Journal of Economic Dynamics & Control 33 (2009) 1577–1592

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Modelling long memory and structural breaks in conditional variances: An adaptive *FIGARCH* approach

Richard T. Baillie [a,b,c], Claudio Morana [d,e,f,*]

[a] Department of Economics, Michigan State University, East Lansing, MI 48824, USA
[b] Department of Finance, Broad School of Business, Michigan State University, East Lansing, MI 48824, USA
[c] Department of Economics, Queen Mary University of London, London E1 4NS, UK
[d] Dipartimento di Scienze Economiche e Metodi Quantitativi (SEMeQ), Università del Piemonte Orientale, Novara, Italy
[e] International Centre for Economic Research (ICER), Torino, Italy
[f] Centre for Research on Pensions and Welfare Policies (CeRP), Torino, Italy

ARTICLE INFO

Article history:
Received 8 July 2007
Accepted 25 February 2009
Available online 14 March 2009

JEL classification:
C15
C22
G1

Keywords:
FIGARCH
Long memory
Structural change
Stock market volatility

ABSTRACT

This paper introduces a new long memory volatility process, denoted by adaptive *FIGARCH*, or *A-FIGARCH* , which is designed to account for both long memory and structural change in the conditional variance process. Structural change is modeled by allowing the intercept to follow the smooth flexible functional form due to Gallant (1984. The Fourier flexible form. American Journal of Agricultural Economics 66, 204–208). A Monte Carlo study finds that the *A-FIGARCH* model outperforms the standard *FIGARCH* model when structural change is present, and performs at least as well in the absence of structural instability. An empirical application to stock market volatility is also included to illustrate the usefulness of the technique.

1. Introduction

The purpose of this paper is to introduce a new long memory volatility process, denoted by Adaptive *FIGARCH*, or *A-FIGARCH*, which is designed to account for both long memory and structural change in the volatility processes of economic and financial time series. It is well known that most daily and high frequency financial time series exhibit quite persistent autocorrelation in their squared returns, power transformations of absolute returns, conditional variances and other measures of volatility. The seminal papers by Ding et al. (1993) and Dacorogna et al. (1993) led to the development of the long memory stochastic volatility models of Breidt et al. (1998) and Harvey (1998), and the long memory ARCH models of Baillie et al. (1996a), Bollerslev and Mikkelsen (1996) and Davidson (2004). While these models appear useful in describing many empirical volatility processes, there is understandably great interest in discerning the reasons and underlying causes for the widespread empirical finding of long memory in volatility. In particular, Granger and Ding (1996) have shown that contemporaneous aggregation of stable *GARCH*(1, 1) processes can result in an aggregate process that exhibits hyperbolically decaying autocorrelations. While this property appears to be consistent with long memory, Zaffaroni (2007) has shown that the autocorrelation function is summable, which is inconsistent with it being classified as a long

* Corresponding author at: Dipartimento di Scienze Economiche e Metodi Quantitativi (SEMeQ), Università del Piemonte Orientale, Novara, Italy.
E-mail address: claudio.morana@eco.unipmn.it (C. Morana).

0165-1889/$ - see front matter
doi:10.1016/j.jedc.2009.02.009

memory process. A related argument of Andersen and Bollerslev (1997) shows how the contemporaneous aggregation of weakly dependent information flow processes can produce the property of long memory in volatility. A further justification is provided by Muller et al. (1997), who suggest that long memory in volatility can arise from the reaction of short-term dealers to the dynamics of a proxy for the expected volatility trend (coarse volatility), which causes persistence in the higher frequency volatility or (fine volatility) process.

While the above papers were concerned with the underlying causes of long memory volatility, other studies have essentially been more skeptical about the validity of the finding of the long memory property in volatility. In particular, it has been suggested that various types of structural change can explain extreme persistence of volatility, and can also generate a series that *appears* to have long memory. In particular, Mikosch and Starica (1998) and Granger and Hyung (2004) have presented theoretical and simulation evidence that spurious long memory can be detected from a time series with breaks. Moreover, while Granger and Hyung (2004) have found that an occasional breaks model provides an inferior forecasting performance than a long memory model for S&P500 absolute returns, for the same series Starica and Granger (2005) have found that a non-stationary model, allowing for breaks in the unconditional variance, can outperform a long memory model in forecasting, but not at short horizons.[1] Furthermore, Diebold and Inoue (2001) have shown how Markov switching processes could generate long memory in the conditional mean, while Granger and Terasvirta (1999) have shown that a process that switches in sign has the characteristics of long memory.

The possible occurrence of structural breaks in conditional variance processes, generating extreme persistence of the *IGARCH* form, appears to have been originally suggested by Lamoreaux and Lastrapes (1990) and Diebold (1986). Theoretical explanations for the above findings have been provided in Morana (2002) and Hillebrand (2005), and studies by Lobato and Savin (1998), Beran and Ocker (1999), Beine and Laurent (2000), Morana and Beltratti (2004) and Martens et al. (2004) have suggested that an appropriate model for the volatility of financial returns should include the joint occurrence of long memory and structural change. These studies are generally consistent with the previous literature such as Hamilton and Susmel (1994),[2] which considered alternating regimes of high and low volatilities, each one being characterized by strong persistence in their fluctuations. Economic explanations of the phenomenon have been suggested by Schwert (1989), who relates alternating volatility regimes to fluctuations in fundamental uncertainty and leverage effects over the business cycle. Also, Beltratti and Morana (2006) relate breaks in stock market volatility to macroeconomic volatility, possibly determined by monetary policy reactions in response to business cycle conditions; while Engle and Rangel (2008) emphasize the role of market size in addition to macroeconomic uncertainty.

Given the above summary of the previous research, this present paper starts from the proposition that both long memory and structural breaks are likely to be present in the volatility processes of many economic and financial time series. Its main contribution is then to present a model which allows for *both long memory and structural change* in a volatility process. The proposed model is named Adaptive *FIGARCH*, or *A-FIGARCH*, and augments the standard *FIGARCH* model of Baillie et al. (1996a) with a deterministic component, following Gallant (1984)'s flexible functional form. Hence, the *A-FIGARCH* model allows for a stochastic long memory component and a deterministic break process component. The approach does not require pre-testing for the number of break points; nor does it require any smooth transition between volatility regimes; and has the advantage of being computationally straightforward.[3]

The rest of this paper is organized as follows. Section 2 introduces the *A-FIGARCH* model and its theoretical properties. Section 3 presents some Monte Carlo evidence for inference in the model and Section 4 presents an empirical application based on equity market returns. The paper ends with a short concluding section.

2. The Adaptive *FIGARCH* process

The Adaptive *FIGARCH*, or *A-FIGARCH* process is formed from two basic components of a long memory volatility process and a deterministic time-varying intercept which allows for breaks, cycles and changes in drift. By definition $\{y_t\}$ is a discrete time, real-valued stochastic process that is serially uncorrelated in its conditional mean, and has long memory type in its conditional variance process. Hence,

$$y_t \equiv \sigma_t z_t, \tag{1}$$

where $E_{t-1}[z_t] = 0$ and $Var_{t-1}[z_t] = 1$; σ_t is a positive, time-varying measurable function with respect to the information set available at time $t-1$, which is denoted as Ω_{t-1}. Hence, σ_t^2 is the time dependent conditional variance defined as $\sigma_t^2 = Var_{t-1}(y_t^2) = Var(y_t^2|\Omega_{t-1})$ and, following Baillie et al. (1996a), is expressed as the long memory *FIGARCH*(p,d,q) process

$$[1-\beta(L)]\sigma_t^2 = w + [1-\beta(L)-\phi(L)(1-L)^d]y_t^2. \tag{2}$$

The process can be most easily motivated from representing $\{y_t^2\}$ as the *ARFIMA*(m,d,q) model

$$\phi(L)(1-L)^d y_t^2 = w + (1-\beta(L))\nu_t, \tag{3}$$

[1] The finding that accounting for structural change may not be relevant for short-term forecasting is a robust finding in the literature. See for instance the discussion in Diebold and Inoue (2001) and the empirical results in Morana and Beltratti (2004).

[2] See also Haas et al. (2004) for a recent contribution in the *GARCH* framework.

[3] Indeed the proposed model is easily estimable with available menu-driven packages as for instance the G@RCH Ox interface.

where $v_t \equiv y_t^2 - \sigma_t^2$ is the innovation in the conditional variance. The fractional differencing parameter is denoted as d, and is constrained to lie in the interval $0<d<1$. The lag polynomials are defined as $\alpha(L) \equiv \alpha_1 L + \cdots + \alpha_q L^q$ and $\beta(L) \equiv \beta_1 L + \cdots + \beta_p L^p$, with $\phi(L) = (1 - \alpha(L) - \beta(L))$. The polynomials $\phi(L)$ and $(1 - \beta(L))$ are assumed to have all their roots lying outside the unit circle, and $m = \max(p, q)$. After rearrangement, an alternative representation for the *FIGARCH*(p, d, q) model is

$$\sigma_t^2 = w[1 - \beta(1)]^{-1} + [1 - \phi(L)(1 - L)^d[1 - \beta(L)]^{-1}]y_t^2 \tag{4}$$

or

$$\sigma_t^2 = w[1 - \beta(1)]^{-1} + \lambda(L)y_t^2, \tag{5}$$

where $\lambda(L) \equiv \lambda_1 L + \lambda_2 L^2 + \cdots$ and $\lambda(1) = 1$ for every d, with $\lambda_i \geqslant 0$, for $i = 1, 2, \ldots$ and $w>0$, for the conditional variance to be well defined, so that it is positive almost surely for all t. A key feature of the *FIGARCH* model is that for high lags, k, the distributed lag coefficients are $\lambda_k \simeq ck^{-d-1}$, where c is a positive constant. Hence, the conditional variance can be expressed as a distributed lag of past squared returns with coefficients that decay at a slow, hyperbolic rate, which is consistent with the long memory property. Davidson (2004) has proposed an alternative definition for the persistence properties of the *FIGARCH* process in terms of hyperbolic memory, which makes more precise the distinction of the *FIGARCH* model from the shorter (geometric) memory cases represented by the *GARCH* and *IGARCH* processes.

Recently, Conrad and Haag (2006) have provided two sets of sufficient conditions for the conditional variance of the *FIGARCH* process to be non-negative almost surely. While the first set immediately implies the above condition, the second set is less restrictive, and in practice requires checking the non-negativity of only a finite number of the impulse response weights λ_is.

It is well known that for $0<d\leqslant 1$ the *FIGARCH*(p, d, q) process has an undefined unconditional variance. However, the process does possess cumulative impulse response weights with a finite sum. This property makes the *FIGARCH* model different from other possible forms of long memory *ARCH* models, such as the class suggested by Karanasos et al. (2004).

As argued in the introduction, there is abundant justification from the literature on financial markets to suspect possible structural instability in the volatility process. A straightforward, but quite powerful approach is to allow the intercept to be time dependent. Hence, the $A-$*FIGARCH*(p, d, q, k) process can be derived from the *FIGARCH*(p, d, q) process by allowing the intercept w in the conditional variance equation to be time-varying according to Andersen and Bollerslev's (1997) flexible functional form. Hence, the model becomes

$$[1 - \beta(L)](\sigma_t^2 - w_t) = [1 - \beta(L) - \phi(L)(1 - L)^d]y_t^2, \tag{6}$$

where

$$w_t = w_0 + \sum_{j=1}^{k}[\gamma_j \sin(2\pi jt/T) + \delta_j \cos(2\pi jt/T)]. \tag{7}$$

The above model reduces to the *FIGARCH* model by setting $w_t = w[1 - \beta(1)]^{-1}$. Although the deterministic process modelled by the flexible functional form is smooth, it has been shown to be able to accurately approximate quite abrupt regime changes, such as discontinuous shifts. Adequate approximations can be achieved with very parsimonious specifications of only $k = 1$ or 2; see Enders and Lee (2004) and the simulation results later in this paper. Hence, the flexible functional form approach allows for a very efficient modelling of structural change, without requiring pretesting to determine the actual location of break points. Furthermore, estimation is relatively straightforward, and the joint presence of long memory and structural change can be assessed by standard hypothesis testing of the fractional differencing parameter and the deterministic trigonometric components.

Analogous to the standard *FIGARCH* model, rearrangement of Eq. (6) produces the alternative representation of the $A-$*FIGARCH*(p, d, q, k) model as

$$\sigma_t^2 = w_t + [1 - \phi(L)(1 - L)^d[1 - \beta(L)]^{-1}]y_t^2 \tag{8}$$

or

$$\sigma_t^2 = w_t + \lambda(L)y_t^2. \tag{9}$$

In order for the conditional variance to be positive almost surely at each point in time, restrictions similar to those holding for the *FIGARCH*(p, d, q) process have to be imposed. In particular, $w_t>0$, for all t, and $\lambda_j \geqslant 0$, for all j. The inclusion of the time-varying intercept component implies that the *A-FIGARCH* process is neither ergodic nor strictly stationary.

One of the great advantages of the above proposed *A-FIGARCH* method concerns the relative simplicity of computation. In fact, the computational burden is only marginally greater than estimating the standard *FIGARCH* model. This is in distinct contrast to the flexible coefficient *GARCH* model of Medeiros and Veiga (2004), or the *spline-GARCH* of Engle and Rangel (2008), or the smooth transition model of Terasvirta and Gonzalez (2006).

2.1. The $A-FIGARCH(1,d,1,k)$ process

A simple version of the model, which appears to be particularly useful in practice, is the $A-FIGARCH(1,d,1,k)$ process

$$[1-\beta L](\sigma_t^2 - w_t) = [1-\beta L - \phi L(1-L)^d]y_t^2 \qquad (10)$$

with w_t as defined in (7). On rearranging, an alternative representation for the $A-FIGARCH(1,d,1,k)$ model is then

$$\sigma_t^2 = w_t + [1-(1-\beta L)^{-1}(1-L)^d(1-\phi L)]y_t^2$$
$$= w_t + \lambda(L)y_t^2 \qquad (11)$$

with $\lambda_0 = 0$, $\lambda_1 = d+\phi-\beta$, and following Conrad and Haag (2006), $\lambda_i = \beta\lambda_{i-1} + (f_i - \phi)(-g_{i-1})$ $i>1$, where $f_j = (j-1-d)/j$, for $j = 1,2,\ldots$ and $g_j = f_j \cdot g_{j-1}$, with $g_0 = 1$. As noted by Baillie et al. (1996a), a sufficient condition for the non-negativity of the conditional variance for the $FIGARCH(1,d,1)$ model requires $w>0$, $0\leqslant\beta\leqslant\phi+d$ and $0\leqslant d\leqslant 1-2\phi$. Less restrictive sufficient conditions have been provided by Bollerslev and Mikkelsen (1996), i.e. $\beta-d\leqslant\phi\leqslant(2-d)/3$ and $d[\phi-(1-d)/2]\leqslant\beta(\phi-\beta+d)$, and Chung (1999), i.e. $0\leqslant\phi\leqslant\beta\leqslant d\leqslant 1$. Finally, Conrad and Haag (2006) have recently proposed alternative and even less restrictive necessary and sufficient conditions, and they show for the case of $0<\beta<1$, either $\lambda_1\geqslant 0$ and $\phi\leqslant f_2$ or $\lambda_{j-1}\geqslant 0$ and $f_{j-1}<\phi\leqslant f_j$ with $j>2$; while for the case of $-1<\beta<0$, either $\lambda_1\geqslant 0$, $\lambda_2\geqslant 0$ and $\phi\leqslant f_2(\beta+f_3)/(\beta+f_2)$ or $\lambda_{j-1}\geqslant 0$, $\lambda_{j-2}\geqslant 0$ and $f_{j-2}(\beta+f_{j-1})/(\beta+f_{j-2})<\phi\leqslant f_{j-1}(\beta+f_j)/(\beta+f_{j-1})$ with $j>3$. Similar restrictions hold for the *A-FIGARCH* model.

2.2. Estimation

Estimation and inference for the parameters of the *A-FIGARCH* process can be facilitated by the familiar method of quasi maximum likelihood estimation (*QMLE*), where the Gaussian log likelihood

$$\ln\{L(\theta, y_1,\ldots,y_T)\} = -0.5T\ln(2\pi) - 0.5\sum_{t=1}^{T}\{\ln(\sigma_t^2) + y_t^2\sigma_t^{-2}\}$$

is numerically maximized with respect to the vector of the parameters $\boldsymbol{\theta} = (d, \boldsymbol{\beta}', \boldsymbol{\phi}', \mathbf{w}', \gamma', \boldsymbol{\delta}')'$. Hence, the procedure simultaneously estimates all the parameters in the model, including those in the flexible functional form of the intercept in the conditional variance process. Under fairly general conditions, the asymptotic distribution of the *QMLE* is

$$T^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \to N\{\mathbf{0}, \mathbf{A}(\boldsymbol{\theta}_0)^{-1}\mathbf{B}(\boldsymbol{\theta}_0)\mathbf{A}(\boldsymbol{\theta}_0)^{-1}\},$$

where $\boldsymbol{\theta}_0$ denotes the true value of the vector of parameters, and where $\mathbf{A}(\boldsymbol{\theta}_0)$ is the Hessian and $\mathbf{B}(\boldsymbol{\theta}_0)$ is the outer product gradient, both of which are evaluated at the true parameter values. Some results for the asymptotic properties of *QMLE* can be established on the basis of available results from the estimation of *GARCH* processes. Jensen and Rahbek (2004) have recently demonstrated that *QMLE* has the properties of consistency and asymptotic normality when applied to the $GARCH(1,1)$ process also when the properties of strict stationarity and ergodicity do not hold. It is worth noting that the conditions required by Jensen and Rahbek (2004) are less stringent than those imposed by Lee and Hansen (1994) and Lumsdaine (1996), where the consistency and asymptotic normality of the *QMLE* was initially shown for the strictly stationary and ergodic case. In particular, Jensen and Rahbek (2004) assume that $z_t \sim i.i.d.(0,1)$, with $Var(z_t^2) = k<\infty$, and that the true parameters satisfy the condition $E\ln(\alpha_0 z_t^2 + \beta_0)\geqslant 0$, where α_0 and β_0 denote the true values of the parameters α and β, i.e. the squared innovation and lagged conditional variance parameters, respectively, in the $GARCH(1,1)$ model. Hence, the requirements do not depend on further higher moment conditions and cover the integrated and explosive cases.[4] Moreover, Jensen and Rahbek (2004) have shown that the asymptotic properties of the estimator still hold for any initial values σ_0^2 and y_0^2, and any value of w. This allows conditioning on the sample mean value of y_t^2, which is $(1/T)\sum_{t=1}^{T} y_t^2$, for σ_0^2 and y_0^2, as is usually implemented in the estimation of *GARCH* models.

It is important to note that the *GARCH*, *IGARCH* and *FIGARCH* models belong to the same family of $ARCH(\infty)$ models. The *FIGARCH* model is essentially a generalization of the *IGARCH* model, since the differencing parameter is allowed to be fractional rather than unity. Albeit the *FIGARCH* model shows relatively more memory than the *IGARCH* model, i.e. hyperbolic memory rather than geometric memory, both models are characterized by summable sequences of lag coefficients for the squared process, converging to unity, independently of the value of the fractional differencing parameter; see Davidson (2004). Recent results of Robinson and Zaffaroni (2006) have established (strong) consistency and asymptotic normality for the *QMLE* estimator of the parameters of the $ARCH(\infty)$ class of models under some general conditions, which also covers the *FIGARCH* model,[5] for which strict stationarity and ergodicity are not established properties.[6] Furthermore, Dahlhaus and Rao (2006) have recently generalized the class of $ARCH(\infty)$ processes to the non-stationary class of $ARCH(\infty)$ processes with time-varying coefficients, establishing consistency and asymptotic normality

[4] Lee and Hansen (1994) assume that $E\ln(\alpha_0 z_t^2 + \beta_0)<0$, which is a necessary and sufficient condition for the stationarity of the $GARCH(1,1)$ process. This latter condition is in fact implied by the condition that $\alpha_0 + \beta_0 \leqslant 1$.

[5] Note that while strong consistency requires $0<d<1$, asymptotic normality requires $d>0.5$.

[6] See for instance Kazakevicius and Leipus (2002).

for the segmented *QMLE* estimator. While these latter results do not directly apply to the *FIGARCH* model, since the existence of the unconditional variance is required in the Dahlhaus and Rao (2006) proof, they do add supporting evidence for the validity of the *QMLE* estimator in non-standard frameworks.[7] Although a formal proof is beyond the scope of this paper, it is therefore conjectured that the optimal asymptotic properties of the *QMLE* estimator hold for the non-strictly stationary and non-ergodic case, and extend to the $A-FIGARCH(1,d,1)$ model.

The numerical maximization of the log likelihood function is implemented by using the asymptotically equivalent method of minimizing the conditional sum of squares function, which neglects starting values. Many previous studies have presented simulation evidence which shows that neglecting initial conditions has minimal effects on parameter estimation of long memory models in either of the first two conditional moments, given a sample size of at least 100 observations. For example, see the results in Baillie et al. (1996b) for the *ARFIMA* model with stable $GARCH(1,1)$ innovations.

3. Simulation results

This section reports some detailed Monte Carlo evidence on the estimation of *A-FIGARCH* models for different data generating process, and comparisons are made with the estimation of corresponding *FIGARCH* models. All of the experiments specify an uncorrelated process y_t for the mean, with various forms of long memory with and without structural breaks, for the conditional variance process. In particular, the data generating process is a martingale difference sequence with $A-FIGARCH(p,d,q)$ model, and $p,q=(0,1)$. Hence,

$$y_t=\sigma_t\varepsilon_t,$$
$$\varepsilon_t \sim NID(0,1)$$

and

$$\sigma_t^2=w_t+(1-L)^d y_t^2 \quad \text{when } p,q=0$$

or

$$[1-\beta L](\sigma_t^2-w_t)=[1-(1-\beta L)^{-1}(1-L)^d(1-\phi L)]y_t^2 \quad \text{when } p,q=1.$$

Three different designs were examined:

Design 1 has a constant intercept of $w_t=w=0.5$, and corresponds to the standard case without structural breaks in the conditional variance.

Design 2 has a step change in the intercept at the midpoint of the sample, where the intercept is doubled at this point. Hence,

$$w_t=\begin{cases}0.5, & t=1,\ldots,T/2.\\ 1, & t=T/2+1,\ldots,T.\end{cases}$$

Design 3 has two step changes equally spaced throughout the sample where the intercept increases eight fold, one-third of the way through the sample and then decreases four fold at a point two-thirds of the length of the sample. Hence,

$$w_t=\begin{cases}0.5, & t=1,\ldots,T/3.\\ 4, & t=T/3+1,\ldots,2T/3.\\ 1, & t=2T/3+1,\ldots,T.\end{cases}$$

These three designs were each simulated for three different values of the long memory parameter, given by $d=(0.15,0.30,0.45)$, and for three values for the short memory parameters $\beta,\phi=(0,0.15,0.30)$. Clearly, the estimation of the *A-FIGARCH* model should prove superfluous in design 1, while the interest in designs 2 and 3 centers on the performance of *QMLE* when the pure *martingale-FIGARCH* process and the new *martingale-A-FIGARCH* models are estimated in the presence of structural breaks in the intercept of the conditional variance. Hence, for designs 1–3 the estimated models were $FIGARCH(p,d,q)$ model with $p,q=(0,1)$; so that

$$y_t=\sigma_t z_t,$$
$$z_t\sim NID(0,1),$$
$$(\sigma_t^2-w)=[1-(1-\beta L)^{-1}(1-L)^d(1-\phi L)]y_t^2 \tag{12}$$

[7] Note that results concerning consistency and asymptotic normality of the *QMLE* have been obtained for the general strictly stationary and ergodic $GARCH(p,q)$ process; see Berkes et al. (2003). However, results for the non-stationary and non-ergodic case currently only exist for the $GARCH(1,1)$ process, which is fortunately the most widely used model in applied econometric work.

and the $A-FIGARCH(p,d,0,k)$ model, with $p,q=(0,1)$,

$$\begin{aligned} &y=\sigma_t z_t,\\ &z_t \sim NID(0,1),\\ &(\sigma_t^2-w_t)=[1-(1-\beta L)^{-1}(1-L)^d(1-\phi L)]y_t^2,\\ &w_t=w_0+\sum_{j=1}^{k}[\gamma_j \sin(2\pi jt/T)+\delta_j \cos(2\pi jt/T)]. \end{aligned} \tag{13}$$

The *A-FIGARCH* models were estimated for each design with one to four pairs of trigonometric terms included, i.e. $k=(1,2,3,4)$. The simulated processes have sample size equal to 3000 observations. Following Baillie et al. (1996a), the order of the truncation in estimation has been set to 1000 observations. Finally, 500 Monte Carlo replications were employed in all of the designs. In Tables 1–3 the Monte Carlo bias (bias), root mean square error (rmse) and the standard error (se) of the estimator, are reported for the various cases.

The simulation experiments reveal several general points concerning the performance of the different estimators of the long memory parameter d. For case 1, where there is no structural change, the application of the *A-FIGARCH* model should clearly be unnecessary since the intercept is a constant. First, the estimate of the long memory parameter obtained from the *A-FIGARCH* estimation has approximately the same degree of small sample bias as the corresponding estimate resulting from the estimation of the *FIGARCH* model. This result appears consistent across all the designs. However, the most interesting result is the reduction in *rmse* of the estimate of the d parameter from using the *A-FIGARCH* model, compared with estimation of the regular *FIGARCH* model. The reduction in *rmse* appears to noticeably increase as the level of persistence (value of d) increases. These results suggest that there is no additional cost from using the *A-FIGARCH* model as opposed to the *FIGARCH* model, even when there is no structural break in the conditional variance. The interpretation of this is intriguing and suggests that the time dependent intercept is also somehow adjusting for parameter uncertainty in the estimation of d.

For cases 2 and 3, where the intercept is subject to structural breaks, apart from the low persistence case ($d=0.15$), the degree of bias in the estimates of d is very small for both estimators. However, the bias is again always smaller for the *A-FIGARCH* model compared to the pure *FIGARCH* model. Moreover, the *rmse* of the estimate of d is generally lower from the *A-FIGARCH* estimation compared to the corresponding *FIGARCH* estimation. Finally, the generally superior performance of the estimate of d from the estimation of the *A-FIGARCH* model, relative to the standard *FIGARCH* model, is robust across the three different values of d used in the designs, with the improvement increasing as the degree of persistence increases. Hence, the Gallant flexible functional form seems to work quite well in the *A-FIGARCH* model estimation framework, and does a good job in terms of modelling the structural change in the intercept.

Table 1
Simulation results for estimation of $A-FIGARCH(0,d,0,k)$ and $FIGARCH(0,d,0)$ models.

		$A\text{-}F(0,0.15,0,k)$			$A\text{-}F(0,0.30,0,k)$			$A\text{-}F(0,0.45,0,k)$		
		$bias_d$	$rmse_d$	se_d	$bias_d$	$rmse_d$	se_d	$bias_d$	$rmse_d$	se_d
$k=0$	m_1	0.001	0.021	0.021	0.005	0.035	0.035	0.013	0.035	0.035
	m_2	0.102	0.020	0.010	0.025	0.022	0.021	−0.012	0.049	0.049
	m_3	0.107	0.023	0.012	0.044	0.030	0.028	0.036	0.040	0.037
$k=1$	m_1	0.036	0.029	0.028	−0.004	0.024	0.024	−0.017	0.024	0.024
	m_2	0.081	0.017	0.010	0.017	0.022	0.021	−0.012	0.034	0.034
	m_3	0.089	0.020	0.012	0.036	0.026	0.025	0.005	0.035	0.035
$k=2$	m_1	−0.004	0.020	0.020	0.002	0.023	0.023	−0.007	0.056	0.042
	m_2	0.075	0.016	0.010	0.003	0.021	0.021	0.003	0.021	0.021
	m_3	0.091	0.020	0.012	0.034	0.027	0.026	0.001	0.027	0.025
$k=3$	m_1	−0.010	0.022	0.022	−0.010	0.023	0.023	−0.020	0.028	0.028
	m_2	0.081	0.016	0.010	0.005	0.021	0.021	−0.030	0.036	0.036
	m_3	0.090	0.020	0.012	0.027	0.028	0.027	−0.003	0.040	0.040
$k=4$	m_1	−0.012	0.021	0.021	−0.011	0.023	0.023	−0.018	0.029	0.029
	m_2	0.081	0.017	0.010	0.004	0.020	0.020	−0.027	0.039	0.038
	m_3	0.090	0.019	0.011	−0.028	0.029	0.028	−0.004	0.037	0.037

Key: The table reports simulation results for the bias, root mean square error (rmse) and standard error (se) for estimation of the fractional differencing parameter d from a sample size of $T=3000$ observations. All the results are based on 500 replications. The simulations are for three different experiments of: no break (m_1), a single break point (m_2) and two break points (m_3). The estimated models use kth order flexible Fourier forms, with $k=0,1,2,3,4$ for the adaptive component. The $k=0$ case corresponds to standard *FIGARCH* estimation.

Table 2
Simulation results for estimation of $A-FIGARCH(1,d,0,k)$ and $FIGARCH(1,d,0)$ models.

		$bias_d$	$rmse_d$	se_d	$bias_\beta$	$rmse_\beta$	se_β
A-F(0.30, 0.45, 0, *k*)							
$k=0$	m_1	0.032	0.079	0.078	0.029	0.079	0.078
	m_2	0.073	0.071	0.066	0.078	0.077	0.071
	m_3	0.108	0.085	0.073	0.112	0.093	0.080
$k=1$	m_1	−0.006	0.042	0.042	−0.008	0.050	0.050
	m_2	0.044	0.051	0.049	0.049	0.060	0.058
	m_3	0.085	0.061	0.053	0.087	0.070	0.063
$k=2$	m_1	−0.005	0.044	0.044	−0.007	0.051	0.051
	m_2	0.022	0.046	0.049	0.031	0.050	0.045
	m_3	0.066	0.060	0.055	0.071	0.067	0.062
k = 3	m_1	−0.021	0.051	0.050	−0.019	0.056	0.056
	m_2	0.023	0.055	0.055	0.033	0.060	0.059
	m_3	0.055	0.061	0.058	0.064	0.070	0.066
$k=4$	m_1	−0.027	0.050	0.049	−0.023	0.054	0.053
	m_2	0.021	0.056	0.055	0.030	0.060	0.059
	m_3	0.062	0.061	0.058	0.072	0.069	0.064
A-F(0.30, 0.30, 0, *k*)							
$k=0$	m_1	−0.002	0.047	0.047	−0.010	0.049	0.049
	m_2	0.149	0.070	0.048	0.138	0.071	0.052
	m_3	0.178	0.106	0.078	0.190	0.106	0.082
$k=1$	m_1	0.002	0.042	0.042	−0.008	0.044	0.044
	m_2	0.130	0.055	0.038	0.118	0.057	0.043
	m_3	0.168	0.070	0.041	0.155	0.071	0.047
$k=2$	m_1	−0.007	0.039	0.039	−0.015	0.041	0.041
	m_2	0.087	0.037	0.027	0.076	0.034	0.031
	m_3	0.137	0.067	0.048	0.123	0.068	0.053
$k=3$	m_1	−0.021	0.039	0.038	−0.028	0.040	0.039
	m_2	0.083	0.031	0.024	0.069	0.033	0.028
	m_3	0.126	0.056	0.040	0.111	0.057	0.045
$k=4$	m_1	−0.027	0.038	0.038	−0.034	0.039	0.038
	m_2	0.095	0.040	0.030	0.085	0.042	0.035
	m_3	0.147	0.067	0.046	0.134	0.068	0.050
A-F(0.15, 0.15, 0, *k*)							
$k=0$	m_1	−0.004	0.030	0.030	−0.011	0.032	0.032
	m_2	0.220	0.071	0.022	0.205	0.074	0.032
	m_3	0.244	0.098	0.047	0.225	0.097	0.038
$k=1$	m_1	−0.008	0.028	0.028	−0.016	0.029	0.029
	m_2	0.202	0.062	0.022	0.186	0.065	0.030
	m_3	0.216	0.082	0.035	0.191	0.085	0.048
$k=2$	m_1	−0.009	0.032	0.032	−0.016	0.033	0.033
	m_2	0.152	0.039	0.016	0.120	0.042	0.025
	m_3	0.158	0.054	0.029	0.124	0.056	0.040
$k=3$	m_1	−0.020	0.029	0.029	−0.014	0.030	0.030
	m_2	0.151	0.037	0.015	0.126	0.041	0.025
	m_3	0.154	0.041	0.018	0.112	0.040	0.026
$k=4$	m_1	−0.027	0.031	0.030	−0.033	0.033	0.032
	m_2	0.167	0.044	0.016	0.147	0.047	0.026
	m_3	0.196	0.072	0.034	0.174	0.075	0.045

Key: As for Table 1; but corresponding results also included for the estimation of the β parameter.

Table 3
Simulation results for estimation of $A-FIGARCH(1,d,1,k)$ and $FIGARCH(1,d,1)$ models.

		$bias_d$	$rmse_d$	se_d	$bias_\beta$	$rmse_\beta$	se_β	$bias_\phi$	$rmse_\phi$	se_ϕ
A-F$(0.30,0.45,0.15,k)$										
$k=0$	m_1	−0.196	0.107	0.069	0.059	0.141	0.138	−0.012	0.128	0.128
	m_2	0.012	0.074	0.073	0.242	0.169	0.110	0.104	0.103	0.092
	m_3	0.045	0.090	0.090	0.261	0.189	0.121	0.085	0.096	0.088
$k=1$	m_1	−0.123	0.117	0.101	−0.010	0.104	0.104	−0.023	0.074	0.074
	m_2	−0.006	0.035	0.035	0.151	0.107	0.084	0.016	0.065	0.064
	m_3	0.047	0.054	0.052	0.230	0.167	0.114	0.056	0.084	0.081
$k=2$	m_1	−0.138	0.103	0.084	0.039	0.121	0.120	0.033	0.091	0.090
	m_2	−0.056	0.041	0.038	0.123	0.090	0.075	0.028	0.056	0.055
	m_3	−0.006	0.068	0.068	0.161	0.148	0.122	0.031	0.083	0.082
$k=3$	m_1	−0.102	0.071	0.060	0.045	0.071	0.060	0.010	0.096	0.094
	m_2	−0.041	0.034	0.032	0.135	0.084	0.066	0.028	0.050	0.049
	m_3	0.016	0.044	0.043	0.210	0.144	0.100	0.056	0.074	0.071
$k=4$	m_1	−0.086	0.049	0.041	0.049	0.025	0.023	−0.012	0.028	0.027
	m_2	−0.054	0.022	0.019	0.059	0.026	0.022	−0.028	0.020	0.019
	m_3	−0.040	0.019	0.017	0.049	0.025	0.023	−0.028	0.018	0.017
A-F$(0.15,0.45,0.30,k)$										
$k=0$	m_1	−0.313	0.134	0.036	−0.097	0.112	0.103	−0.085	0.110	0.102
	m_2	0.053	0.159	0.154	0.494	0.472	0.214	0.114	0.147	0.130
	m_3	0.122	0.191	0.173	0.533	0.505	0.21	0.108	0.130	0.116
$k=1$	m_1	−0.276	0.191	0.115	−0.063	0.090	0.086	−0.074	0.070	0.065
	m_2	−0.049	0.037	0.034	0.134	0.090	0.072	−0.078	0.045	0.039
	m_3	−0.040	0.037	0.035	0.156	0.137	0.112	−0.047	0.062	0.060
$k=2$	m_1	−0.079	0.098	0.092	0.036	0.063	0.061	−0.055	0.040	0.037
	m_2	−0.061	0.032	0.028	0.086	0.038	0.030	−0.084	0.025	0.018
	m_3	−0.060	0.050	0.047	0.086	0.077	0.070	−0.080	0.047	0.041
$k=3$	m_1	−0.178	0.141	0.109	−0.008	0.086	0.086	−0.053	0.067	0.065
	m_2	−0.112	0.040	0.027	0.090	0.060	0.052	−0.073	0.023	0.017
	m_3	−0.067	0.033	0.029	0.108	0.099	0.087	0.072	0.049	0.043
$k=4$	m_1	−0.168	0.117	0.089	0.021	0.050	0.050	−0.042	0.035	0.034
	m_2	−0.080	0.018	0.014	0.049	0.016	0.014	−0.052	0.015	0.012
	m_3	−0.072	0.018	0.013	0.067	0.022	0.017	−0.069	0.018	0.014

Key: As for Table 2; but corresponding results also included for the estimation of the ϕ parameter.

Interestingly, from Tables 1–3, it can also be noted that neglecting structural breaks does not only lead to an upward biased estimate of the fractional differencing parameter, as already found for the $p=0$ case, but also in the estimates of the β and ϕ parameters in the conditional variance equation. This latter finding is particularly evident when the degree of persistence is low, as in the $d=0.15$ case, or when β is low ($\beta=0.15$). The upward bias in the estimate of d from the regular *FIGARCH* estimation appears to be mitigated by the inclusion of the trigonometric components in the *A-FIGARCH* estimation. The improved performance of the estimation of d tends to increase with the degree of persistence of the series. Hence, estimation of the *A-FIGARCH* shows a superior performance relatively to the *FIGARCH* model in terms of bias and *rmse* in all the designs. Interestingly, the greatest improvement is in the $d=0.45$ case, which is the one mostly relevant for financial applications. In this case there is a 145% reduction in bias and a 60% reduction in *rmse* obtained from using the *A-FIGARCH* model, relatively to the *FIGARCH* model.

Overall, the above results indicate potentially significant gains from using the *A-FIGARCH* specification, and certainly no perceptible losses, even in the absence of structural breaks. The possible loss of efficiency in using an unnecessary, over-parameterized *A-FIGARCH* model specification does not appear to be an issue. It may be that the estimation from smaller sample sizes would find losses in efficiency of the estimation of d. Since a sample size of $T=3000$ is quite common for finance applications, the situation from smaller sample sizes has not been investigated in this study.

Therefore, in the light of the Monte Carlo evidence, it seems preferable to include the adaptive non-linear trend component in the specification for the conditional variance equation at the out set, since no negative consequences for estimation may be expected, apart from the case of weak long memory, which however does not seem to be relevant for financial returns. Then, following a general to specific methodology, the best fitting parsimonious model may be obtained. Moreover, for the cases investigated in the Monte Carlo exercise, there is no evidence of an improvement in the performance of the model by the inclusion of polynomial terms beyond the second or third order. Yet, for actual data more profligate parameterizations may be needed.

4. Applications to stock market volatility

This section of the paper reports estimation of $A-FIGARCH(1,d,1,k)$ and $FIGARCH(1,d,1)$ models for the S&P500 returns, as well as comparison with competing models such as the standard $GARCH(1,1)$ model, the Engle and Rangel *Spline-GARCH* ($S-GARCH(1,1,p)$) model, an adaptive version of the *GARCH* model, i.e. the $A-GARCH(1,1)$ model,[8] and a spline version of the *FIGARCH* model, i.e. the *S-FIGARCH*$(1,d,1,p)$ model.

The time span is from January 3, 1928 to February 15, 2007, which realizes a total of $T=20,863$ daily observations, or 4172 weekly observations,[9] and is a long enough period for the likely occurrence of multiple structural breaks in volatility. For the practical implementation of the *A-FIGARCH* and *A-GARCH* methods, an important consideration is the determination of the order of the trigonometric terms in the Gallant flexible functional form, in addition to the order of the specification of the stationary components in the conditional mean and conditional variance equations. Similarly, for the implementation of the spline models, in addition to the specification of the conditional mean and variance equations, a smoothing parameter, controlling the trade off between minimizing the residual error and minimizing local variation, needs to be selected for the estimation of the long-term volatility component. In the reported results the Schwartz *BIC* information criterion is used for model selection in all cases. For computational convenience, estimation has been carried out using weekly observations.

Since the conditional mean did not exhibit any significant autocorrelation, only an intercept was included in the mean equation. Hence, the following long memory models of the $FIGARCH(1,d,1)$ family

$$\begin{aligned}
&y_t = \mu + \varepsilon_t \\
&\varepsilon_t = \sigma_t z_t \\
&z_t \sim NID(0,1) \\
&[1-\beta L]\sigma_t^2 = w_t + [1-\beta L - \phi L(1-L)^d]\varepsilon_t^2 \\
&w_t = \begin{cases} w & FIGARCH(1,d,1), \\ w_0 + \sum_{j=1}^{k}[\gamma_j \sin(2\pi jt/T) + \delta_j \cos(2\pi jt/T)] & A-FIGARCH(1,d,1,k) \\ f_p(t) & S-FIGARCH(1,d,1,p) \end{cases},
\end{aligned}$$

and short memory models of the $GARCH(1,1)$ family

$$\begin{aligned}
&y_t = \mu + \varepsilon_t \\
&\varepsilon_t = \sigma_t z_t \\
&z_t \sim NID(0,1) \\
&[1-\beta L]\sigma_t^2 = w_t + \phi L \varepsilon_t^2 \\
&w_t = \begin{cases} w & GARCH(1,d,1), \\ w_0 + \sum_{j=1}^{k}[\gamma_j \sin(2\pi jt/T) + \delta_j \cos(2\pi jt/T)] & A-GARCH(1,d,1,k), \\ f_p(t) & S-GARCH(1,d,1,p), \end{cases}
\end{aligned}$$

where $f_p(t)$ is the spline volatility component, were estimated for the S&P500 returns series. All the above models have been estimated by imposing non-negativity constraints, using, as in Engle and Rangel (2008), an exponential specification at the log-likelihood function maximization steps.[10]

[8] Related and independent work by Pascalau (2007) has also dealt with the short memory, exponentially decaying, version of the proposed adaptive model.

[9] Neglecting the first three daily observations of the sample.

[10] The estimated w, γ_j and δ_j parameters reported in Table 4 actually refer to the exponential specification. See Engle and Rangel (2008) for details.

A cubic spline smoother was used to model the long-term volatility component. The objective function for the determination of the smoothing parameter p can then be written as

$$S(p) = p\sum_t (l_t - f(x_t))^2 + (1-p)\int f''(x)^2,$$

where $t = 1,\dots,T$, l_t is the generic process to be smoothed, x_i defines the position of knots, $\int f''(t)$ is the integrated squared second derivative of the cubic spline function $f(x) = a_i + b_i x + c_i x^2 + d_i x^3$. See Silverman (1985) for details.

Estimation results are reported in Table 4. First, the Nyblom (1989) test indicates clear rejection at the 0.05 level of the null hypothesis of no breaks in variance for the estimates of the standard *GARCH*(1, 1) model. A consequence of neglecting structural breaks is that the estimation of the *GARCH*(1, 1) model tends to produce results consistent with the data being generated by an *IGARCH* process. The tests for serial correlation in the mean of the residuals do not suggest any further evidence of mis-specification. Accounting for structural breaks in the spline or adaptive framework leads to non-significant Nyblom (1989) stability statistics. The estimation of the *A-FIGARCH* model also corrects for the upward bias in the persistence parameters that is a feature of the basic *GARCH* model. The above findings are fully consistent with evidence on the presence of structural breaks previously detected for S&P500 returns, as reported by Lobato and Savin (1998), Granger and Hyung (2004), Starica and Granger (2005) and Beltratti and Morana (2006).

Table 4
Estimation of *GARCH* and *FIGARCH* models for S&P500 returns.

	GARCH	*S-GARCH*	*A-GARCH*	*FIGARCH*	*S-FIGARCH*	*A–FIGARCH*
μ	0.201 (0.029)	0.210(0.029)	0.118 (0.006)	0.209 (0.029)	0.211 (0.029)	0.252 (0.000)
w	0.102 (0.037)	–	0.112 (0.036)	0.101 (0.037)	–	0.101 (0.036)
β	0.879 (0.015)	0.789 (0.028)	0.795 (0.027)	0.593 (0.060)	0.401 (0.212)	0.487 (0.180)
ϕ	0.106 (0.012)	0.120 (0.014)	0.120 (0.013)	0.322 (0.057)	0.331 (0.208)	0.307 (0.134)
d	–	–	–	0.446 (0.032)	0.242 (0.033)	0.340 (0.063)
p	–	1e-008	–	–	1e-008	–
γ_1	–	–		–	–	–
δ_1	–	–	0.405 (0.050)	–	–	0.443 (0.050)
γ_2	–	–	0.477 (0.050)	–	–	0.476 (0.050)
δ_2	–	–	0.145 (0.050)	–	–	0.119 (0.050)
γ_3	–	–	0.183 (0.050)	–	–	0.188 (0.050)
δ_3	–	–	0.185 (0.050)	–	–	0.215 (0.050)
γ_4	–	–	–	–	–	–
δ_4	–	–	−0.181 (0.050)	–	–	−0.202 (0.050)
γ_5	–	–	–	–	–	–
δ_5	–	–	−0.154 (0.050)	–	–	−0.147 (0.050)
γ_6	–	–	–	–	–	–
δ_6	–	–	−0.403 (0.050)	–	–	−0.384 (0.050)
γ_7	–	–	0.158 (0.050)	–	–	0.137 (0.050)
δ_7	–	–	−0.196 (0.050)	–	–	−0.177(0.050)
γ_8	–	–	0.249 (0.050)	–	–	0.227 (0.050)
δ_8	–	–	−0.219 (0.050)	–	–	−0.226 (0.050)
LB_{65}	0.170	0.278	0.044	0.171	0.227	0.029
LB^2_{65}	0.958	0.870	0.792	0.988	0.978	0.933
Sk	−0.583	−0.504	−0.495	−0.478	−0.452	−0.492
Ku	2.635	2.018	1.81	1.592	1.441	1.638
S-b	0.670	0.633	0.540	0.431	0.462	0.534
NS-b	0.000	0.005	0.002	0.168	0.181	0.154
PS-b	0.055	0.019	0.024	0.002	0.001	0.002
Nyb	0.547*	0.297	0.417	0.098	0.036	0.083
SBC	−0.0045	0.0004	0.0093	0.0053	0.0170	−0.0050
AIC	−0.0039	0.0008	0.0064	0.0058	0.0173	−0.0084
L	−2.1798	−2.1689	−2.1702	−2.1746	−2.1676	−2.1690

The sample is from January 3, 1928 to February 15, 2007, for a total of $T = 4172$ weekly observations. The asymptotic standard errors are reported in parenthesis beside corresponding parameter estimates. The diagnostic statistics are LB which denotes the Ljung–Box test for serial correlation in the standardized residuals, LB^2 is the Ljung–Box test for serial correlation in the squared standardized residuals, Sk is the index of skewness and Ku is the index of excess kurtosis. The Ljung–Box statistics are computed from the first $T^{\wedge 0.5}$ sample autocorrelations. In addition, S-b denotes the p-value of the sign bias t-test, NS-b the p-value of the negative size bias t-test, PS-b the p-value of the positive size bias t-test, while SBC is the Schwarz–Bayes information criterion and AIC is the Akaike information criterion. Finally, L is the value of the average log-likelihood function and Nyb is the Nyblom stability test for the unconditional variance carried out on the standardized residuals (* denotes rejection of the null of stability at the 5% significance level). The estimated models are the *GARCH*(1,1) model (*GARCH*), the *Spline-GARCH*(1,1) model (*S-GARCH*), the *Adaptive-GARCH*(1,1) model (*A-GARCH*), the *FIGARCH*(1, d, 1) model (*FIGARCH*), the *Spline-FIGARCH*(1, d, 1) model (*S-FIGARCH*) and the *Adaptive-FIGARCH*(1, d, 1, 8) model (*A–FIGARCH*).

Accounting for long memory yields an additional improvement in specification in all the cases considered for the *GARCH* family of models. Firstly, stronger evidence of stability in variance can be detected in all the cases. Interestingly, as found for the *FIGARCH* model, no evidence of instability in variance can be detected once long memory is allowed for. This finding is consistent with the view that long memory and structural breaks are features which can be easily confounded. Yet, once structural breaks and long memory are jointly modelled, in the spline or in the adaptive framework, an improvement in fit can be noted, as well as a reduction in the persistence parameter. As already found for the *GARCH* model, by comparing the estimated fractional differencing and autoregressive parameters, it can be noted that an upward bias in persistence is imparted by neglecting structural breaks in both cases. Differently, the moving average parameter does not seem to be sizably affected. Overall, the evidence is consistent with the Monte Carlo evidence provided in this study. Finally, by comparing the spline and adaptive long memory specifications it can be concluded that both strategies lead to very satisfactory results, albeit the adaptive specification possibly shows some computational advantages over the spline model.

In Table 5 some descriptive statistics for the estimated conditional variance processes from the estimated models are reported for comparison. As is presented in the table, the estimated conditional variance processes are very similar in terms of overall mean variance level (in the range 6.38–6.69) and standard deviations (8.04–9.15). More noticeable differences

Table 5
Descriptive statistics for the conditional variance processes for S&P500 returns.

	GARCH	*S-GARCH*	*A-GARCH*	*FIGARCH*	*S-FIGARCH*	*A–FIGARCH*
Mean	6.670	6.382	6.543	6.690	6.548	6.658
Std. dev.	9.012	8.041	8.446	9.071	8.455	9.145
Min	1.511	0.886	0.698	1.037	0.850	0.736
Max	100.48	97.651	100.40	129.68	123.19	123.51
Correlations						
GARCH	1.000					
S-GARCH	0.978	1.000				
A-GARCH	0.978	0.997	1.000			
FIGARCH	0.981	0.985	0.983	1.000		
S-FIGARCH	0.959	0.986	0.983	0.990	1.000	
A–FIGARCH	0.977	0.992	0.993	0.993	0.996	1.000

The sample is from January 3, 1928 to February 15, 2007, for a total of $T = 4172$ weekly observations. In the table descriptive statistics for the estimated conditional variance processes are reported. The estimated models are the *GARCH*(1,1) model (*GARCH*), the *Spline-GARCH*(1,1) model (*S-GARCH*), the *Adaptive-GARCH*(1,1) model (*A-GARCH*), the *FIGARCH*$(1, d, 1)$ model (*FIGARCH*), the *Spline-FIGARCH*$(1, d, 1)$ model (*S-FIGARCH*) and the *Adaptive-FIGARCH*$(1, d, 1, 8)$ model (*A–FIGARCH*).

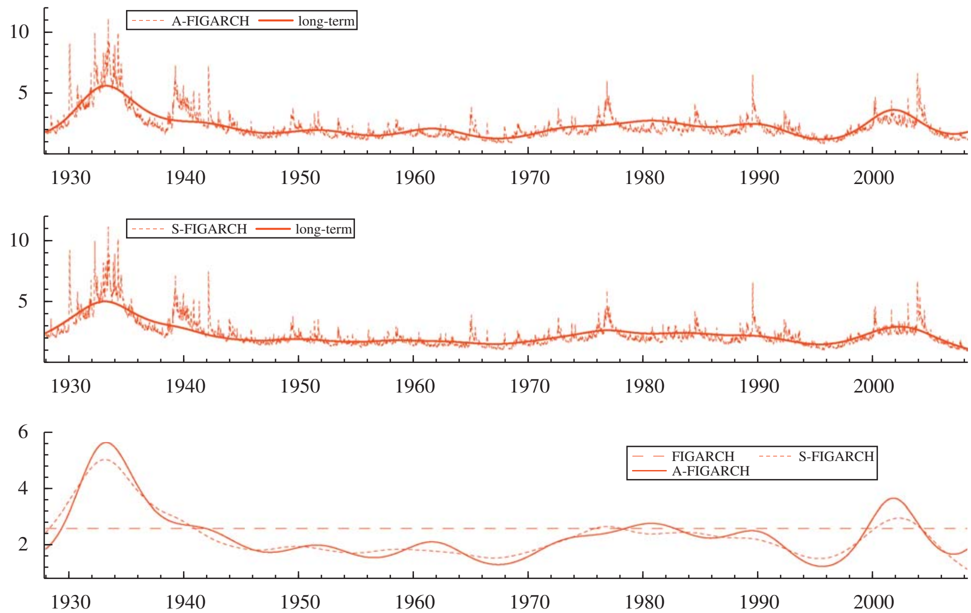

Fig. 1. S&P 500 conditional standard deviation and long-term volatility processes, *A–FIGARCH* and *S-FIGARCH* models.

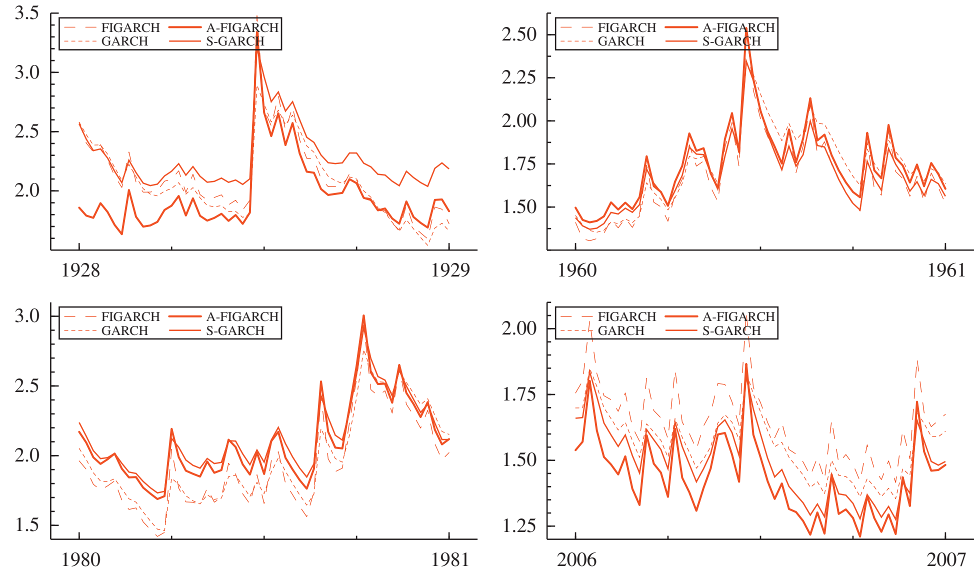

Fig. 2. Estimated conditional standard deviations from the *FIGARCH*, *A–FIGARCH*, *GARCH* and *S-GARCH* models.

concern the minimum (0.70–1.51) to maximum (97.65–129.68) value range. The processes are also strongly correlated (in the range 0.985–0.997), with the *GARCH* model showing in general the lowest correlations with all the other models.

In Fig. 1 the estimated conditional standard deviation by the preferred *A-FIGARCH* and *S-FIGARCH* models are plotted, and the corresponding long-term volatility components are contrasted with the constant long-term level estimated by the standard *FIGARCH* model as well. As is shown in the plot, the long-term variance components estimated by means of the adaptive and spline specifications are very close, taking values in general different from the estimated long-term value from the standard *FIGARCH* specification. The latter would largely underestimate the long-term variance value in the 1930s and the 1940s, and the overestimate it up to the end of the 1970s. Underestimation can then be observed again over the 1990s, while both overestimation and underestimation over the 2000s.

Finally, the consequences of neglecting structural change can be clearly noted in Fig. 2, where the conditional standard deviations from the *FIGARCH*, the *A-FIGARCH*, the *GARCH* and the *S-GARCH* models have been plotted for four years in the sample. As shown in the plots, modelling structural change in conditional variance can make the difference, particularly in terms of overall volatility level, as volatility fluctuations always look strictly synchronized across models. Hence, sizable bias, both upward or downward can be imparted by neglecting structural change. Yet, as shown in the first sub-plot in Fig. 2, results can also be sensitive to the way the time-varying long-term volatility is modelled.

4.1. *Out of sample forecasting analysis*

In order to further assess the consequences of the specification and also mis-specification of the conditional variance equation, out of sample forecasting experiments were conducted. All the models were estimated recursively, with *ex ante* out of sample forecasts generated for 1–12-week horizons. When generating predictions for the *A-FIGARCH* and *A-GARCH* models, forecasts of the deterministic component were obtained by extrapolating the flexible functional form specification. Given the specification of the latter, it then follows:

$$E[w_{t+s|t}] = w_t, \quad s = 1, \ldots, 12.$$

A naive forecast of the intercept has then also been employed for the non-adaptive models, including the spline models. A total of 940 point forecasts, for each forecasting horizon, were computed from the last 20 years of data in the sample. The benchmark volatility process for assessing forecast accuracy was computed from the realized variance estimator of Andersen et al. (2001), using daily data. Forecasts were assessed on the basis of the root mean square forecast error criterion (*RMSFE*), and also with the West and Cho (1995) test to determine the statistical significance of the *MSFE* based model ranking.

Since 12 different forecasting horizons were considered, it is worth using a criteria that assesses the overall *MSFE* based forecasts. Clements and Hendry (1993) have suggested a generalized forecast error second moment (*GFESM*) statistic, given

Table 6
Out of sample forecasting analysis for the *GARCH* and *FIGARCH* models.

	GARCH	*S-GARCH*	*A-GARCH*	*FIGARCH*	*S-FIGARCH*	*A–FIGARCH*
RMSFE						
1	6.851	6.838	6.703	6.441	6.513	6.388
2	11.395	11.377	11.007	10.713	10.903	10.551
3	15.553	15.539	14.865	14.731	15.074	14.443
4	19.548	19.543	18.505	18.374	18.908	17.944
8	34.344	34.669	31.522	32.307	34.003	31.258
12	48.854	49.807	43.783	45.238	48.588	43.491
L-GFESM						
1	3.849	3.845	3.805	3.726	3.748	3.709
2	5.175	5.172	5.113	5.052	5.083	5.025
3	6.037	6.035	5.959	5.922	5.962	5.888
4	6.686	6.684	6.592	6.567	6.615	6.526
8	8.349	8.359	8.207	8.228	8.308	8.171
12	9.373	9.398	9.191	9.238	9.351	9.169
R^2						
1	0.142	0.163	0.155	0.212	0.221	0.224
2	0.187	0.216	0.205	0.242	0.259	0.261
3	0.216	0.251	0.240	0.254	0.278	0.279
4	0.231	0.271	0.259	0.271	0.298	0.300
8	0.267	0.322	0.313	0.288	0.328	0.329
12	0.267	0.337	0.329	0.293	0.344	0.344
θ						
1	1.688 (0.399)	1.510 (0.355)	1.214 (0.424)	0.968 (0.448)	1.079 (0.399)	0.884 (0.430)
2	3.541 (0.801)	3.152 (0.701)	2.515 (0.849)	2.368 (0.795)	2.571 (0.709)	2.168 (0.750)
3	5.322 (1.176)	4.712 (1.038)	3.627 (1.230)	3.949 (1.115)	4.219 (1.034)	3.601 (1.044)
4	7.270 (1.563)	6.406 (0.593)	4.846 (1.547)	5.347 (1.469)	5.763 (1.414)	4.859 (1.360)
8	14.68 (3.660)	13.03 (3.610)	8.518 (3.608)	11.27 (3.614)	12.77 (3.605)	10.31 (3.325)
12	22.80 (6.369)	20.45 (6.230)	12.06 (6.201)	16.81 (6.362)	20.17 (6.238)	15.59 (5.822)
δ						
1	0.633 (0.087)	0.623 (0.076)	0.728 (0.095)	0.800 (0.103)	0.713 (0.086)	0.817 (0.100)
2	0.614 (0.083)	0.606 (0.073)	0.716 (0.093)	0.752 (0.090)	0.667 (0.078)	0.773 (0.088)
3	0.611 (0.076)	0.602 (0.071)	0.722 (0.086)	0.723 (0.077)	0.640 (0.074)	0.748(0.078)
4	0.600 (0.068)	0.593 (0.068)	0.718 (0.075)	0.716 (0.069)	0.630 (0.075)	0.743 (0.072)
8	0.588 (0.069)	0.574 (0.090)	0.737 (0.085)	0.694 (0.072)	0.591 (0.094)	0.724 (0.083)
12	0.568 (0.067)	0.551 (0.103)	0.740 (0.095)	0.690 (0.082)	0.568 (0.108)	0.719 (0.098)

The table reports forecast evaluation statistics for various out of sample horizons, i.e. 1-week up to 12-weeks. The statistics are as follows: root mean square forecast error (RMSFE), log generalized forecast error second moment (L-GFESM) and the coefficient of determination (R^2), the intercept (θ) and slope (δ) parameters for the Mincer–Zarnowitz regression of realized variance on conditional variance forecasted by various models. The sample cover a total of 940 forecasts for each of the horizons considered. The forecasting models are the *GARCH*(1,1) model (*GARCH*), the *Spline-GARCH*(1,1) model (*S-GARCH*), the *Adaptive-GARCH*(1,1) model (*A-GARCH*), the *FIGARCH*$(1,d,1)$ model (*FIGARCH*), the *Spline-FIGARCH*$(1,d,1)$ model (*S-FIGARCH*) and the *Adaptive-FIGARCH*$(1,d,1,8)$ model (*A–FIGARCH*).

by the determinant of the complete (stacked) forecast error second moment matrix, which is

$$GFESM = |E[\mathbf{uu}']|,$$

where $\mathbf{u}$ is the vector of forecasts errors, i.e. the vector constructed by stacking the forecast errors for each of the 12 horizons. In the current framework the *GFESM* statistic can then be simply computed by summing the *MSFE* statistics over the different forecasting horizons considered.[11]

Mincer and Zarnowitz (1969) regressions have also been employed, i.e. by regressing realized variance (σ_t^2), at a give horizon k, on the conditional (to time $t-1$ information) forecast of model i, $f_{t,i}$, at the same horizon

$$\sigma_{t,k}^2 = \theta + \delta f_{t,i,k} + \varepsilon_{i,t}.$$

An accurate forecasting model would then show $\theta = 0$ and $\delta = 1$. Moreover, the R^2 of the regression may be used to assess the ability of the model to track the variability of the forecasting target over time. Newey–West standard errors have been computed to account for the $MA(k-1)$ error process, due to the use of overlapping intervals.

[11] To avoid very large numbers in Table 6 the log transformation of the GFESM statistic (L-GFESM) has been reported.

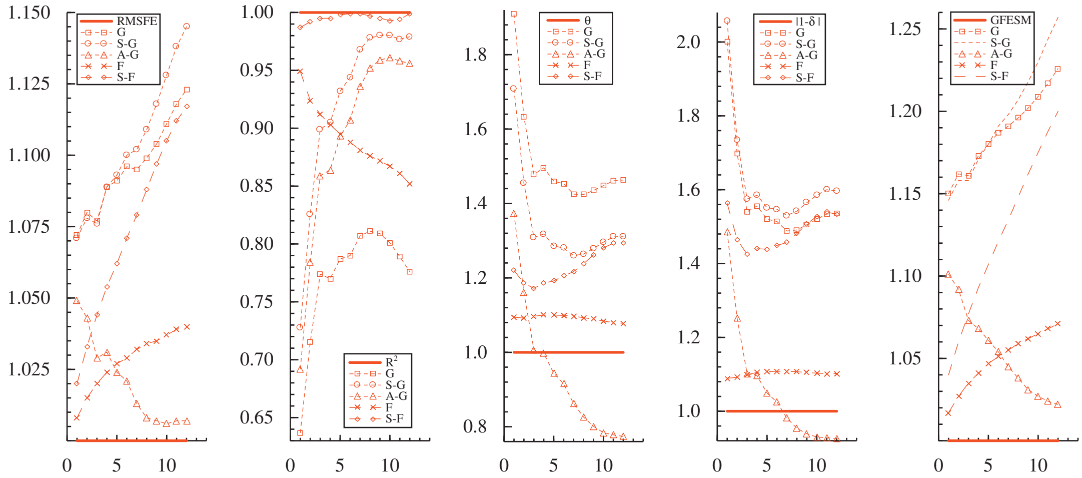

Fig. 3. Forecasting comparison: *GARCH*(1, 1) (G), *Spline* – *GARCH*(1, 1) (*S-G*), *Adaptive* – *GARCH*(1, 1, *k*) (*A-G*), *FIGARCH*(1, *d*, 1) (*F*) and *Spline* – *FIGARCH*(1, *d*, 1) (*S-F*).

Table 6 presents statistics for the various models for selected horizons, while in Fig. 3 the same information is displayed in relative terms (to the *A-FIGARCH* model) for the whole period. Values larger than unity for the relative *RMSFE*, *GFESM*, θ and $|1-\delta|$ statistics for a given competitor model then denote that the *A-FIGARCH* model yields a superior forecasting performance than the competitor model at the selected horizon. On the other hand, a relative R^2 statistic smaller than unity denotes that the *A-FIGARCH* model is performing better than the competing model.

As presented in Table 6, the results of the forecasting analysis are clear-cut, pointing to the *A-FIGARCH* model as the best forecasting model across different horizons and criteria. As is shown in Fig. 3, the *A-FIGARCH* model yields the lowest *RMSFE* across all the horizons, as well as the lowest *GFESM* statistic.[12] According to the West and Cho (1995) test, albeit very small in some cases, the difference in the *RMSFE* statistics is always statistically significant. Moreover, the *A-FIGARCH* model also performs best in terms of Mincer–Zarnowitz regression, particularly at short forecasting horizons. For instance, at the 1-step horizon, only the *FIGARCH* and *A-FIGARCH* models satisfy all the requirements for unbiased forecasting, with the *A-FIGARCH* model always showing superior forecasts in relative terms.

Some interesting conclusions can then be drawn from the forecasting experiments. Firstly, jointly accounting for long memory and structural breaks is important for stock market volatility forecasting, particularly at horizons usually of interest, i.e. the one-week and two-week horizons. In the case just one of the features had to be modelled, then, according to the results, at short-term horizons a pure long memory model, i.e. the *FIGARCH* model, should be preferred to any of the short memory models. Hence, the findings are consistent with the literature, i.e. modelling structural breaks may not be fundamental for short-term forecasting, as a pure long memory model can provide a good approximation to a long memory process subject to breaks in the short-term.

However, as the forecasting horizon increases, modelling structural change can make important differences, since the stationary long memory part of the process reverts towards the long-term level modelled by the break component. Hence, the latter component is important in determining the forecasting ability of the model. The noteworthy performance of the *A-GARCH* model at long-term forecasting horizons can then be understood on the basis of the above considerations: for a long memory process subject to structural change the modelling of the break component is of fundamental importance for accurate long-term forecasting, while the modelling of the long memory part is much less important, and could actually be neglected according to the results of the Mincer–Zarnowitz regression, pointing to non-statistically different results for the two models. Yet, according to the *MSFE* based criteria, the modelling of both features yields to statistically different results, favoring the *A-FIGARCH* model over all the other models, including the *A-GARCH* model. How the break process is modelled can however make the difference, as the trigonometric approach leads to superior results to the spline based approach.

The above evidence refers to the case in which the realized variance is used as a benchmark for forecasting evaluation. However, qualitatively similar conclusions can be reached when the squared range estimator of Alizadeh et al. (2002) is employed for the estimation of the benchmark.[13]

[12] The latter has been computed in a cumulative way. Hence, it is the final value reported in the plot, i.e. the one for $k = 12$, which corresponds to the GFESM statistic formula reported.

[13] The results are available upon request from the authors.

5. Conclusions

This paper has introduced the new Adaptive *FIGARCH* or *A-FIGARCH* process to model volatility, which is designed to account for both long memory and structural change in the conditional variance process. Structural change is modeled by allowing the intercept to follow a slowly varying function, specified by Gallant (1984)'s flexible functional form. A detailed simulation experiment finds that the *A-FIGARCH* model outperforms the standard *FIGARCH* model when structural change is present, and performs at least as well in the absence of structural instability. Overall, there appear to be significant gains in terms of bias and efficiency from using the *A-FIGARCH* specification. An empirical application to stock market volatility is also included to illustrate the usefulness of the technique, as well as the superiority relatively to available alternative modelling strategies.

Acknowledgments

Part of the work reported in this paper was completed while Morana was a Fulbright scholar at Michigan State University. The generosity of the Fulbright Commission and the hospitality of Michigan State University is gratefully acknowledged. The authors are grateful to Andrea Beltratti, Marcelo Fernandes and Christian Conrad for constructive comments. Part of the computations exploited the grid architecture provided to the SEMeQ Department by Avanade of Italy. Claudio Morana also gratefully acknowledges funding by CRT foundation to the project "Grid Computing and Financial Applications".

References

Alizadeh, S., Brandt, M.W., Diebold, F.X., 2002. Range-based estimation of stochastic volatility models. Journal of Finance 57, 1047–1092.
Andersen, T.G., Bollerslev, T., 1997. Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. Journal of Finance 52, 975–1005.
Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American Statistical Association 96, 42–55.
Baillie, R.T., Bollerslev, T., Mikkelsen, H.-O., 1996a. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3–30.
Baillie, R.T., Chung, C.-F., Tieslau, M.A., 1996b. Analysing inflation by the fractionally integrated ARFIMA-GARCH model. Journal of Applied Econometrics 11, 23–40.
Beine, M., Laurent, S., 2000. Structural change and long memory in volatility: new evidence from daily exchange rates. Working Paper, University of Liege.
Beltratti, A., Morana, C., 2006. Breaks and persistency: macroeconomic causes of stock market volatility. Journal of Econometrics 131, 151–171.
Beran, J., Ocker, D., 1999. Volatility of stock market indices—an analysis based on SEMIFAR models. Working Paper, University of Kostanz.
Berkes, I., Horvath, L., Kokoszka, P., 2003. GARCH processes: structure and estimation. Bernoulli 9, 201–227.
Breidt, F.J., Crato, N., de Lima, P., 1998. The detection and estimation of long memory in stochastic volatility. Journal of Econometrics 83, 325–348.
Bollerslev, T., Mikkelsen, H.-O., 1996. Modelling and pricing long memory in stock market volatility. Journal of Econometrics 73, 151–184.
Chung, C.-F., 1999. Estimating the fractionally integrated GARCH model. Working Paper, National Taiwan University.
Clements, M.P., Hendry, D.F., 1993. On the limitations of comparing mean square forecast errors. Journal of Forecasting 12, 617–637.
Conrad, C., Haag, B.R., 2006. Inequality constraints in the fractionally integrated GARCH model. Journal of Financial Econometrics 4, 413–449.
Dacorogna, M.M., Muller, U.A., Nagler, R.J., Olsen, R.B., Pictet, O.V., 1993. A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. Journal of International Money and Finance 12, 413–438.
Dahlhaus, R., Rao, S.S., 2006. Statistical inference for time-varying ARCH processes. Annals of Statistics 34, 1075–1114.
Davidson, J.E.H., 2004. Conditional heteroskedasticity models and a new model. Journal of Business and Economic Statistics 22, 16–29.
Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83–106.
Diebold, F.X., 1986. Comment on "Modeling the persistence of conditional variance" by R. Engle, T, Bollerslev. Econometric Reviews 5, 51–56.
Diebold, F.X., Inoue, A., 2001. Long memory and regime switching. Journal of Econometrics 105, 131–159.
Enders, W., Lee, J., 2004. Testing for a unit-root with a non linear Fourier function. Working Paper, University of Alabama.
Engle, R.F., Rangel, J.G., 2008. The spline-GARCH model for low frequency volatility and its global macroeconomic causes. Review of Financial Studies 21, 1187–1222.
Gallant, A.R., 1984. The Fourier flexible form. American Journal of Agricultural Economics 66, 204–208.
Granger, C.W.J., Ding, Z., 1996. Modeling volatility persistence of speculative returns. Journal of Econometrics 73, 185–215.
Granger, C.W.J., Hyung, N., 2004. Occasional structural breaks and long memory with an application to the S&P500 absolute returns. Journal of Empirical Finance 11, 399–421.
Granger, C.W.J., Terasvirta, T., 1999. A simple nonlinear time series model with misleading linear properties. Economics Letters 62, 161–165.
Hamilton, J.D., Susmel, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics 64, 307–333.
Harvey, A.C., 1998. Long memory in stochastic volatility. In: Knight, J., Satchell, S. (Eds.), Forecasting Volatility in Financial Markets. Butterworth-Heineman, Oxford, pp. 307–320.
Haas, M., Mittnik, S., Paolella, M.S., 2004. A new approach to switching GARCH models. Journal of Financial Econometrics 2, 493–530.
Hillebrand, E., 2005. Neglecting parameter changes in GARCH models. Journal of Econometrics 129, 121–138.
Jensen, S.T., Rahbek, A., 2004. Asymptotic inference for nonstationary GARCH. Econometric Theory 20, 1203–1226.
Karanasos, M., Psaradakis, Z., Sola, M., 2004. On the autocorrelation properties of long memory GARCH processes. Journal of Time Series Analysis 25, 265–281.
Kazakevicius, V., Leipus, R., 2002. On stationarity in the ARCH (∞) model. Econometric Theory 18, 1–16.
Lamoreaux, C.G., Lastrapes, W.D., 1990. Persistence in variance, structural change and the GARCH model. Journal of Business and Economic Statistics 8, 225–234.
Lee, S.-W., Hansen, B.E., 1994. Asymptotic theory for the GARCH(1, 1) quasi-maximum likelihood estimator. Econometric Theory 10, 29–52.
Lobato, I.N., Savin, N.E., 1998. Real and spurious long memory properties of stock market data. Journal of Business and Economic Statistics 16, 261–268.
Lumsdaine, R.L., 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.

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.
Medeiros, M., Veiga, A., 2004, Modelling multiple regimes in financial volatility with a flexible coefficient GARCH$(1,1)$ model, Testo para Discussao 486, Pontificia Universidade Catolica do Rio de Janeiro.
Mikosch, T., Starica, C. 1998, Change of structure in financial time series, long range dependence and the GARCH model. Working Paper, University of Groningen.
Mincer, J., Zarnowitz, V., 1969. The evaluation of economic forecasts. In: Mincer, J. (Ed.), Economic Forecasts and Expectations. National Bureau of Economic Research, New York.
Morana, C., 2002. IGARCH effects: an interpretation. Applied Economics Letters 9, 745–748.
Morana, C., Beltratti, A., 2004. Structural change and long range dependence in volatility of exchange rates: either, neither or both? Journal of Empirical Finance 11, 629–658.
Muller, U.A., Dacorogna, M.M., Davé, R.D., Olsen, R.B., Pictet, O.V., von Weizsacker, J.E., 1997. Volatilities of different time resolutions—analyzing the dynamics of market components. Journal of Empirical Finance 4, 213–239.
Nyblom, J., 1989. Testing for the constancy of parameters over time. Journal of the American Statistical Association 84, 223–230.
Pascalau, R., 2007. Unconditional mean, volatility and the Fourier-GARCH representation. Mimeo, University of Alabama.
Robinson, P.M., Zaffaroni, P., 2006. Pseudo-maximum likelihood estimation of ARCH(∞) models. Annals of Statistics 34, 1049–1074.
Schwert, G.W., 1989. Business cycles, financial crises, and stock volatility? Carnegie–Rochester Conference Series on Public Policy 31, 83–126.
Silverman, B.W., 1985. Some aspects of the spline smoothing approach to non-parametric regression curve fitting. Journal of the Royal Statistical Society B 47, 1–52.
Starica, C., Granger, C., 2005. Non-stationarities in stock returns. Review of Economics and Statistics 87, 503–522.
Terasvirta, T., Gonzalez, A., 2006. Modelling autoregressive processes with a shifting mean. Working Paper no. 637, Stockholm School of Economics.
West, K.D., Cho, D., 1995. The predicitive ability of several models of exchange rate volatility. Journal of Econometrics 69, 367–391.
Zaffaroni, P., 2007. Contemporaneous aggregation of GARCH processes. Journal of Time Series Analysis 28, 521–544.