that the non-negativity of these coefficients is not necessary. They derived necessary and sufficient conditions for the nonnegativity of GARCH(p, q) models with $p \leq 2$ and sufficient conditions for $p > 2$. For example, the conditions for the GARCH(1, 2) are $\alpha_1 \geq 0$, $\beta_1\alpha_1 + \alpha_2 \geq 0$ and $0 \leq \beta_1 < 1$. Clearly, these restrictions allow α_2 to be negative. Recently, Tsai and Chan (2008) showed that the conditions of Nelson and Cao (1992) for $p > 2$ are not only sufficient, but also necessary. By rearranging the terms in the equation of the conditional variance in (2), it follows that

$$[1 - \alpha(L) - \beta(L)]\varepsilon_t^2 = \omega + [1 - \beta(L)]\upsilon_t \tag{3}$$

which is the so-called "ARMA in squares" representation of the GARCH equation, where $\upsilon_t = (\varepsilon_t^2 - h_t)$. All the roots of $[1-\alpha(L)-\beta(L)]$ and $[1-\beta(L)]$ are assumed to lie outside the unit circle. However, many empirical applications of the GARCH(p, m) model for volatility indicate the existence of unit root in the estimated lag polynomial $[1 - \hat{\alpha}(L) - \hat{\beta}(L)]$. To solve this problem, Engle and Bollerslev (1986) proposed the Integrated GARCH, or IGARCH(p, q), model, in a way that the autoregressive polynomial in (3) has one unit root. This assumes that $[1-\alpha(L)-\beta(L)] \equiv (1-\phi(L))(1-L)$ where all the roots of $(1-\phi(L))$ lie outside the unit circle. $\phi(L) = \phi_1 L + \cdots + \phi_q L^q$ where $q = \max\{p, m\} - 1$. Then the IGARCH(p, q) model may be defined as

$$(1 - \phi(L))(1 - L)\varepsilon_t^2 = \omega + [1 - \beta(L)]\upsilon_t \tag{4}$$

The IGARCH model implies infinite persistence of the conditional variance to a shock in squared returns. However, empirical evidence suggests that, in most situations, the volatility process is mean-reverting and the IGARCH model seems to be too restrictive as it implies infinite persistence of a volatility shock. Such a feature stands in sharp opposition to the observed behaviour of agents and does not match more closely the persistence in observed volatility. (see Bollerslev and Engle 1993; Baillie et al. 1996). As a consequence, many researchers have proposed extensions of GARCH models which can produce such long-memory behaviour; Robinson (1991) introduced the ARCH(∞), as the first model permitting long memory in the conditional variance. Baillie et al. (1996) introduced the Fractionally Integrated GARCH (FIGARCH) model that allows for fractional order of integration to describe the long memory properties in the volatility. The study of fractional integration in time series processes was introduced to econometrics by Granger and Joyeux (1980) and Granger (1981) using the autoregressive fractionally integrated moving average (ARFIMA) model. It is based on the fractional differencing operator $(1-L)^d$ where d can take value other than 0 or 1.

The FIGARCH(p, d, q) model of Baillie et al. (1996) is simply obtained by replacing the operator $(1-L)$ in the IGARCH model in (4) by the fractional operator $(1-L)^d$;

$$(1-\phi(L))(1-L)^d\varepsilon_t^2 = \omega + [1-\beta(L)]v_t \tag{5}$$

where $\phi(L)$ and $\beta(L)$ are as before, such that $(1-\phi(L))$ and $[1-\beta(L)]$ are assumed to have all their roots lying outside the unit circle, and the fractional differencing parameter d lies between 0 and 1. Note that (5) is the ARFIMA representation of the squared errors ε_t^2. The fractional difference operator $(1-L)^d$ can be defined as

$$\begin{aligned}(1-L)^d &= \sum_{j=0}^{\infty}\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}L^j \\ &= 1 - dL - \frac{d(1-d)}{2!}L^2 - \frac{d(1-d)(2-d)}{3!}L^3 - \cdots \\ &= 1 - \sum_{j=1}^{\infty} b_j(d)L^j,\end{aligned} \tag{6}$$

where $\Gamma(\cdot)$ denotes the gamma function.

The FIGARCH processes are not covariance stationary, in contrast to the hyperbolic GARCH (HYGARCH) and long-memory GARCH (LMGARCH) models which have been examined in Davidson (2004) and Karanasos et al. (2004) respectively. However, some results seemingly suggest that the FIGARCH processes are indeed strictly stationary for $0 \leq d \leq 1$ (see Baillie et al. (1996) and Davidson (2004)). For $d < 1$, the process is mean reverting, with the effect of the shocks dying away in the long run. If d is higher than 0, the process is said to be long memory, so-named because of the slow hyperbolic rate of decay after a volatility shock instead of the faster geometric rate of the GARCH model or the permanence of the IGARCH model. When $d = 0$, then (5) reduces to the standard GARCH model; and when $d = 1$, then (5) becomes the Integrated GARCH, or IGARCH model, and implies infinite persistence of the conditional variance to a shock in squared returns.

An alternative representation of (5) is given by

$$h_t = \omega + \left[1-\beta(L)-(1-\phi(L))(1-L)^d\right]\varepsilon_t^2 + \beta(L)h_t \tag{7}$$

Thus, the FIGARCH model implies the following ARCH(∞) representation

$$h_t = \frac{\omega}{(1-\beta(1))} + \left[1 - \frac{(1-\phi(L))(1-L)^d}{(1-\beta(L))}\right]\varepsilon_t^2 \tag{8}$$

$$= \tilde{\omega} + \psi(L)\varepsilon_t^2 \tag{9}$$

where $\psi(L) = \psi_1 L + \psi_2 L^2 + \cdots$. To guarantee the non-negativity of the conditional variance as surely for all t, all the coefficients in the ARCH (∞) representation in (8) must be non-negative, i.e. $\tilde{\omega} > 0$, and $\psi_i \geq 0$ for all $i \geq 1$. For a FIGARCH $(1, d, 1)$ model, various different sets of sufficient parameter constraints for the conditional variance to be strictly positive are discussed in the literature. Baillie et al. (1996) imposed the conditions $\omega > 0$, $0 \leq \beta_1 \leq \phi_1 + d$ and $0 \leq d \leq 1 - 2\phi_1$. Bollerslev and Mikkelsen (1996) state the following sufficient conditions $\beta_1 - d \leq \phi_1 \leq (2-d)/3$ and $d(\phi_1 - (1-d)/2) \leq \beta_1 (\phi_1 - \beta_1 + d)$. Chung (1999) suggests another set of sufficient constraints which is given by $0 \leq \phi_1 \leq \beta_1 \leq d < 1$. More recently, Conrad and Haag (2006) have proposed less restrictive constraints. As denoted by them, the parameters of the polynomial $\psi(L)$ can be derived recursively as $\psi_1 = \phi_1 - \beta_1 + d$ and $\psi_j = \beta_1 \psi_{j-1} + \left(f_j - \phi_1\right)\left(-g_{j-1}\right)$ for all $j > 1$, where $f_j = (j - 1 - d)/j$ and $g_j = f_j g_{j-1}$ with $g_0 = 1$. The set of necessary and sufficient conditions of Conrad and Haag (2006) are stated as: case (1) $0 < \beta_1 < 1$, either $\psi_1 \geq 0$ and $\phi_1 \leq f_2$ or $\psi_{j-1} \geq 0$ and $f_{j-1} < \phi_1 \leq f_j$ for $j > 2$; case (2) $-1 < \beta_1 < 0$, either $\psi_1 \geq 0$, $\psi_2 \geq 0$ and $\phi_1 \leq f_2(\beta_1 + f_3)/(\beta_1 + f_2)$ or $\psi_{j-1} \geq 0$, $\psi_{j-2} \geq 0$ and $f_{j-2}(\beta_1 + f_{j-1})/(\beta_1 + f_{j-2}) < \phi_1 \leq f_{j-1}(\beta_1 + f_j)/(\beta_1 + f_{j-1})$ for $j > 3$. The conditions of Conrad and Haag (2006) are necessary and sufficient for the non-negativity of the process and allow the model for more flexibility to capture the dynamics in the conditional variance.

3 Fractionally integrated time varying GARCH model

As mentioned above, there are motivations from the recent econometric literature on financial time series to allow for both long memory and structural change in the volatility process. Hence, the main focus of this study is to take account for the co-existence of long memory and structural change in the conditional volatility. To this end, we extend the FIGARCH model of Baillie et al. (1996) by allowing the conditional variance parameters to change over time. Changing of the parameters can be done using smooth transition function. Smooth transition has been used in recent studies to describe nonlinearity or structural change phenomena in the volatility processes. Hagerud (1997), Gonzalez-Rivera (1998), and Anderson et al. (1999) have discussed the smooth transition GARCH (ST-GARCH), as a nonlinear version of the GARCH process, to model the asymmetric behaviour of the conditional variance. Amado and Teräsvirta (2008) considered the time varying parameter GARCH (TV-GARCH) model for modeling structural change in the volatility process. The TV-GARCH(p, q) model of Amado and Teräsvirta (2008) allows the parameters of the GARCH(p, q) model to change with the time. It is defined as

$$
\begin{aligned}
y_t &= \mu_t + \varepsilon_t \\
\varepsilon_t &= \eta_t \sigma_t, \quad \eta_t \sim N(0, 1)
\end{aligned}
$$

where μ_t is the conditional mean of the innovations ε_t, assumed to be equal to zero in Amado and Teräsvirta (2008) in order to focus only on the conditional variance σ_t^2 which can have either an additive form as

$$
\sigma_t^2 = h_t + g_t \tag{10}
$$

or a multiplicative form as

$$\sigma_t^2 = h_t g_t \tag{11}$$

where h_t follows the standard GARCH(p, q) model

$$h_t = \omega_1 + \alpha_1(L)\varepsilon_t^2 + \beta_1(L)h_t$$

with $\alpha_1(L) = \alpha_{1,1}L + \cdots + \alpha_{1,q}L^q$, $\beta_1(L) = \beta_{1,1}L + \cdots + \beta_{1,p}L^p$, and g_t is the time-varying component of the conditional variance σ_t^2. For the additive structure in (10), the function g_t can be defined as

$$g_t = \left(\omega_2 + \alpha_2(L)\varepsilon_t^2 + \beta_2(L)h_t\right) F(t^*; \gamma, c)$$

with $\alpha_2(L) = \alpha_{2,1}L + \cdots + \alpha_{2,q}L^q$, $\beta_2(L) = \beta_{2,1}L + \cdots + \beta_{2,p}L^p$, $F(t^*; \gamma, c)$ is a logistic smooth transition function defined as

$$F(t^*; \gamma, c) = \left(1 + \exp\left\{-\gamma \prod_{k=1}^{K}(t^* - c_k)\right\}\right)^{-1}, \tag{12}$$
$$\gamma > 0, \quad c_1 \leq c_2 \leq \cdots \leq c_K.$$

where the transition variable is the standardized time variable $t^* = t/T$ and T is the sample size. The transition function $F(t^*; \gamma, c)$ is a continuous function bounded between 0 and 1. The parameter γ corresponds to the speed of transition between the two regimes, while the parameter c_k, known as the threshold parameter, indicates when, in the range of t, the transition takes place.

The most common choices of K in the logistic transition function (12) are $K = 1$ and $K = 2$. For $K = 1$, the logistic function F changes from 0 to 1 as t increases. The smoothness parameter, γ, measures the slope of the logistic function and, therefore, governs the speed with which the transition between regimes takes place. As γ increases, the logistic function approaches to the indicators function $I[t^* > c_1]$ that takes up basically two values; $I(\cdot) = 1$ if argument is true and $I(\cdot) = 0$ otherwise. As a result the transition between regimes happens instantaneously when $t^* = c_1$. For $K = 2$, the logistic function F changes symmetrically around the mid-point $(c_1 + c_2)/2$ where this logistic function attains its minimum value. The minimum lies between 0 and 1/2, reaching 0 when $\gamma \to \infty$ and equaling 1/2 when $c_1 = c_2$ and $\gamma < \infty$.

For the multiplicative case, Amado and Teräsvirta (2008) note that when assuming $\omega_2 = \delta\omega_1, \alpha_{2,i} = \delta\alpha_{1,i}$ and $\beta_{2,j} = \delta\beta_{1,j}$ for $i = 1, \ldots, q$ and $j = 1, \ldots, p$, and setting $g_t = (1 + F(t^*; \gamma, c))$ in (11), the model can be seen as a particular case of the additive model in (10).

Recent studies have used the smooth transition approach for jointly modeling long memory and nonlinearity in time series, see van Dijk et al. (2002) and Ajmi et al. (2008) among others. Following Amado and Teräsvirta (2008), we extend the

FIGARCH model by assuming that the conditional variance parameters of this model vary smoothly over time. The main objective of this model is to take account for both long memory and smooth structural change in the volatility process. To this end, we adopt an additive decomposition of the conditional variance in the FIGARCH model. The conditional variance will be decomposed into two components. The first follows the standard FIGARCH process and the second is a time varying component. Assume for simplicity that $\mu = 0$, the new Fractionally Integrated Time Varying GARCH(p, d, q) (FITVGARCH(p, d, q)) model can be defined as

$$y_t = \varepsilon_t \tag{13}$$

$$\varepsilon_t = \eta_t h_t^{1/2}, \quad \eta_t \sim N(0, 1) \tag{14}$$

$$[1 - \phi_t(L)](1 - L)^d \varepsilon_t^2 = \omega_t + [1 - \beta_t(L)]v_t \tag{15}$$

where $\omega_t = \omega_1 + \omega_2 F(t^*; \gamma, c)$, $\phi_t(L) = \phi_1(L) + \phi_2(L)F(t^*; \gamma, c)$; $\phi_1(L) = \phi_{1,1}L + \cdots + \phi_{1,q}L^q$, $\phi_2(L) = \phi_{2,1}L + \cdots + \phi_{2,q}L^q$, $\beta_t(L) = \beta_1(L) + \beta_2(L)F(t^*; \gamma, c)$; $\beta_1(L) = \beta_{1,1}L + \cdots + \beta_{1,p}L^p$ and $\beta_2(L) = \beta_{2,1}L + \cdots + \beta_{2,p}L^p$. We assume that the roots of the polynomials $[1 - \phi_t(L)]$ and $[1 - \beta_t(L)]$ are outside the unit circle for all t. This implies that $[1 - \phi_t(1)] > 0$ and $[1 - \beta_t(1)] > 0$. With $K = 1$, the parameters of the FIGARCH model change smoothly over time from $(\omega_1, \phi_{1,i}, \beta_{1,j})$ to $(\omega_1 + \omega_2, \phi_{1,i} + \phi_{2,i}, \beta_{1,j} + \beta_{2,j})$, $i = 1, \ldots, q$, $j = 1, \ldots, p$. The transition between regimes happens instantaneously when $t^* = c_1$. When $\gamma \to 0$, the FITVGARCH(p, d, q) model in (15) nests the FIGARCH(p, d, q) model in (5) since the logistic transition function becomes constant and equal to 1/2. As for the FIGARCH model, the fractional parameter d is assumed to be in the interval [0,1]. However, unlike the FIGARCH process, the FITVGARCH process is not strictly stationary, due to the time varying parameters. Noting that (15) is the ARFIMA representation of the squared errors ε_t^2, and it is the time varying version of (5). After Rearrangement of terms in (15), an alternative representation for the FITVGARCH(p, d, q) model is

$$[1 - \beta_t(L)]h_t = \omega_t + \left[1 - \beta_t(L) - (1 - \phi_t(L))(1 - L)^d\right]\varepsilon_t^2 \tag{16}$$

Then, the conditional variance of the FITVGARCH(p, d, q) model is given by the following ARCH representation

$$\begin{aligned} h_t &= \omega_t[1 - \beta_t(L)]^{-1} + \left\{1 - [1 - \beta_t(L)]^{-1}(1 - \phi_t(L))(1 - L)^d\right\}\varepsilon_t^2 \\ &= \tilde{\omega}_t + \sum_{j=1}^{\infty} \psi_{j,t}\varepsilon_{t-j}^2 \end{aligned} \tag{17}$$

Similarly to the FIGARCH model, conditions on the parameters of the FITVGARCH process have to be imposed to guarantee that the conditional variance is positive almost surely for all t. To this end, we assume that all the time varying parameters in the infinite ARCH representation (17) are positive for all t; i.e., $\tilde{\omega}_t \geq 0$ and $\psi_{j,t} \geq 0$, for $j = 1, 2, \ldots$, and for $t = 1, \ldots, T$. For a FITVGARCH$(1, d, 1)$

model, and following Conrad and Haag (2006) for the FIGARCH$(1, d, 1)$ model, we can derive recursively the ARCH(∞) time varying coefficients in (17) as $\psi_{1,t} = \phi_{1,t} - \beta_{1,t} + d$ and $\psi_{j,t} = \beta_{1,t}\psi_{j-1,t} + \left(f_j - \phi_{1,t}\right)\left(-g_{j-1}\right)$ for all $j > 1$. $\psi_{1,t} \geq 0$ implies $\phi_{1,t} - \beta_{1,t} + d \geq 0$. If $\beta_{1,t} \geq 0$ and $\phi_{1,t} \leq f_2$ it follows recursively that $\psi_{j,t} \geq 0$ for all $j > 1$ since $g_{j-1} \leq 0$ and $\phi_{1,t} \leq f_j$ because f_j is increasing. Thus, to guarantee that the conditional variance remains non-negative almost surely for all t, it is sufficient to impose the conditions $\omega_t \geq 0$, $0 \leq \beta_{1,t} \leq \phi_{1,t} + d$ and $\phi_{1,t} \leq (1-d)/2$, i.e. $\omega_1 + \omega_2 F(t^*; \gamma, c) \geq 0$, $0 \leq \beta_{1,1} + \beta_{2,1}F(t^*; \gamma, c) \leq \phi_{1,1} + \phi_{2,1}F(t^*; \gamma, c) + d$ and $\phi_{1,1} + \phi_{2,1}F(t^*; \gamma, c) \leq (1-d)/2$. As $F(t^*; \gamma, c)$ is increasing function and bounded between 0 and 1, the sufficient conditions can be stated as $\omega_1 \geq 0$, $\omega_1 + \omega_2 \geq 0$, $0 \leq \beta_{1,1} \leq \phi_{1,1} + d$, $0 \leq \beta_{1,1} + \beta_{2,1} \leq \phi_{1,1} + \phi_{2,1} + d$, $\phi_{1,1} \leq (1-d)/2$ and $\phi_{1,1} + \phi_{2,1} \leq (1-d)/2$.

The conditional variance of the FITVGARCH(p, d, q) model in (16) can be written as

$$h_t = \omega_1 + \left[1 - \beta_1(L) - (1 - \phi_1(L))(1-L)^d\right]\varepsilon_t^2 + \beta_1(L)h_t + \left(\omega_2 + [\phi_2(L)(1-L)^d - \beta_2(L)]\varepsilon_t^2 + \beta_2(L)h_t\right)F(t^*; \gamma, c) \tag{18}$$

Note that if $d = 0$, (18) is reduced to the TV-GARCH model discussed in Amado and Teräsvirta (2008).

The FITVGARCH model can be considered as a regime switching model characterizing two extreme regimes, each associated with one of the two extreme values of the transition function $F(\cdot) = 0$ and $F(\cdot) = 1$. The transition between these two regimes is allowed to be smooth and is governed by the transition variable t^*. For $K = 1$, it is easy to see that if $\gamma = 0$, the transition function $F(\cdot)$ become equal to $\frac{1}{2}$ and the FITVGARCH model in (18) reduces to the FIGARCH model in (7) where $(\omega, \phi', \beta')' = (\omega_1, \phi_1', \beta_1')' + \frac{1}{2}(\omega_2, \phi_2', \beta_2')'$. This new long memory model is capable of generating instability in the volatility structure, which makes it an interesting tool for modeling financial market time series, exhibiting jointly long memory and structural change in their dynamic properties over time. Before fitting a specific time varying equation to the conditional variance in the FIGARCH model for volatility, it is common practice first to test whether this specification can be suitable for the data.

4 Testing parameter constancy

Testing parameter constancy is an important tool to check the adequacy of a model with parameter stability. The assumption of parameter constancy implies that the model's parameters remain constant across the estimation period. In the statistical and econometric literature, the maintained hypothesis of parameter stability has been tested both against specified and unspecified forms of alternative hypothesis. From unspecified alternative tests, see for example the CUSUM tests of Brown et al. (1975). Alternatively, parameter stability tests can be designed against a specified form. An example of specific alternative is that developed by Eitrheim and Teräsvirta (1996) for the STAR model. In this section, we propose an LM-type test for parameter constancy in the FIGARCH model, which explicitly allows the parameters to change smoothly over time based on the additive structure as specified in Sect. 3. If the null hypothesis of parameter constancy against smoothly changing parameters is rejected, one can conclude that the structure of the conditional variance of the process is changing over time.

The null hypothesis of parameter constancy can be expressed as equality of the FIGARCH parameters in the two regimes. As in Lin and Teräsvirta (1994); Eitrheim and Teräsvirta (1996), the alternative hypothesis is that the parameters may change smoothly over time. Thus, the null hypothesis can be stated as $H_0 : \gamma = 0$ against alternative hypothesis $H_1 : \gamma > 0$. Testing for parameter constancy is complicated because of the existence of unidentified nuisance parameters under the null hypothesis H_0. More explicitly, when $\gamma = 0$, $F(t^*; \gamma, c) = 1/2$. This makes the parameters γ and c not identified in (18) when the null hypothesis is valid. The identification problem is circumvented, following Luukkonen et al. (1988), using a Taylor expansion for the transition function about $\gamma = 0$. From a first-order Taylor approximation we obtain

$$
\begin{aligned}
F_1(t^*; \gamma, c) &= F(t^*; 0, c) + \gamma \left. \frac{\partial F(t^*; \gamma, c)}{\partial \gamma} \right|_{\gamma=0} + R_1(t^*; \gamma, c) \\
&= \frac{1}{2} + \frac{1}{4}\gamma \prod_{k=1}^{K} (t^* - c_k) + R_1(t^*; \gamma, c) \qquad (19)
\end{aligned}
$$

where $R_1(t^*; \gamma, c)$ is a remainder term. It is easy to see that (19) can be expressed as

$$
F_1(t^*; \gamma, c) = \frac{1}{2} + \sum_{k=0}^{K} \gamma c_k^* t^{*k} + R_1(t^*; \gamma, c) \qquad (20)
$$

where c_k^*, $k = 0, \ldots, K-1$, are functions of the parameters $c_1, \ldots, c_K$, $c_K^* = \frac{1}{4}$.

Replacing $F(t^*; \gamma, c)$ in (18) by $F_1(t^*; \gamma, c)$ in (20) gives

$$
\begin{aligned}
h_t &= \omega^* + [1 - \beta^*(L)]\varepsilon_t^2 - [1 - \phi^*(L)]e_t + \beta^*(L)h_t \\
&\quad + \sum_{k=1}^{K} \left(\delta_k t^{*k} + \varphi_k(L) e_t t^{*k} + \lambda_k(L)(h_t - \varepsilon_t^2) t^{*k} \right) + R_{1,t}^* \qquad (21)
\end{aligned}
$$

where

$$
\begin{aligned}
e_t &= (1 - L)^d \varepsilon_t^2; \\
\omega^* &= \omega_1 + \left(\frac{1}{2} + \gamma c_0^* \right) \omega_2; \\
\beta^*(L) &= \beta_1(L) + \left(\frac{1}{2} + \gamma c_0^* \right) \beta_2(L); \\
\phi^*(L) &= \phi_1(L) + \left(\frac{1}{2} + \gamma c_0^* \right) \phi_2(L);
\end{aligned}
$$

$$\delta_k = \gamma c_k^* \omega_2, \qquad k = 1, \ldots, K.$$
$$\varphi_k(L) = \gamma c_k^* \phi_2(L), \quad k = 1, \ldots, K.$$
$$\lambda_k(L) = \gamma c_k^* \beta_2(L), \quad k = 1, \ldots, K.$$
$$R_{1,t}^* = \left(\omega_2 + \phi_2(L)e_t + \beta_2(L)(h_t - \varepsilon_t^2)\right) R_1(t^*; \gamma, c)$$

with $\varphi_k = (\varphi_{k,1}, \ldots, \varphi_{k,q})'$ and $\lambda_k = (\lambda_{k,1}, \ldots, \lambda_{k,p})'$, $k = 1, \ldots, K$. It is clear that under $H_0 : \gamma = 0$, the parameters δ_k, $\varphi_{k,i}$ and $\lambda_{k,j}$, for $i = 1, \ldots, q$ and $j = 1, \ldots, p$, become equal to zero. Thus, the null hypothesis H_0 can be written equivalently as $H_0' : \delta_k = \varphi_{k,i} = \lambda_{k,j} = 0$. Under the null hypothesis of parameter constancy, $R_{1,t}^* = 0$, such that this remainder does not affect the distribution theory. Under H_0 the FITV-GARCH model reduces to a simple FIGARCH model.

Given that $\varepsilon_t | I^{t-1} \sim N(0, h_t)$, in (21), where I^{t-1} is the information set at time $t-1$, the conditional normal quasi log-likelihood function for observation t is given by

$$l_t(\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(h_t) - \frac{1}{2}\frac{\varepsilon_t^2}{h_t} \tag{22}$$

where $\theta = (\theta_1', \theta_2')'$, such that under H_0', $\theta_2 = 0$. Let $\theta_1 = (\omega^*, \phi^{*\prime}, \beta^{*\prime}, d)'$ and $\theta_2 = (\theta_{2,1}', \ldots, \theta_{2,K}')'$ where $\theta_{2,k} = (\delta_k, \varphi_k', \lambda_k')'$, $k = 1, \ldots, K$. The first order partial derivative of the conditional quasi log-likelihood function in (22), at time t, with respect to θ is

$$\frac{\partial l_t(\theta)}{\partial \theta} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{1}{h_t}\begin{pmatrix} z_{1,t} \\ z_{2,t} \end{pmatrix} \tag{23}$$

where $z_{1,t} = \frac{1}{h_t}\frac{\partial h_t}{\partial \theta_1}$ and $z_{2,t} = \frac{1}{h_t}\frac{\partial h_t}{\partial \theta_2}$.

This test can be carried out using standard LM test. Under H_0, the LM type statistic is defined as

$$LM = \frac{1}{2}\left(\sum_{t=1}^{T} \hat{u}_t \hat{z}_{2,t}'\right)\mathcal{A}_T^{-1}\left(\sum_{t=1}^{T} \hat{u}_t \hat{z}_{2,t}\right), \tag{24}$$

where

$$\mathcal{A}_T = \sum_{t=1}^{T} \hat{z}_{2,t}\hat{z}_{2,t}' - \sum_{t=1}^{T} \hat{z}_{2,t}\hat{z}_{1,t}'\left(\sum_{t=1}^{T} \hat{z}_{1,t}\hat{z}_{1,t}'\right)^{-1}\sum_{t=1}^{T} \hat{z}_{1,t}\hat{z}_{2,t}',$$

$\hat{u}_t = \left(\frac{\hat{\varepsilon}_t^2}{\hat{h}_t^0} - 1\right)$, $\hat{z}_{1,t} = \frac{1}{\hat{h}_t^0}\left.\frac{\partial \hat{h}_t}{\partial \theta_1}\right|_{H_0}$ and $\hat{z}_{2,t} = \frac{1}{\hat{h}_t^0}\left.\frac{\partial \hat{h}_t}{\partial \theta_2}\right|_{H_0}$.

Under H_0', the partial derivatives of the conditional variance h_t in (21) with respect to θ_1 are

$$\left.\frac{\partial \hat{h}_t}{\partial \theta_1}\right|_{H_0} = (\hat{w}_t', \hat{x}_t)' + \sum_{j=1}^{p} \hat{\beta}_j^* \left.\frac{\partial \hat{h}_{t-j}}{\partial \theta_1}\right|_{H_0} \tag{25}$$

where $\hat{w}_t = \left(1, \hat{e}_{t-1}, \ldots, \hat{e}_{t-q}, \left(\hat{h}_{t-1} - \hat{\varepsilon}_{t-1}^2\right), \ldots, \left(\hat{h}_{t-p} - \hat{\varepsilon}_{t-p}^2\right)\right)'$, $\hat{e}_t = (1 - L)^{\hat{d}} \varepsilon_t^2$ and $\hat{x}_t = \sum_{j=1}^{t-1} \frac{\hat{e}_{t-j}}{j} - \sum_{i=1}^{q} \sum_{n=1}^{t-i-1} \phi_i^* \frac{\hat{e}_{t-i-n}}{n}$.

The partial derivatives of h_t, under H_0', with respect to the parameters vector θ_2 are given by a vector of the partial derivative of h_t with respect to each parameter in θ_2

$$\left.\frac{\partial \hat{h}_t}{\partial \theta_2}\right|_{H_0} = \left(\left(\hat{w}_t t^*\right)', \left(\hat{w}_t t^{*2}\right)', \ldots, \left(\hat{w}_t t^{*K}\right)'\right)' + \sum_{j=1}^{p} \hat{\beta}_j^* \left.\frac{\partial \hat{h}_{t-j}}{\partial \theta_2}\right|_{H_0} \tag{26}$$

The proofs of (24)–(26) are provided in the Appendix A.

The partial derivatives of l_t with respect to θ, under the null hypothesis H_0, are given by

$$\left.\frac{\partial \hat{l}_t(\theta)}{\partial \theta}\right|_{H_0} = \frac{1}{2}\left(\frac{\hat{\varepsilon}_t^2}{\hat{h}_t^0} - 1\right)\begin{pmatrix}\hat{z}_{1,t}\\ \hat{z}_{2,t}\end{pmatrix} \tag{27}$$

In practice, the LM test of parameter constancy may be carried out using the following steps:

1. Estimate the parameters of the conditional variance under the null hypothesis and compute $\hat{u}_t = \left(\frac{\hat{\varepsilon}_t^2}{\hat{h}_t^0} - 1\right)$, $t = 1, \ldots, T$, and the sum of squares $SSR_0 = \sum_{t=1}^{T} \hat{u}_t^2$.
2. Regress $\hat{u}_t$ on $\hat{z}_{1,t}$ and $\hat{z}_{2,t}$, $t = 1, \ldots, T$ and compute the sum of squared residuals SSR_1.
3. Compute the LM test statistic as

$$LM_K = T\frac{SSR_0 - SSR_1}{SSR_0}$$

Under the null hypothesis, the statistic LM_K of parameter constancy test is χ^2 distribution with $K(p+q+1)$ degrees of freedom.

5 Size and power

In order to examine the empirical size and the power of the LM type test for parameter constancy in the finite sample we perform a set of Monte-Carlo experiments. We generate time series from FIGARCH model in (7) and time series from FITVGARCH model

Table 1 Empirical size of the parameter constancy test

d	β_1	ϕ_1	Nominal size (%)	$T = 1{,}000$			$T = 2{,}000$			$T = 3{,}000$		
				LM_1	LM_2	LM_3	LM_1	LM_2	LM_3	LM_1	LM_2	LM_3
0.25	0.35	0.15	1	0.9	1.6	2.1	1.3	1.3	1.8	1.9	1.3	2.0
			5	6.7	7.6	8.5	7.2	6.7	6.9	7.3	7.6	7.1
			10	12.9	12.9	13.9	12.1	12.8	11.8	12.5	15.4	12.8
0.5	0.2	−0.15	1	1.7	2.3	1.5	2.2	2.1	1.7	1.4	1.6	1.9
			5	7.4	7.2	8.0	6.2	6.8	8.1	6.3	7.1	6.5
			10	14.2	13.8	13.7	12.8	12.0	13.0	13.8	14.2	14.2
0.75	0.2	−0.15	1	1.0	1.6	1.2	1.6	1.4	1.7	0.9	1.4	1.3
			5	5.0	7.1	6.5	5.6	5.9	5.6	5.4	5.7	5.1
			10	11.5	12.8	12.2	11.8	10.5	11.5	13.0	11.1	10.5

Note: The table contains rejection frequencies of the null hypothesis of parameter constancy of the LM-type test where LM_K, $K = 1, 2$ and 3, denotes the LM-type test based on the Kth order logistic smooth transition regression. The data are generated by the DGP 1

in (18). The sample sizes for each model are $T = 1{,}000, T = 2{,}000$ and $T = 3{,}000$. The number of replications for each model is set equal to 1,000. The testing procedure discussed above is made for these simulated time series with $K = 1, 2$ and 3.

The data generating processes are:

DGP 1: FIGARCH$(1, d, 1)$ model in (5) with $\mu = 0$, $d = \{0.25, 0.5, 0.75\}$, $\omega = 0.1$, $\beta_1 = \{0.2, 0.35\}$, $\phi_1 = \{-0.15, 0.15\}$.

DGP 2: FITVGARCH$(1, d, 1)$ model in (18) with $d = \{0.25, 0.5, 0.75\}$, $\omega_1 = 0.1$, $\beta_{1,1} = \{0.2, 0.35, 0.7\}$, $\phi_{1,1} = \{-0.15, 0.1, 0.15\}$, $\omega_2 = \{-0.05, 0, 0.2\}$ $\beta_{2,1} = \{-0.3, -0.1, 0\}$, $\phi_{2,1} = \{-0.1, 0, 0.2, 0.25\}$, $c = 0.5$, $\gamma = \{10, 20, 50\}$.

The empirical sizes of the test are reported in Table 1 which contains the rejection frequencies of the null hypothesis of parameter constancy by the χ^2 version of the LM-type test, at nominal sizes 1, 5, and 10%, where the data are generated by the DGP 1. It is clear from the simulation results that the empirical sizes of the test are reasonably close to the nominal levels for all parameter combinations examined, for all sample sizes T and for all K. It is interesting to notice that there is no general tendency for the empirical size to deviate from the nominal size when increasing the sample size or changing K. However, we can note that the empirical sizes are improved when increasing d from 0.5 to 0.75.

Table 2 shows rejection frequencies, for series generated by the DGP 2. To conserve space, we report the empirical power only at 5% of significance. For $d = 0.25$, we investigate the cases of a change in one parameter, in two parameters or in the three parameters of the conditional variance. The power of the test when there is a change in β is higher than the cases of a change in ω or in ϕ. It seems that the test is very powerful when there is a change in the two parameters β and ϕ and that the power is negatively affected by allowing for a change in ω. For $d = 0.5$, the test is less powerful but when increasing d to 0.75, the power increases. It is clear that the power

Table 2 Empirical power of the parameter constancy test

ω_2	β_2	ϕ_2	γ	$T = 1{,}000$			$T = 2{,}000$			$T = 3{,}000$		
				LM_1	LM_2	LM_3	LM_1	LM_2	LM_3	LM_1	LM_2	LM_3
$d = 0.25,\ \beta_{1,1} = 0.35,\ \phi_{1,1} = 0.15$												
0.2	0	0	10	35.2	27.5	23.7	60.5	47.9	43.3	75.4	63.8	59.6
			20	38.6	31.2	27.6	65.7	53.6	53.2	80.4	70.7	70.8
			50	39.1	30.7	28.5	65.3	52.6	53.5	80.4	70.5	72.2
0	−0.3	0	10	60.7	48.9	45.9	90.2	81.7	77.6	98.4	95.0	93.7
			20	71.0	57.4	55.4	95.3	90.7	87.6	99.7	98.2	97.7
			50	73.9	60.3	60.6	96.4	92.0	90.8	99.7	99.0	99.0
0	0	0.2	10	31.8	24.8	24.2	54.8	41.4	38.8	74.4	61.8	57.0
			20	36.9	28.0	27.8	62.2	50.1	48.1	82.4	72.2	69.3
			50	37.8	29.5	29.5	64.9	53.2	51.9	84.2	74.6	73.6
0.2	−0.3	0	10	56.4	43.5	38.5	85.6	75.2	70.8	96.7	91.9	89.5
			20	64.8	51.4	47.9	91.1	85.0	83.2	98.5	96.4	96.6
			50	67.0	54.3	51.8	91.9	86.8	87.1	98.7	97.1	97.8
0.2	0	0.2	10	50.5	40.4	35.5	79.9	68.8	63.6	93.1	85.8	82.2
			20	58.7	47.6	43.5	86.8	77.7	74.8	96.3	92.0	91.5
			50	59.2	47.4	46.0	88.3	78.4	79.6	96.3	92.7	93.8
0	−0.3	0.2	10	94.4	89.5	86.2	100	100	99.6	100	100	100
			20	97.5	94.3	93.7	100	100	100	100	100	100
			50	98.0	95.4	95.5	100	100	100	100	100	100
0.2	−0.3	0.2	10	84.9	74.3	70.1	99.2	98.0	96.6	100	99.9	99.9
			20	91.4	83.2	83.9	99.8	99.6	99.6	100	100	100
			50	92.7	85.1	87.2	100	99.7	99.7	100	100	100
$d = 0.5,\ \beta_{1,1} = 0.2,\ \phi_{1,1} = -0.15$												
0.2	−0.1	0.25	10	58.9	46.6	41.6	88.3	78.6	74.4	96.9	92.4	90.5
			20	68.5	55.3	50.7	94.2	87.3	85.7	99.1	96.7	95.7
			50	72.1	57.3	56.0	95.1	88.8	88.5	99.4	97.9	96.8
$d = 0.75,\ \beta_{1,1} = 0.2,\ \phi_{1,1} = -0.15$												
0.2	−0.1	0.25	10	64.8	52.3	47.7	92.8	83.8	81.0	98.7	96.4	95.5
			20	74.2	61.4	58.3	96.6	93.0	91.2	99.9	98.9	98.6
			50	76.5	63.9	63.3	97.3	94.3	93.9	100	99.4	99.0
$d = 0.75, \beta_{1,1} = 0.7, \phi_{1,1} = 0.1$												
−0.05	−0.3	−0.1	10	77.9	69.3	64.7	97.3	93.0	91.8	99.4	99.1	98.5
			20	84.5	78.3	76.8	98.3	96.6	97.1	99.8	99.7	99.7
			50	86.1	80.5	80.5	98.7	97.9	98.0	99.8	99.7	99.7

Note: The table contains rejection frequencies of the null hypothesis of parameter constancy of the LM-type test where LM_K, $K = 1, 2$ and 3, denotes the LM-type test based on the Kth order logistic smooth transition regression. The data are generated by the DGP 2 of the test depends on the values of the parameters. The results prove also that the test becomes more powerful as the sample size and γ increase. For example, the results in the last panel of Table 2 indicate that with $K = 1$ and $\gamma = 20$, and by increasing the sample size T from 1,000 to 2,000 and to 3,000 improves the power from 84.5 to 98.3 and to 99.8%, respectively. Similarly, for $K = 1$ and $T = 2,000$, the power of the test increases from 97.3% for $\gamma = 10$ to 98.3% for $\gamma = 20$ and to 98.7% for $\gamma = 50$. It should be noted that as expected LM_1 is more powerful than LM_2 and LM_3 versions. This is explained by the fact that the parameter change is monotonic since $K = 1$ in the DGP 2. It is to note that the DGP produces nonmonotonic symmetrical change when $K = 2$ and nonmonotonic and nonsymmetrical change when $K = 3$. In summary, the simulations indicate that both empirical size and power of the constancy test are quite good.

6 Estimation

The most widely used estimation method for FIGARCH model is the Quasi Maximum Likelihood (QML) estimation procedure. As shown by Baillie et al. (1996) and Bollerslev and Wooldridge (1992), the QML estimates obtained with the assumption that the innovations are normally distributed behave relatively well. Following Bollerslev and Wooldridge (1992) and Baillie et al. (1996), we propose the QML procedure to estimate the FITVGARCH model. Consequently, the estimates for the parameters may be obtained by maximizing the following Gaussian log-likelihood function

$$l(\theta) = -\frac{T}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\left[\log(h_t) + \frac{\varepsilon_t^2}{h_t}\right] \tag{28}$$

where $\theta = (\omega_1, \phi_1', \beta_1', \omega_2, \phi_2', \beta_2', c, \gamma, d)'$ is the parameter vector of the FITVGARCH model defined in (18). Because of the positive value of the fractional differencing parameter d, it is required to use a sufficiently high truncation lag order. Indeed, as shown by Teyssière (1997) through Monte Carlo simulations, using a too low order induces severe biases. To keep this estimation problem aside, the truncation order of the infinite polynomial $(1 - L)^d$ is set to 1000 lags. Following Baillie et al. (1996), the pre-sample values of squared innovations, for $t = 0, -1, -2, \ldots, -1{,}000$, are set equal to the sample unconditional variance of the process. Baillie et al. (1996) suggest that pre-sample values might be expected to have a bigger impact than with stationary GARCH processes. Finally, starting values for the parameters must be fixed in the optimization procedure. As mentioned above, the most common choices of K in the logistic transition function (12) are $K = 1$ and $K = 2$. For simplicity we only focus in our study on the case of $K = 1$. Then, the logistic transition function in (12) becomes

$$F(t^*; \gamma, c) = \left[1 + \exp\left\{-\gamma(t^* - c)\right\}\right]^{-1}, \quad \gamma > 0 \tag{29}$$

Let us now move to the estimation of the variance–covariance matrix of the parameters. Under certain regularity conditions, the asymptotic distribution of the QMLE is

$$T^{1/2}(\hat{\theta} - \theta_0) \sim N(0, C) \tag{30}$$

where θ_0 denotes the true parameter vector. In the literature, there are three most widely used estimation methods of the asymptotic covariance matrix C. The first, as suggested by Efron and Hinkley (1978), is based on the estimates of the Hessian matrix $\hat{A}_T$ evaluated at the maximum likelihood estimated parameter vector $\hat{\theta}$, namely

$$\hat{C}_{HE} = \hat{A}_T^{-1} \tag{31}$$

where

$$\hat{A}_T = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 l_t(\hat{\theta})}{\partial\hat{\theta}\partial\hat{\theta}'} \tag{32}$$

with $l_t(\theta)$ is the log likelihood function at time t.

The second is based on the outer product of the gradient $\hat{B}_T$ suggested by Berndt et al. (1974).

$$\hat{C}_{OP} = \hat{B}_T^{-1} \tag{33}$$

where

$$\hat{B}_T = \frac{1}{T}\sum_{t=1}^{T}\hat{g}_t\hat{g}_t' \tag{34}$$

where $\hat{g}_t$ is the gradient of the log likelihood function evaluated at $\hat{\theta}$.

$$\hat{g}_t = \frac{\partial l_t(\hat{\theta})}{\partial\hat{\theta}} \tag{35}$$

The last estimator is known as the Quasi Maximum Likelihood Estimator (QMLE) (White (1982))

$$\hat{C}_{QMLE} = \hat{A}_T^{-1}\hat{B}_T\hat{A}_T^{-1} \tag{36}$$

As it is clear from the different estimation methods of the covariance matrix of the parameters, first and second partial derivatives of the log likelihood function of the FITVGARCH model in (18) must be computed. Consider the log likelihood function at time t.

$$l_t(\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(h_t) - \frac{1}{2}\frac{\varepsilon_t^2}{h_t}, \tag{37}$$

the first partial derivative of $l_t(\theta)$ with respect to the parameter vector θ is given by

$$\frac{\partial l_t(\theta)}{\partial \theta} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{1}{h_t}\frac{\partial h_t}{\partial \theta} \tag{38}$$

The partial derivatives of the conditional variance to build the outer product of gradient matrix are contained in Appendix B.

The asymptotic properties of the QMLE for the ARCH and GARCH models have been studied under various conditions. Lee and Hansen (1994) and Lumsdaine (1996) proved consistency and asymptotic normality of the QMLE for the GARCH(1,1) where the process is strictly stationary and ergodic. Berkes et al. (2003) shows consistency and asymptotic normality of QMLE for the general strictly stationary and ergodic GARCH(*p*,*q*) model.

Jensen and Rahbek (2004) proved consistency and asymptotic normality of the GARCH(1,1) even when the process is nonstationary and nonergodic. Formal results of the asymptotic properties of QMLE for the FIGARCH model are not yet available. However, empirical evidence for the FIGARCH process, based on Monte Carlo simulation, suggests that QMLE is consistent and asymptotically normal (see, for example, Baillie et al. (1996)). The formal proofs of the asymptotic properties of QMLE for the FIGARCH process as well as for the FITVGARCH process are beyond the scope of this paper. However, in order to assess the adequacy of this estimation method for the FITVGARCH process, we performed a simulation study where the model is simulated with FITVGARCH$(1, d, 1)$ model in (18) with $d = \{0.25, 0.5, 0.75\}, \omega_1 = 0.1, \beta_{1,1} = \{0.2, 0.35\}, \phi_{1,1} = \{-0.15, 0.15\}, \omega_2 = 0.2, \beta_{2,1} = \{-0.3, -0.1\}, \phi_{2,1} = \{0.2, 0.25\},\ c = 0.5, \gamma = 20$. We have generated 1,000 replications for each design. The sample sizes are $T = 1{,}000, 2{,}000$, and 3,000. We report in Table 3 the simulation results of estimating the long memory parameter and the threshold parameter, which determines the date of the structural change, from the FITVGARCH model. We present the average bias, the root mean squared errors (RMSE) as well as the average of the standard error (SE) for each parameter. The simulations indicate that both long memory parameter and threshold parameter are estimated very well by the QMLE method. Indeed, the results suggest a very small bias in both parameters. The bias in the differencing parameter is negative only for the first parameter design, that is, when d is equal to 0.25 and positive when d increases to 0.5 and 0.75. Interestingly, the simulation results indicate that the long memory parameter is slightly overestimated, especially when is large and that the bias is decreasing, in absolute value, with the sample size T. The bias of c is usually negative which suggests that the QMLE method slightly underestimates the threshold parameter.

Both RMSE and SE of the differencing parameter estimates tend to decrease as the sample size increases. For the threshold parameter, there is slight improvement in term of RMSE when T increases. Hence, the quality of the application of the QMLE is generally very satisfactory, in that the degree of persistence and de location of the change are correctly estimated. To check the effect of ignoring structural change on long memory parameter estimates, we report in Table 4 the simulation results of estimating the long memory parameter from the standard FIGARCH model when the data are generated using the FITVGARCH process. For $d = 0.25$, we investigate the cases of a change in one parameter, in two parameters or in the three parameters of the conditional variance. The results indicate that d frequently exhibits a negative bias, except when there is only a change in the constant parameter or in the ω and ϕ parameters. In these cases, the bias is positive, increasing with the sample size, and is more important in the case of only a change in ω. This upward bias in long memory parameter estimates seems to be caused by neglecting structural change in the constant parameter. This finding is consistent with the results in Baillie and Morana (2009) in that their simulations also suggest an upward bias in the differencing parameter estimated from the FIGARCH model when generating the data by the A-FIGARCH model, in which only the constant parameter is subject to structural change. However, it is important to note that no additional persistence is detected when allowing for structural change in the constant parameter jointly with a change in the β parameter. The same result holds when allowing for a change only in β parameter and/or in the ϕ parameter. In these cases the bias is very small and decreasing, in absolute value, with T. It is also to be mentioned that when allowing for structural change in all parameters leads to a downward bias in estimates of the differencing parameter, but the bias is decreasing, in absolute value, as the sample size increases. For $d = 0.5$ and $d = 0.75$, the simulation results suggest a very small positive bias for $d = 0.5$ and negative one for $d = 0.75$. Comparing the simulation results of estimating d from Tables 3 and 4 reveals that the bias is very small for both models, with slightly higher bias, in absolute value, for the FIGARCH model than for the FITVGARCH model when d is equal to 0.25. In term of RMSE, the results indicate that both models give almost the same RMSE of long memory parameter estimates in all cases. However, the SE of the estimate of d is generally lower from the estimation of the FITVGARCH model compared to the corresponding FIGARCH model. Hence, the FITVGARCH model seems to work quite well in modeling both long memory and structural change in time series.

Table 3 Simulation results of estimating the FITVGARCH(1,d,1) model

d	$\beta_{1,1}$	$\phi_{1,1}$	ω_2	β_2	ϕ_2	T	$\hat{d}$			$\hat{c}$		
							Bias	RMSE	SE	Bias	RMSE	SE
0.25	0.35	0.15	0.2	−0.3	0.2	1,000	−0.033	0.103	0.077	−0.016	0.199	0.090
						2,000	−0.015	0.062	0.052	−0.026	0.193	0.131
						3,000	−0.010	0.048	0.042	−0.017	0.182	0.111
0.5	0.2	−0.15	0.2	−0.1	0.25	1,000	0.039	0.126	0.103	−0.007	0.215	0.140
						2,000	0.025	0.078	0.067	−0.025	0.204	0.128
						3,000	0.019	0.060	0.051	−0.012	0.206	0.143
0.75	0.2	−0.15	0.2	−0.1	0.25	1,000	0.035	0.088	0.082	−0.019	0.214	0.152
						2,000	0.016	0.055	0.052	−0.029	0.210	0.129
						3,000	0.011	0.043	0.041	−0.012	0.210	0.145

Note: The table reports the average bias, the root mean squared errors (RMSE) and the average of the standard error (SE) of the QMLE of the estimates of d and c from the FITVGARCH(1,d,1) model. Simulations are based on 1,000 replications generated from the FITVGARCH(1,d,1) model in (18). In all cases, $\omega_1 = 0.1, c = 0.5$, and $\gamma = 20$. The sample sizes are $T = 1{,}000$, 2,000 and 3,000

Table 4 Simulation results: effects of ignoring structural change

d	$\beta_{1,1}$	$\phi_{1,1}$	ω_2	β_2	ϕ_2	$T = 1{,}000$			$T = 2{,}000$			$T = 3{,}000$		
						Bias	RMSE	SE	Bias	RMSE	SE	Bias	RMSE	SE
0.25	0.35	0.15	0.2	0	0	0.062	0.109	0.112	0.069	0.088	0.056	0.071	0.083	0.040
0.25	0.35	0.15	0	−0.3	0	−0.011	0.090	0.216	−0.004	0.060	0.171	0.000	0.048	0.230
0.25	0.35	0.15	0	0	0.2	−0.017	0.086	0.250	−0.011	0.056	0.193	−0.008	0.044	0.113
0.25	0.35	0.15	0.2	−0.3	0	−0.014	0.079	0.642	−0.004	0.051	0.323	0.001	0.039	0.418
0.25	0.35	0.15	0.2	0	0.2	0.036	0.083	0.242	0.044	0.065	0.112	0.047	0.059	0.066
0.25	0.35	0.15	0	−0.3	0.2	−0.024	0.101	0.128	−0.020	0.071	0.091	−0.016	0.058	0.096
0.25	0.35	0.15	0.2	−0.3	0.2	−0.060	0.097	0.205	−0.053	0.071	0.134	−0.048	0.061	0.081
0.5	0.2	−0.15	0.2	−0.1	0.25	0.028	0.111	0.483	0.018	0.074	0.180	0.011	0.058	0.204
0.75	0.2	−0.15	0.2	−0.1	0.25	−0.004	0.082	0.209	−0.012	0.057	0.073	−0.017	0.048	0.046

Note: The table reports the average bias, the root mean squared errors (RMSE) and the average of the standard error (SE) of the QMLE of the estimates of d from the FIGARCH(1,d,1) model. Simulations are based on 1,000 replications generated from the FITVGARCH(1, d, 1) model in (18). In all cases, $\omega_1 = 0.1$, $c = 0.5$, and $\gamma = 20$. The sample sizes are $T = 1{,}000$, 2,000, and 3,000

7 Misspecification tests

The final stage of specifying the FITVGARCH model consists on evaluating it using misspecification tests. In this section, we focus on testing serial dependence in the squared standardized errors. We use the test proposed by Lundbergh and Teräsvirta (2002) as a parametric LM type tests of no ARCH in standardized errors. To describe the test we assume that the misspecification structure in the FITVGARCH model has the additive form

$$\begin{aligned} h_t &= \omega_1 + [1 - \beta_1(L) - (1 - \phi_1(L))(1 - L)^d]\varepsilon_t^2 + \beta_1(L)h_t \\ &\quad + \left(\omega_2 + [\phi_2(L)(1 - L)^d - \beta_2(L)]\varepsilon_t^2 + \beta_2(L)h_t\right) F(t^*; \gamma, c) \\ &\quad + \pi' \nu_t \end{aligned} \tag{39}$$

where $\nu_t = (\eta_{t-1}^2, \ldots, \eta_{t-r}^2)'$ and $\pi = (\pi_1, \ldots, \pi_r)'$. The null hypothesis of no serial dependence in η_t^2 up to the rth order is defined as $H_0 : \pi_1 = \pi_2 = \cdots = \pi_r = 0$. Under H_0, the LM type statistic is defined as in (24) where $\hat{u}_t = \left(\frac{\hat{\varepsilon}_t^2}{\hat{h}_t^0} - 1\right)$, $\hat{z}_{1,t} = \frac{1}{\hat{h}_t^0} \left.\frac{\partial \hat{h}_t}{\partial \theta}\right|_{H_0}$, $\theta = (\theta_1', \theta_2', c, \gamma, d)'$, $\theta_1 = (\omega_1, \phi_1', \beta_1')'$, $\theta_2 = (\omega_2, \phi_2', \beta_2')'$ and

$$\begin{aligned} \hat{z}_{2,t} &= \frac{1}{\hat{h}_t^0} \left.\frac{\partial \hat{h}_t}{\partial \pi}\right|_{H_0} \\ &= \frac{1}{\hat{h}_t^0} \left((\hat{\eta}_{t-1}^2, \ldots, \hat{\eta}_{t-r}^2)' + \sum_{j=1}^{p} \left[\hat{\beta}_{1,j} + \hat{\beta}_{2,j} F(t^*; \hat{\gamma}, \hat{c})\right] \left.\frac{\partial \hat{h}_{t-j}}{\partial \pi}\right|_{H_0} \right) \end{aligned}$$

The *LM* test for rth order serial dependence in the squared standardized errors can be performed in three stages as follows:

1. Estimate the parameters of the FITVGARCH model under the null hypothesis and compute $\hat{u}_t = (\frac{\hat{\varepsilon}_t^2}{\hat{h}_t^0} - 1)$, $t = 1, \ldots, T$, and the sum of squares $SSR_0 = \sum_{t=1}^{T} \hat{u}_t^2$.
2. Regress $\hat{u}_t$ on $\hat{z}_{1,t}$ and $\hat{z}_{2,t}$, $t = 1, \ldots, T$ and compute the sum of squared residuals SSR_1.
3. Compute the LM test statistic as

$$LM = T \frac{SSR_0 - SSR_1}{SSR_0}$$

which under the null hypothesis is approximately χ^2 distributed with r degrees of freedom.

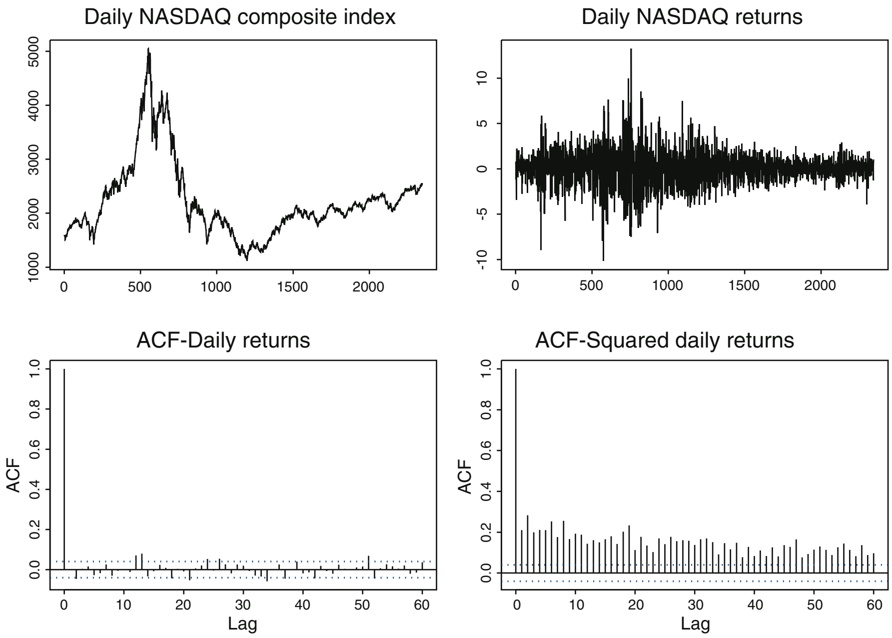

Fig. 1 Graphical presentation of time series for daily NASDAQ composite index. *Upper panel* displays time series plots for daily index and for daily index returns. *Lower panel* displays autocorrelation functions for daily returns and for daily squared returns

8 Empirical evidence

8.1 Data

In this section, The FITVGARCH model is performed for US stock market. The data we consider is the daily NASDAQ composite index spanning the period from January 2, 1998 to May 2, 2007, with a total of 2,346 observations. In Fig. 1, the upper panel shows the level and the daily returns time series of the NASDAQ composite index. Daily returns are obtained by taking 100 times the first difference of the natural log of the index. A visual inspection of this plot, gives the impression that the structure of the daily returns time series exhibits two different periods of volatility. That is, the volatility displays structural change over time. More precisely, the first period of daily returns time series is characterized by very high volatility whereas the second period is described by a very low volatility. This can be explained by the fact that the dynamic structure of the volatility has changed over time. Another remark which appears interesting is that the change was not carried out instantaneously. This can be preliminary evidence in favour of the use of smooth transition model to these data. The lower panel of Fig. 1 displays autocorrelation functions for daily returns and for daily squared returns. It is clear that daily returns seem to be not autocorrelated. However, for the squared returns, autocorrelations show strong temporal dependence and exhibit a hyperbolic rate of decay. Those features may suggest that long-range dependence

Table 5 LM-type tests for parameter constancy of FIGARCH model against FITVGARCH model for NASDAQ composite daily index returns

LM_k	LM statistic	p-Value
LM_1	19.702	2×10^{-4}
LM_2	34.134	6×10^{-6}
LM_3	42.164	3×10^{-6}

Note: The table contains LM statistics and p-values of parameter constancy test where LM_K, $K = 1, 2$, and 3, denotes the LM-type test for parameter constancy based on the Kth order logistic smooth transition regression

of squared returns may be modelled by a fractionally integrated process. In view of all this, we propose our new FITVGARCH model to describe the volatility structure of the NASDAQ composite daily index returns. In fact, this model is able to capture both long memory and structural change in the volatility process.

8.2 Testing for parameter constancy

We begin the modelling procedure by testing parameter constancy in the standard FIGARCH $(1, d, 1)$ model against smoothly changing parameters (FITVGARCH (1,d,1) model) using the test discussed above.

Table 5 contains the statistics and the corresponding p-values of the LM test for constancy of the FIGARCH parameters against the alternative of time dependent parameters as specified by the FITVGARCH model. The test is performed based on the Kth order logistic smooth transition regression; $K = 1, 2$ and 3. The results indicate that there is strong evidence of time varying FIGARCH parameters since the null hypothesis of parameter constancy is strongly rejected for all orders K. Those results are in accord with the idea that volatility structure of the NASDAQ composite index returns is changing over time.

8.3 Estimation results

The second stage of the analysis is the estimation of FITVGARCH$(1, d, 1)$ model in (18) for the data based on the quasi maximum likelihood estimation procedure discussed above. As the estimation of the ful model indicated that $\hat{\phi}_{1,1}$ is not significant even at 10% level and does not contribute to the explanatory power of the model, we removed this parameter from the conditional variance equation and we repeated the estimation. The results are reported in the columns 1 and 2 of Table 6. Columns 1 shows parameter estimates while standard errors are given in columns 2. Standard errors are estimated using the outer product of gradient method. For comparison purposes, we also estimate a FIGARCH$(1, d, 1)$ model where the parameter estimates and their asymptotic standard errors are reported, respectively, in columns 3 and 4 of Table 6. The long memory parameter $\hat{d}$ is significantly different from zero for both models; of about 0.51 for the FITVGARCH model and 0.43 for the FIGARCH. The

Table 6 Summary of estimated models for NASDAQ composite daily index returns

Parameters	FITVGARCH(1, d, 1)		FIGARCH(1, d, 1)	
$\hat{d}$	0.505	(0.061)	0.435	(0.064)
$\hat{c}$	0.354	(0.100)	–	–
$\hat{\gamma}$	15.09	(16.09)	–	–
$\hat{\omega}_1$	0.190	(0.076)	0.044	(0.013)
$\hat{\phi}_{1,1}$	–	–	0.052	(0.041)
$\hat{\beta}_{1,1}$	0.415	(0.084)	0.483	(0.075)
$\hat{\omega}_2$	−0.187	(0.076)	–	–
$\hat{\phi}_{2,1}$	0.237	(0.060)	–	–
$\hat{\beta}_{2,1}$	0.327	(0.070)	–	–
$Q(20)$	16.268	[0.700]	15.431	[0.751]
$Q^2(20)$	16.839	[0.663]	13.382	[0.860]
LM_{SC}	26.329	[0.155]	–	–
AIC	8,552.57		8,567.29	
Log Lik	−4,268.29		−4,279.65	

Note: Standard errors are given in parentheses

fact that the estimated value of the long memory parameter decreases when estimating it by the FIGARCH model should not be surprising given the simulation results reported in Sect. 6. It has been noted that only a change in the constant parameter leads to an upward bias in the estimate of d. However, when all conditional variance parameters are subject to structural change, the parameter d is biased downward. The significance of d indicates strong evidence of long memory in the squared returns. The conditional variance parameters of the FITVGARCH model are highly significant in the two regimes. For the FIGARCH model, the parameter $\hat{\phi}_1$ is significant only at 10% level while the parameters $\hat{\omega}$ and $\hat{\beta}_1$ are highly significant at 5% significance level. Looking now at the transition function parameters in the FITVGARCH model, the threshold parameter $\hat{c}$ is highly significant with a value about 0.354 indicating that the structural change of the volatility process was happen at time $\hat{t} = 0.354 \times T$, ($\hat{t} \simeq 830$), where T is the number of observations. Nevertheless, this structural change is not instantaneously at time of the turning point because of the smoothness of the transition function. In fact, this structural change can be instantaneous at turning points if the smoothness parameter γ is very large. This is not the case for our data since the estimates parameter $\hat{\gamma}$ is equal to 15.09 that is not sufficiently high to imply a quick change between regimes. This can be observed from the first panel of Fig. 2 which plots the logistic transition function in (29) as a function of time. It is clear that the transition between the extreme regimes is rather smooth. The long term volatility level depends on the estimates of the constant parameter in the conditional variance. In the second panel of Fig. 2, we contrast the estimated time varying parameter $\hat{\omega}_t$ from the FITVGACRH model with the constant one, $\hat{\omega}$, from the FIGARCH model. It is clear from the figure that, using the FIGARCH model, the long term volatility is largely underestimated for the first 830 observations and is overestimated for the latter observations.

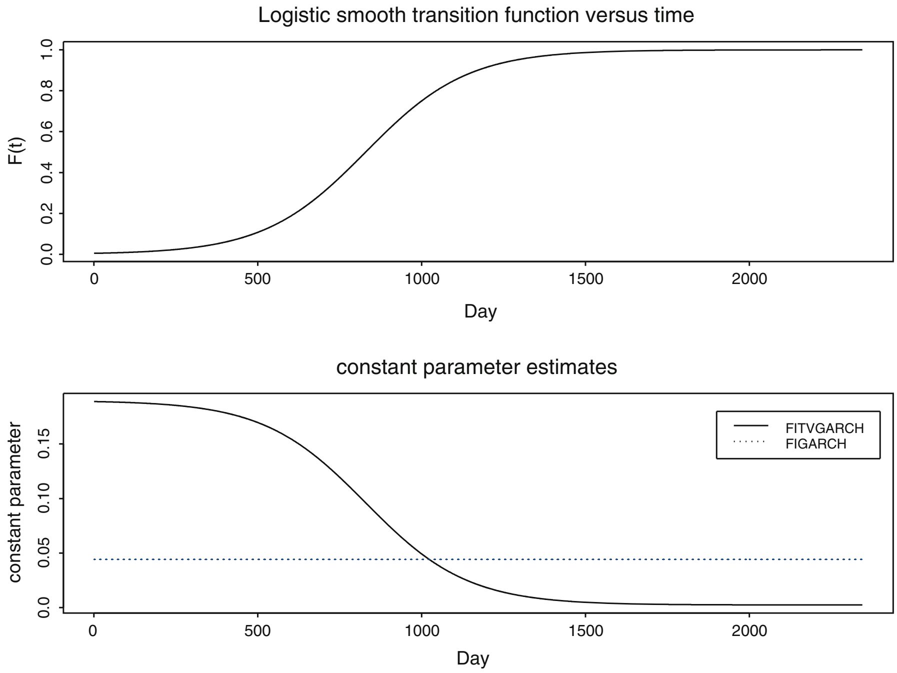

Fig. 2 Estimated logistic transition function from the FITVGARCH model and estimated constant parameter in the conditional variance from FIGARCH and FITVGARCH models

Comparing the FIGARCH $(1, d, 1)$ model with FITVGARCH$(1, d, 1)$ model, the first panel of Fig. 3 indicates that both models tend to have high and increasing volatility estimates in the first regime and decreasing volatility estimates in the second regime. It seems also that the volatility estimates have persistence property which confirms the evidence of long memory and structural change behaviours in the volatility. The residuals obtained from both models, showed in the second panel of Fig. 3, cannot reject the null of white noise series according to the Ljung-Box portmanteau test statistic $Q(20)$. Similarly, Ljung-Box statistic $Q^2(20)$ indicates that the hypothesis of serial dependence, up to order 20, in squared standardized errors is strongly rejected for both models. The misspecification test discussed in section 7 also indicates, according to LM_{SC} statistic, that the squared standardized errors from the FITVGARCH model seem to not be autocorrelated up to the 20th order. However, it is not difficult to reach a conclusion that the empirical evidences are in favor of FITVGARCH$(1, d, 1)$ model for NASDAQ index returns time series according to the Akaike information criterion (AIC) and to the log-likelihood values.

9 Conclusions

In this paper, a new FITVGARCH model was proposed to capture both long memory and structural change in the volatility process. The model allows for time

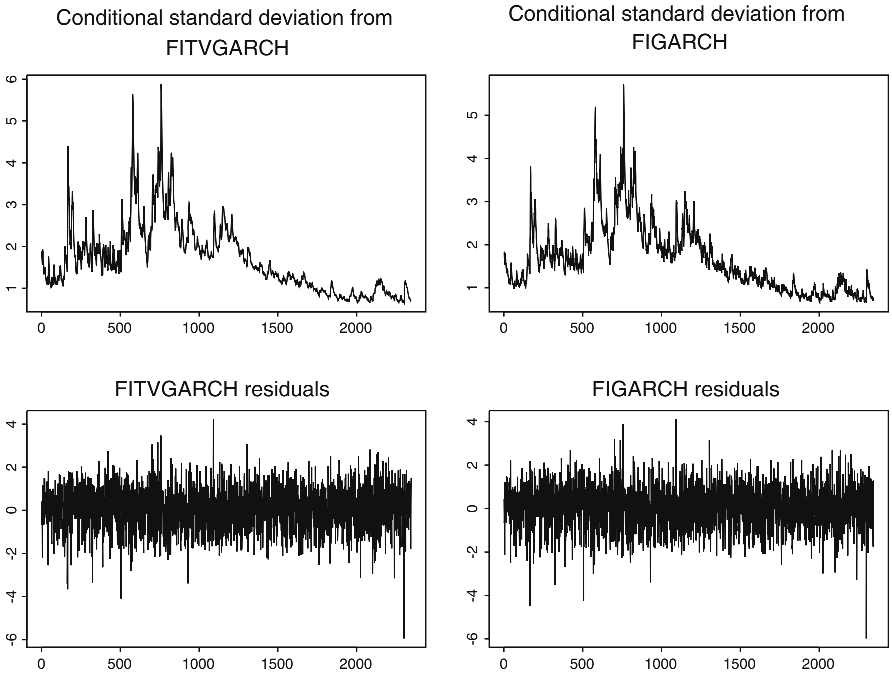

Fig. 3 Estimation results of FITVGARCH and FIGARCH models for daily NASDAQ index returns. *Upper panel* displays the estimated conditional standard deviation. *Lower panel* displays the estimated residuals

varying dynamic structure in the conditional variance of the process. The structural change is assumed to be smooth between regimes. More precisely, the conditional variance parameters of the FIGARCH model are allowed to change smoothly over time. We have derived an LM-type test for parameter constancy of the FIGARCH model against the alternative of time dependent parameters (FITVGARCH model). Simulation analysis shows that both empirical size and power of the constancy test are quite good. The quality of the application of the QMLE for the FITVGARCH model, examined by simulation study, is generally very satisfactory. Our application has been to NASDAQ stock market volatility. Results indicate that this new class of model seems to outperform the FIGARCH model in the description of the daily NASDAQ composite index returns. The volatility of these data seems to be characterized by both long memory and structural change. Indeed, the volatility structure changes over time where the transition between the extreme regimes seems to be smooth. For instance, we have assumed that the volatility structure, with long memory property, changes between two regimes. Further research should also examine the feasibility of considering more than one structural change.

Acknowledgments The authors wish to thank two anonymous referees for helpful suggestions resulting in a much-improved paper.

Appendix

A Analytical derivatives for parameter constancy test

In order to derive the parameter constancy test, partial derivatives of the log likelihood function of the FITVGARCH model, with respect to the parameters are considered. The first order partial derivative of the conditional quasi log-likelihood function in (22), at time t, with respect to θ is

$$\frac{\partial l_t(\theta)}{\partial \theta} = \frac{1}{2}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\frac{1}{h_t}\frac{\partial h_t}{\partial \theta} \tag{A.1}$$

The average score vector is defined as

$$g = \frac{1}{T}\sum_{t=1}^{T}\frac{\partial l_t(\theta)}{\partial \theta} \tag{A.2}$$

The average score vector can be partitioned as $g = (g_1', g_2')'$ such that $g_1 = (1/T)\sum_{t=1}^{T} \partial l_t(\theta)/\partial \theta_1$ and $g_2 = (1/T)\sum_{t=1}^{T} \partial l_t(\theta)/\partial \theta_2$. Under the null hypothesis $\hat{g}_1 = 0$ and the Lagrange multiplier statistic is given by

$$LM = T\hat{g}' I(\hat{\theta})^{-1}\hat{g} \tag{A.3}$$

where $I(\hat{\theta})$ is the information matrix evaluated under the null hypothesis. The information matrix is defined as the expected negative value of the average Hessian matrix A_T

$$I(\theta) = E\left[-A_T\right]$$

where

$$\begin{aligned} A_T &= \frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 l_t(\theta)}{\partial \theta \partial \theta'} \\ &= \frac{1}{T}\sum_{t=1}^{T}\frac{1}{2h_t}\left(\frac{\varepsilon_t^2}{h_t} - 1\right)\left(\frac{\partial^2 h_t}{\partial \theta \partial \theta'} - \frac{1}{h_t}\frac{\partial h_t}{\partial \theta}\frac{\partial h_t}{\partial \theta'}\right) - \frac{\varepsilon_t^2}{2h_t^3}\frac{\partial h_t}{\partial \theta}\frac{\partial h_t}{\partial \theta'} \end{aligned}$$

which gives

$$I(\theta) = E\left[\frac{1}{T}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{2h_t^3}\frac{\partial h_t}{\partial \theta}\frac{\partial h_t}{\partial \theta'}\right] = \frac{1}{2T}\sum_{t=1}^{T}E(z_t z_t')$$

where $z_t = (z'_{1,t}, z'_{2,t})'$, $z_{1,t} = \frac{1}{h_t}\frac{\partial h_t}{\partial \theta_1}$ and $z_{2,t} = \frac{1}{h_t}\frac{\partial h_t}{\partial \theta_2}$. So that, the information matrix can be estimated consistently as

$$I(\hat{\theta}) = \frac{1}{2T}\sum_{t=1}^{T} \hat{z}_t \hat{z}'_t$$

where $\hat{z}_t = (\hat{z}'_{1,t}, \hat{z}'_{2,t})'$, $\hat{z}_{1,t} = \frac{1}{\hat{h}_t^0}\left.\frac{\partial h_t}{\partial \theta_1}\right|_{H_0}$ and $\hat{z}_{2,t} = \frac{1}{\hat{h}_t^0}\left.\frac{\partial h_t}{\partial \theta_2}\right|_{H_0}$. Let $\hat{u}_t = \left(\frac{\hat{\varepsilon}_t^2}{\hat{h}_t^0} - 1\right)$, the average score vector evaluated under the null hypothesis is obtained by

$$\hat{g} = \frac{1}{2T}\sum_{t=1}^{T} \hat{u}_t (0, \hat{z}'_{2,t})'$$

Thus, the Lagrange multiplier test statistic in (A.3) becomes

$$\begin{aligned} LM &= T\left(\frac{1}{2T}\sum_{t=1}^{T} \hat{u}_t(0, \hat{z}'_{2,t})'\right)'\left(\frac{1}{2T}\sum_{t=1}^{T} \hat{z}_t \hat{z}'_t\right)^{-1}\left(\frac{1}{2T}\sum_{t=1}^{T} \hat{u}_t(0, \hat{z}'_{2,t})'\right) \\ &= \frac{1}{2}\left(\sum_{t=1}^{T} \hat{u}_t \hat{z}'_{2,t}\right)\mathcal{A}_T^{-1}\left(\sum_{t=1}^{T} \hat{u}_t \hat{z}_{2,t}\right), \end{aligned}$$

where

$$\mathcal{A}_T = \sum_{t=1}^{T} \hat{z}_{2,t}\hat{z}'_{2,t} - \sum_{t=1}^{T} \hat{z}_{2,t}\hat{z}'_{1,t}\left(\sum_{t=1}^{T} \hat{z}_{1,t}\hat{z}'_{1,t}\right)^{-1}\sum_{t=1}^{T} \hat{z}_{1,t}\hat{z}'_{2,t}.$$

The partial derivatives of the conditional variance h_t with respect to the parameter vector θ_1 are

$$\begin{aligned} \frac{\partial h_t}{\partial \omega^*} &= 1 + \beta^*(L)\frac{\partial h_t}{\partial \omega^*} + \sum_{k=1}^{K}\left(\lambda_k(L)\frac{\partial h_t}{\partial \omega^*}t^{*k}\right) + \frac{\partial R^*_{1,t}}{\partial \omega^*} \\ \frac{\partial h_t}{\partial \phi^*} &= (e_{t-1}, \ldots, e_{t-q})' + \beta^*(L)\frac{\partial h_t}{\partial \phi^*} + \sum_{k=1}^{K}\left(\lambda_k(L)\frac{\partial h_t}{\partial \phi^*}t^{*k}\right) + \frac{\partial R^*_{1,t}}{\partial \phi^*} \\ \frac{\partial h_t}{\partial \beta^*} &= ((h_{t-1} - \varepsilon_{t-1}^2), \ldots, (h_{t-p} - \varepsilon_{t-p}^2))' + \beta^*(L)\frac{\partial h_t}{\partial \beta^*} + \sum_{k=1}^{K}\left(\lambda_k(L)\frac{\partial h_t}{\partial \beta^*}t^{*k}\right) \\ &\quad + \frac{\partial R^*_{1,t}}{\partial \beta^*} \end{aligned}$$

$$\frac{\partial h_t}{\partial d} = -[1-\phi^*(L)]\frac{\partial e_t}{\partial d} + \beta^*(L)\frac{\partial h_t}{\partial d} + \sum_{k=1}^{K}\left(\varphi_k(L)\frac{\partial e_t}{\partial d}t^{*k} + \lambda_k(L)\frac{\partial h_t}{\partial d}t^{*k}\right) + \frac{\partial R^*_{1,t}}{\partial d}$$

Under H'_0, $\varphi_k = \lambda_k = 0$ and $R^*_{1,t} = 0$, then the partial derivative of h_t with respect to d is

$$\left.\frac{\partial h_t}{\partial d}\right|_{H_0} = -[1-\phi^*(L)] \left.\frac{\partial e_t}{\partial d}\right|_{H_0} + \beta^*(L) \left.\frac{\partial h_t}{\partial d}\right|_{H_0}$$

where $e_t = (1-L)^d\varepsilon_t^2$. The derivative of e_t with respect to d is

$$\begin{aligned}\frac{\partial e_t}{\partial d} &= \frac{\partial(1-L)^d}{\partial d}\varepsilon_t^2\\ &= \ln(1-L)(1-L)^d\varepsilon_t^2\\ &= -\sum_{n=1}^{\infty}\frac{L^n}{n}e_t\end{aligned}$$

Under H_0,

$$\left.\frac{\partial e_t}{\partial d}\right|_{H_0} = -\sum_{n=1}^{t-1}\frac{\hat{e}_{t-n}}{n}$$

where $\hat{e}_t = (1-L)^{\hat{d}}\varepsilon_t^2$, then

$$\left.\frac{\partial h_t}{\partial d}\right|_{H_0} = [1-\phi^*(L)]\sum_{n=1}^{t-1}\frac{\hat{e}_{t-n}}{n} + \beta^*(L) \left.\frac{\partial h_t}{\partial d}\right|_{H_0} \tag{A.4}$$

$$= \sum_{j=1}^{t-1}\frac{\hat{e}_{t-j}}{j} - \sum_{i=1}^{q}\sum_{n=1}^{t-i-1}\phi_i^*\frac{\hat{e}_{t-i-n}}{n} + \sum_{j=1}^{p}\beta_j^* \left.\frac{\partial h_{t-j}}{\partial d}\right|_{H_0} \tag{A.5}$$

Let $w_t = (1, e_{t-1}, \ldots, e_{t-q}, (h_{t-1}-\varepsilon_{t-1}^2), \ldots, (h_{t-p}-\varepsilon_{t-p}^2))'$ and

$$x_t = \sum_{j=1}^{t-1}\frac{\hat{e}_{t-j}}{j} - \sum_{i=1}^{q}\sum_{n=1}^{t-i-1}\phi_i^*\frac{\hat{e}_{t-i-n}}{n}.$$

Under H'_0, $\lambda_k = 0$ and $R^*_{1,t} = 0$, then the partial derivatives of h_t with respect to θ_1 is equal to

$$\left.\frac{\partial h_t}{\partial \theta_1}\right|_{H_0} = (w_t', x_t)' + \sum_{j=1}^{p} \beta_j^* \left.\frac{\partial h_{t-j}}{\partial \theta_1}\right|_{H_0} \tag{A.6}$$

The partial derivatives of h_t with respect to the parameters vector θ_2 are given by a vector of the partial derivative of h_t with respect to each parameter in θ_2 where $\theta_2 = (\theta_{2,1}', \ldots, \theta_{2,K}')', \theta_{2,m} = (\delta_m, \varphi_m', \lambda_m')$ for $m = 1, \ldots, K$ with $\varphi_m = (\varphi_{m,1}, \ldots, \varphi_{m,q})'$ and $\lambda_m = (\lambda_{m,1}, \ldots, \lambda_{m,p})'$.

$$\begin{aligned}
\frac{\partial h_t}{\partial \delta_m} &= t^{*m} + \beta^*(L)\frac{\partial h_t}{\partial \delta_m} + \sum_{k=1}^{K}\left(\lambda_k(L)\frac{\partial h_t}{\partial \delta_m}t^{*k}\right) + \frac{\partial R_{1,t}^*}{\partial \delta_m} \\
\frac{\partial h_t}{\partial \varphi_m} &= (e_{t-1}, \ldots, e_{t-q})' t^{*m} + \beta^*(L)\frac{\partial h_t}{\partial \varphi_m} + \sum_{k=1}^{K}\left(\lambda_k(L)\frac{\partial h_t}{\partial \varphi_m}t^{*k}\right) + \frac{\partial R_{1,t}^*}{\partial \varphi_m} \\
\frac{\partial h_t}{\partial \lambda_m} &= ((h_{t-1} - \varepsilon_{t-1}^2), \ldots, (h_{t-p} - \varepsilon_{t-p}^2))' t^{*m} + \beta^*(L)\frac{\partial h_t}{\partial \lambda_m} \\
&\quad + \sum_{k=1}^{K}\left(\lambda_k(L)\frac{\partial h_t}{\partial \lambda_m}t^{*k}\right) + \frac{\partial R_{1,t}^*}{\partial \lambda_m}
\end{aligned}$$

We obtain

$$\frac{\partial h_t}{\partial \theta_{2,m}} = w_t t^{*m} + \beta^*(L)\frac{\partial h_t}{\partial \theta_{2,m}} + \sum_{k=1}^{K}\left(\lambda_k(L)\frac{\partial h_t}{\partial \theta_{2,m}}t^{*k}\right) + \frac{\partial R_{1,t}^*}{\partial \theta_{2,m}}$$

Under H_0', $\lambda_k = 0$ and $R_{1,t}^* = 0$. Then, the partial derivative of h_t with respect to θ_2 is equal to

$$\left.\frac{\partial h_t}{\partial \theta_2}\right|_{H_0} = \left((w_t t^*)', (w_t t^{*2})', \ldots, (w_t t^{*K})'\right)' + \sum_{j=1}^{p} \beta_j^* \left.\frac{\partial h_{t-j}}{\partial \theta_2}\right|_{H_0} \tag{A.7}$$

B Analytical derivatives for FITVGARCH model

Partial derivative of the conditional variance h_t in (18) with respect to the parameters vector $\theta = (\omega_1, \phi_1', \beta_1', \omega_2, \phi_2', \beta_2', c, \gamma, d)'$ are given as follows

$$\begin{aligned}
\frac{\partial h_t}{\partial \omega_1} &= 1 + [\beta_1(L) + \beta_2(L)F(t^*; \gamma, c)]\frac{\partial h_t}{\partial \omega_1} \\
\frac{\partial h_t}{\partial \phi_1} &= ((1-L)^d \varepsilon_{t-1}^2, \ldots, (1-L)^d \varepsilon_{t-q}^2)' + [\beta_1(L) + \beta_2(L)F(t^*; \gamma, c)]\frac{\partial h_t}{\partial \phi_1} \\
\frac{\partial h_t}{\partial \beta_1} &= (h_{t-1} - \varepsilon_{t-1}^2, \ldots, h_{t-1} - \varepsilon_{t-q}^2)' + [\beta_1(L) + \beta_2(L)F(t^*; \gamma, c)]\frac{\partial h_t}{\partial \beta_1}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial h_t}{\partial \omega_2} &= F(t^*;\gamma,c) + [\beta_1(L) + \beta_2(L)F(t^*;\gamma,c)]\frac{\partial h_t}{\partial \omega_2} \\
\frac{\partial h_t}{\partial \phi_2} &= ((1-L)^d \varepsilon_{t-1}^2, \ldots, (1-L)^d \varepsilon_{t-q}^2)' F(t^*;\gamma,c) \\
&\quad + [\,\beta_1(L) + \beta_2(L)F(t^*;\gamma,c)]\frac{\partial h_t}{\partial \phi_2} \\
\frac{\partial h_t}{\partial \beta_2} &= (h_{t-1} - \varepsilon_{t-1}^2, \ldots, h_{t-1} - \varepsilon_{t-q}^2)' F(t^*;\gamma,c) \\
&\quad + [\beta_1(L) + \beta_2(L)F(t^*;\gamma,c)]\frac{\partial h_t}{\partial \beta_2} \\
\frac{\partial h_t}{\partial c} &= \left[\omega_2 + \phi_2(L)(1-L)^d \varepsilon_t^2 + \beta_2(L)(h_t - \varepsilon_t^2)\right] F_c'(t^*;\gamma,c) \\
&\quad + \left[\beta_1(L) + \beta_2(L)F(t^*;\gamma,c)\right]\frac{\partial h_t}{\partial c}
\end{aligned}$$

where

$$\begin{aligned}
F_c'(t^*;\gamma,c) &= \frac{\partial F(t^*;\gamma,c)}{\partial c} \\
&= -\gamma \exp(-\gamma(t^* - c))[1 + \exp -\gamma(t^* - c)]^{-2} \\
&= -\gamma \exp(-\gamma(t^* - c))F(t^*;\gamma,c)^2 \\
\frac{\partial h_t}{\partial \gamma} &= \left[\omega_2 + \phi_2(L)(1-L)^d \varepsilon_t^2 + \beta_2(L)(h_t - \varepsilon_t^2)\right] F_\gamma'(t^*;\gamma,c) \\
&\quad + \left[\beta_1(L) + \beta_2(L)F(t^*;\gamma,c)\right]\frac{\partial h_t}{\partial \gamma}
\end{aligned}$$

where

$$\begin{aligned}
F_\gamma'(t^*;\gamma,c) &= \frac{\partial F(t^*;\gamma,c)}{\partial \gamma} \\
&= (t^* - c)\exp(-\gamma(t^* - c))[1 + \exp -\gamma(t^* - c)]^{-2} \\
&= (t^* - c)\exp(-\gamma(t^* - c))F(t^*;\gamma,c)^2 \\
\frac{\partial h_t}{\partial d} &= [\phi_1(L) + \phi_2(L)F(t^*;\gamma,c) - 1]\frac{\partial (1-L)^d}{\partial d}\varepsilon_t^2 \\
&\quad + [\beta_1(L) + \beta_2(L)F(t^*;\gamma,c)]\frac{\partial h_t}{\partial d} \\
&= [1 - \phi_1(L) - \phi_2(L)F(t^*;\gamma,c)]\left(\sum_{j=1}^{t-1}\frac{(1-L)^d \varepsilon_t^2}{j}\right) \\
&\quad + [\beta_1(L) + \beta_2(L)F(t^*;\gamma,c)]\frac{\partial h_t}{\partial d}.
\end{aligned}$$

References

Ajmi AN, Ben Nasr A, Boutahar M (2008) Seasonal nonlinear long memory model for the US inflation rates. Comput Econ 31:243–254

Amado C, Teräsvirta T (2008) Modelling conditional and unconditional heteroskedasticity with smoothly time-varying structure. SSE/EFI Working Paper Series in Economics and Finance No. 691

Andersen TG, Bollerslev T (1997) Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. J Finance 52:975–1005

Andreou E, Ghysels E (2002) Detecting multiple breaks in financial market volatility dynamics. J Appl Econom 17:579–600

Anderson HM, Nam K, Vahid F (1999) Asymmetric nonlinear smooth transition GARCH models. In: Rothman P (ed) Nonlinear time series analysis of economic and financial data. Kluwer, Boston pp 191–207

Baillie RT, Morana C (2009) Modeling long memory and structural breaks in conditional variances: an adaptive FIGARCH approach. J Econ Dyn Control 33:1577–1592

Baillie RT, Bollerslev T, Mikkelsen H (1996) Fractionally integrated generalized autoregressive conditional heteroskedasticity. J Econom 74:3–30

Baillie RT, Han YW, Kwon T (2002) Further long memory properties of inflationary shocks. South Econ J 68:496–510

Beine M, Laurent S (2001) Structural changes and long memory in volatility: new evidence from daily exchange rates. In: Dunis C, Timmerman A, Moody J (eds) Developments in forecast combination and portfolio choice, Wiley series in quantitative analysis, chap. 6. Wiley, pp 145–157

Berkes I, Horvath L, Kokoszka P (2003) GARCH processes: structure and estimation. Bernoulli 9:201–228

Berndt E, Hall B, Hall R, Hausman J (1974) Modelling the persistence of conditional variances. Ann Econ Soc Meas 3:653–665

Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econom 31:307–327

Bollerslev T, Engle RF (1993) Common persistence in conditional variances. Econometrica 61:167–186

Bollerslev T, Mikkelsen HO (1996) Modelling and pricing long memory in stock market volatility. J Econom 73:151–184

Bollerslev T, Wooldridge JM (1992) Quasi maximum likelihood estimation and inference in dynamic models with time varying covariances. Econom Rev 11:143–172

Bos CS, Franses PH, Ooms M (1999) Long memory and level shifts: re-analyzing inflation rates. Empir Econ 24:427–449

Breidt FJ, Hsu NJ (2002) A class of nearly long-memory time series models. Int J Forecast 18:265–281

Breidt FJ, Crato N, de Lima P (1998) Modeling long memory stochastic volatility. J Econom 83:325–348

Brown RL, Durbin J, Evans JM (1975) Techniques for testing the constancy of regression relationships over time with comments. J R Stat Soc B37:149–192

Chung CF (1999) Estimating the fractionally integrated GARCH model. National Taiwan University, Taipei

Cai J (1994) A Markov model of regime-switching ARCH. J Bus Econ Stat 12:309–316

Conrad C, Haag BR (2006) Inequality constraints in the fractionally integrated GARCH model. J Financ Econom 4:413–449

Davidson J (2004) Moment and memory properties of linear conditional heteroscedasticity models, and a new model. J Bus Econ Stat 22:16–29

Diebold FX (1986) Comment on "Modeling the persistence of conditional variance" by Engle R, Bollerslev T. Econom Rev 5:51–56

Ding Z, Granger CWJ, Engle RF (1993) A long memory property of stock market returns and a new model. J Empir Finance 1:83–106

Dueker MJ (1997) Markov switching in GARCH processes and mean-reverting stock market volatility. J Bus Econ Stat 12:309–316

Efron B, Hinkley D (1978) Assessing the accuracy of the maximum likelihood estimator: observed versus expected Fisher information. Biometrika 65:457–482

Eitrheim Ø, Teräsvirta T (1996) Testing the adequacy of smooth transition autoregressive models. J Econom 74:59–75

Engle RF (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. Econometrica 50:987–1008

Engle RF, Bollerslev T (1986) Modeling the persistence of conditional variances. Econom Rev 5:1–50

Engle RF, Rangel JG (2008) The spline-GARCH model for low frequency volatility and its global macroeconomic causes. Rev Financ Stud 21:1187–1222

Gonzalez-Rivera G (1998) Smooth transition GARCH models. Stud Nonlinear Dyn Econom 3:161–178

Granger CWJ (1981) Some properties of time series data and their use in econometric model specification. J Econom 16:121–130

Granger CWJ, Hyung N (2004) Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. J Empir Finance 11:399–421

Granger CWJ, Joyeux R (1980) An introduction to long-memory time series and fractional differencing. J Time Ser Anal 1:15–39

Hagerud G (1997) A new non-linear GARCH model. EFI Economic Research Institute, Stockholm

Hamilton JD, Susmel R (1994) Autoregressive conditional heteroscedasticity and changes in regime. J Econom 64:307–333

Jensen ST, Rahbek A (2004) Asymptotic inference for nonstationary GARCH. Econom Theory 20:1203–1226

Karanasos M, Psaradakis Z, Sola M (2004) On the autocorrelation properties of long-memory GARCH processes. J Time Ser Anal 25:265–281

Lamoureux CG, Lastrapes WD (1990) Persistence in variance, structural change and the GARCH model. J Bus Econ Stat 8:225–234

Lee SW, Hansen BE (1994) Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. Econom Theory 10:29–52

Lin CFJ, Teräsvirta T (1994) Testing the constancy of regression parameters against continuous structural change. J Econom 62:211–228

Lobato IN, Savin NE (1998) Real and spurious long memory properties of stock market data. J Bus Econ Stat 16:261–268

Lumsdaine RL (1996) Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models. Econometrica 64:575–596

Lundbergh S, Teräsvirta T (2002) Evaluating GARCH models. J Econ 110:417–435

Luukkonen R, Saikkonen P, Teräsvirta T (1988) Testing linearity against smooth transition autoregressive models. Biometrika 75:491–499

Martens M, van Dijk D, de Pooter M (2004) Modeling and forecasting S&P 500 volatility: long memory, structural breaks and nonlinearity. Tinbergen Institute Discussion Paper 04-067/4

Mikosch T, Stărică C (2004) Changes of structure in financial time series and the GARCH model. Revstat Stat J 2(1):41–73

Morana C, Beltratti A (2004) Structural change and long-range dependence in volatility of exchange rates: either, neither or both? J Empir Finance 11:629–658

Nelson DB (1991) Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59: 347–370

Nelson DB, Cao CQ (1992) Inequality constraints in the univariate GARCH model. J Bus Econ Stat 10:229–235

Perron P, Qu Z (2007) An analytical evaluation of the log-periodogram estimate in the presence of level shifts. Unpublished Manuscript, Department of Economics, Boston University

Perron P, Qu Z (2009) Long-memory and level shifts in the volatility of stock market return indices. J Bus Econ Stat (forthcoming)

Robinson PM (1991) Testingfor strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. J Econom 47:67–84

So MKP, Lam K, Li NK (1998) A stochastic volatility model with Markov switching. J Bus Econ Stat 16(2):244–253

Stărică C, Granger C (2005) Nonstationarities in stock returns. Rev Econ Stat 87:503–522

Teyssière G (1997) Double long-memory financial time series. Paper presented at the ESEM, Toulouse

Tsai H, Chan KS (2008) A note on inequality constraints in the GARCH model. Econom Theory 24:823–828

Tse YK (1998) The conditional heteroskedasticity of the Yen-Dollar exchange rate. J Appl Econom 13:49–55

van Dijk D, Franses PH, Paap R (2002) A nonlinear long memory model for US unemployment. J Econom 102:135–165

White H (1982) Maximum likelihood estimation of misspecified models. Econometrica 50:1–25