

## Fractionally integrated generalized autoregressive conditional heteroskedasticity

Richard T. Baillie<sup>a</sup>, Tim Bollerslev<sup>\*,b</sup>, Hans Ole Mikkelsen<sup>c</sup>

<sup>a</sup>*Department of Economics, Michigan State University, East Lansing, MI 48824, USA*

<sup>b</sup>*Department of Economics, University of Virginia, Charlottesville, VA 22901, USA*

<sup>c</sup>*Department of Finance and Business Economics, University of Southern California, Los Angeles, CA 90089, USA*

---

### Abstract

The new class of Fractionally Integrated Generalized AutoRegressive Conditionally Heteroskedastic (FIGARCH) processes is introduced. The conditional variance of the process implies a slow hyperbolic rate of decay for the influence of lagged squared innovations. Unlike  $I(d)$  processes for the mean, Maximum Likelihood Estimates (MLE) of the FIGARCH parameters are argued to be  $T^{1/2}$ -consistent. The small-sample behavior of an approximate MLE procedure is assessed through a simulation study, which also documents how the estimation of a standard GARCH model tends to produce integrated, or IGARCH, like estimates. An empirical example with daily Deutschmark–U.S. dollar exchange rates illustrates the practical relevance of the new FIGARCH specification.

*Key words:* Fractional integrated ARCH; FIGARCH; IGARCH; Mean-reversion; Monte Carlo simulations; Exchange rate volatility

*JEL classification:* C15; C22; F31

---

\*Corresponding author.

The second author is also affiliated with the NBER. Most of this research was completed while the third author was visiting Northwestern University. We gratefully acknowledge the helpful comments received from three anonymous referees, Robin Brenner, Yin-Wong Cheung, Miguel Delgado, Francis X. Diebold, Andrew Harvey, Campbell Harvey, Daniel B. Nelson, Peter M. Robinson, participants at the 1993 workshop on 'Modern Time Series Analysis in Finance' at the University of Aarhus, Denmark, the 1994 conference on 'Asymmetries and Non-Linearities in Dynamic Economic Models' in Madrid, the 1994 NBER Summer Institute, as well as seminar audiences at the University of Arizona, University of California at Santa Barbara, University of Iowa, University of Minnesota, University of Montreal, University of Texas at Austin, and University of Virginia.

## 1. Introduction

Much recent experience in modeling the conditional mean of macroeconomic and financial time series data has emphasized the role of persistence of shocks, and a large literature has emerged on testing for and estimating unit roots in the autoregressive representations of univariate and vector processes. Related advances have occurred in the area of common stochastic trends and cointegration. However, the knife-edge distinction between  $I(0)$  and  $I(1)$  processes can be far too restrictive. In the discrete time long-memory fractionally integrated  $I(d)$  class of processes, proposed by Adenstedt (1974), Granger (1980, 1981), Granger and Joyeux (1980), and Hosking (1981), the propagation of shocks to the mean occurs at a slow hyperbolic rate of decay when  $0 < d < 1$ , as opposed to the extremes of  $I(0)$  exponential decay associated with the stationary and invertible ARMA class of processes, or the infinite persistence resulting from an  $I(1)$  process. Important theoretical work on this class of models includes the recent advances by Dahlhaus (1989), Lo (1991), Robinson (1991, 1994a), Sowell (1990, 1992a), and Yajima (1988), among others; see also Baillie (1996) and Robinson (1994b) for recent surveys of the literature on  $I(d)$  processes in economics.

It has long been recognized that asset returns determined in speculative markets are approximately uncorrelated but not independent through time, as most return processes tend to exhibit temporal bursts of volatility; see Mandelbrot (1963) for some of the earliest evidence along these lines. In particular, following the introduction of the ARCH model by Engle (1982), and the popular GARCH and EGARCH parameterizations proposed by Bollerslev (1986) and Nelson (1991), many empirical studies have noted the extreme degree of persistence of shocks to the conditional variance process; for a recent survey of the extensive literature on ARCH modeling in finance see Bollerslev, Chou, and Kroner (1992). Analogous to the issues pertaining to the proper modeling of the long-run dependencies in the conditional mean of economic time series, similar questions therefore become relevant in the modeling of conditional variances. This observation led Engle and Bollerslev (1986) to the formulation of the Integrated GARCH, or IGARCH, class of models, which possesses many of the features of the unit root, or  $I(1)$ , processes for the mean.<sup>1</sup> For instance, the implied effect of a shock for the optimal forecast of the future conditional variance will be to make the corresponding cumulative impulse response weights tend to a nonzero constant, so that the forecasts will increase linearly with the forecast horizon. This implies that the pricing of risky securities, including long-term options and futures contracts, may show extreme

---

<sup>1</sup> The analogy to  $I(1)$  processes for the conditional mean is far from complete, however; see, e.g., Bollerslev, Engle, and Nelson (1994), Gallant, Rossi, and Tauchen (1993), Nelson (1990a), and the discussion in Section 2 below.

dependence on the initial conditions, or the current state of the economy. However, this extreme degree of dependence seems contrary to observed pricing behavior. Also, recent studies by deLima, Breidt, and Crato (1994), Dacorogna et al. (1993), Ding, Granger, and Engle (1993), and Harvey (1993) all report the presence of apparent long-memory in the autocorrelations of squared or absolute returns of various financial asset prices. Motivated by these observations, this paper introduces the Fractionally Integrated Generalized AutoRegressive Conditionally Heteroskedastic, or FIGARCH, class of processes. The primary purpose of this new approach is to develop a more flexible class of processes for the conditional variance that are more capable of explaining and representing the observed temporal dependencies in financial market volatility.

The plan for the rest of the paper is as follows. Section 2 formally defines the process and discusses some of the most important population characteristics. The FIGARCH process combines many of the features of the fractionally integrated process for the mean together with the regular GARCH process for the conditional variance. In particular, the FIGARCH model implies a slow hyperbolic rate of decay for the lagged squared innovations in the conditional variance function, although the cumulative impulse response weights associated with the influence of a volatility shock on the optimal forecasts of the future conditional variance eventually tend to zero; a property the model shares with weakly stationary GARCH processes.<sup>2</sup> An approximate Maximum Likelihood Estimation (MLE) procedure is discussed in Section 3, which also presents the results of a detailed simulation study. Unlike the estimation of ARFIMA processes for the mean, the MLE of the parameters for the FIGARCH process appears  $T^{1/2}$ -consistent. Detailed simulation evidence reveals that for the sample sizes typically encountered with financial data, this approximate MLE works extremely well in terms of estimating both the parameters of the process and their asymptotic standard errors. Section 4 compares and contrasts the FIGARCH and IGARCH models. The results of estimating regular GARCH processes, when the true data generating process is FIGARCH, are also examined. Interestingly, the estimated autoregressive parameters in the misspecified GARCH models are found to be very close to unity, indicative of IGARCH type behavior. These results lead us to conjecture that the apparent widespread IGARCH property so frequently reported with high-frequency asset pricing data may well be spurious, and that the IGARCH process provides a poor diagnostic for distinguishing between integrated, as opposed to long-memory

---

<sup>2</sup> In a related development, Harvey (1993) and deLima, Breidt, and Crato (1994) have recently considered stochastic volatility models which imply a similarly slow hyperbolic rate of decay. The much easier inferential procedures for ARCH-type models is one obvious advantage of this approach; for additional discussion of these issues see Andersen (1994), Harvey, Ruiz, and Shephard (1994), Harvey and Shephard (1993), Jacquier, Polson, and Rossi (1994), and Taylor (1986, 1994).

formulations of the conditional variance process. Section 5 of the paper applies the FIGARCH model in estimating the volatility process for a long time series of daily Deutschmark–U.S. dollar spot exchange rates. The empirical evidence favors the FIGARCH formulation, indicating clear rejections of the stable GARCH and IGARCH null hypotheses. The corresponding impulse response weights derived from the FIGARCH model also appear to be more realistic from an economic perspective when compared to the fairly rapid rate of decay associated with the estimated covariance stationary GARCH model or the infinite persistence for the IGARCH formulation. The final section provides a brief conclusion, together with several suggestions for extensions and applications of the ideas developed in the paper.

## 2. Definition and properties of the FIGARCH process

Following Engle (1982), consider the discrete time real-valued stochastic ARCH process,  $\{\varepsilon_t\}$ ,

$$\varepsilon_t \equiv z_t \sigma_t, \quad (1)$$

where  $E_{t-1}(z_t) = 0$  and  $\text{VAR}_{t-1}(z_t) = 1$ , and  $\sigma_t$  is a positive time-varying and measurable function with respect to the information set available at time  $t - 1$ . Throughout,  $E_{t-1}(\cdot)$  and  $\text{VAR}_{t-1}(\cdot)$  refer to the conditional expectation and variance with respect to this same information set. Thus, by definition, the  $\{\varepsilon_t\}$  process is serially uncorrelated with mean zero, but the conditional variance of the process,  $\sigma_t^2$ , is changing over time.

In the classic ARCH( $q$ ) model of Engle (1982), the conditional variance  $\sigma_t^2$  is postulated to be a linear function of the lagged squared innovations implying Markovian dependence dating back only  $q$  periods; i.e.,  $\varepsilon_{t-i}^2$  for  $i = 1, 2, \dots, q$ . The GARCH( $p, q$ ) specification of Bollerslev (1986) provides a more flexible lag structure. Formally, this model is defined by

$$\sigma_t^2 = \omega + \alpha(L)\varepsilon_t^2 + \beta(L)\sigma_t^2, \quad (2)$$

where  $L$  denotes the lag or backshift operator, and  $\alpha(L) \equiv \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_q L^q$  and  $\beta(L) \equiv \beta_1 L + \beta_2 L^2 + \dots + \beta_p L^p$ . For stability and covariance stationarity of the  $\{\varepsilon_t\}$  process, all the roots of  $[1 - \alpha(L) - \beta(L)]$  and  $[1 - \beta(L)]$  are constrained to lie outside the unit circle. The GARCH( $p, q$ ) process may be rewritten as the infinite-order ARCH process,

$$\begin{aligned} \sigma_t^2 &= \omega[1 - \beta(1)]^{-1} + \alpha(L)[1 - \beta(L)]^{-1} \varepsilon_t^2 \\ &\equiv \omega[1 - \beta(1)]^{-1} + \lambda(L)\varepsilon_t^2. \end{aligned} \quad (3)$$

The before mentioned stationarity condition implies that the effect of the past squared innovations on the current conditional variance decays exponentially

with the lag length. Alternatively, the GARCH( $p, q$ ) process in Eq. (2) may also be expressed as an ARMA( $m, p$ ) process in  $\varepsilon_t^2$ ,

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

where  $m \equiv \max\{p, q\}$ , and  $v_t \equiv \varepsilon_t^2 - \sigma_t^2$  is mean zero serially uncorrelated. Thus, the  $\{v_t\}$  process is readily interpreted as the ‘innovations’ for the conditional variance. When the autoregressive lag polynomial,  $1 - \alpha(L) - \beta(L)$ , contains a unit root, the GARCH( $p, q$ ) process is defined by Engle and Bollerslev (1986) to be integrated in variance. The corresponding Integrated GARCH( $p, q$ ), or IGARCH( $p, q$ ), class of models is given succinctly by

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

where  $\phi(L) \equiv [1 - \alpha(L) - \beta(L)](1 - L)^{-1}$  is of order  $m - 1$ . The Fractionally Integrated GARCH, or FIGARCH, class of models is simply obtained by replacing the first difference operator in Eq. (5) with the fractional differencing operator.

In order to motivate this development, it is worth briefly considering the fractionally integrated process for the mean. The concept of long-memory and fractional Brownian motion was originally developed by Hurst (1951) and Mandelbrot and Van Ness (1968). However, the ideas were essentially operationalized for applications with discrete time representations by Granger (1980, 1981), Granger and Joyeux (1980), and Hosking (1981). In particular, the ARFIMA( $k, d, l$ ) class of models for the discrete time real-valued process  $\{y_t\}$  is defined by

$$a(L)(1 - L)^d y_t = b(L)\varepsilon_t, \tag{6}$$

where  $a(L)$  and  $b(L)$  are polynomials in the lag operator of orders  $k$  and  $l$  respectively, and  $\{\varepsilon_t\}$  is a mean-zero, serially uncorrelated process. The fractional differencing operator,  $(1 - L)^d$ , has a binomial expansion which is most conveniently expressed in terms of the hypergeometric function,

$$\begin{aligned} (1 - L)^d &= F(-d, 1, 1; L) \\ &= \sum_{k=0, \infty} \Gamma(k - d)\Gamma(k + 1)^{-1}\Gamma(-d)^{-1}L^k \\ &\equiv \sum_{k=0, \infty} \pi_k L^k, \end{aligned} \tag{7}$$

where  $\Gamma(\cdot)$  denotes the Gamma function.<sup>3</sup> Provided that  $\text{var}(\varepsilon_t) < \infty$  and  $-0.5 < d < 0.5$ , the  $\{y_t\}$  process in Eq. (6) is weakly stationary and invertible,

<sup>3</sup>The hypergeometric function is formally defined by

$$F(m, n, s; x) = \Gamma(s)\Gamma(m)^{-1}\Gamma(n)^{-1} \sum_{j=0, \infty} \Gamma(m + j)\Gamma(n + j)\Gamma(s + j)^{-1}\Gamma(j + 1)^{-1}x^j.$$

and will possess unique infinite moving average and autoregressive representations. However, for  $d > 0$  the process is long memory in the sense that  $\lim_{k \rightarrow \infty} \sum_{j=-k, k} |\rho_j|$ , where  $\rho_j$  denotes the autocorrelation of the process at lag  $j$ , does not converge to a finite limit. As argued forcefully by Sowell (1992b), the ARFIMA model essentially disentangles the short-run and the long-run dynamics, by modelling the short-run behavior through the conventional ARMA lag polynomials,  $a(L)$  and  $b(L)$ , while the long-run characteristic is captured by the fractional differencing parameter,  $d$ .

Analogously to the ARFIMA( $k, d, l$ ) process for the mean, the FIGARCH( $p, d, q$ ) process for  $\{\varepsilon_t\}$  is naturally defined by

$$\phi(L)(1-L)^d \varepsilon_t^2 = \omega + [1 - \beta(L)]v_t, \quad (8)$$

where  $0 < d < 1$ , and all the roots of  $\phi(L)$  and  $[1 - \beta(L)]$  lie outside the unit circle.<sup>4</sup> Rearranging the terms in Eq. (8), an alternative representation for the FIGARCH( $p, d, q$ ) model is

$$[1 - \beta(L)]\sigma_t^2 = \omega + [1 - \beta(L) - \phi(L)(1-L)^d]\varepsilon_t^2. \quad (9)$$

Thus, the conditional variance of  $\varepsilon_t$  is simply given by

$$\begin{aligned} \sigma_t^2 &= \omega[1 - \beta(1)]^{-1} + \{1 - [1 - \beta(L)]^{-1}\phi(L)(1-L)^d\}\varepsilon_t^2 \\ &\equiv \omega[1 - \beta(1)]^{-1} + \lambda(L)\varepsilon_t^2, \end{aligned} \quad (10)$$

where  $\lambda(L) = \lambda_1 L + \lambda_2 L^2 + \dots$ . Of course, for the FIGARCH( $p, d, q$ ) process in Eq. (8) to be well-defined and the conditional variance to be positive almost surely for all  $t$ , all the coefficients in the infinite ARCH representation in Eq. (10) must be nonnegative; i.e.,  $\lambda_k \geq 0$  for  $k = 1, 2, \dots$ . As for the GARCH( $p, q$ ) class of models analyzed by Nelson and Cao (1992), general conditions to ensure nonnegativity of all the lag coefficients in  $\lambda(L)$  have proven elusive. Fortunately, as illustrated below, sufficient conditions are fairly easy to establish on a case-by-case basis.

For  $0 < d \leq 1$  the hypergeometric function evaluated at  $L = 1$  equals zero,  $F(-d, 1, 1; 1) = 0$ , so that  $\lambda(1) = 1$ . The  $\omega > 0$  term therefore has the same interpretation as in the IGARCH model. Consequently, the second moment of the unconditional distribution of  $\varepsilon_t$  is infinite, and the FIGARCH process is clearly not weakly stationary; a feature it shares with the IGARCH class of processes.<sup>5</sup> However, as shown by Nelson (1990a) for the IGARCH(1, 1) model

<sup>4</sup> A similar specification with  $\omega = 0$  has been used by Robinson (1991) in formulating tests for  $d > 0$ .

<sup>5</sup> Note, this differs from the corresponding analog for ARFIMA-type models. With dynamic models for the conditional mean that does not include a drift, summability of the squared coefficients of the infinite-order MA representation is required for the model to be covariance-stationary. This holds true for the ARFIMA model with  $d < 0.5$ . In contrast, for ARCH-type models the stationarity condition depends on the actual cumulated coefficients from the MA representation.

and extended to the general IGARCH( $p, q$ ) model by Bougerol and Picard (1992), IGARCH models are strictly stationary and ergodic. Since the high-order lag coefficients in the infinite ARCH representation of any FIGARCH model may be dominated in an absolute value sense by the corresponding IGARCH coefficients from Eq. (5), it follows by a direct extension of the proofs for the IGARCH case that the FIGARCH( $p, d, q$ ) class of processes is strictly stationary and ergodic for  $0 \leq d \leq 1$ .

As highlighted by this discussion, considerable care should be exercised in interpreting persistence in nonlinear models. Formally, Bollerslev and Engle (1993) define a process to be persistent in variance if  $\limsup_{k \rightarrow \infty} |E_{t+s}(\varepsilon_{t+k}^2) - E_t(\varepsilon_{t+k}^2)| > 0$  for some  $s > 0$ . This same notion of infinite dependence on the initial conditions for the optimal forecasts of the future conditional variances also underlies the conditional moment profiles analyzed by Gallant, Rossi, and Tauchen (1993). However, in the present context in which the conditional variance is parameterized as a linear function of the past squared innovations, the persistence of the conditional variance is most simply characterized in terms of the impulse response coefficients for the optimal forecast of the future conditional variance as a function of the time  $t$  innovation,  $v_t$ ,

$$\gamma_k \equiv \partial E_t(\varepsilon_{t+k}^2) / \partial v_t - \partial E_t(\varepsilon_{t+k-1}^2) / \partial v_t. \tag{11}$$

Of course, in more general conditional variance models the  $\gamma_i$ 's will depend on the time  $t$  information set. However, for the FIGARCH class of models analyzed here, the impulse response coefficients are independent of  $t$ , and the persistence as measured by the  $\gamma_i$  coefficients corresponds directly to the generalization of the linear impulse response analysis to nonlinear models developed by Gallant, Rossi, and Tauchen (1993). Specifically, the impulse response coefficients may be found from the coefficients in the  $\gamma(L)$  lag polynomial,

$$\begin{aligned} (1 - L)\varepsilon_t^2 &= (1 - L)^{1-d} \phi(L)^{-1} \omega + (1 - L)^{1-d} \phi(L)^{-1} [1 - \beta(L)] v_t, \\ &\equiv \zeta + \gamma(L) v_t, \end{aligned} \tag{12}$$

where the first equality follows directly from the definition of the FIGARCH( $p, d, q$ ) process in Eq. (8). Analogously to conventional impulse response analysis for the mean, the long-run impact of past shocks for the volatility process may now be assessed in terms of the limit of the cumulative impulse response weights, i.e.,

$$\begin{aligned} \gamma(1) &= \lim_{k \rightarrow \infty} \sum_{i=0, k} \gamma_i = \lim_{k \rightarrow \infty} \lambda_k \\ &= F(d - 1, 1, 1; 1) \phi(1)^{-1} [1 - \beta(1)]. \end{aligned} \tag{13}$$

As noted above, for  $0 \leq d < 1$ ,  $F(d - 1, 1, 1; 1) = 0$ , so that for the covariance-stationary GARCH( $p, q$ ) model and the FIGARCH( $p, d, q$ ) model with  $0 < d$

$< 1$ , shocks to the conditional variance will ultimately die out in a forecasting sense. There are important differences in the shock dissipation for  $d = 0$  and  $0 < d < 1$ , however. Whereas shocks to the GARCH process die out at a fast exponential rate, for the FIGARCH model  $\lambda_k$  will eventually be dominated by a hyperbolic rate of decay; see, e.g., Diebold, Husted, and Rush (1991). Thus, even though the cumulative impulse response function converges to zero for  $0 \leq d < 1$ , the fractional differencing parameter provides important information regarding the pattern and speed with which shocks to the volatility process are propagated. In contrast, for  $d = 1$ ,  $F(d - 1, 1, 1; 1) = 1$ , and the cumulative impulse response weights will converge to the nonzero constant  $\gamma(1) = \phi(1)^{-1} \times [1 - \beta(1)]$ . Thus, from a forecasting perspective shocks to the conditional variance of the IGARCH model persist indefinitely. For values of  $d > 1$ ,  $F(d - 1, 1, 1; 1) = \infty$ , resulting in an unrealistic explosive conditional variance process and  $\gamma(1)$  being undefined.

In most practical applications relatively simple first-order models have been found to provide good representations of the conditional variance processes. To illustrate the ideas developed above consider therefore the simple GARCH(1,1) model,

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

rewritten in ARMA(1,1) form as

$$(1 - \phi_1 L)\varepsilon_t^2 = \omega + (1 - \beta_1 L)v_t,$$

where  $\phi_1 \equiv \alpha_1 + \beta_1$ . The impulse response weights for this model are given by the coefficients in the polynomial,  $\gamma(L) = (1 - L)(1 - \phi_1 L)^{-1}(1 - \beta_1 L)$ , so that  $\gamma_0 = 1$ ,  $\gamma_1 = \phi_1 - \beta_1 - 1$ , and  $\gamma_k = (\phi_1 - \beta_1)(\phi_1 - 1)\phi_1^{k-2}$  for  $k > 2$ ; see also Engle and Bollerslev (1986). The cumulative impulse response weights for the process equals  $\lambda_k = (\phi_1 - \beta_1)\phi_1^{k-1}$  for  $k > 1$ , and in the limit  $\gamma(1) = 0$  provided that  $0 < \phi_1 < 1$ . Hence, the effect of a shock for the forecast of the future conditional variance tend to zero at a fast exponential rate. The IGARCH(1,1) model occurs when  $\phi_1 = 1$ ,

$$(1 - L)\varepsilon_t^2 = \omega + (1 - \beta_1 L)v_t.$$

In this situation,  $\lambda_k = (1 - \beta_1)$  for all lags  $k > 1$ , and all the cumulative impulse response weights are equal to the nonzero constant  $\gamma(1) = 1 - \beta_1$ . The corresponding FIGARCH(1,  $d$ , 0) model is

$$(1 - L)^d \varepsilon_t^2 = \omega + (1 - \beta_1 L)v_t.$$

By analogy to the properties for the ARFIMA(0,  $d$ , 1) model developed in Hosking (1981), it is possible to show that the cumulative impulse response coefficients in the infinite ARCH representation for the FIGARCH(1,  $d$ , 0) model,  $\lambda(L) \equiv 1 - (1 - \beta_1 L)^{-1}(1 - L)^d$ , equal

$$\lambda_k = [1 - \beta_1 - (1 - d)k^{-1}] \cdot \Gamma(k + d - 1)\Gamma(k)^{-1}\Gamma(d)^{-1},$$

for  $k > 1$ , and  $\lambda_0 = 1$ . Thus, provided that  $\omega > 0$ , the condition  $0 \leq \beta_1 < d \leq 1$  is both necessary and sufficient to ensure that the conditional variance in the FIGARCH(1,  $d$ , 0) model is positive almost surely for all  $t$ . Furthermore, it follows by a straightforward application of Sterling's formula, that for high lags,  $k$ ,

$$\lambda_k \approx [(1 - \beta_1)\Gamma(d)^{-1}]k^{d-1}.$$

In contrast to the covariance stationary GARCH(1,1) model or the IGARCH(1,1) model, where shocks to the conditional variance either dissipates exponentially or persist indefinitely; for the FIGARCH(1,  $d$ , 0) model the response of the conditional variance to past shocks decays at a slow hyperbolic rate.

### 3. Estimation and inference

The most common approach for estimation of ARCH models assumes conditional normality of the process. Under this assumption, Maximum Likelihood Estimates (MLE) for the parameters of the FIGARCH( $p$ ,  $d$ ,  $q$ ) process based on the sample  $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T\}$  may therefore be obtained by maximizing the expression

$$\begin{aligned} \log L(\theta; \varepsilon_1, \varepsilon_2, \dots, \varepsilon_T) \\ = -0.5 \cdot T \cdot \log(2\pi) - 0.5 \cdot \sum_{t=1, T} [\log(\sigma_t^2) + \varepsilon_t^2 \sigma_t^{-2}], \end{aligned} \quad (14)$$

where  $\theta' \equiv (\omega, d, \beta_1, \dots, \beta_p, \phi_1, \dots, \phi_q)$ .<sup>6</sup>

In practice some initial conditions are required to start up the recursions for the conditional variance function. The approach taken here is to maximize the likelihood function conditional on these start-up values. In particular, as is standard in the estimation of both stable GARCH and IGARCH models, we simply fix all the pre-sample values of  $\varepsilon_t^2$  for  $t = 0, -1, -2, \dots$  in the infinite ARCH representation in Eq. (10) at the unconditional sample variance. Of course, for the FIGARCH( $p$ ,  $d$ ,  $q$ ) model with  $d > 0$  the population variance does not exist. However, subject to the regularity conditions specified below, conditioning on the pre-sample values will not affect the asymptotic

<sup>6</sup> In most practical applications  $\varepsilon_t$  will correspond to the innovations for the mean of some other process of interest, e.g.,  $y_t = E_{t-1}(y_t) + \varepsilon_t = f_{t-1}(\mu) + \varepsilon_t$ , where  $\mu$  refers to the parameters for the conditional mean function. The likelihood function for the augmented parameter vector  $(\theta', \mu')$  and the sample  $\{y_1, y_2, \dots, y_T\}$  is identical to Eq. (14).

distributions of the resulting estimators and test statistics.<sup>7</sup> In most practical applications with high-frequency financial data the standardized innovations  $z_t = \varepsilon_t \sigma_t^{-1}$  are leptokurtic and not i.i.d. normally distributed through time. In these situations the robust Quasi-MLE (QMLE) procedures discussed by Weiss (1986) and Bollerslev and Wooldridge (1992) may be invoked to allow for asymptotically valid inference.

Unfortunately, the consistency and asymptotic normality of the QMLE based ARCH estimators and test statistics have only been formally established for the IGARCH(1,1) case to date. In particular, following Lee and Hansen (1994), and assuming that (i)  $z_t$  is stationary and ergodic, (ii)  $z_t^2$  is nondegenerate, (iii)  $E_{t-1}(z_t^4) \leq \kappa < \infty$  almost surely, and (iv)  $\sup E_{t-1}[\log(\beta_1 + \alpha_1 z_t^2)] < \infty$  almost surely, it is possible to show that the quasi-likelihood function and the corresponding score vector and Hessian, are all strictly stationary and ergodic.<sup>8</sup> It follows therefore by a central limit theorem, that the QMLE obtained by maximizing (14), say  $\hat{\theta}_T$ , is both consistent and asymptotically normally distributed,

$$T^{1/2}(\hat{\theta}_T - \theta_0) \rightarrow N(0, A(\theta_0)^{-1} B(\theta_0) A(\theta_0)^{-1}), \quad (15)$$

where  $A(\cdot)$  and  $B(\cdot)$  represent the Hessian and the outer product of the gradients respectively, both evaluated at the true parameters,  $\theta_0$ . While the above result for the IGARCH(1,1) case extends directly to the FIGARCH(1,  $d$ , 0) model through a dominance-type argument, a formal proof of consistency and asymptotic normality of  $\hat{\theta}_T$  for the general IGARCH( $p$ ,  $q$ ) case, with obvious implications for the FIGARCH( $p$ ,  $d$ ,  $q$ ) model, remains an important issue for future research.

In order to assess the practical applicability and small sample performance of the QMLE procedure for the estimation of FIGARCH processes, we now turn to the results from a detailed simulation study. As discussed above, the implementation of this approach necessitates conditioning on pre-sample values and a truncation of the infinite lag polynomial in Eq. (10). Given the long memory and relatively slow decay of a response to a lagged squared innovation, the effect of pre-sample values might be expected to have a bigger impact than with stationary GARCH processes.<sup>9</sup> Thus, truncating at too low a lag may destroy

<sup>7</sup> Diebold and Schuermann (1996) use nonparametric density estimation techniques in evaluating the exact likelihood function for low-order covariance-stationary ARCH models. For sample sizes of 50 or larger, the exact results are almost identical to the estimates based on the conditional likelihood function.

<sup>8</sup> An alternative approach, due to Lumsdaine (1996), assumes that the first 32 moments of the conditional distribution of  $z_t$  exist.

<sup>9</sup> The analogous effect of pre-sample values in the estimation of long-memory ARFIMA process for the mean has been documented by Cheung and Diebold (1994) and Chung and Baillie (1993). For sample sizes of  $T = 100$  or larger the effect of pre-sample values appears to have negligible effect on the MLE of the ARFIMA parameters. Furthermore, with financial applications the time series of interest often consist of several thousand observations.

important long-run dependencies. To mitigate these effects the truncation lag was set at 1,000 for all of the simulation and estimation results reported below.<sup>10</sup>

The true model parameters for the different Data Generation Processes (DGP) are reported in the first five columns of Tables 1 and 2, with the orders of the estimated FIGARCH( $p, d, q$ ) models indicated in the ( $p, d, q$ ) column. FIGARCH models with  $d = 0.0, 0.5, 0.75$ , and  $1.0$ , and sample sizes of 1,500 and 3,000 were investigated. A constant mean,  $\mu$ , was estimated jointly with the FIGARCH parameters; i.e.,  $y_t = \mu + \varepsilon_t$  for  $t = 1, 2, \dots, T$ . In order to concentrate on the estimation of conditional variance parameters,  $\mu$  was fixed at zero across all the experiments. Similarly, the intercept in the conditional variance equation was set at  $\omega = 0.1$ . In addition to the simulations with conditionally normally distributed errors, we also report the results for a FIGARCH(1,  $d$ , 0) model with  $z_t$  i.i.d. Student- $t$  distributed with seven degrees of freedom. This latter distribution has conditional kurtosis equal to five, as often observed in empirical work with GARCH models for high-frequency data; see, e.g., Bollerslev (1987). For all of the different DGP's a total of 500 replications were generated.<sup>11</sup> The Bias, RMSE, and Std. columns report the simulated bias, root mean squared error, and average estimated standard error of the QMLE across the 500 replications. The standard errors are based on the finite-sample approximation to Eq. (15); i.e.,  $\hat{\theta}_T \sim N(\theta_0, T^{-1}A(\hat{\theta}_T)^{-1}B(\hat{\theta}_T)A(\hat{\theta}_T)^{-1})$ .

Turning to the results for  $\hat{d}$  in Table 1, it is immediately clear from the Bias column, that from a practical perspective the approximate MLE method works extremely well in terms of estimating the true  $d$  parameters for all of the eight different designs. Even though the simulated finite-sample bias is significantly different from zero at conventional significance levels for some of the models, no systematic bias is apparent.<sup>12</sup> Interestingly, from the first two rows of the table, the estimation of the extra  $\phi_1$  parameter in the FIGARCH(1,  $d$ , 1) model, which nests the standard GARCH(1,1) specification, adds only little to the sampling variability of  $\hat{d}$  when compared to the estimates from the true FIGARCH(1,  $d$ , 0) model. Also, when comparing the estimates of  $d$  from the two FIGARCH(1,  $d$ , 0) models with  $T = 3,000$  and  $d = 0.75$  or  $d = 0.5$  to the results for the same two

<sup>10</sup> To gauge an idea about the magnitude of the truncation bias note that for  $0 < d$  the infinite sum in Eq. (7) evaluated at  $L = 1$  equals zero; i.e.,  $\sum_{k=0, \dots, \infty} \Gamma(k-d)\Gamma(k+1)^{-1}\Gamma(-d)^{-1} = 0$ . Truncating this expansion at  $k = 1,000$  for  $d = 0.75$  yields 0.00155.

<sup>11</sup> To avoid start-up problems, the first 7,000 realizations were discarded for each replication. The normal random variables,  $z_t$ , were generated by RNDNS in the GAUSS computer language. The  $t$ -distributed errors were generated as  $z_t \equiv 5^{1/2}z_{1,t}(z_{2,t}^2 + z_{3,t}^2 + \dots + z_{8,t}^2)^{-1/2}$  by drawing the  $z_{i,t}$ 's as i.i.d. standard normal.

<sup>12</sup> Let  $\bar{d}_N = N^{-1}\sum_{i=1, \dots, N} \hat{d}_i$  denote the mean estimate of  $d$  across the  $N = 500$  replications. By a central limit theorem argument,  $N^{1/2}(\bar{d}_N - E(\hat{d})) \rightarrow N(0, \sigma^2(\hat{d}))$ , where  $\sigma^2(\hat{d})$  denotes the variance of  $\hat{d}$ . Thus, if  $\hat{d}$  is unbiased the Monte Carlo standard error is consistently estimated by  $N^{-1/2}$  RMSE.

Table 1  
Finite-sample distributions of the QMLE for  $d$  in FIGARCH(1,  $d$ , 1) models

$$y_t = \mu + \varepsilon_t, \quad \varepsilon_t \sigma_t^{-1} \text{ i.i.d. } t_v(0, 1), \quad \sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + [1 - \beta_1 L - (1 - \phi_1 L)(1 - L)^d] \varepsilon_t^2, \\ t = 1, 2, \dots, T$$

$d$	$\beta_1$	$\phi_1$	$v$	$T$	$(p, d, q)$	$\hat{d}$			$t_{\hat{d}=d}$			
						Bias	RMSE	Std.	0.050	0.100	0.900	0.950
0.75	0.70	0.000	$\infty$	3000	(1, $d$ , 1)	0.014	0.091	0.081	0.058	0.088	0.786	0.858
0.75	0.70	0.000	$\infty$	3000	(1, $d$ , 0)	0.009	0.075	0.062	0.056	0.096	0.792	0.860
0.75	0.70	0.000	$\infty$	1500	(1, $d$ , 0)	-0.003	0.109	0.091	0.074	0.092	0.810	0.852
0.75	0.70	0.000	7	3000	(1, $d$ , 0)	0.005	0.092	0.074	0.058	0.102	0.792	0.846
0.50	0.45	0.000	$\infty$	3000	(1, $d$ , 0)	0.013	0.075	0.065	0.070	0.122	0.882	0.936
0.50	0.45	0.000	$\infty$	1500	(1, $d$ , 0)	0.018	0.123	0.092	0.104	0.168	0.846	0.894
0.00	0.85	0.975	$\infty$	3000	(1, $d$ , 1)	-0.008	0.056	0.049	0.052	0.120	0.898	0.948
1.00	0.85	0.000	$\infty$	3000	(1, $d$ , 0)	-0.005	0.051	0.043	0.050	0.092	0.892	0.948

For all of the simulated models  $\mu = 0.0$  and  $\omega = 0.1$ .  $v = \infty$  corresponds to conditional normality. The orders of the estimated models are indicated in the  $(p, d, q)$  column. The bias, the root mean square error, and the average of the standard error estimate for  $\hat{d}$  over the 500 simulated QMLE estimates are reported in the  $\hat{d}$  columns. The  $t_{\hat{d}=d}$  columns report the empirical rejection frequencies for the indicated nominal significance levels for a robust  $t$ -test of the true null hypothesis.

models based on only  $T = 1,500$  observations, there is an increase in the RMSE and the Std. of approximately  $\sqrt{(3,000/1,500)} = \sqrt{2}$  as expected. This increase in the estimation error uncertainty for the smaller sample size is also evident from Fig. 1, which graphs the  $\hat{d}$  densities for the two FIGARCH(1,  $d$ , 0) models with  $d = 0.75$ ,  $\beta_1 = 0.7$ , and  $T = 3,000$  and 1,500, respectively.<sup>13</sup> Note, that even for  $T = 3,000$ , the distribution of  $\hat{d}$  is slightly skewed to the left. Very interestingly, the estimate of  $d$  from the FIGARCH(1,  $d$ , 1) model for the GARCH(1,1) DGP with  $d = 0.0$ , reported in the second to last row, and the estimate of  $d$  from the FIGARCH(1,  $d$ , 0) model for the IGARCH(1,1) DGP with  $d = 1.0$ , reported in the last row, are the most accurate  $d$  estimates in Table 1. Also, as Fig. 2 clearly illustrates, the simulated finite-sample distributions of  $\hat{d}$  for  $d = 0.0, 0.5$ , and 1.0 have virtually no common support, so that a conventional nested testing procedure for  $d$  should be able to differentiate between the corresponding GARCH, FIGARCH, or IGARCH models.

Although the average robust standard errors systematically underestimate the true sampling variation in  $\hat{d}$ , the numbers in the RMSE and Std. columns are

<sup>13</sup>The smooth densities were calculated by an Epanechnikov kernel,  $\hat{f}(d) = 0.75 \cdot (N \cdot h)^{-1} \cdot \sum_{i=1, N} [1 - ((d - \hat{d}_i)h^{-1})^2] \cdot I(|(d - \hat{d}_i)h^{-1}| \leq 1)$ , with the bandwidth,  $h$ , chosen by formula (3.31) in Silverman (1986). We are grateful to Bo Honoré for sharing his GAUSS computer program used in this kernel estimation.

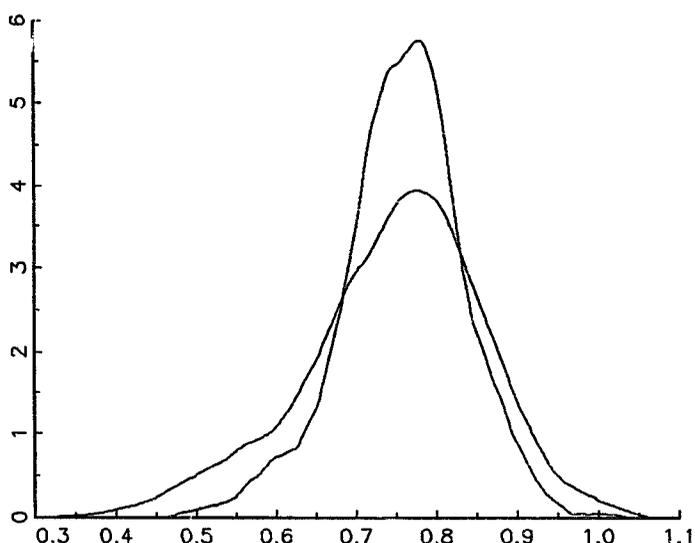


Fig. 1. Finite-sample distributions of  $\hat{d}$ .

The figure graphs the kernel estimates of the simulated small-sample densities of the Quasi Maximum Likelihood Estimates (QMLE) for  $d$ . The parameters for the two different FIGARCH  $(1, d, 0)$  DGP's are  $\mu = 0.0$ ,  $\omega = 0.1$ ,  $d = 0.75$ ,  $\beta_1 = 0.7$ ,  $T = 1,500$ , and  $T = 3,000$ , respectively.

generally close. Of course, in practice any inference regarding the FIGARCH parameters must be based on the actual test statistics computed from the before mentioned finite-sample approximation to (15). The simulated distributions of the resulting  $t$ -statistics for the true null hypotheses regarding the fractional differencing parameter are summarized in the last four columns of the table, which report the empirical rejection frequencies for the indicated nominal significance levels.<sup>14</sup> Even though the QMLE for  $d$  is approximately unbiased, because of the left skewness in the distribution of  $\hat{d}$  evident in Fig. 1, there is a tendency for the corresponding  $t$ -statistic to overreject in the right tail of the distribution when  $d = 0.5$  or  $d = 0.75$ , whereas the left-tail rejection frequencies are always close to the nominal levels. A comparison of the results for the FIGARCH(1,  $d$ , 0) model with  $d = 0.5$  and the two different sample sizes suggests that this skewness in the distribution eventually disappears, but that very large sample sizes may be required. It is interesting to note that the robust covariance matrix estimator seems to do a good job of adjusting for conditional nonnormality in the errors, as the results for the FIGARCH(1,  $d$ , 0)

<sup>14</sup> Let  $\bar{p}_N \equiv N^{-1} \sum_{i=1, N} p_i$  denote the simulated rejection frequency, where  $p_i$  equals one if the  $t$ -statistic in the  $i$ th replication exceeds the nominal significance level and zero otherwise. The Monte Carlo sampling error may then be assessed by  $N^{1/2}(\bar{p}_N - p) \rightarrow N(0, p[1 - p])$ .

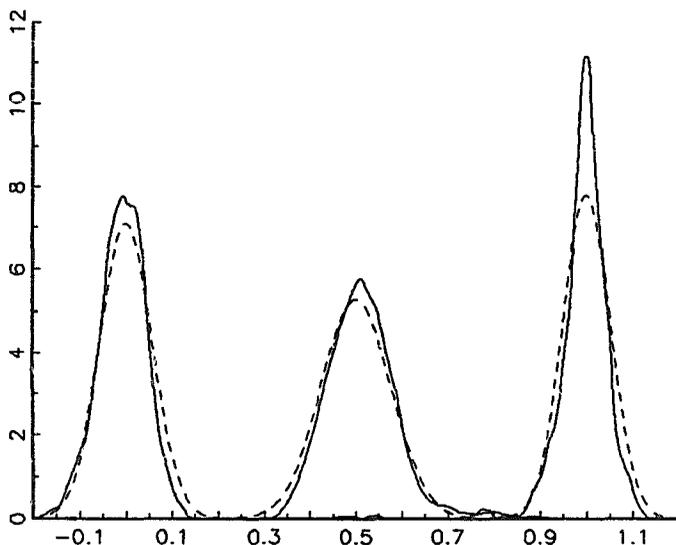


Fig. 2. Finite-sample distributions of  $\hat{d}$ .

The figure graphs the kernel estimates of the simulated small-sample densities of the Quasi Maximum Likelihood Estimates (QMLE) for  $d$ . For all three simulated FIGARCH(1,  $d$ , 1) DGP's  $\mu = 0.0$ ,  $\omega = 0.1$ , and  $T = 3,000$ . The other model parameters are  $d = 0.0$ ,  $\beta_1 = 0.85$ ,  $\phi_1 = 0.975$ ;  $d = 0.5$ ,  $\beta_1 = 0.45$ ,  $\phi_1 = 0.0$ ; and  $d = 1.0$ ,  $\beta_1 = 0.85$ ,  $\phi_1 = 0.0$ , respectively. The figure also graphs the corresponding normal densities with mean equal to the true value of  $d$  and variance equal to the simulated variance of  $\hat{d}$ .

model with conditionally normal and  $t$ -distributed errors are very close. It is also interesting to note, that despite the apparent nonnormal distribution of  $\hat{d}$  for the FIGARCH(1,  $d$ , 0) model with  $d = 1.0$  graphed in Fig. 2, the simulated rejection frequencies for the  $t$ -tests for both  $d = 1.0$  and  $d = 0.0$ , reported in the last two rows of the table, are both extremely well-behaved in either tail of the distribution. This is also apparent from Fig. 3 which plots the distribution of  $t_{d=1}$  from the estimated FIGARCH(1,  $d$ , 0) model with the true value of  $d = 1.0$ . Thus, in terms of a conventional nested testing procedure for GARCH versus FIGARCH, or IGARCH versus FIGARCH, the actual size of the standard  $t$ -test is very reliable for the sample sizes and models analyzed here.

While the discussion of the simulation results above have centered on the distribution of the fractional differencing parameter, equally satisfactory results for the approximate QMLE are available for the other FIGARCH parameters. In particular, Table 2 reports the same set of summary statistics for the estimation of  $\beta_1$  from each of the eight different DGP's.<sup>15</sup> We shall not discuss these

<sup>15</sup>The simulation results for the other model parameters are available upon request.

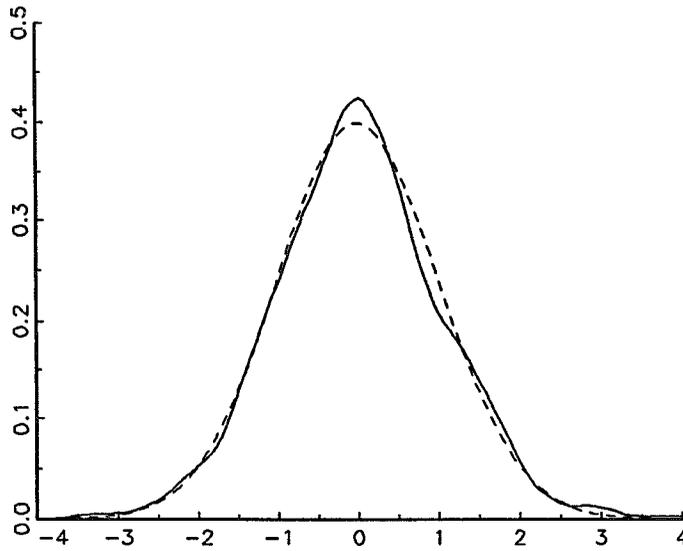


Fig. 3. Finite-sample distribution of the  $t$ -test for  $\hat{d} = 1$ .

The figure graphs the kernel estimate of the simulated small-sample density of the robust  $t$ -test for  $d = 1$ . The parameters for the simulated FIGARCH(1,  $d$ , 0) DGP are  $\mu = 0.0$ ,  $\omega = 0.1$ ,  $d = 1.0$ ,  $\beta_1 = 0.85$ , and  $T = 3,000$ . The figure also graphs the density for the standard normal distribution.

Table 2

Finite-sample distributions of the QMLE for  $\beta_1$  in FIGARCH(1,  $d$ , 1) models

$$y_t = \mu + \varepsilon_t, \quad \varepsilon_t \sigma_t^{-1} \text{ i.i.d. } t_v(0, 1), \quad \sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + [1 - \beta_1 L - (1 - \phi_1 L)(1 - L)^d] \varepsilon_t^2, \\ t = 1, 2, \dots, T$$

$d$	$\beta_1$	$\phi_1$	$v$	$T$	$(p, d, q)$	$\hat{\beta}_1$		$t\hat{\beta} = \beta$				
						Bias	RMSE	Std.	0.050	0.100	0.900	0.950
0.75	0.70	0.000	$\infty$	3000	(1, $d$ , 1)	0.008	0.066	0.061	0.040	0.072	0.810	0.854
0.75	0.70	0.000	$\infty$	3000	(1, $d$ , 0)	0.003	0.065	0.056	0.056	0.088	0.790	0.854
0.75	0.70	0.000	$\infty$	1500	(1, $d$ , 0)	-0.012	0.097	0.082	0.084	0.142	0.792	0.858
0.75	0.70	0.000	7	3000	(1, $d$ , 0)	-0.002	0.085	0.067	0.064	0.108	0.804	0.848
0.50	0.45	0.000	$\infty$	3000	(1, $d$ , 0)	0.011	0.077	0.068	0.076	0.126	0.876	0.924
0.50	0.45	0.000	$\infty$	1500	(1, $d$ , 0)	0.015	0.122	0.094	0.092	0.154	0.830	0.898
0.00	0.85	0.975	$\infty$	3000	(1, $d$ , 1)	-0.008	0.039	0.034	0.056	0.100	0.894	0.942
1.00	0.85	0.000	$\infty$	3000	(1, $d$ , 0)	-0.007	0.037	0.029	0.054	0.120	0.902	0.952

See footnote to Table 1.

results in any detail here. It is worth noting, however, that like  $\hat{d}$ ,  $\hat{\beta}_1$  also appears to be unbiased, but that there is a similar tendency for the  $t$ -statistic for  $\beta_1$  to over-reject in the right tail of the distribution when  $0 < d < 1$ . In summary, however, the simulations indicate that the proposed QMLE procedure performs very well for the sample sizes typically encountered with high-frequency financial data.

#### 4. The occurrence of integrated GARCH

A striking empirical regularity that emerges from numerous studies of high-frequency, say daily, asset pricing data with ARCH-type models, concerns the apparent widespread finding of integrated GARCH behavior. This property has been found in stock returns, exchange rates, commodity prices and interest rates; see Bollerslev, Chou, and Kroner (1992). Yet, unlike  $I(1)$  processes for the mean, there is less theoretical motivation for truly integrated behavior in the conditional variance.

Nelson (1990b, 1992) and Nelson and Foster (1994) have argued that the apparent IGARCH property of high-frequency data may arise as an artefact from empirically approximating continuous time diffusions or semi-martingale processes. In particular, a GARCH(1,1) model with parameters  $\omega^{(h)} = \omega h$ ,  $\alpha_1^{(h)} = \alpha(h/2)^{1/2}$ , and  $\beta_1^{(h)} = 1 - \alpha(h/2)^{1/2} - \phi h$ , where  $\omega \geq 0$ ,  $\alpha > 0$  and  $\phi > 0$  are fixed, provides a consistent approximation to the true continuous time variance process as the sampling interval,  $h$ , goes to zero, for a wide variety of continuous time processes. Note, that in the limit, the sum of the two GARCH(1,1) parameters,  $\phi^{(h)} \equiv \alpha^{(h)} + \beta^{(h)} = 1 - \phi h$ , tends to one, indicative of IGARCH behavior. However, it is unclear for what 'small' sampling interval,  $h$ , the GARCH(1,1) model would actually provide a good approximation. Thus, from a practical perspective, considerable interest remains in determining which parameterization for the discrete time process provides the most parsimonious representation and best forecasting performance for a given fixed sampling interval.

To that end, the IGARCH process has forecasting properties that are radically different to those of the stable GARCH model or the new FIGARCH class of models. As discussed above, the occurrence of a shock to the IGARCH volatility process will persist for an infinite prediction horizon. This extreme behavior of the IGARCH process may reduce its attractiveness for asset pricing purposes, where the IGARCH assumption could make the pricing functions for long-term contracts very sensitive to the initial conditions. This seems contrary to perceived behavior of agents who typically do not frequently and radically change their portfolio compositions; see Bollerslev and Engle (1993) for some further discussion on the effects of persistence of volatility in this context.

Also, as noted by Engle and Mustafa (1992) and Schwert (1990), the IGARCH model is not compatible with the persistence observed after large shocks such as

Table 3  
Finite-sample distributions of the GARCH(1, 1) QMLE for  $\phi_1$  under FIGARCH(1,  $d$ , 1) DGP's

$$y_t = \mu + \varepsilon_t, \quad \varepsilon_t \sigma_t^{-1} \text{ i.i.d. } t_r(0, 1), \quad \sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + [1 - \beta_1 L - (1 - \phi_1 L)(1 - L)^d] \varepsilon_t^2, \\ t = 1, 2, \dots, T$$

$d$	$\beta_1$	$\phi_1$	$\nu$	$T$	$\hat{\phi}$			$t_{\hat{\phi} = 1}$			
					Mean	RMSE	Std.	0.050	0.100	0.900	0.950
0.75	0.70	0.000	$\infty$	3000	0.996	0.007	0.006	0.172	0.288	0.926	0.960
0.75	0.70	0.000	$\infty$	1500	0.993	0.010	0.009	0.196	0.342	0.964	0.976
0.75	0.70	0.000	7	3000	0.996	0.011	0.008	0.160	0.268	0.916	0.958
0.50	0.45	0.000	$\infty$	3000	0.983	0.011	0.007	0.724	0.814	0.998	0.998
0.50	0.45	0.000	$\infty$	1500	0.976	0.016	0.012	0.640	0.764	0.998	1.000
0.00	0.85	0.975	$\infty$	3000	0.973	0.009	0.009	0.988	0.998	1.000	1.000
0.00	0.85	1.000	$\infty$	3000	0.998	0.006	0.006	0.068	0.138	0.948	0.974

See footnote to Table 1. For all of the simulated FIGARCH(1,  $d$ , 1) models  $\mu = 0.0$  and  $\omega = 0.1$ . The remaining model parameters for the true DGP are given in columns one through five. The mean column reports the mean of the GARCH(1, 1) QMLE for  $\hat{\phi}_1$  across the 500 Monte Carlo replications. The corresponding root mean square error and the average of the QMLE standard error estimate for  $\hat{\phi}_1$  are given in the RMSE and Std. columns, respectively. The  $t_{\hat{\phi} = 1}$  columns give the empirical rejection frequencies for the indicated nominal significance levels for the null of IGARCH(1, 1); i.e.,  $\phi_1 = 1.0$  and  $d = 0.0$ , or equivalently,  $\phi_1 = 0.0$  and  $d = 1.0$ .

the Crash of October 1987. Along these lines Cai (1994), Hamilton and Susmel (1994), and Lamoureux and Lastrapes (1990) have argued that deterministic or stochastic regime shifts in the unconditional variance may easily be mistaken for IGARCH-type behavior.<sup>16</sup>

A further reason to doubt the empirical reasonableness of IGARCH models relates to issues of temporal aggregation. As shown by Drost and Nijman (1993), a data generating process of IGARCH at high frequencies would also imply a properly defined weak IGARCH model at low frequencies of observation. However, this theoretical result seems at odds with reported empirical findings for most asset categories. Also, while studies of daily asset returns data have almost uniformly found IGARCH behavior, studies with even higher-frequency data over shorter time spans have often uncovered less persistence; see, e.g., Baillie and Bollerslev (1991) for an analysis of hourly exchange rates over a 2½-month period.<sup>17</sup>

<sup>16</sup> This is closely related to the arguments for I(1) processes for the mean in Perron (1989).

<sup>17</sup> This phenomenon may be directly analogous to the results reported by Shiller and Perron (1985) for the mean, where the consequence of having relatively short spans of high-frequency data is to reduce the power of detecting low-frequency, unit root behavior. The pronounced intraday seasonal patterns observed in most financial markets also serve to obscure the estimated degree of volatility persistence; see, e.g., Andersen and Bollerslev (1996).

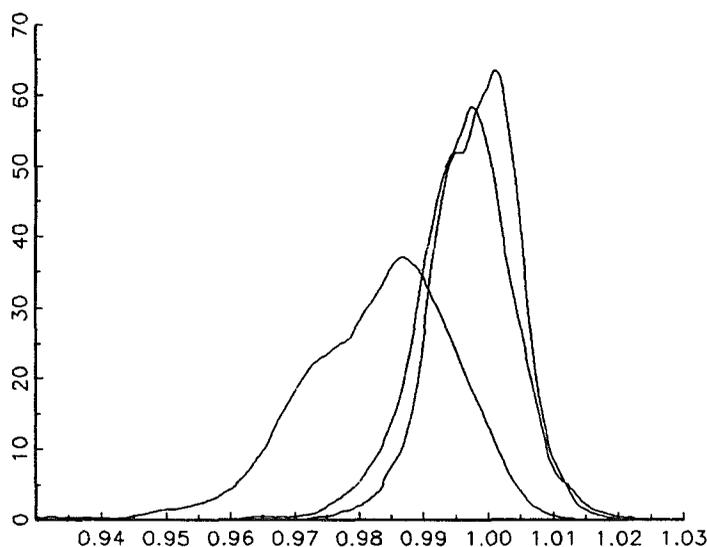


Fig. 4. Finite-sample distributions of  $\hat{\phi}_1$ .

The figure graphs the kernel estimates of the simulated small-sample density of the Quasi Maximum Likelihood Estimates (QMLE) for  $\phi_1$  based on a GARCH(1,1) parameterization. For all three simulated FIGARCH(1,  $d$ , 0) DGP's  $\mu = 0.0$ ,  $\omega = 0.1$ , and  $T = 3,000$ . The other model parameters are  $d = 0.5$  and  $\beta_1 = 0.45$ ,  $d = 0.75$  and  $\beta_1 = 0.7$ , and  $d = 1.0$  and  $\beta_1 = 0.85$ , respectively.

To shed further light on the apparent widespread empirical IGARCH behavior in view of the above mentioned criticisms, the possibility that IGARCH may be an artefact of a mean-reverting long-memory FIGARCH data generating process was therefore investigated by estimating GARCH(1,1) models for data generated by the same FIGARCH(1,  $d$ , 0) models analyzed in Section 3. Additionally, a covariance-stationary GARCH(1,1) model with  $\phi_1 = 0.975$  and  $\beta_1 = 0.85$ , and an IGARCH(1,1) model with  $\beta_1 = 0.85$  were also simulated. Table 3 shows that across all the different FIGARCH(1,  $d$ , 0) designs, the mean of the estimated 'persistence' parameter in the misspecified GARCH(1,1) model,  $\hat{\phi}_1$ , averages between 0.976 to 0.996; all very close to unity as in the IGARCH(1,1) model. The finite sample distributions of  $\hat{\phi}_1$  are further illustrated in Fig. 4, which plots the densities for the three FIGARCH(1,  $d$ , 0) DGP's with  $T = 3,000$ , and  $d = 0.5$ , 0.75, and 1.0. When the true DGP is  $d = 1.0$ , and  $\beta_1 = 0.85$ , or equivalently IGARCH(1,1) with  $\beta_1 = 0.85$ ,  $\phi_1 = 1.0$ , and  $d = 0.0$ , the finite-sample distribution of  $\hat{\phi}_1$  is heavily concentrated around unity as expected. Interestingly, however, the distribution of  $\hat{\phi}_1$  for the FIGARCH(1,  $d$ , 0) DGP with  $d = 0.75$  and  $\beta_1 = 0.70$  is very similar to the results for the IGARCH(1,1) model. Even though the distribution for the FIGARCH(1,  $d$ , 0) DGP with  $d = 0.50$  and  $\beta_1 = 0.45$  is more dispersed, the mode of  $\hat{\phi}_1$  from the GARCH(1,1) model still equals 0.983. Thus, an investigator confined to the

conventional GARCH paradigm might falsely conclude that the appropriate model is an IGARCH process. This is further underscored by the distribution of the  $t$ -tests for  $\phi_1 = 1.0$  reported in the last four columns of Table 3. For instance, for the FIGARCH(1,  $d$ , 0) model with  $d = 0.75$  and  $T = 3,000$ , a one-sided nominal 5 percent  $t$ -tests for  $\phi_1 < 1$  would only reject 17.2 percent of the times. With conditional  $t$ -distributed errors the power drops even further to 16.0 percent. Interestingly, the rejection frequencies are somewhat higher for  $T = 1,500$  observations. At the same time, the results in the last row of the table indicate that the empirical size of the  $t$ -test for  $\phi_1 = 1.0$  is quite good. Also, from the second to last row the test has considerable power against the covariance-stationary GARCH(1,1) model with  $\phi_1 = 0.975$ ; for additional evidence on the small-sample behavior of the GARCH(1,1) and IGARCH(1,1) parameter estimates see Lumsdaine (1995).

Hence, data generated from a process exhibiting long-memory FIGARCH volatility may be easily mistaken for IGARCH behavior. This is analogous to the results of Diebold and Rudebusch (1991), who demonstrate that the conventional unit root testing procedures for the mean have low power for distinguishing against fractional white noise alternatives. For both the mean and the variance, being confined to only considering the extreme cases of  $I(0)$  and  $I(1)$ , or stable GARCH and IGARCH processes, can be very misleading when long-memory, but eventual mean-reverting processes are generating the observed data.

## 5. A long-memory model for the Deutschmark–U.S. dollar volatility

A large number of studies have documented the persistence of volatility in nominal exchange rates; see, e.g., Baillie and Bollerslev (1989), Bollerslev (1987), Hsieh (1989), and McCurdy and Morgan (1988). Using daily data many of these studies have concluded that the volatility process is very persistent and appears to be well approximated by an IGARCH process. To shed further light on this issue, this section presents estimation results for a long time series of daily Deutschmark–U.S. dollar spot exchange rate from March 13, 1979 through December 30, 1992, for a total of 3,454 observations. It is immediately obvious from the plot in Fig. 5, that the raw spot exchange rate series,  $s_t$ , is nonstationary. Following standard practice, we shall therefore concentrate on modeling the daily nominal percentage returns; i.e.,  $y_t \equiv 100 \cdot \log(s_t/s_{t-1})$  for  $t = 1, 2, \dots, 3,453$ .<sup>18</sup> Even though the return series is approximately uncorrelated

<sup>18</sup> This is consistent with the outcome of numerous unit root tests reported in the existing literature, although Cheung (1993) and Tschernig (1995) both report weak evidence for the existence of long memory type behavior in the mean of U.S. dollar exchange rates. The arguments for a unit root in  $\log(s_t)$  is not just statistical, however. Compelling market efficiency reasons dictate that the spot rates should be  $I(1)$ , so that the corresponding returns on open positions in the foreign exchange market, i.e., the  $y_t$  series, would be  $I(0)$ .



Fig. 5. Deutschmark–U.S. dollar spot rate.

The figure graphs the daily nominal Deutschmark–U.S. dollar spot exchange rate from March 13, 1979 though December 1992.

through time, the plot in Fig. 6 clearly indicates the occurrence of tranquil and volatile periods. This is also borne out by the Ljung and Box (1978) portmanteau test for up to 20th-order serial correlation in the returns,  $Q(20)$ , and the squared returns,  $Q^2(20)$ , reported in the first column of Table 4. Judged by the same portmanteau tests for the standardized residuals,  $\hat{\varepsilon}_t \hat{\sigma}_t^{-1}$ , from the QMLE for the standard GARCH(1,1) model reported in the second column of the table, it appears that this simple model does a very good job of tracking the short-run volatility dependencies. Also, entirely consistent with previous findings, the estimated autoregressive parameter in the conditional variance equation,  $\hat{\phi}_1$ , is very close to unity, suggestive of IGARCH behavior. Indeed, the estimates for the restricted IGARCH(1,1) model in the third column of the table are very similar to the results for the GARCH(1,1) model. Hence, any researcher operating within the conventional paradigm might realistically conclude that the IGARCH model provides a satisfactory description of the volatility process. However, from the FIGARCH(1,  $d$ , 1) model estimates, reported in the fourth column, it appears that the long-run dynamics are better modeled by the fractional differencing parameter.<sup>19</sup> Whereas  $\hat{d}$  is between zero and one, the

<sup>19</sup> It is fairly easy to show that  $0 < \omega$ ,  $0 \leq d \leq 1 - 2\phi_1$  and  $0 \leq \beta_1 \leq \phi_1 + d$  are sufficient to ensure that the conditional variance of the FIGARCH(1,  $d$ , 1) model is positive almost surely for all  $t$ . These conditions are trivially satisfied by the QMLE in Table 4.

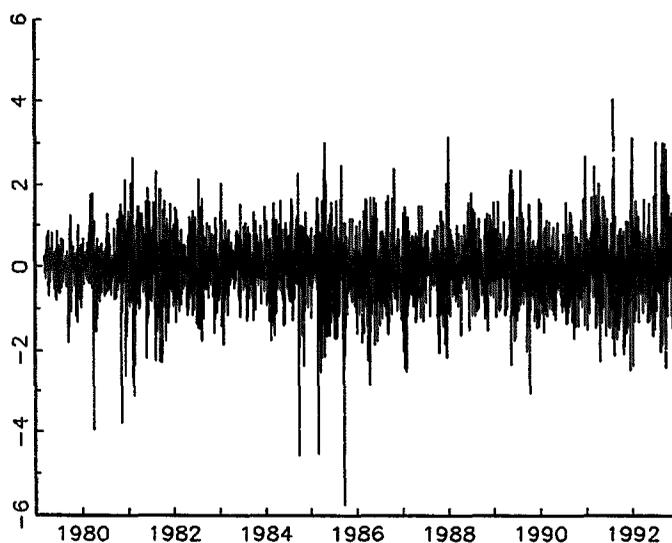


Fig. 6. Deutschmark–U.S. dollar returns.

The figure graphs the daily percentage returns for the Deutschmark–U.S. dollar spot exchange rate from March 14, 1979 through December 1992.

estimate of the autoregressive conditional variance parameter,  $\phi_1$ , is no longer significant at conventional levels. The final column therefore gives the results for the preferred FIGARCH(1,  $d$ , 0) specification. A one-sided  $t$ -test for  $d = 1.0$  against  $d < 1.0$  also clearly rejects the IGARCH null hypothesis against this FIGARCH model.<sup>20</sup>

Thus, purely from the perspective of searching for a model that best describes the volatility in the exchange rate return series, the FIGARCH(1,  $d$ , 0) model in column five of Table 4 appears the most satisfactory representation. However, the true importance of the new FIGARCH specification only becomes fully apparent when considering and analyzing the effect and propagation of volatility shocks. To illustrate, Fig. 7 plots the cumulative impulse response weights for the influence of an innovation on the optimal forecasts of the future conditional variances for the FIGARCH(1,  $d$ , 0) model estimates in Table 4; i.e.,  $\lambda_k(\hat{\theta}_T)$  for  $k = 1, 2, \dots$  as discussed in Section 2. To get an idea about the associated parameter estimation error uncertainty, the figure also graphs the asymptotic

<sup>20</sup> Interestingly, the estimate of  $d$  from this preferred specification is in close accordance with the recent results in Harvey (1993), who on estimating a simple fractional white noise stochastic volatility process for the daily Deutschmark–U.S. dollar exchange rate over the much shorter sample period from October 1, 1981 through June 28, 1985, reports  $\hat{d} = 0.868$ .

Table 4

FIGARCH(1,  $d$ , 1) models for the Deutschmark–U.S. dollar exchange rate

$$y_t \equiv 100 \cdot \log(s_t/s_{t-1}) = \mu + \varepsilon_t, \quad \varepsilon_t \sigma_t^{-1} \text{ i.i.d. } N(0, 1), \quad \sigma_t^2 = \omega + \beta_1 \sigma_{t-1}^2 + [1 - \beta_1 L - (1 - \phi_1 L)(1 - L)^d] \varepsilon_t^2, \quad t = 1, 2, \dots, 3453$$

$(p, d, q)$	(0, 0, 0)	(1, 0, 1)	(1, 1, 0)	(1, $d$ , 1)	(1, $d$ , 0)
$\mu$	-0.04 (0.013)	-0.003 (0.011)	-0.002 (0.011)	-0.004 (0.011)	-0.004 (0.011)
$\omega$	0.574 (0.014)	0.016 (0.004)	0.010 (0.003)	0.018 (0.006)	0.017 (0.005)
$\beta_1$	—	0.867 (0.015)	0.872 (0.015)	0.675 (0.104)	0.762 (0.061)
$\phi_1$	—	0.980 (0.010)	—	0.094 (0.075)	—
$d$	—	—	1.000 (—)	0.652 (0.160)	0.823 (0.077)
$b_3$	-0.26	0.11	-0.11	-0.10	-0.10
$b_4$	6.12	4.76	4.76	4.77	4.72
$Q(20)$	24.91	34.17	36.69	35.52	35.68
$Q^2(20)$	262.50	19.30	19.77	16.26	19.96

The table reports the Quasi Maximum Likelihood Estimates (QMLE) for various FIGARCH ( $p, d, q$ ) models for the percentage return on the daily Deutschmark–U.S. dollar spot exchange rate from March 14, 1979 through December 30, 1992, for a total of 3,453 observations. The QMLE are calculated under the assumption of conditional normality. Robust standard errors are reported in parentheses. The sample skewness and kurtosis for the standardized residuals,  $\hat{\varepsilon}_t \hat{\sigma}_t^{-1}$ , are denoted by  $b_3$  and  $b_4$ .  $Q(20)$  and  $Q^2(20)$  refer to the Ljung–Box portmanteau tests for up to 20th-order serial correlation in the standardized and the squared standardized residuals, respectively.

5 and 95 percent confidence bands. These intervals were calculated from the finite-sample approximation to the asymptotic distribution for  $\hat{\theta}_T$  in Eq. (15) and a standard first-order Taylor series expansion of  $\lambda_k(\hat{\theta}_T)$  around  $\theta_0$ ; i.e.,

$$\lambda_k(\hat{\theta}_T) \sim N(\lambda_k(\theta_0), T^{-1} r_k(\hat{\theta}_T)' A(\hat{\theta}_T)^{-1} B(\hat{\theta}_T) A(\hat{\theta}_T)^{-1} r_k(\hat{\theta}_T)),$$

where  $r_k(\theta) = \partial \lambda_k(\theta) / \partial \theta$ . From the figure, the influence of past shocks are quite persistent. Even for up to 160 lags, or more than seven months, the confidence bands do not include zero. The contrast to the conventional volatility specifications is clearly seen from Fig. 8, which plots the cumulative impulse response weights from the estimated GARCH(1, 1) and IGARCH(1, 1) models together with the FIGARCH(1,  $d$ , 0) responses. Although the estimated GARCH(1, 1) model with  $\hat{\phi}_1 = 0.980$  is virtually indistinguishable from the IGARCH(1, 1) model in terms of the statistics in Table 4, the interpretation of its long-run characteristics are entirely different. While the IGARCH model imposes complete persistence in its impulse response weights, the covariance-stationary

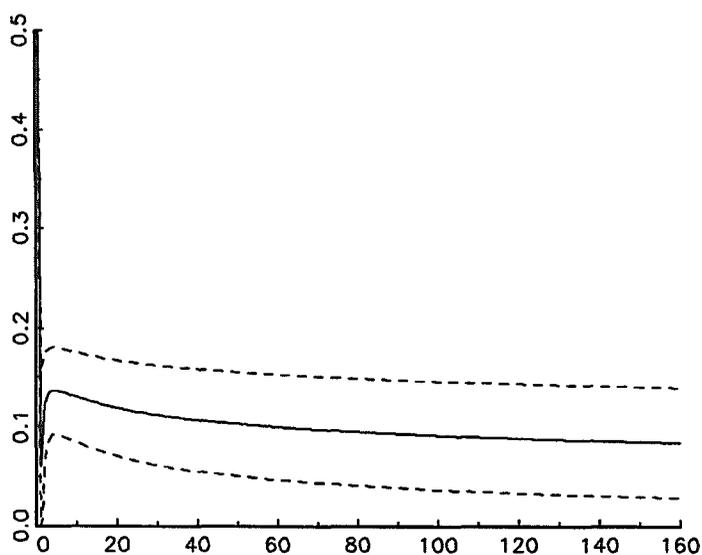


Fig. 7. Cumulative impulse response function.

The figure graphs the cumulative impulse response function for the conditional variance for the estimated FIGARCH(1,  $d$ , 0) model for the daily Deutschmark-U.S. dollar returns. The dashed lines give the 5 and 95 percent confidence bands for the impulse response weights, taking into account the model parameter estimation error uncertainty.

GARCH model implies relatively rapid exponential decay. The preferred FIGARCH model, however, imposes a more realistic slow hyperbolic decay so that the effect of a volatility shock is very persistent, but eventually mean-reverting.

## 6. Conclusion

This paper has proposed the new class of Fractionally Integrated Generalized AutoRegressive Conditionally Heteroskedastic, or FIGARCH, processes. This new model has many attractive features that seem consistent with recently documented long-run dependencies in absolute and squared asset returns. Although mean-reverting, shocks to the conditional variance will die out at a slow hyperbolic rate of decay determined by a fractional differencing parameter, while the short-run dynamics are modeled by the conventional GARCH parameters. The QMLE of the FIGARCH model parameters are argued to be  $T^{1/2}$ -consistent, with a limiting normal distribution that provides very good finite-sample approximations for the sample sizes typically encountered with financial data. Estimation and subsequent analysis and inference are

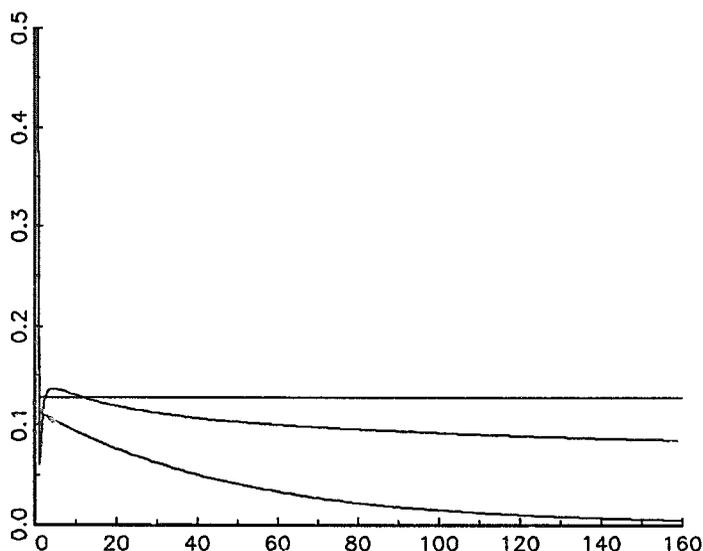


Fig. 8. Cumulative impulse response functions.

The figure graphs the cumulative impulse response functions for the conditional variances for the estimated GARCH(1, 1), FIGARCH(1,  $d$ , 0), and IGARCH(1,1) models for the daily Deutschmark–U.S. dollar returns.

correspondingly straightforward. The application of the model to a set of daily exchange rates is strongly suggestive of long-range volatility dependence.

The FIGARCH model developed here may obviously be extended directly to other parametric ARCH formulations, including the asymmetric EGARCH model developed by Nelson (1991) for modeling stock return volatility and the permanent-transitory components model recently proposed by Engle and Lee (1993). The empirical relevance of this approach has already been documented by Bollerslev and Mikkelsen (1996), who report FIEGARCH estimates for the daily return on the Standard and Poor's 500 composite stock index with daily data from 1953 through 1990. Interestingly, the estimated fractional differencing parameter equals 0.633, with an asymptotic standard error of 0.063, indicating highly significant long-memory components in aggregate U.S. stock market volatility. The evidence in Bollerslev and Mikkelsen (1996) also suggests that for simulated long-term synthetic options the differences between the implied EGARCH and FIEGARCH based Hull and White (1987) options prices may be quite substantial. Work is currently under way on trying to extend these empirical results to other asset categories, and the pricing of actual long-term options contracts.

The results developed in this paper have concentrated solely on univariate models. Yet, for most modern asset pricing theories and optimal asset allocation

decisions, a multivariate modeling framework is necessarily called for. Thus, extending the present ideas to a multivariate setting, allowing for the possibility of co-persistence in variance as discussed by Bollerslev and Engle (1993), would also be important.

Recent findings in Baillie and Bollerslev (1994) suggest that the forward premium in the foreign exchange market,  $s_t - f_t$ , where  $s_t$  and  $f_t$  denote the logarithmic spot rate and the corresponding forward rate for delivery  $l$  periods from now, is best characterized as an  $I(d)$  process. However, by a standard asset pricing argument the difference between the expected future spot rate and the current forward rate,  $E_t(s_{t+l}) - f_t$ , is directly related to the conditional variance of the spot rate; see, e.g., Hodrick (1987). Thus, the long-term dependence in the conditional variances of daily spot exchange rates documented here might therefore be directly related to the apparent long-memory in the forward premium, which in turn might help to explain the systematic rejections of the unbiasedness hypothesis as an artefact due to the unbalanced regression of  $s_{t+l} - s_t$  on  $f_t - s_t$ .

Following Granger (1980) and Haubrich and Lo (1992), who demonstrate how fractional integration in the mean can occur through the aggregation of  $I(0)$  time series, the overall volatility for any speculative asset may naturally be interpreted as the result of the markets aggregation of several different autocorrelated 'news' arrival processes, thus providing a possible reason behind the findings of long-memory characteristics in financial market volatility. A more formal analysis of this idea could proceed along the lines of Nijman and Sentana (1996). We leave further investigations of all these important topics for future research.

## References

- Adenstedt, R., 1974, On large-sample estimation for the mean of a stationary random sequence, *Annals of Statistics* 2, 1095–1107.
- Andersen, T.G., 1994, Stochastic autoregressive volatility: A framework for volatility modeling, *Mathematical Finance* 4, 75–102.
- Andersen, T.G. and T. Bollerslev, 1996, Intraday seasonality and volatility persistence in financial markets, *Journal of Empirical Finance*, forthcoming.
- Baillie, R.T., 1996, Long memory processes and fractional integration in econometrics, *Journal of Econometrics*, forthcoming.
- Baillie, R.T. and T. Bollerslev, 1989, The message in daily exchange rates: A conditional variance tale, *Journal of Business and Economic Statistics* 7, 297–305.
- Baillie, R.T. and T. Bollerslev, 1991, Intra day and inter market volatility in foreign exchange rates, *Review of Economic Studies* 58, 565–585.
- Baillie, R.T. and T. Bollerslev, 1994, The long-memory of the forward premium, *Journal of International Money and Finance* 13, 565–571.
- Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307–327.

- Bollerslev, T., 1987, A conditional heteroskedastic time series model for speculative prices and rates of return, *Review of Economics and Statistics* 69, 542–547.
- Bollerslev, T. and R.F. Engle, 1993, Common persistence in conditional variances, *Econometrica* 61, 167–186.
- Bollerslev, T. and H.O. Mikkelsen, 1996, Modeling and pricing long-memory in stock market volatility, *Journal of Econometrics*, forthcoming.
- Bollerslev, T. and J.M. Wooldridge, 1992, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, *Econometric Reviews* 11, 143–172.
- Bollerslev, T., R.Y. Chou, and K.F. Kroner, 1992, ARCH modeling in finance: A review of the theory and empirical evidence, *Journal of Econometrics* 52, 5–59.
- Bollerslev, T., R.F. Engle, and D.B. Nelson, 1994, ARCH models, in: R.F. Engle and D. McFadden, eds., *Handbook of econometrics*, Vol. 4 (Elsevier Science B.V., Amsterdam).
- Bougerol, P. and N. Picard, 1992, Stationarity of GARCH processes and of some nonnegative time series, *Journal of Econometrics* 52, 115–128.
- Cai, J., 1994, A Markov model of unconditional variance in ARCH, *Journal of Business and Economic Statistics* 12, 309–316.
- Cheung, Y.W., 1993, Long memory in foreign exchange rates, *Journal of Business and Economic Statistics* 11, 93–101.
- Cheung, Y.W. and F.X. Diebold, 1994, On maximum likelihood estimation of the degree of fractional integration when the mean is unknown, *Journal of Econometrics* 62, 301–316.
- Chung, C.F. and R.T. Baillie, 1993, Small sample bias in conditional sum of squares estimators of fractionally integrated ARMA models, *Empirical Economics* 18, 791–806.
- Dacorogna, M.M., U.A. Miller, R.J. Nagler, R.B. Olsen, and O.V. Pictet, 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., 1989, Efficient parameter estimation for self-similar processes, *Annals of Statistics* 17, 1749–1766.
- deLima, P., F.J. Breidt, and N. Crato, 1994, Modeling long-memory stochastic volatility, Manuscript (Department of Economics, Johns Hopkins University, Baltimore, MD).
- Diebold, F.X. and G.D. Rudebusch, 1991, On the power of Dickey–Fuller tests against fractional alternatives, *Economics Letters* 35, 155–160.
- Diebold, F.X. and T. Schuermann, 1996, Exact maximum likelihood estimation of observation driven econometric models, in: R.S. Mariano, M. Waecks, and T. Schuermann, eds., *Simulation based inference in econometrics: Methods and applications* (Cambridge University Press, Cambridge) forthcoming.
- Diebold, F.X., S. Husted, and M. Rush, 1991, Real exchange rates under the gold standard, *Journal of Political Economy* 99, 1252–1271.
- Ding, Z., C.W.J. Granger, and R.F. Engle, 1993, A long memory property of stock market returns and a new model, *Journal of Empirical Finance* 1, 83–106.
- Drost, F.C. and T. Nijman, 1993, Temporal aggregation of GARCH processes, *Econometrica* 61, 909–927.
- Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica* 50, 987–1008.
- Engle, R.F. and T. Bollerslev, 1986, Modelling the persistence of conditional variances, *Econometric Reviews* 5, 1–50.
- Engle, R.F. and G.G.J. Lee, 1993, A permanent and transitory component model of stock return volatility, Manuscript (Department of Economics, University of California, San Diego, CA).
- Engle, R.F. and C. Mustafa, 1992, Implied ARCH models from options prices, *Journal of Econometrics* 52, 289–311.
- Gallant, A.R., P.E. Rossi, and G. Tauchen, 1993, Nonlinear dynamic structures, *Econometrica* 61, 871–907.

- Granger, C.W.J., 1980, Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics* 14, 227–238.
- Granger, C.W.J., 1981, Some properties of time series data and their use in econometric model specification, *Journal of Econometrics* 16, 121–130.
- Granger, C.W.J. and R. Joyeux, 1980, An introduction to long memory time series models and fractional differencing, *Journal of Time Series Analysis* 1, 15–39.
- Hamilton, J.D. and R. Susmel, 1994, Autoregressive conditional heteroskedasticity and changes in regime, *Journal of Econometrics* 64, 307–333.
- Harvey, A.C., 1993, Long memory in stochastic volatility, Manuscript (London School of Economics, London).
- Harvey, A.C. and N. Shephard, 1993, Estimation and testing of stochastic variance models, Manuscript (London School of Economics, London).
- Harvey, A.C., E. Ruiz, and N. Shephard, 1994, Multivariate stochastic variance models, *Review of Economic Studies* 61, 247–264.
- Haubrich, J.G. and A.W. Lo, 1992, The sources and nature of long-term dependence in the business cycle, Manuscript (Department of Finance, MIT, Cambridge, MA).
- Hodrick, R.J., 1987, The empirical evidence on the efficiency of forward and futures foreign exchange markets (Harwood Academic Publishers, Zürich).
- Hosking, J.R.M., 1981, Fractional differencing, *Biometrika* 68, 165–176.
- Hsieh, D.A., 1989, Modeling heteroskedasticity in daily foreign exchange rates, *Journal of Business and Economic Statistics* 7, 307–317.
- Hull, J. and A. White, 1987, The pricing of options on assets with stochastic volatilities, *Journal of Finance* 42, 381–400.
- Hurst, H., 1951, Long term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers* 116, 770–799.
- Jacquier, E., N.G. Polson, and P.E. Rossi, 1994, A Bayesian analysis of stochastic volatility models, *Journal of Business and Economic Statistics* 12, 371–389.
- Lamoureux, C.G. and W.D. Lastrapes, 1990, Persistence in variance, structural change and the GARCH model, *Journal of Business and Economic Statistics* 8, 225–234.
- Lee, S.W. and B.E. Hansen, 1994, Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator, *Econometric Theory* 10, 29–52.
- Ljung, G.M. and G.E.P. Box, 1978, On a measure of lack of fit in time series models, *Biometrika* 65, 297–303.
- Lo, A.W., 1991, Long term memory in stock market prices, *Econometrica* 59, 1279–1313.
- Lumsdaine, R.L., 1995, Finite sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: A Monte Carlo investigation, *Journal of Business and Economic Statistics* 13, 1–10.
- 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*, forthcoming.
- Mandelbrot, B.B., 1963, The variation of certain speculative prices, *Journal of Business* 36, 394–419.
- Mandelbrot, B.B. and J.W. Van Ness, 1968, Fractional Brownian motions, fractional noises and applications, *SIAM Review* 10, 422–437.
- McCurdy, T. and I. Morgan, 1988, Testing the martingale hypothesis in Deutschmark futures with models specifying the form of heteroskedasticity, *Journal of Applied Econometrics* 3, 187–202.
- Nelson, D.B., 1990a, Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory* 6, 318–334.
- Nelson, D.B., 1990b, ARCH models as diffusion approximations, *Journal of Econometrics* 45, 7–38.
- Nelson, D.B., 1991, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica* 59, 347–370.

- Nelson, D.B., 1992, Filtering and forecasting with misspecified ARCH models I: Getting the right variance with the wrong model, *Journal of Econometrics* 52, 61–90.
- Nelson, D.B. and C.Q. Cao, 1992, Inequality constraints in the univariate GARCH model, *Journal of Business and Economic Statistics* 10, 229–235.
- Nelson, D.B. and D.P. Foster, 1994, Asymptotic filtering theory for univariate ARCH models, *Econometrica* 62, 1–41.
- Nijman, T. and E. Sentana, 1996, Marginalization and contemporaneous aggregation in multivariate GARCH processes, *Journal of Econometrics*, forthcoming.
- Perron, P., 1989, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica* 57, 1361–1401.
- Robinson, P.M., 1991, Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression, *Journal of Econometrics* 47, 67–84.
- Robinson, P.M., 1994a, Semiparametric analysis of long-memory time series, *Annals of Statistics* 22, 515–539.
- Robinson, P.M., 1994b, Time series with strong dependence, in: C.A. Sims, ed., *Advances in econometrics – Sixth world congress of the econometric Society* (Cambridge University Press, Cambridge).
- Schwert, G.W., 1990, Stock volatility and the crash of 87, *Review of Financial Studies* 3, 77–102.
- Shiller, R.J. and P. Perron, 1985, Testing the random walk hypothesis: Power versus frequency of observations, *Economics Letters* 18, 381–386.
- Silverman, B.W., 1986. *Density estimation for statistics and data analysis* (Chapman and Hall, New York, NY).
- Sowell, F.B., 1990, Fractional unit root distribution, *Econometrica* 58, 495–506.
- Sowell, F.B., 1992a, Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics* 53, 165–188.
- Sowell, F.B., 1992b, Modeling long-run behavior with the fractional ARIMA model, *Journal of Monetary Economics* 29, 277–302.
- Taylor, S., 1986, *Modelling financial time series* (Wiley, Chichester).
- Taylor, S., 1994, Modelling stochastic volatility, *Mathematical Finance* 4, 183–204.
- Tschernig, R., 1995, Long memory in foreign exchange rates revisited, *Journal of International Financial Markets, Institutions, and Money* 5, 53–78.
- Weiss, A.A., 1986, Asymptotic theory for ARCH models: Estimation and testing, *Econometric Theory* 2, 107–131.
- Yajima, Y., 1988, On estimation of a regression model with long-memory stationary errors, *Annals of Statistics* 16, 791–807.