EVALUATING THE LYAPOUNOV EXPONENT AND EXISTENCE OF MOMENTS FOR THRESHOLD AR-ARCH MODELS

BY DAREN B. H. CLINE

Texas A&M University

First Version received July 2006

Abstract. We demonstrate a reliable and computationally feasible method for determining whether a given threshold autoregression autoregressive conditional heteroscedastic (AR-ARCH) model is ergodic, and for determining which moments exist when it is ergodic. This method may be used to delineate the parameter space of the model. We show (for an order 2 model) that the parameter space is much less constrained than commonly is assumed.

Keywords. Threshold models; ARCH; GARCH; Lyapounov exponent; stationary.

1. INTRODUCTION

Modelling the stochastic volatility of econometric and other time series with autoregressive conditional heteroscedastic (ARCH) and generalized autoregressive conditional heteroscedastic (GARCH)-type models has proved to be very successful. This effort has been extended to include additional nonlinearity such as threshold (G)ARCH models (e.g. Rabemananjara and Zakoïan, 1993; Zakoïan, 1994), as well as the addition of autoregression components, which also may be nonlinear (Li and Li, 1996; Liu *et al.*, 1997; Lu, 1998; Ling, 1999; Hwang and Woo, 2001; Lu and Jiang, 2001; Lanne and Saikkonen, 2005). ARCH and GARCH models with regime switching also have been suggested and analysed (see references in Francq *et al.*, 2001; Francq and Zakoïan, 2005).

Although it is a standard practice to impose second-order stationarity when modelling a (G)ARCH time series, this is not so easily done when the model includes an autoregression component or is a threshold model. Standard assumptions prove to be far too restrictive. As these more general models come into use, theoreticians and analysts will need tools to help circumscribe each model's parameter space or, at the very least, to determine whether a given set of parameters is within the parameter space. Unfortunately, the very shape of these parameter spaces is unknown, and they are quite possibly complicated and surprising. Indeed, as we will show, there are models for which a moderate amount of stochastic volatility is *required* for stationarity.

0143-9782/07/02 241-260

JOURNAL OF TIME SERIES ANALYSIS Vol. 28, No. 2

© 2006 The Author

Journal compilation © 2006 Blackwell Publishing Ltd., 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

The condition for second-order stationarity can be somewhat stronger than that for the existence of a stationary distribution (ergodicity). Since ergodicity is required for most statistical limit theorems, verifying when it holds is of equal interest. Indeed, the estimation of parameters usually depends on it, irrespective of any moment requirements.

Still, the moment question itself is quite important. These models are known to have heavy tails and so, beyond verifying that the model is stable, it is useful to ascertain the nature of those tails. Ideally, this includes determining not only which moments exist but also the relationships among consecutive extreme values in the series.

As a general rule, it can be difficult to check for ergodicity, especially in the context of stochastic volatility. This is both a statistical question for actual time series and an analytical one for specific models. Although we direct our attention only to the analytical question, statistical issues arise when simulating a model to evaluate its properties. Ordinarily, testing ergodicity and the existence of moments requires long series and multiple instances of the more 'extraordinary' behaviour possible. The latter requirement makes direct simulation of a model quite unwieldy for these purposes.

Instead, our aim here is to provide a stable and efficient method for identifying ergodicity of a given model and, if ergodic, for determining which moments exist for the stationary distribution. Objectively, the problem is to check analytical conditions which are implicit by necessity and, in doing so, to sidestep the simulation of a stochastically volatile process. If workable, then this method could ultimately be used statistically to appraise the properties of an actual time series.

This paper is organized as follows. First, in section 2, we describe the threshold AR-ARCH models we study. These have natural, easily evaluated, *sufficient* conditions for second-order stationarity or existence of a stationary distribution which are well known but ultimately are too strong. An alternative approach, motivated by the work of Bougerol and Picard (1992a,b) and Cline and Pu (2004), is to determine the process's Lyapounov exponent. The very useful result is that this constant may be evaluated by simulating or numerically analysing a simpler, *bounded and uniformly ergodic* process. These points we discuss in sections 3 and 4. The question of moments is closely related as existence of a moment also may be determined numerically from the simpler process.

In section 5, we discuss numerical computations and the issues involved and, in section 6, we implement the methods for a threshold AR-ARCH model of order 2 and plot cross-sections of its parameter space. In section 7, we briefly discuss extending the method to GARCH and related models.

2. THRESHOLD AR-ARCH MODELS

Throughout we assume e_1, e_2, \ldots are independent and identically distributed (i.i.d.) random errors with density f, symmetric about 0 and positive on the entire

real line. We also presume that $E(|e_t|^r) < \infty$ for r sufficiently large. These assumptions allow us to focus on parameter values.

An ARCH model with autoregression is

$$\xi_t = a_0 + \sum_{i=1}^p a_i \xi_{t-i} + \left(b_0^2 + \sum_{i=1}^p b_i^2 \xi_{t-i}^2 \right)^{1/2} e_t, \tag{1}$$

where $b_0 \neq 0$ and either $a_p \neq 0$ or $b_p \neq 0$. The first part, $a_0 + \sum_{i=1}^p a_i \xi_{t-i}$ is the autoregression or AR component. The second part is the ARCH component. The conditional 'variance' or 'volatility', $b_0^2 + \sum_{i=1}^p b_i^2 \xi_{t-i}^2$, is a regression on the squared process.

The model (1) differs from an autoregression model with ARCH errors in that the conditional variance of the latter is a regression on the past noise (the ARCH component above). Another related model is one with GARCH errors. (Ling, 2004, calls (1) a 'double autoregression model'). The main advantage of our formulation here is the simplicity of the state vector $X_t = (\xi_t, \dots, \xi_{t-p+1})$.

This model is ergodic and has a stationary distribution if

$$\left(\sum_{i=1}^{p} |a_i|\right)^2 + \left(\sum_{i=1}^{p} b_i^2\right) E(e_t^2) < 1$$
(2)

(cf. Lu, 1998). This is a standard assumption that ensures $E(\xi_t^2) < \infty$, but it also is stronger than the condition for ergodicity. (See Masry and Tjøstheim, 1995 for an alternative condition.). Precise conditions for ergodicity are known when p = 1(Borkovec and Klüppelberg, 2001) or when all AR coefficients are 0 (many different authors such as Bougerol and Picard, 1992a,b).

The autoregression component need not be linear, however. For example, it can be piecewise linear, as is the case with a threshold AR model. Tong (1990) called this combination of threshold AR with ARCH, a second generation model. More generally, both the autoregression and the conditional variance may be nonlinear, including having piecewise characteristics. This is what we call the nonlinear AR-ARCH model. Others who have investigated these or similar models include Rabemanajara and Zakoian (1993), Li and Li (1996), Liu *et al.* (1997), Lu (1998), Ling (1999), Hwang and Woo (2001), Lu and Jaing (2001) and Lanne and Saikkonen (2005).

The state space (Markov chain) representation of the model is

$$\xi_t = a(X_{t-1}) + b(X_{t-1})e_t + c(X_{t-1}, e_t),$$

$$X_{t-1} = (\xi_{t-1}, \dots, \xi_{t-p}).$$
(3)

We will assume that a and b are piecewise continuous and homogeneous:

$$a(x) = a\left(\frac{x}{\|x\|}\right)\|x\| \quad \text{and} \quad b(x) = b\left(\frac{x}{\|x\|}\right)\|x\|,\tag{4}$$

and that c(x,u) = o(||x||)O(|u|) as $||x||, |u| \to \infty$.

© 2006 The Author

A special case is the threshold AR-ARCH model (TAR-ARCH) with order p and delay lag $k \le p$, expressed as follows.

$$\xi_{t} = \begin{cases} a_{10} + \sum_{i=1}^{p} a_{1i}\xi_{t-i} + \left(b_{10}^{2} + \sum_{i=1}^{p} b_{1i}^{2}\xi_{t-i}^{2}\right)^{1/2}e_{t}, & \text{if } \xi_{t-k} \leq 0, \\ a_{20} + \sum_{i=1}^{p} a_{2i}\xi_{t-i} + \left(b_{20}^{2} + \sum_{i=1}^{p} b_{2i}^{2}\xi_{t-i}^{2}\right)^{1/2}e_{t}, & \text{if } \xi_{t-k} > 0. \end{cases}$$
(5)

Here, with $x = (x_1, \ldots, x_p)$, we have

$$a(x) = \sum_{i=1}^{p} (a_{1i} \mathbf{1}_{x_k \le 0} + a_{2i} \mathbf{1}_{x_k > 0}) x_i$$

and

$$b(x) = \left(\sum_{i=1}^{p} (b_{1i}^2 \mathbf{1}_{x_k \le 0} + b_{2i}^2 \mathbf{1}_{x_k > 0}) x_i^2\right)^{1/2}.$$

If we assume, for the general model in (3), that

$$|a(x_1,...,x_p)| \le a_1|x_1| + \dots + a_p|x_p|$$
 and $b(x_1,...,x_p) \le (b_1^2 x_1^2 + \dots + b_p^2 x_p^2)^{1/2}$, (6)

then a blanket application of the standard assumption (2) for an AR-ARCH model gives this condition (Lu, 1998):

$$\left(\sum_{i=1}^{p} a_{i}\right)^{2} + \left(\sum_{i=1}^{p} b_{i}^{2}\right) E(e_{t}^{2}) < 1.$$
(7)

Lu and Jiang (2001) give a similar condition that guarantees existence of the first moment:

$$\sum_{i=1}^{p} (a_i + b_i E(|e_i|)) < 1.$$
(8)

(see also Cline and Pu, 2004). These conditions, however, are quite a bit stronger than the necessary condition for second-order stationarity, and they are extremely limiting conditions for ergodicity of even the simplest models. The exact condition for ergodicity (see section 4), determined by Cline and Pu (2004), must be computed numerically if p > 1.

The generalized ARCH (GARCH) model is

$$\xi_t = \sigma_t e_t, \sigma_t = \left(b_0^2 + \sum_{i=1}^p b_i^2 \xi_{t-i}^2 + \sum_{i=1}^q c_i^2 \sigma_{t-i}^2\right)^{1/2}.$$
(9)

The ARCH model is a special case with each $c_j = 0$. Assuming $E(e_t^2) = 1$, it is possible to give an exact condition for second-order stationarity of the GARCH model, namely

$$\sum_{i=1}^{p} b_i^2 + \sum_{i=1}^{q} c_i^2 < 1 \tag{10}$$

(Bollerslev, 1986). If, however, an autoregression component is included or nonlinearity allowed or both then once again exact conditions are not so easily computed (see section 7).

3. THE LYAPOUNOV EXPONENT

The results given in this section and the next are from Cline and Pu (2004), unless otherwise attributed.

To express precise conditions for ergodicity, we turn to the Lyapounov exponent, a concept well known by those studying stability of dynamical systems. In the context of nonlinear time series, nonstability means explosive behaviour. The Lyapounov exponent, as we define it for the state space model $\{X_t\}$ of a time series, is

$$\gamma = \liminf_{n \to \infty} \limsup_{||x|| \to \infty} \frac{1}{n} E\left(\log\left(\frac{||X_n||}{||X_0||}\right) \mid X_0 = x\right),\tag{11}$$

which measures the drift of the process when large. When $\gamma < 0$, a geometric drift condition holds:

$$E(||X_n||^r |X_0 = x) \le \rho ||x||^r + K,$$
(12)

for some $n \ge 1$, r > 0, $\rho < 1$, $K < \infty$. If the process is irreducible, this a standard condition for geometric ergodicity (Meyn and Tweedie, 1993). Geometric ergodicity implies mixing and the drift condition implies existence of an *r*th moment for the stationary distribution. Consequently, statistical limit theorems such as the strong law of large numbers and the central limit theorem will hold if applied to averages of functions of the series with the required first or second moments.

On the other hand, $\gamma > 0$ (plus some regularity) implies

$$P(||X_n|| \to \infty | X_0 = x) > 0 \quad \text{for any } x, \tag{13}$$

and hence the process is transient. For the boundary case, $\gamma = 0$, the process can be positive recurrent, null recurrent or transient, as is the case of threshold AR models (Cline and Pu, 1999).

Because the squared GARCH process is effectively linear, it is possible to reexpress it as a random coefficients model. The stability properties of such a model have been known for some time. The random coefficients embedding is as follows. Let

$$Y_t = (\xi_t^2, \dots, \xi_{t-p+1}^2, \sigma_{t+1}^2, \dots, \sigma_{t-q+2}^2).$$
(14)

Then there exists a fixed matrix L and random matrices $B_t = (0, ..., 0, b_0^2 e_t^2, 0, ..., 0)$ and

$$C_{t} = \begin{pmatrix} 0 \cdots 0 & e_{t}^{2} & 0 \cdots 0 \\ L & 0 & 0 \cdots 0 \\ b_{2}^{2} \cdots b_{p}^{2} & b_{1}^{2} e_{t}^{2} + c_{1}^{2} & c_{2}^{2} \cdots c_{q}^{2} \\ 0 \cdots 0 & 1 & 0 \cdots 0 \\ 0 \cdots 0 & 0 & I \end{pmatrix}$$
(15)

such that

$$Y_t = B_t + C_t Y_{t-1}.$$
 (16)

The random coefficients model (and hence $\{\xi_t\}$) is ergodic iff

$$2\gamma = \lim_{n \to \infty} \frac{1}{n} E(\log(||C_n C_{n-1} \cdots C_1||)) < 0$$
(17)

(Bougerol and Picard, 1992a, b). The limit is the Lyapounov exponent of the squared GARCH process and hence is twice that of the GARCH process itself. Furthermore, the GARCH model has finite *r*th moment (and $E(|\xi_t|^r) < \infty$) if

$$\rho_r \stackrel{\text{def}}{=} \lim_{n \to \infty} \left(E(||C_n C_{n-1} \cdots C_1||^{r/2}) \right)^{1/n} < 1$$
(18)

(Basrak *et al.*, 2002). In fact, the stationary distribution of linear combinations of the squared GARCH process have multivariate regularly varying tails (see also Borkovec, 2000; Borkovec and Klüppelberg, 2001). Klüppelberg and Pergamenchtchikov (2004) have a similar result for random coefficient models. Francq and Zakoïan (2005) have looked at GARCH models with Markov regime switching, developing precise conditions for existence of moments.

The constants γ and ρ_r in eqns (17) and (18) are defined in terms of products of random matrices which makes both, and especially the latter, difficult to estimate efficiently by direct simulation. Nevertheless, these results show that the stability features of the GARCH process can be expressed in terms of a somewhat simpler process. The successful attainment of analytic results for the GARCH model immediately suggests questions for the extended and nonlinear models under study here.

- Are there similar results (and critical constants γ and ρ_r) if AR terms are included, such as in the AR-ARCH model?
- What happens if the model is 'nonlinear' in either the AR part or in the conditional heteroscedasticity?
- Can γ and ρ_r be evaluated more easily than by direct simulation of the model?

Note that if the model either has an AR component or is a threshold model, then it cannot be expressed as a random coefficients model; so, our approach must be more subtle than evaluating asymptotic behaviour of random matrices as in eqn (17).

4. EXACT STABILITY CONDITIONS VIA THE PIGGYBACK METHOD

The now standard approach to proving geometric ergodicity of a Markov chain $\{X_i\}$ involves verifying the *Foster–Lyapounov drift condition*: there is a *test function* $V(x) \ge 1$, $K < \infty$ and $\rho < 1$ such that

$$E\Big(V(X_1)\Big|X_0=x\Big) \le K + \rho V(x).$$
(19)

(see Meyn and Tweedie, 1993). The difficult part is choosing the optimal test function. This condition is equivalent to equation (12) when $c_1||x||^r \le V(x) \le c_2 + c_3||x||^r$ for finite positive c_1, c_2, c_3 . The difficulty in proving eqn (12) is that an 'optimal' *n* typically is too large to make analytic computation feasible for nonlinear models. The piggyback method we developed, however, makes it possible to choose an optimal, or near optimal, test function *V* to verify eqn (19).

To understand the piggyback method, recall eqn (3) and let

$$\zeta(x, e_t) = (a(x) + b(x)e_t, x_1, \dots, x_{p-1}),$$
(20)

for $x = (x_1, \dots, x_p)$. Because of the homogeneity of a(x) and b(x) (see eqn 4),

$$\zeta\left(\frac{x}{\|x\|}, e_t\right) = \frac{\zeta(x, e_t)}{\|x\|}.$$
(21)

Thus, when $||X_{t-1}||$ is very large [and $c(X_{t-1}, e_t)$ is negligible],

$$\frac{X_t}{\|X_{t-1}\|} \doteq \zeta \left(\frac{X_{t-1}}{\|X_{t-1}\|}, e_t \right).$$
(22)

Accordingly, if $\theta_t = X_t / ||X_t||$ and $||X_{t-1}||$ is large, then

$$\theta_t \doteq \zeta(\theta_{t-1}, e_t) / \|\zeta(\theta_{t-1}, e_t)\| \quad \text{and} \quad \|X_t\| / \|X_{t-1}\| \doteq \|\zeta(\theta_{t-1}, e_t)\|.$$
(23)

We introduce a new process as follows. Let $\Theta = \{||x|| \in \mathbb{R}^p : ||x|| = 1\}$ be the unit sphere in \mathbb{R}^p and define the Markov chain on Θ

$$\theta_t^* = \frac{\zeta(\theta_{t-1}^*, e_t)}{\|\zeta(\theta_{t-1}^*, e_t)\|}.$$
(24)

Since this process mimics the first part of eqn (23), we call $\{\theta_t^*\}$ the *collapsed* chain associated with $\{X_t\}$. Its utility lies in the fact that it is uniformly ergodic with some stationary distribution, say π .

Also from the second part of eqn (23), we can see that it is the behaviour of $\|\zeta(\theta_{t-1}^*, e_t)\|$ under stationarity that is paramount in determining the drift of $\{X_t\}$.

The key to our method, therefore, is that we piggyback a drift condition on the known ergodicity of the collapsed chain.

It is worth noting that this approach is both necessary for threshold models and distinct from that of a Markov switching model such as those studied by Francq and Zakoïan (2005). For such a model, the regime switching depends directly on a hidden Markov chain and only indirectly on the current state of the process itself.

We can now state the principal result.

THEOREM 1 (STABILITY OF THE THRESHOLD AR-ARCH MODEL). Let π be the stationary distribution of the collapsed process $\{\theta_t^*\}$.

(a) The Lyapounov exponent for $\{X_t\}$ is

$$\gamma = \int_{\Theta} E(\log \|\zeta(\theta, e_t)\|) \pi(\mathrm{d}\theta).$$
⁽²⁵⁾

- (b) If $\gamma < 0$ then $\{X_t\}$ is geometrically ergodic.
- (c) For any $\epsilon > 0$, there exists a piecewise continuous function $v(\theta)$ that satisfies the near equilibrium equation

$$\sup_{\theta \in \Theta} \left| E(\nu(\theta_1^*) - \nu(\theta) + \log \|\zeta(\theta, e_1)\| \|\theta_0^* = \theta) - \gamma \right| < \epsilon.$$
(26)

(d) $\{X_t\}$ is geometrically ergodic if there exists bounded $v(\theta)$ such that

$$\sup_{\theta \in \Theta} E(\nu(\theta_1^*) - \nu(\theta) + \log \|\zeta(\theta, e_1)\| \|\theta_0^* = \theta) < 0.$$
(27)

Note that *v* plays the role of a test function and γ is no longer defined as a limit (as it was in the GARCH condition). Indeed, γ is no longer computed in terms of random matrices. One may think of $\|\zeta(\theta_{t-1}^*, e_t)\|$ as playing the role of $\|C_t\|$ in eqn (17).

A corollary to Theorem 4.1 yields the means to verify existence of moments.

COROLLARY 1 (MOMENTS OF THE THRESHOLD AR-ARCH MODEL). The following are equivalent conditions for $\{X_t\}$ to have a stationary distribution with finite rth moment.

(a) There exists $\lambda_r(\theta)$, bounded and bounded away from 0, such that

$$\sup_{\theta \in \Theta} E\left(\frac{\lambda_r(\theta_1^*)}{\lambda_r(\theta)} \|\zeta(\theta, e_1)\|^r |\theta_0^* = \theta\right) < 1.$$
(28)

(b)
$$\rho_r = \limsup_{n \to \infty} \sup_{\theta \in \Theta} \left(E\left(\prod_{t=1}^n \|\zeta(\theta_{t-1}^*, e_t)\|^r \mid \theta_0^* = \theta \right) \right)^{1/n} < 1.$$
(29)

© 2006 The Author Journal compilation © 2006 Blackwell Publishing Ltd. JOURNAL OF TIME SERIES ANALYSIS Vol. 28, No. 2

248

Here, λ_r plays the role of a test function and its optimal choice is the eigenfunction corresponding to a maximal eigenvalue (ρ_r). The limit ρ_r in (b) corresponds to that given earlier for the GARCH model.

Note that (a) and (b) are *not* the same as $\int_{\Theta} E(||\zeta(\theta, e_1)||^r) \pi(d\theta) < 1$.

5. NUMERICAL EVALUATION OF THE CONDITIONS

When p = 1, the conditions in both Theorem 1 and Corollary 1 can be determined explicitly (Goldie, 1991; Cline, 2005), so we are concerned here only with the case p > 1.

Simulating or numerically analysing a uniformly ergodic process has two very crucial advantages over simulating the time-series model itself. First, the critical questions we are studying concern the most extreme behaviour of the process and, consequently, relatively rare behaviour if the model is stable. It is also a highly variable behaviour. Second, a uniformly ergodic process will converge to stationarity quite rapidly so that the properties we desire may be calculated quickly and with confidence.

Evaluating the Lyapounov exponent γ may be accomplished by either of two straightforward schemes. The first is to estimate γ as the mean in eqn (25). That is, by simulating $\{(\theta_t^*, e_t)\}$, we obtain the estimator

$$\hat{\gamma} = \frac{1}{n} \sum_{t=m}^{m+n} \log \|\zeta(\theta_{t-1}^*, e_t)\|,\tag{30}$$

where *m* is the 'warm-up' time for the collapsed Markov chain. Statistical properties for such an estimator are known (because we know the process is ergodic) and our experience suggests that this approach is quite reliable, even when accounting for correlation in the sequence. Because the collapsed process converges to stationarity so fast, we found that a warm-up time of m = 30 was quite sufficient for our examples.

By regressing $\log \|\zeta(\theta_{t-1}^*, e_t)\|$ on other simulated values with known means, the method can easily be improved. For example, the first component of θ_t^* is

$$\theta_{t,1}^* = \frac{a(\theta_{t-1}^*) + b(\theta_{t-1}^*)e_t}{\|\zeta(\theta_{t-1}^*, e_t)\|}.$$

Since $\log(|\theta_{t,1}^*|/|\theta_{t-1,1}^*|)$ has mean 0 and is correlated with $\log \|\zeta(\theta_{t-1}^*, e_t)\|$, it proved to be a useful regressor. Likewise, for a reasonable function g,

$$g(a(\theta_{t-1}^*) + b(\theta_{t-1,1}^*)e_t) - \int_{-\infty}^{\infty} g(a(\theta_{t-1}^*) + b(\theta_{t-1}^*)z)f(z)dz$$

has mean 0 and can be used if it may be computed precisely. In our simulations, we chose f to be the normal density and we took g(z) to be a piecewise linear approximation of $\log |z|$. We also used

D. B. H. CLINE

$$h(|a(\theta_{t-1}^*) + b(\theta_{t-1}^*)e_t| / |\theta_{t-1,1}^*|) - \int_{-\infty}^{\infty} h(|a(\theta_{t-1}^*) + b(\theta_{t-1}^*)z| / |\theta_{t-1}^*|)f(z)dz$$

where h(z) is a piecewise linear approximation to $\log (1 + z^2)$. With these as regressors (and with an estimated covariance matrix), we obtained an estimator that decreased the variance by a factor of 5–10.

The second scheme is to solve the near equilibrium equation eqn (26). Ideally, one may even attempt to solve an exact equilibrium equation given by

$$v(\theta) = E\Big(v(\theta_1^*) + \log \|\zeta(\theta, e_1)\| \mid \theta_0^* = \theta\Big) - \gamma \quad \text{s.t.} \int_{\Theta} v(\theta) d\theta = 0.$$
(31)

We cannot claim that the 'optimal' function $v(\theta)$ will be smooth, however, thereby presenting possible difficulties in the numerical integration. Since eqn (26) indicates we only need a piecewise continuous function to get close; an iterative procedure can still lead to a satisfactory approximation for γ .

This can proceed in either of two ways. One option is to evaluate the expectation in eqn (31) using the transition kernel of the collapsed Markov chain. But in addition to the intricacies of integrating on the unit sphere, the transition kernel is singular in general. The second (and maybe more obvious) option is to integrate with respect to the error distribution. This requires interpolating each approximation of the test function, however, and avoiding interpolating across discontinuities in particular. That our numerical solutions are somewhat 'jittery' looking is probably because of an interaction between the quadrature (numerical integration) and the interpolation.

For the TAR-ARCH model (5) of order p = 2 (see section 6), we found that the latter method works well if one makes educated guesses about the discontinuity points of the optimal function. In fact, it may suffice to use a test function discontinuous only at the threshold, but we cannot rule out the existence of other discontinuities.

Let f be the error density and let $\eta(\theta, u) = \zeta(\theta, u)/||\zeta(\theta, u)||$ so that

$$\theta_t^* = \eta(\theta_{t-1}^*, e_t). \tag{32}$$

A numerical integration method obtains γ by iterating the following:

$$\tilde{v}_{j}(\theta) = \int \left(v_{j-1}(\eta(\theta, u)) + \log \|\zeta(\theta, u)\| \right) f(u) du,$$

$$\gamma_{j} = \sup_{\theta \in \Theta} \left(\tilde{v}_{j}(\theta) - v_{j-1}(\theta) \right),$$

$$v_{j}(\theta) = \tilde{v}_{j}(\theta) - \int_{\Theta} \tilde{v}_{j}(\tilde{\theta}) d\tilde{\theta}.$$
(33)

The sequence $\{\gamma_j\}$ is decreasing, and to confirm stability one may stop as soon as $\gamma_j < 0$, even without knowing all the possible discontinuities of the optimal v. It may be shown mathematically in fact that

$$\inf_{\theta \in \Theta} \left(\tilde{v}_j(\theta) - v_{j-1}(\theta) \right) \le \gamma \le \sup_{\theta \in \Theta} \left(\tilde{v}_j(\theta) - v_{j-1}(\theta) \right), \tag{34}$$

© 2006 The Author

Journal compilation © 2006 Blackwell Publishing Ltd.

JOURNAL OF TIME SERIES ANALYSIS Vol. 28, No. 2

assuming the integration is correct. Therefore, the principle source of error here is the quadrature. If the trapezoid method is used then the quadrature error is known to be bounded by $0.25\delta^2 \int |g''(u)| du$ where g is the function being integrated and δ is the increment used for the integration variable. Since convergence is monotone and (in our experience) does not fail, the initial function is not critical. We used $v_0(\theta) \equiv 0$.

Applying either scheme with estimated parameters enables one to determine if the estimates correspond to a valid (stable) model. A possible hypothesis test procedure might be to vary the estimates according to their estimated sampling behaviour and to check stability accordingly. We leave this question to others.

To check for existence of the *r*th moment in the stationary distribution of $\{\xi_t\}$, one needs to estimate ρ_r . To do this by simulation is, at best, a delicate operation. In principle, the estimator

$$\hat{\rho}_{r} = \left(\frac{1}{N} \sum_{k=0}^{N-1} \prod_{i=0}^{n-1} \|\zeta(\theta_{kn+i}^{*}, e_{kn+i+1})\|^{r}\right)^{1/n}$$
(35)

is consistent for ρ_r as $n, N \to \infty$ (see eqn 4.11). Just how to balance n and N is uncertain, and the sampling properties of the estimator are equally cumbersome to estimate. Because we found the next approach very suitable, we did not consider simulation any further.

Like the second scheme for evaluating γ , the approach is to solve a criterion involving a test function, namely eqn (28). In fact, solving eqn (28) is tantamount to solving an eigenvalue problem:

$$\rho_r \lambda_r(\theta) = E \Big(\lambda_r(\theta_1^*) \| \zeta(\theta, e_1) \|^r \Big| \theta_0^* = \theta \Big) \quad \text{s.t.} \int_{\Theta} \lambda_r(\theta) d\theta = 1,$$
(36)

and checking if $\rho_r < 1$. For quadrature purposes, we express eqn (36) as

$$\rho_r \lambda_r(\theta) = \int \lambda_r(\eta(\theta, u)) \|\zeta(\theta, u)\|^r f(u) du \quad \text{s.t.} \int_{\Theta} \lambda_r(\theta) d\theta = 1.$$
(37)

This may also be solved iteratively. The procedure is

$$\tilde{\lambda}_{j}(\theta) = \int \lambda_{j-1}(\eta(\theta, u)) \|\zeta(\theta, u)\|^{r} f(u) du,$$

$$\rho_{r,j} = \sup_{\theta \in \Theta} \left(\tilde{\lambda}_{j}(\theta) / \lambda_{j-1}(\theta) \right),$$

$$\lambda_{j}(\theta) = \frac{\tilde{\lambda}_{j}(\theta)}{\int_{\Theta} \tilde{\lambda}_{j}(\tilde{\theta}) d\tilde{\theta}}$$
(38)

Again, to verify $\rho_r < 1$ one may stop as soon as the estimate is less than 1. It suffices to initialize the procedure with a constant function $\lambda_0(\theta)$.

See Figure 1 for an example of computed test functions v, λ_1 and λ_2 .



FIGURE 1. Test functions for the TAR-ARCH(2) model with $a_{11} = 0.3$, $a_{12} = 0.2$, $a_{21} = -0.4$, $a_{22} = 0.1$, $b_{11} = 0.7$, $b_{12} = 0.2$, $b_{21} = 0.3$, $b_{22} = 0.1$. Functions shown are unique up to an additive constant.

6. AN EXAMPLE

We implemented these methods to help delineate the parameter space for the threshold AR-ARCH model of order 2 and delay 1. This model is expressed as

$$\xi_{t} = \begin{cases} a_{10} + a_{11}\xi_{t-1} + a_{12}\xi_{t-2} + \left(b_{10}^{2} + b_{11}^{2}\xi_{t-1}^{2} + b_{12}^{2}\xi_{t-2}^{2}\right)^{1/2}e_{t}, & \text{if } \xi_{t-1} \le 0, \\ a_{20} + a_{21}\xi_{t-1} + a_{22}\xi_{t-2} + \left(b_{20}^{2} + b_{21}^{2}\xi_{t-1}^{2} + b_{22}^{2}\xi_{t-2}^{2}\right)^{1/2}e_{t}, & \text{if } \xi_{t-1} > 0. \end{cases}$$

$$(39)$$

We may assume that the errors have variance equal to 1. The state vector is $X_t = (\xi_t, \xi_{t-1})$ and the collapsed Markov chain $\{\theta_t^*\}$ 'lives' on the unit circle in \mathbb{R}^2 . The thresholds on the unit circle are at $\operatorname{arc}(\theta) = \pm \pi/2$ and are probable points of discontinuity for the test functions (as confirmed in Figure 1).

The eight parameters critical for the stability of $\{\xi_t\}$ are $a_{11}, a_{12}, a_{21}, a_{22}, b_{11}, b_{12}, b_{21}$ and b_{22} . The criterion commonly used for second-order stationarity (which we call 'naive') is

$$\left(\max(|a_{11}|,|a_{21}|) + \max(|a_{12}|,|a_{22}|)\right)^2 + \max(b_{11}^2,b_{21}^2) + \max(b_{12}^2,b_{22}^2) < 1.$$
(40)

Clearly, this requires all the parameters to be less than one in magnitude and most of them to be quite small. Although sufficient, this condition is unduly harsh.

With proper programming [using R (R Development Core Team, 2004) on a standard desktop computer], either scheme for evaluating γ was reliable and

© 2006 The Author

Journal compilation © 2006 Blackwell Publishing Ltd.

JOURNAL OF TIME SERIES ANALYSIS Vol. 28, No. 2

AR parameters				ARCH parameters				Simulation scheme			Equilibrium scheme		
a_{11}	<i>a</i> ₁₂	a_{21}	<i>a</i> ₂₂	b_{11}	b_{12}	b_{21}	<i>b</i> ₂₂	Ŷ	SE	CPU	Ŷ	Quad. Err.	CPU
-0.2	-1.0	0.0	-0.2	0.4	0.0	0.2	0.9	-0.09663	0.00082	9.3	-0.09773	0.00064	2.1
0.0	0.5	-0.4	-0.7	0.5	0.3	0.1	0.3	-0.25415	0.00080	9.2	-0.25358	0.00035	2.2
0.2	-0.2	-0.5	0.2	0.7	0.6	0.1	0.5	-0.27846	0.00057	9.0	-0.27780	0.00074	0.5
-0.2	0.3	0.1	0.6	0.5	0.8	0.7	0.3	-0.17624	0.00059	9.2	-0.17708	0.00045	1.3
-0.5	-0.7	0.3	-1.0	0.0	0.9	0.8	0.2	0.00963	0.00055	9.3	0.00966	0.00028	2.4
-0.8	0.0	0.5	0.7	0.0	1.1	0.3	0.7	0.02333	0.00082	9.1	0.02371	0.00031	1.2
0.2	-0.4	-0.5	-0.2	0.6	1.0	0.6	0.3	-0.20844	0.00117	8.9	-0.20816	0.00083	0.7
0.0	-0.2	0.4	0.6	0.5	0.3	0.0	0.2	-0.00852	0.00006	9.3	-0.00853	0.00306	2.9
0.3	-0.1	0.0	0.4	0.3	0.4	0.1	0.2	-0.52412	0.00074	9.2	-0.52411	0.00058	2.9
-0.4	0.0	-0.5	1.0	0.4	0.5	0.2	0.2	-0.01304	0.00147	9.1	-0.01108	0.00103	2.7

TABLE 1 Example Results from the Two Schemes for Estimating the Lyapounov Exponent γ

reasonably fast. Solving the equilibrium equation appears to be more efficient, by a factor of 5 or more. Table 1 shows, for a few randomly chosen parameter values, estimates of γ by simulation (using sequences of length 50,000) and by solving the equilibrium equation (evaluating *v* at 200 points on the unit circle and using 200 points between -5 and 5 for the quadrature). The error shown for the latter is an estimate of the quadrature error. CPU time is in seconds.

With 200 such trials, we found that the average standard error of the simulation method and the average quadrature error of the equilibrium method were both 0.001, but the simulation method averaged 9.1 seconds of CPU time while the equilibrium method averaged 1.6 seconds of CPU time.

Using the equilibrium/eigenvalue methods to evaluate γ , ρ_1 and ρ_2 , while varying two parameters at a time, we mapped out regions for ergodicity, finite first moment and finite second moment respectively. These are shown in Figures 2–9. The parameter region defined by the 'naive' criterion (40) is also shown, and is clearly much smaller even than the region for finite second moment. In particular, some parameters may have magnitude much bigger than one and the second moment still is finite.

As is quite common for threshold models, the parameter regions can be quite asymmetric. For example, Figures 2 and 6 show parameter regions for the AR coefficients that apply when $\xi_{t-1} < 0$. The leading coefficient, a_{11} can be quite negative because that causes the process to shift to the positive side where it very likely will shrink dramatically. Figures 3 and 7 also demonstrate that one leading coefficient may be large and negative as long as the other leading coefficient is not too big. Similarly, Figures 4, 5, 8 and 9 illustrate that one volatility (ARCH) parameter can be large because that presumably makes it more likely the process will switch to a regime where it will quickly shrink.

Figures 8 and 9 are interesting because the parameter regions do not include a neighborhood of the origin. Note that the AR parameters a_{11} , a_{12} correspond to a non-stationary AR(2) model. Consequently, a moderate amount of stochastic volatility is required for the threshold AR-ARCH model to be stable in such a case.



FIGURE 2. Stability regions for the TAR-ARCH(2) model with $a_{21} = -0.4$, $a_{22} = 0.1$, $b_{11} = 0.7$, $b_{12} = 0.2$, $b_{21} = 0.3$, $b_{22} = 0.1$.



FIGURE 3. Stability regions for the TAR-ARCH(2) model with $a_{12} = 0.2$, $a_{22} = 0.1$, $b_{11} = 0.7$, $b_{12} = 0.2$, $b_{21} = 0.3$, $b_{22} = 0.1$.



FIGURE 4. Stability regions for the TAR-ARCH(2) model with $a_{11} = 0.3$, $a_{12} = 0.2$, $a_{21} = -0.4$, $a_{22} = 0.1$, $b_{21} = 0.3$, $b_{22} = 0.1$.



FIGURE 5. Stability regions for the TAR-ARCH(2) model with $a_{11} = 0.3$, $a_{12} = 0.2$, $a_{21} = -0.4$, $a_{22} = 0.1$, $b_{12} = 0.2$, $b_{22} = 0.1$.



FIGURE 6. Stability regions for the TAR-ARCH(2) model with $a_{21} = -0.8$, $a_{22} = 0.1$, $b_{11} = 0.3$, $b_{12} = 0.2$, $b_{21} = 0.3$, $b_{22} = 0.1$.



FIGURE 7. Stability regions for the TAR-ARCH(2) model with $a_{12} = -0.75$, $a_{22} = 0.1$, $b_{11} = 0.3$, $b_{12} = 0.2$, $b_{21} = 0.3$, $b_{22} = 0.1$.



FIGURE 8. Stability regions for the TAR-ARCH(2) model with $a_{11} = -0.27$, $a_{12} = -0.75$, $a_{21} = -0.8$, $a_{22} = 0.1$, $b_{21} = 0.3$, $b_{22} = 0.1$.



FIGURE 9. Stability regions for the TAR-ARCH(2) model with $a_{11} = -0.27$, $a_{12} = -0.75$, $a_{21} = -0.8$, $a_{22} = 0.1$, $b_{12} = 0.2$, $b_{22} = 0.1$.

D. B. H. CLINE

7. A THRESHOLD AR-GARCH MODEL

The theory in Cline and Pu (2004) did not cover nonlinear GARCH models, strictly speaking, but the piggyback method should nevertheless be applicable. We outline the concept here for an order 1 threshold AR-GARCH model and suggest how its stability properties may be ascertained.

The model is defined by

$$\xi_t = a_{0j} + a_{1j}\xi_{t-1} + \sigma_t e_t, \quad \sigma_t^2 = b_{0j}^2 + b_{1j}^2\xi_{t-1}^2 + c_{1j}^2\sigma_{t-1}^2, \text{ if } (-1)^j\xi_{t-1} > 0, j = 1, 2.$$
(41)

If we define the homogeneous functions on $\mathbb{R} \times \mathbb{R}_+$

$$\alpha(x,s) = \frac{a_{11}x\mathbf{1}_{x<0} + a_{12}x\mathbf{1}_{x>0}}{(x^2 + s^2)^{1/2}}$$
(42)

and

$$\beta(x,s) = \left(\frac{(b_{11}^2 x^2 + c_{11}^2 s^2)\mathbf{1}_{x<0} + (b_{12}^2 x^2 + c_{12}^2 s^2)\mathbf{1}_{x>0}}{x^2 + s^2}\right)^{1/2}$$
(43)

then we have

$$X_{t} \stackrel{\text{def}}{=} \begin{pmatrix} \xi_{t} \\ \sigma_{t} \end{pmatrix} = \begin{pmatrix} \alpha(\xi_{t-1}, \sigma_{t-1}) + \beta(\xi_{t-1}, \sigma_{t-1})e_{t} \\ \beta(\xi_{t-1}, \sigma_{t-1}) \end{pmatrix} (\xi_{t}^{2} + \sigma_{t}^{2})^{1/2} + h(\xi_{t-1}, \sigma_{t-1}) \begin{pmatrix} e_{t} \\ 1 \end{pmatrix},$$
(44)

where $0 \le h(x,s) \le \max(b_{01}, b_{02})$. Analogous to the approach in section 4, we now let $\theta = (x, s)$ be in the half unit circle $\Theta = \{(x, s) : x^2 + s^2 = 1, s > 0\}$ and define

$$\zeta(\theta, u) = \begin{pmatrix} \alpha(\theta) + \beta(\theta)u\\ \beta(\theta) \end{pmatrix}$$
(45)

and $\eta(\theta, u) = \zeta(\theta, u)/||\zeta(\theta, u)||$. The collapsed Markov chain will again be

$$\theta_t^* = \eta(\theta_{t-1}^*, e_t), \tag{46}$$

and the results and methods of sections 4 and 5 will follow exactly as described there.

ACKNOWLEDGEMENTS

The author thanks Lian Liu and R. Jason Reed for early efforts in the computations, and Huay-min Pu for substantial collaboration in the theory underlying the approach given here.

NOTE

Corresponding author: Daren B. H. Cline, Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA. E-mail: dcline@stat.tamu.edu

REFERENCES

- BASRAK, B., DAVIS, R. and MIKOSCH, T. (2002) Regular variation of GARCH processes. *Stochastic Processes and Their Applications* 99, 95–115.
- BOLLERSLEV, T. (1986) Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–27.
- BORKOVEC, M. (2000) Extremal behavior of the autoregressive process with ARCH(1) errors. Stochastic Processes and thier Applications 85, 189–207.
- BORKOVEC, M. and KLÜPPELBERG, C. (2001) The tail of the stationary distribution of an autoregressive process with ARCH(1) errors. *Annals of Applied Probability* 11, 1220–41.

BOUGEROL, P. and PICARD, N. (1992a) Strict stationarity of generalized autoregressive processes. Annals of Probability 20, 1714–30.

- BOUGEROL, P. and PICARD, N. (1992b) Stationarity of GARCH processes and some nonnegative time series. *Journal of Econometrics* 52, 115–27.
- CLINE, D. B. H. (2005) Regular variation of order 1 nonlinear AR-ARCH models. Preprint, Statistics Department, Texas A&M University.
- CLINE, D. B. H. and PU, H. H. (1999) Stability of nonlinear AR(1) time series with delay. *Stochastic Processes and Their Applications* 82, 307–33.
- CLINE, D. B. H. and Pu, H. H. (2004) Stability and the Lyapounov exponent of threshold AR-ARCH models. Annals of Applied Probability 14, 1920–49.
- FRANCQ, C. and ZAKOÏAN, J.-M. (2005) The L^2 -structures of standard and switching regime GARCH models. Stochastic Processes and Their Applications 115, 1557–82.
- FRANCQ, C., ROUSSIGNOL, M. and ZAKOïAN, J.-M. (2001) Conditional heteroskedasticity driven by hidden Markov chains. *Journal of Time Series Analysis* 22, 197–220.
- GOLDIE, C. M. (1991) Implicit renewal theory and tails of solutions of random equations. *Annals of Applied Probability* 1, 126–66.
- HWANG, S. Y. and Woo, M.-J. (2001) Threshold ARCH(1) processes: asymptotic inference. *Statistics & Probability Letters* 53, 11–20.
- KLÜPPELBERG, C. and PERGAMENCHTCHIKOV, P. (2004) The tail of the stationary distribution of a random coefficient AR(q) model. *Annals of Applied Probability* 14, 971–1005.
- LANNE, M. and SAIKKONEN, P. (2005) Nonlinear GARCH models for highly persistent volatility. *Econometrics* 8, 251–76.
- LI, C. W. and LI, W. K. (1996) On a double-threshold autoregressive heteroscedastic time series model. *Journal of Applied Economics* 11, 253–74.
- LING, S. (1999) On the probabilistic properties of a double threshold ARMA conditional heteroskedastic model. *Journal of Applied Probability*, 36, 688–705.
- LING, S. (2004) Estimation and testing stationarity for double-autoregressive models. *Journal of the Royal Statistical Society* B 66, 63–78.
- LIU, J., LI, W. K. and LI, C. W. (1997) On a threshold autoregression with conditional heteroscedastic variances. *Journal of Statistical Planning and Inference* 62, 279–300.
- LU, Z. (1998) On the geometric ergodicity of a non-linear autoregressive model with an autoregressive conditional heteroscedastic term. *Statistica Sinica* 8, 1205–17.
- LU, Z. and JIANG, Z. (2001) L_1 geometric ergodicity of a multivariate nonlinear AR model with an ARCH term. *Statistics & Probability Letters* 51, 121–30.
- MASRY, E. and TJØSTHEIM, D. (1995) Nonparametric estimation and identification of nonlinear ARCH time series: strong convergence and asymptotic normality. *Econometric Theory* 11, 258–89.
- MEYN, S. P. and TWEEDIE, R. L. (1993) Markov Chains and Stochastic Stability. London: Springer-Verlag.
- RABEMANAJARA, R. and ZAKOÏAN, J. M. (1993) Threshold ARCH models and asymmetries in volatility. *Journal of Applied Econometrics* 8, 31–49.

R DEVELOPMENT CORE TEAM (2004) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

TONG, H. (1990) Non-linear Time Series Analysis: A Dynamical System Approach. London: Oxford University Press.

ZAKOïAN, J-M. (1994) Threshold heteroskedastic models. *Journal of Econometic Dynamics and Control* 18, 931–55.