WORKING PAPER n.03.09

July 2003

Stationarity, Memory and Parameter Estimation of FIGARCH Models

M. Caporin[1]

[1] GRETA, Venice

Stationarity, Memory and Parameter Estimation of FIGARCH Models

Massimiliano Caporin

Università Ca' Foscari Venezia
Dipartimento di Scienze Economiche
Cannaregio 873
30121 Venezia
ITALY

and

GRETA Associati
San Marco 3870
30124 Venezia
ITALY

Proofs should be sent to the personal address of the author: Massimiliano Caporin - Casella Postale n. 8 - 31020 Villorba (TV) - ITALY

Abstract: this work reviews the current literature on the stationarity and on the memory properties of FIGARCH processes showing that a theorem of Zaffaroni (2000) can be applied to prove strict stationarity. Moreover it verifies that consistency of quasi-maximum likelihood estimators cannot be obtained by the Lee and Hansen (1994) approach as claimed by Baillie et al. (1996).

Keywords: FIGARCH, long memory, stationarity, parameter estimation.

This work analyses the basic issues of stationarity and estimator consistency for the long memory GARCH models. After the introduction of the FIGARCH models, in section 1, section 2 recalls the current definitions of memory and stationarity of FIGARCH models, showing that a recent theorem of Zaffaroni (2000) can be used in proving stationarity. In section 3 the focus is on the estimation problem and shows that the consistency result expected by Baillie, Bollerslev and Mikkelsen (1996) cannot be obtained following Lee and Hansen (1994). A Montecarlo approach is therefore suggested to verify the asymptotic properties of the Quasi-Maximum Likelihood estimator. This last section is concluded with a note on a problem found in the estimation procedure, due to the parameter constraints and that can be solved with an appropriate optimisation algorithm.

1 Long memory GARCH models

Assume, unless differently specified, that the following representation holds for the mean process: $y_t = \mu_t + \varepsilon_t$, where, for simplicity, μ_t is set equal to zero, I^{t-1} represents the information set up to time $t-1$ and $\varepsilon_t | I^{t-1} \sim iid\left(0, \sigma_t^2\right)$, that is the residuals, conditionally to the information set up to time $t-1$, are identically distributed with zero mean and time-dependent variances.

Following Engle (1982) and Bollerslev (1986) a GARCH(p,q) model for the variance is specified as: $\varepsilon_t = z_t \sigma_t$, with $E\left[z_t | I^{t-1}\right] = 0$, $Var\left[z_t | I^{t-1}\right] = 1$ and

σ_t is defined by

$$\sigma_t^2 = \omega + \alpha(L)\,\varepsilon_t^2 + \beta(L)\,\sigma_t^2 \tag{1}$$

where L is the lag operator, $\alpha(L) = \sum_{i=1}^{q} \alpha_i L^i$, $\beta(L) = \sum_{j=1}^{p} \beta_j L^j$. The stationarity of this process is achieved when the following restriction is satisfied: $\alpha(1) + \beta(1) < 1$. Defining $\upsilon_t = \varepsilon_t^2 - \sigma_t^2$ this process may be conveniently rewritten as an ARMA(m,p) process

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

with $m = \max\{p, q\}$. Starting from this formulation and allowing for the presence of a unit root in $[1 - \alpha(L) - \beta(L)]$, Engle and Bollerslev (1986) defined the IGARCH(p,q) process:

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

where $\phi(L) = \sum_{i=1}^{m-1} \phi_i L^i$ and it is of order $m - 1$. For a comprehensive survey on GARCH processes, refer to Bollerslev, Engle and Nelson (1994). Even if flexible, and with numerous extensions to include particular characteristics found in the markets, such as asymmetric behavior, switching regime and news impact, the GARCH model is not able to adequately explain the various finding of persistence (or long memory) in the volatility of financial instrument returns. Using a parallel with ARMA and ARFIMA processes Baillie et al. (1996) extended the IGARCH process allowing the integration coefficient (here previously restricted to the usual dichotomy $\{0, 1\}$) to vary in the range $[0, 1]$. The FIGARCH(p,d,m) process is then defined as follow:

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

where $\phi(L) = \sum_{i=1}^{m-1} \phi_i L^i$ is of order $m - 1$. Baillie et al. (1996) claimed that, extending the arguments of Nelson (1990), the FIGARCH(p,d,m) process, even if not weakly stationary was ergodic and strictly stationary. Unfortunately, this is not so easy to verify, this problem will be analysed in a following section. The major feature of the FIGARCH model is connected with the impulse response analysis, which have in this case an hyperbolic decay, typical of long memory models. This mean that the impact of the innovations lies between the exponential decaying, typical of any GARCH, and the infinite persistence, typical of any IGARCH. Davidson (2001) gave some insight on the memory properties of the FIGARCH, pointing out that the degree of persistence of the FIGARCH model operates in the opposite direction of the ARFIMA one: as the d parameter gets closer to zero, the memory of the process increases. This is due to the inverse relation between the integration coefficient and the conditional variance: the memory parameter acts directly on the squared errors, not on the σ_t^2, this particular behavior may also influence the stationarity properties of the process, again Davidson (2001). These observations are strictly related to the impulse response analysis on the effects of a shock on a system driven by a FIGARCH process. In such a system, a shock in time t (υ_t), should be interpreted as the difference between the squared mean-residuals in time t (ε_t^2) and the one-step-ahead forecast of the variance of time t (σ_t^2), made in time $t-1$, $\upsilon_t = \varepsilon_t^2 - \sigma_t^2$. This shock is exactly the innovation in the ARMA representation of the FIGARCH process

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

The shock may be also interpreted as an unexpected volatility variation, or, as the forecast error of the variance (remember that the squared residuals are a proxy for the variance and that the time t variance depends on time $t-1$ information set and may be viewed as a one-step-ahead forecast). Rearranging the FIGARCH equation as in Baillie et al. (1996), expanding then the polynomial in the lag operator, it is easy to see that the coefficients of this polynomial converge to zero at a rate $O\left(j^{-d-1}\right)$: this mean that the memory of the process increases ad d gets closer to 1 (Baillie et al. (1996) obtained the opposite sign claiming the same memory property valid for the ARFIMA).

Analyzing in detail ARFIMA and FIGARCH processes Chung (2001) noted that the claimed parallel between the two was not complete: in the ARFIMA case the long memory operator is applied to the constant but this is not true in the FIGARCH model, moreover in ARFIMA processes $d \in \left(-\frac{1}{2}, \frac{1}{2}\right)$, while in FIGARCH $d \in [0,1]$. In his work Chung suggested an alternative parameterization: starting from (2) we can rewrite the GARCH(p,q) using the value of the unconditional variance $\sigma^2 = \omega / \left(1 - \alpha\left(1\right) - \beta\left(1\right)\right)$ as:

$$[1 - \alpha(L) - \beta(L)] \left(\varepsilon_t^2 - \sigma^2\right) = [1 - \beta(L)] \upsilon_t$$

and from this equation the alternative formulation is straightforward:

$$(1 - L)^d \phi(L) \left(\varepsilon_t^2 - \sigma^2\right) = [1 - \beta(L)] \upsilon_t \qquad (6)$$

In this formulation, however, the interpretation of the parameter σ^2 is not clear: does it represent the unconditional variance as claimed by Chung, or is it simply a constant for the squared observations? This work will not pursues this point, the motivation will become clear in the next section. In the remainder (4) will be referred as FIGARCH I or simply FIGARCH and (6) as FIGARCH II. Exploiting the relation $\upsilon_t = \varepsilon_t^2 - \sigma_t^2$ the two processes can be conveniently rewritten, respectively as:

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

$$\sigma_t^2 = \sigma^2 + \left\{1 - [1 - \beta(L)]^{-1} (1 - L)^d \phi(L)\right\} \left(\varepsilon_t^2 - \sigma^2\right) \qquad (8)$$

sometimes these equations are referred to as the ARCH(∞) representation. In both FIGARCH I and II, parameters have to fulfill some restrictions to ensure positivity of conditional variances. Two different sets of sufficient conditions, valid for the FIGARCH(1,d,1), are available, the first was suggested by Baillie et al. (1996), the second by Chung (2001). As noted by by the last author, both sets are admissible for FIGARCH I and II, however they are not equivalent and there may exist a set of parameters value that satisfy one set of conditions and not the other. Baillie et al. (1996) derived a group of two sets of inequalities

$$\beta - d \leq \phi \leq \frac{2-d}{3} \tag{9}$$

$$d\left(\phi - \frac{1-d}{2}\right) \leq \beta\left(d - \beta + \phi\right)$$

while Chung (2001) express the restriction with a unique set

$$0 \leq \phi \leq \beta \leq d \leq 1 \tag{10}$$

Restrictions for lower order models can be derived directly from the previously presented while for higher order models parameters restrictions cannot be so easily represented and are not included in this work.

2 Stationarity of FIGARCH I and II

Baillie et al. (1996) were quoting Nelson (1990) for proving the stationarity of the FIGARCH model they proposed, but only limited to the case where $p = 1$ and $m = 0$. They claimed that stationarity could be verified with a dominance type argument between the sequence of coefficients of the ARCH(∞) representations of the FIGARCH(1,d,0) and of an appropriately chosen IGARCH(1,1). However it has been noted, independently from Mikosch and Starica (2001), that this "proof" is questionable: how can we bound an hyperbolically decaying sequence of coefficients with an exponential one? This way seems therefore inapplicable. Some insight on the stationarity of this model is due to Davidson (2001) who pointed out that some of the particular relations that hold for FIGARCH may be due to the inverse memory relation. Again referring to Mikosch and Starica (2001) an ambiguous point in the Baillie et al. (1996) work should be stressed: they were defining the FIGARCH model using the ARMA formulation of a general GARCH and then imposing a long memory integration operator $(1-L)^d$. However, this methodology is not completely correct since in this derivation the innovation process v_t depends on the process we are trying to define, therefore we are building a noise sequence that depends on a process defined using that noise sequence! Moreover, the ARMA formulation of a FIGARCH process can be derived once a stationary solution is given. The best way to define a FIGARCH model goes therefore through the use of a much more general approach such as the ARCH(∞) model, as defined by Robinson (1991):

$$\sigma_t^2 = \tau + \sum_{k=1}^{\infty} \psi_k \varepsilon_{t-k}^2 \tag{11}$$
$$\tau > 0 \qquad \psi_k \geq 0$$

The FIGARCH structure can be imposed with an adequate formulation of the coefficients in the infinite ARCH expansion. Given this representation, the stationarity of the FIGARCH process can be proved recalling the stationarity conditions needed by a generic ARCH(∞) process and trying to figure out if the coefficient structure of the FIGARCH can meet these requirements via its ARCH(∞) formulation. The main works in this area are the one of Giraitis, Kokoszka and Leipus (2000), Kazakevicius and Leipus (1999 and 2001), and Zaffaroni (2000).

The first paper, Giraitis et al. (2000) presents a condition for the existence of a stationary solution of an ARCH(∞) process, giving the following theorem:

Theorem 2.1 *(rearranged from Giraitis et al. (2000), page 6, theorem 2.1): given $\varepsilon_t = z_t \sigma_t$ and (11), a stationary solution with finite first moment $E(\varepsilon_t)$ exist if $E(z_t^2) < \infty$ and $E(z_t^2)\sum_{k=1}^{\infty}\psi_k < 1$.*

If the constant $\tau = 0$ unique stationary solution is $\varepsilon_t = 0$.

If $E(z_t^4) < \infty$ and $\left[E(z_t^4)\right]^{1/2}\sum_{k=1}^{\infty}\psi_k < 1$ the stationary solution is unique.

(See the cited paper for the proof).

The stationary solution proposed follow a Volterra series expansion of the form

$$\varepsilon_t^2 = \tau z_t^2 \sum_{l=0}^{\infty} \sum_{h_1 < h_2 < \dots\dots < h_l < l}^{\infty} \psi_{l-h_1}\psi_{h_1-h_2}\dots\dots\psi_{h_{l-1}-h_l} z_l^2 z_{h_1}^2 \dots\dots z_{h_l}^2 \tag{12}$$

This formulation imposes a moment condition on the square of the observations and rules out long memory a priori, in fact for any value of d, we have: $(1-L)^d = \sum_{j=0}^{\infty} \pi_j(d) = 1 + \sum_{j=1}^{\infty} \pi_j(d) = 0$. Therefore, this result is inapplicable in the FIGARCH case. An extension of this methodology is due to Kazakevicius and Leipus (1999 and 2001), who reformulate the existence and stationarity conditions for an ARCH(∞) in a form similar to the one given by Bougerol and Picard (1992) for the GARCH(p,q) model, that is using a top Lyapunov exponent γ, which is defined as follows:

$$\gamma = \lim_{n\to\infty} n^{-1} \log \|A_1 A_2 .....A_n\| \tag{13}$$

where the matrices A_j depend on the parameters and on the structure of the process (see the cited papers for an example). The main result of Kazakevicius and Leipus (1999) is summarized in the following theorem:

Theorem 2.2 *(adapted and rearranged from Kazakevicius and Leipus (1999)): given $\varepsilon_t = z_t\sigma_t$ and (11), if $E\left(\log z_t^2\right)$ is well defined $\gamma \leq 0$ is a necessary condition and $\gamma < 0$ is a sufficient condition for the existence of an ARCH(∞) process.*

If for any strictly stationary sequence $(h_i, i \geq 1)$ of non-negative random variables such that $\sum_{i=1}^{\infty} \psi_i h_i < \infty$ we have

$$\lim_{n\to\infty} \sum_{i=1}^{\infty} \psi_{i+n} h_i = 0 \qquad a.s.$$

and the top Lyapunov exponent γ is negative then (12) is the unique strictly stationary solution.

If $\gamma = 0$ there is no solution at all.

(See the cited paper for the proof).

In this theorem there are no moment conditions on the standardized errors but there is an integrability condition and a limit condition on the coefficients of the ARCH(∞) expansion. This result was then used by Kazakevicius and

Leipus (2000) to assess the existence and stationarity of the FIGARCH model. The main point is in the following theorem:

Theorem 2.3 *(adapted and rearranged from Kazakevicius and Leipus (2001)): If a)* $E\left|\log z_t^2\right| < \infty$ *and b) for some* $k > 1$ *we have* $\sum_{i=1}^{\infty} \psi_i k^i < \infty$ *then the top Lyapunov exponent* γ *is strictly negative, therefore the ARCH(*∞*) exist as well as a stationary solution.*

If assumption b) is not satisfied the Lyapunov exponent is identically equal to zero.

Assumption b) simply require that the coefficients of the ARCH(∞) decay at an exponential rate, when this is not the case, as in FIGARCH, the existence of the ARCH(∞) as well as of a stationary solution become questionable. At the end of this excursus among this first group of papers a point is stressed: the condition for the existence of a stationary solution, imposed through a Lyapunov exponent is a necessary one, therefore a possible less restrictive condition, a sufficient one, may exist. The results of the previous papers did not considered a general approach but came to the FIGARCH analysis only indirectly, imposing conditions that are not fulfilled by FIGARCH processes.

Focus now on the work of Zaffaroni (2000) in order to verify the strict stationarity and ergodicity of FIGARCH(p,d,m). The main result is a corollary to the following theorem of Zaffaroni. Consider the following setup:

given the ARCH(∞) formulation (11), then assuming that $\gamma = E\left(\ln z_t^2\right)$ is well defined (even unbounded) and setting

$$\lambda = \begin{cases} \frac{\gamma}{2} & \gamma < 0 \\ \frac{3(\gamma+\delta)}{2} & \gamma \geq 0 \end{cases}$$

for any constant $\delta > 0$ the following result holds

Theorem 2.4 *(Zaffaroni 2000, Theorem 2, page 6) Let* $\underline{\sum}_M = \sum_{k=1}^{M} \psi_k$, $\overline{\sum}^M = \sum_{k=M+1}^{\infty} \psi_k$, $\kappa = E\left(z_t^2\right)$. *Assume that a)* $0<\tau<\infty$ *and b) for at least* *one* $0 < M < \infty$

$$\max\left[e^{\lambda}\underline{\sum}_{M}+\kappa\overline{\sum}^{M},e^{\lambda}\overline{\sum}^{M}+\kappa\underline{\sum}_{M}\right]<1$$

then for the ARCH(∞) model, for any t, $\tau \leq \sigma_t^2 < \infty$ *a.s. and* σ_t^2 *is strictly stationary and ergodic, with a well-defined nondegenerate probability measure on* $[\tau, \infty)$.

Sufficient conditions to satisfy assumption b) are

$$e^{\lambda}\sum_{k=1}^{\infty}\psi_k < 1 \quad , \quad \kappa\sum_{k=1}^{\infty}\psi_k \leq 1$$

and $\psi_k\psi_j > 0$ *for at least two* $k \neq j$.

Proof. See Zaffaroni (2000). ■

The power of this theorem is that it does not require any moment condition, apart the integrability condition on the squared residuals as in Kazakevicius an Leipus (1999, 2000), moreover it does not require any strict condition on coefficients allowing mild explosive behaviors as well as hyperbolic decaying. Given this result the following corollary can be obtained:

Corollary 2.1 *(adapted from Zaffaroni (2000) Remark 2.2) For 0<d≤1, q≥0, p≥0 and with adequate restrictions on coefficients that ensure positivity of conditional variances, the FIGARCH(p,d,q) I is strictly stationary and ergodic if* $\gamma = E\left(\ln z_t^2\right) < 0$.

Proof. In the ARCH(∞) representation of the FIGARCH model (7) we use the following polynomial to represent the coefficient structure

$$\lambda(L) = 1 - [1-\beta(L)]^{-1}(1-L)^d\phi(L) = \sum_{i=1}^{\infty}\lambda_k L^k$$

if the coefficients satisfy the restrictions that ensure positivity of conditional variances $\lambda_k \geq 0\ \forall k$, and the inequality is strictly positive for at least one $k \geq 0$, we have $0 < \omega/[1-\beta(1)] < \infty$ and $0 < \sigma^2 < \infty$ by the previous theorem. Then noting that

$$(1-L)^d\Big|_{L=1} = \sum_{i=0}^{\infty} \pi_i L^i\Big|_{L=1} = \sum_{i=0}^{\infty} \pi_i = 0 \tag{14}$$

since $\pi_0 = 1$, $\pi_i < 0\ \forall i > 0$ and $\lim_{k\to\infty} \sum_{i=1}^{k} \pi_i = -1$ we can write

$$\lambda(1) = \sum_{i=1}^{\infty} \lambda_k = 1 - [1-\beta(1)]^{-1} \phi(1) \sum_{i=0}^{\infty} \pi_i = 1$$

Using then the fact that $\gamma < 0$ and plugging in the condition of Zaffaroni $\gamma/2 < 0$ we have

$$e^{\gamma/2} \sum_{i=1}^{\infty} \lambda_k = e^{\gamma/2} < 1$$

QED. ■

It can be noted that the FIGARCH(p,d,m) is strictly stationary and ergodic under the assumption of normality of the standardized residuals, this can be easily verified given the strict concavity of the logarithm function and using Jensen inequality. It has to be stressed another point: in GARCH processes it is of common use the assumption that the standardized residuals follow a Student T distribution, this to capture the fact that the tails of the empirical distributions of financial market returns are thicker than in the normal case. Under the assumption of a T-distribution for z_t, in order to prove the strict stationarity and ergodicity of the FIGARCH the condition $E\left(\ln z_t^2\right) < 0$ has to be checked. The square of a T distribution with n degrees of freedom follow an $F(1,n)$ distribution. The evaluation of the expected value was carried out numerically, and the results show that increasing the degrees of freedom, the expected value converge to zero but from above. From this it can be stated that the FIGARCH(p,d,m) is not strictly stationary under the assumption of a T-distribution for the standardized residuals.

Turning now to the analysis of the FIGARCH specification suggested by Chung (2001): in this case, given the structure of the model it can be rewritten as

$$\sigma_t^2 = [1-\beta(L)]^{-1}(1-L)^d \phi(L)\sigma^2 + \left\{1-[1-\beta(L)]^{-1}(1-L)^d \phi(L)\right\}\varepsilon_t^2 =$$
$$= \left\{1-[1-\beta(L)]^{-1}(1-L)^d \phi(L)\right\}\varepsilon_t^2$$

given the relation (14). This violate one of the assumptions of the Zaffaroni's theorem, the presence of a positive constant and the result cannot be applied. In this situation the previous work of Nelson (1990) shows that the only stationary solution when the constant is null is that the conditional variance itself is null. This result was also derived by Giraitis, Kokoszka and Leipus (2000). This was not noted by Chung (2001), but probably could be observed in a well defined Montecarlo experiment, simulating a long time series reducing in such a way the effect of truncation in the ARCH(∞) expansion. Probably this depends on the approximation induced by the truncation which induces a stationary solution as in an ARCH(p) model with very high p.

Again referring to Zaffaroni (2000), a direct application of Theorem 3, page 9, shows, using previous results, that FIGARCH(p,d,m) is not covariance stationary as the IGARCH process.

3 On the consistency of FIGARCH QML estimates

In the estimation of FIGARCH processes the mainly used technique is the Quasi Maximum Likelihood, maximizing with respect to the parameters of interest the following log-likelihood function:

$$Q\left(\theta;\{\varepsilon_t\}_{t=1\ldots T}\right) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left[\log\sigma_t^2 + \varepsilon_t^2/\sigma_t^2\right] \qquad (15)$$

where T is the sample size, σ_t^2 follow a FIGARCH (7), and θ represent the set of parameters. As normal practice in this field $\varepsilon_t/\sigma_t = z_t$ are called the standardized residuals. Baillie et al. (1996) claimed that the result of Lee and Hansen (1994), which shows the consistence and asymptotic normality of the QMLE for IGARCH(1,1) processes, "...extends directly to the FIGARCH(1,d,0) model through a dominance-type argument...", unfortunately this is not correct. The cornerstone of Lee and Hansen's proof is in the possibility of bounding the ratio between the conditional volatility computed with the true parameters and the one computed with the estimated parameters. This is established in their Lemma 4.(4) and 4.(5) (Lee and Hansen, 1994, pag.). This result is then repeatedly used to assess boundness of other ratios between conditional variances and then of their expected value who enter in the proof of the boundness of the likelihood function. The point is that Lemma 4.(4) is no more valid with a FIGARCH(p,d,q) DGP, resulting in an unbounded ratio. In the following the Lee and Hansen proof will be reconsidered to verify this claim, and in order to avoid confusion we will maintain their notation

Lemma 3.1 *(Lee and Hansen Lemma 4. page 34)*
(4) If $\beta \leq \beta_0$, $\frac{{}_{\epsilon}\sigma_t^2}{{}_0\sigma_t^2} \leq K_l \equiv \frac{\omega_u}{\omega_0} + \frac{\alpha_u}{\alpha_0} < \infty$ *a.s.*
(5) If $\beta \geq \beta_0$, $\frac{{}_{\epsilon}\sigma_t^2}{{}_0\sigma_t^2} \leq H_u \equiv \frac{\omega_0}{\omega l} + \frac{\alpha_0}{\alpha l} < \infty$ *a.s.*

where ${}_{\epsilon}\sigma_t^2$ represents the conditional variance with the estimated parameters and ${}_0\sigma_t^2$ the true conditional variance, whose parameters are denoted by ω_0 and α_0. For estimated parameters Lee and Hansen derived a bound that depends on the upper (lower) limits of the compact parameter space ω_u and α_u (ω_l and α_l). Moreover they also splitted the parameter space for the β deriving two bounds depending on the relation between the estimated and the true value. This result is then used in deriving the bounds needed in the verification of the boundness of likelihood function and then for consistence and asymptotic normality. This result is therefore necessary for all the proof, and will be now reconsidered plugging in the FIGARCH(1,d,0) instead of the GARCH(1,1).

Proof. the proof can be obtained both using the standard FIGARCH representation or with the ARCH(∞) formulation. Both formulation are equivalent, here will be presented the first one, the other is available form the author upon request. Start plugging FIGARCH in Lemma 4.(4)

$$\frac{{}_{\epsilon}\sigma_t^2}{{}_0\sigma_t^2} = \frac{\omega + \beta\sigma_{t-1}^2 + (d-\beta)\,\varepsilon_{t-1}^2 + \sum_{i=2}^{\infty}(-\pi_i)\,\varepsilon_{t-i}^2}{\omega_0 + \beta\sigma_{t-1}^2 + (d_0-\beta_0)\,\varepsilon_{t-1}^2 + \sum_{i=2}^{\infty}[-\pi_i\,(d_0)]\,\varepsilon_{t-i}^2}$$

where it has been dropped for convenience the subscripts of the conditional variance. Repeatedly substituting the conditional variance with its expression, back to the past infinity, the following representation is obtaines

$$\frac{{}_{\epsilon}\sigma_t^2}{{}_0\sigma_t^2} = \frac{\frac{\omega}{1-\beta} + (d-\beta)\sum_{i=0}^{\infty}\beta^i\varepsilon_{t-1-i}^2 + \sum_{j=0}^{\infty}\beta^j A_j}{\frac{\omega_0}{1-\beta_0} + (d_0-\beta_0)\sum_{i=0}^{\infty}\beta_0^i\varepsilon_{t-1-i}^2 + \sum_{j=0}^{\infty}\beta_0^j \hat{A}_j}$$

where $A_j = \sum_{i=2}^{\infty}(-\pi_i)\,\varepsilon_{t-i-j}^2$ and $\hat{A}_j = \sum_{i=2}^{\infty}[-\pi_i\,(d_0)]\,\varepsilon_{t-i-j}^2$. Using the fact that all quantities are positive it can be rewritten

$$\frac{{}_{\epsilon}\sigma_t^2}{{}_0\sigma_t^2} \leq \frac{\omega}{1-\beta}\frac{1-\beta_0}{\omega_0} + \frac{d-\beta}{d_0-\beta_0}\sum_{i=0}^{\infty}\left(\frac{\beta}{\beta_0}\right)^i + \sum_{j=0}^{\infty}\left(\frac{\beta}{\beta_0}\right)^j\frac{A_j}{\hat{A}_j}$$

focus now on the last term in the formula

$$\frac{A_j}{\hat{A}_j} = \frac{\sum_{i=2}^{\infty}(-\pi_i)\,\varepsilon_{t-i-j}^2}{\sum_{i=2}^{\infty}[-\pi_i\,(d_0)]\,\varepsilon_{t-i-j}^2} \leq \sum_{i=2}^{\infty}\frac{\pi_i}{\pi_i\,(d_0)}$$

using again the fact that all terms are positive. Noting that for large M the Stirling approximation on the coefficients can be used, therefore

$$\pi_k = \frac{\Gamma\,(k-d)}{\Gamma\,(-d)\,\Gamma\,(k+1)} \sim k^{-d-1} \quad \text{for } k > M$$

then from the last summation

$$\sum_{i=2}^{\infty}\frac{\pi_i}{\pi_i\,(d_0)} \geq \sum_{i=M}^{\infty}\frac{\pi_i}{\pi_i\,(d_0)} \sim \sum_{i=M}^{\infty}\frac{i^{-d-1}}{i^{-d_0-1}} = \sum_{i=M}^{\infty} i^{d_0-d} = \infty$$

Last equality follow from the fact that for $d_0 > d$ we have a succession of term greater than 1, diverging to infinity, while for $d_0 < d$ we have a generalized harmonic succession again diverging to infinity. The approximation may be taken as closed as required, but the important point is that this implies

$$\frac{A_j}{\hat{A}_j} = \infty \Rightarrow \frac{{}_{\epsilon}\sigma_t^2}{{}_0\sigma_t^2} \leq \infty$$

and the ratio cannot be so easily bounded as in Lee and Hansen (1994). This does not ensure that the inequality is always strict nor that an upper bound exist and could be found in this way. This result may be interpreted reasoning on the asymptotic decaying of the coefficients. In the GARCH(1,1) case coefficients decay exponentially to zero, while in the FIGARCH(1,d,0) the convergence is hyperbolic, so a dominance type argument, such as in the claim of Baillie, Bollerslev and Mikkelsen (1996) cannot be used, an hyperbolic decaying (to zero) succession cannot be dominated by an exponentially decaying (again to zero) one, since there always exists a point in which the exponential convergent sequence crosses the hyperbolic convergent one and stays below in the infinity. ■

Given that Lemma 4.(4) of Lee and Hansen (1994) is not consistent with FIGARCH(1,d,0), then following their proof also Lemma 4.(5) and Lemma 6.(1) give non-bounded relations. Therefore the parameter space cannot be splitted as in page 35 and Lemma (5), (7) and (8) are no more valid, breaking down all the proof for the consistence. Moreover, also the proof of asymptotic normality break down because is built on the bounds used to prove the consistence. A similar result can be obtained also for the FIGARCH(1,d,1) with a non-bounded solution for likelihood function ratios. Therefore consistence and asymptotic normality of the Quasi maximum likelihood estimator should be verified with a Montecarlo experiment. A limited analysis can be found in Baillie et al. (1996) and in Bollerslev and Mikkelsen (1996), even if their claim was not completely correct, the reported Montecarlo experiment is still valid and shows consistence and asymptotic normality of the QML estimators, but only for the FIGARCH(1,d,0), the only parametrization they considerd. Caporin (2002) showes that the QML estimator is consistent and converge to the normality also for the FIGARCH(1,d,1) and the FIGARCH(0,d,0). At the moment a formal proof of the asymptotic properties of the QMLE for long memory GARCH model is not available. It is worth to mention that the results of Jeantheau

(1998) could be used to assess consistency in a restricted parameter space, see Caporin (2002), however this methodology verifies only a poitwise convergence of the parameters estimates and not a uniform convergence.

This work is concluded with a note on optimization methods. As previously noted, with the orders p and m greater or equal 1 a set of nonlinear constraint is needed to impose positivity on conditional variances. The focus is on the FIGARCH(1,d,1) case: in this model the nonlinear constraint in (9) may be represented with the following graph, where attention is restricted on the parameters β, d and ϕ all bounded between 0 and 1.

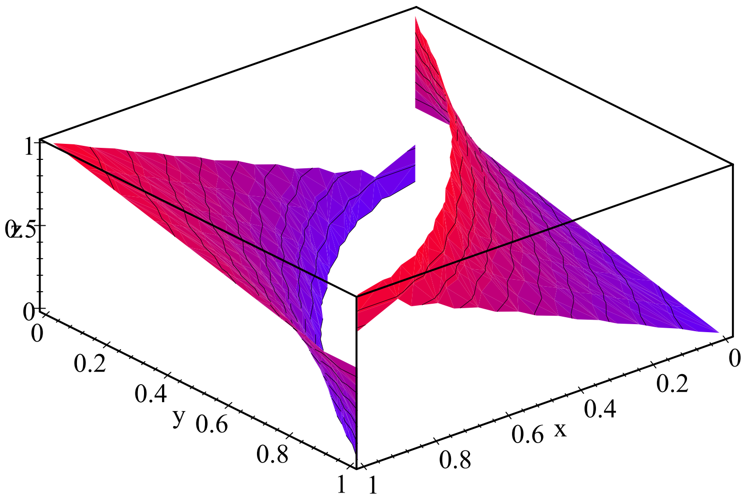

Figure 1: surface of non-linear constraints for FIGARCH(1,d,1)

Figure 1 reports the surface of the equality constraint, where $x = \beta$, $y = d$ and $z = \phi$. This surface is transalted in a space considering the inequality, precisely on the left on an ideal plane $x = y$ ($\beta = d$) all points above the surface while, on the right all points below the surface. Remenber that the admissible region of parameter combination is also affected by the other linear constraints, all taken into account drawing Figure 1. The main thing that appear from the graph is that it is not defined for $\beta = d$: in this case all values of ϕ are admissible since the inequality is always satisfied. In principle the space of admissible parameter combination can be divided in two subspaces, separated by the plane $\beta = d$, and define them as L ($d < \beta$) and R ($d > \beta$). A particular behavoiur of optimization algorithms was noted: they optimize with respect to the parameter combinations that belong to only to one of these subspaces, and are not able to switch between the two. Therefore it may happen that given a simulated series with parameters ($\beta = 0.5$, $d = 0.8$ and $\phi = 0.3$) $\in$ R, and then estimateing on that series the FIGARCH(1,d,1) it can be obtained as a result a parameter combination taht belongs to L. This effect is evident even using as starting values a parameter combination on the correct subspace, therefore this behaviour of optimization algorithm must be solved with other methods. Two solutions are proposed: optimize on the two distinc subspaces and then choose the parameter combination that lead to the higher loglikelihood or use an optimization algorithm that switch randomly in the parameter space.

References

[1] Baillie, R.T and T. Bollerslev, H.O. Mikkelsen, 1996, Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 74, 3-30

[2] Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307-327

[3] 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

[4] Bollerslev, T and H.O. Mikkelsen, 1996, Modeling and pricing long memory in stock market volatility, *Journal of Econometrics* 73, 151-184

[5] Bougerol, P. and N. Picard, 1992, Stationarity of GARCH processes and some non-negative time series, *Journal of Econometrics* 52, 115-127

[6] Caporin, M., 2002, *Long memory conditional heteroskedasticity and second order causality*, Università di Venezia, PhD dissertation

[7] Chung, Ching-Fan, 2001, Estimating the fractionally integrated GARCH model, *National Taiwan University discussion paper*

[8] Davidson, J., 2001, *Moment and memory properties of linear conditional heteroskedastic models*, working paper

[9] Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, *Econometrica* 50, 987-1008

[10] Engle, R.F. and T. Bollerslev, 1986, Modelling the persistence of conditional variances, *Econometric reviews* 5, 1-50

[11] Giraitis, L., Kokoszka, P. and Leipus R., 2000, Stationary ARCH models: dependence structure and central limit theorem, *Econometric theory*, 16, 3-22

[12] Kazakevicius, V. and Leipus, R. 1999, On stationarity in the ARCH(∞) model, *Econometric theory*, forthcoming

[13] Kazakevicius, V. and Leipus, R. 2001, *Stationarity in the integrated ARCH models*, working paper

[14] Jeantheau T. 1998, Strong consistency of estimators for multivariate ARCH models, *Econometric theory*, 14, 70-86

[15] Lee, Sang-Won and B.E. Hansen, 1994, Asymptotic theory fir the GARCH(1,1) quasi-maximum likelihood estimator, *Econometric Theory* 10, 29-52

[16] Mikosch, T. and Starica, C., 2000, *Long range dependence effects and ARCH modelling*, Preprint

[17] Nelson D.B., 1990, Stationarity and persistence in the GARCH(1,1) model, *Econometric theory*, 6, 318-334

[18] Robinson, P. M., 1991, Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression, *Journal of Econometrics*, 47, 67-84

[19] Zaffaroni P., 2000, Stationarity and memory of ARCH models, *STICERD publication EM/2000/383*