ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Krämer, Walter

Working Paper

Long memory with Markov-Switching GARCH

Technical Report / Universität Dortmund, SFB 475 Komplexitätsreduktion in Multivariaten Datenstrukturen, No. 2006,35

Provided in cooperation with:
Technische Universität Dortmund

Suggested citation: Krämer, Walter (2006) : Long memory with Markov-Switching GARCH, Technical Report / Universität Dortmund, SFB 475 Komplexitätsreduktion in Multivariaten Datenstrukturen, No. 2006,35, http://hdl.handle.net/10419/22679

Mitglied der

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Leibniz-Gemeinschaft

Long memory with
Markov-Switching GARCH[1]

by

Walter Krämer

Fachbereich Statistik, Universität Dortmund, Germany
Phone xx231/755-3125, Fax: xx231/755-5284
e-mail: walterk@statistik.uni-dortmund.de

version October 2006

Summary

The paper considers the Markov-Switching GARCH(1,1)-model with time-varying transition probabilities. It derives sufficient conditions for the square of the process to display long memory and provides some additional intuition for the empirical observation that estimated GARCH-parameters often sum to almost one.

Keywords: Markov switching, GARCH, long memory

JEL-numbers: C13, C22

1 Introduction

The GARCH(1,1) - model

$$
\begin{aligned}
\epsilon_t &= \sigma_t \eta_t \\
\sigma_t^2 &= \omega + \alpha \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{1}
$$

is still the main workhorse in all areas of applied economics whenever conditional heteroskedasticity is seen to be a problem. Almost from the moment it was born, it was however plagued by the observation that in many applications, the observed empirical autocorrelations of the ε_t^2 were much larger than the theoretical autocorrelations implied by the estimated model parameters. In particular, the empirical autocorrelations of the ε_t^2 often seem to indicate long memory, which is not possible in the GARCH-model; in fact, in all standard GARCH-models, theoretical autocorrelations must eventually decrease exponentially, so long memory is ruled out.

The same holds for Markov-Switching GARCH models with constant transition probabilities, as suggested, among others, by Cai (1994), Hamilton and Susmel (1994), Francq et al. (2001), Klaassen (2002) or Haas et al. (2004). It is easily seen (for a proof see e.g. Francq and Zakoian (2005)) that theoretical autocorrelations must likewise decrease exponentially in such models. The present paper therefore allows for transition probabilities that change with sample size, along the lines of Diebold and Inoue (2001), and derives the limiting behaviour of the variance of the sum of the ϵ_t^2 as sample size increases. If the staying probabilities of the underlying Markov-process tend to 1, these autocorrelations lead to a variance of the sum of the ϵ_t^2 which grows faster than sample size, and thus induce the appearance of long memory.

2 Structural breaks and sample size

Most models that allow for changes in the coefficients of (1) do so by letting ω, α and β depend on the unobserved state of a finite - dimensional irreducible homogeneous Markov chain

$$\sigma_t^2 = \omega(\Delta_t) + \alpha(\Delta_t)\epsilon_{t-1}^2 + \beta(\Delta_t)\sigma_{t-1}^2, \quad \Delta_t \in \{1, ..., M\}, \tag{2}$$

$$P(\Delta_t = j | \Delta_{t-1} = i) = p_{ij} = \text{constant}. \tag{3}$$

Recent examples and variants thereof, with useful surveys of the literature, are Francq et al. (2001), Klaassen (2002) or Haas et al. (2004). Although theoretical autocorrelations of the ϵ_t^2 are notoriously hard to derive exactly for such models, it is clear from the Markov-structure that they must eventually decrease exponentially (if they exist; for details see Francq and Zakoian 2005), so these models cannot explain long memory in the squared ϵ_t's. Intuitively speaking, the reason is that the number of realized regimes in such models is roughly proportional to sample size. The present paper considers another type of asymptotics where the expected number of realized regimes remains bounded as sample size increases. The most simple example is the one considered by, among others, Mikosch and Starica (2004) or Hilebrand (2005), who divide the sample $\{1, ..., T\}$ into $K + 1$ subsamples

$$\{1, ..., [Td_1]\}, \{[Td_1 + 1], ..., [Td_2]\}, ..., \{[Td_K], ..., T\}, \tag{4}$$

where $0 < d_1 < ... < d_K < 1$ are fixed and $[Td_i]$ denotes the integer part of Td_i, and where different GARCH models hold in each subsample.

The present paper considers the Markov switching model (2), where however the transition probabilities p_{ij} depend on sample size. A similar set-up is investigated by Diebold and Inoue (2001), who explore Markov switching in the

expectation μ_t in the model $y_t = \mu_t + \epsilon_t$, while the present paper explores Markov switching with time dependent transition probabilities in the dynamics of the ϵ_t - process itself.

It is easily seen that with Markov-switching the expected number of regimes in a given time interval 1,...,T depends only on the staying probabilities p_{ii} :

$$
\begin{aligned}
E(\sum_{t=1}^{T-1} I_{\Delta_{t+1} \neq \Delta_t}) &= \sum_{t=1}^{T-1} P(\Delta_{t+1} \neq \Delta_t) \\
&= \sum_{t=1}^{T-1} \sum_{i=1}^{M} \pi(i) \sum_{\substack{j=1 \\ j \neq l}}^{M} p_{ij} = (T-1) \sum_{i=1}^{M} \pi_i(1 - p_{ii}),
\end{aligned}
\tag{5}
$$

where $\pi(i) = P(\Delta_t = i)$ denotes the stationary distribution of the Markov chain. In particular, if

$$
1 - p_{ii}^{(T)} = \frac{K}{(T-1)}
\tag{6}
$$

for all i and for some natural number K, we have

$$
E(\text{number of regimes}) = K.
$$

If $1 - p_{ii}^{(T)}$ tends to zero more slowly than in (6), the expected number of regimes will increase with sample size, but will still be $o(T)$, which provides the intuition behind our result below that time-dependent staying probabilities of type (6) imply the appearance of long memory in the squares of the ε_t - process.

A similar set-up was studied by Hilebrand (2005), who shows by different arguments that, with finitely many nonstochastic switches in regime along the lines of (4), the estimated persistence parameters $\hat{\lambda} = \hat{\alpha} + \hat{\beta}$ must tend to unity as sample size increases.

4

3 Structural change and long memory

There are various definitions of long memory in ϵ_t^2 (see e.g. Diebold and Inoue 2001, p. 133, for an overview). The most general of these simply requires that $E(\epsilon_t^2) < \infty$, but

$$\frac{1}{T} \; var(\sum_{t=1}^{T} \epsilon_t^2) \to \infty \;\; as \; T \to \infty. \tag{7}$$

For the case of a stationary ϵ_t^2 - process, this is equivalent to the spectral density of ϵ_t^2 tending to infinity for small frequencies.

Below we consider a triangle sequence of models

$$\epsilon_1^{(1)}$$
$$\epsilon_1^{(2)}, \epsilon_2^{(2)}$$
$$\vdots$$
$$\epsilon_1^{(T)}, ..., \epsilon_T^{(T)}$$

where the sequence $\epsilon_1^{(T)}, ..., \epsilon_T^{(T)}$ is generated by a Markov switching model (2) and where the transition matrices depend on T in such a way that the expected number of regimes remains bounded away from both zero and infinity (The superscript T will be omitted in the sequel whenever there is no danger of confusion). From Francq and Zakoian (2005, theorem 3.1) we obtain the covariance matrix of $(\epsilon_1^2, ..., \epsilon_T^2)'$ as a function of the GARCH parameters $\omega(i), \alpha(i), \beta(i)$ and of the transition probabilities p_{ij} $(i, j = 1, ..., M)$. In particular, Francq and Zakoian (2005) show that $\{\epsilon_t^2\}$, which follows an ARMA(1,1) process

$$\epsilon_t^2 = \omega + (\alpha + \beta)\epsilon_{t-1}^2 + u_t - \beta u_{t-1} \tag{8}$$

when there is no Markov-switching, where $u_t := \epsilon_t^2 - \sigma_t^2$ continues to follow an ARMA-process, albeit with different parameters and different orders, when

Markov-switching in the GARCH - parameters is introduced. The autocorrelation function $\gamma(l)$ of this process satisfies a linear difference equation of the form

$$\sum_{i=0}^{n-n_0} a_{n-i}\gamma(l-i) = 0 \; , \; (l > n - n_0), \tag{9}$$

where n_0, $0 \leq n_0 \leq n$, is the index of the first nonzero coefficient of the polynomial

$$f(x) = \prod_{k=1}^{K}(\lambda_k - x)g(x), \tag{10}$$

where the $\lambda'_k s$ ($k = 1, ..., K < M$) are the eigenvalues - in increasing order - of the transition matrix of the Markov process which are different from 0 and 1, and where $g(x)$ is the characteristic polynomial of the matrix

$$\tilde{P} := \begin{pmatrix} p_{11}[\alpha(1) + \beta(1)] & \ldots & p_{M1}[\alpha(1) + \beta(1)] \\ \vdots & & \vdots \\ p_{1M}[\alpha(M) + \beta(M)] & \ldots & p_{MM}[\alpha(M) + \beta(M)] \end{pmatrix} \tag{11}$$

(see Francq and Zakoian (2005), formula 3.10). This implies that one can express the autocorrelations $\gamma(l)$ of the ε_t^2 - process as

$$\gamma(l) = \sum_{k=1}^{K} c_k \lambda_k^l + \sum_{k=1}^{M} \tilde{c}_k \tilde{\lambda}_k^l \tag{12}$$

for suitable coefficients c_k and $\tilde{c}_k$, where the $\tilde{\lambda}_k$ are the eigenvalues of the matrix (11). From $\alpha(i) + \beta(i) < 1$ ($i = 1, ..., M$), these eigenvalues are all smaller than the largest eigenvalue (different from unity) λ_K of the transition matrix of the underlying Markov process, which implies that

$$\gamma(l) = c_K \lambda_K^l + o([\gamma(l)]). \tag{13}$$

Therefore, the limiting behavior of $var(\sum_{t=1}^{T} \epsilon_t^2)$ depends only on λ_K and is, for a fixed λ_K, given by

$$
\begin{aligned}
var(\sum_{t=1}^{T} \epsilon_t^2) &= O(\sum_{i,j=1}^{T} \lambda_K^{|i-j|}) = O(T[1 + 2\sum_{i=1}^{T-i} \frac{T-i}{T} \lambda_K^i]) \\
&= O(T\frac{1}{1-\lambda_K}).
\end{aligned}
\tag{14}
$$

However, from elementary calculus, it is easily seen that (14) continues to apply if λ_K depends on T, as long as $(1 - \lambda_K)/T = O(1)$. In particular,

$$
var(\sum_{t=1}^{T} \epsilon_t^2) = O(T^{2d+1})
\tag{15}
$$

whenever

$$
1 - \lambda_K = O(T^{-2d}).
\tag{16}
$$

In the conventional notation of the long-memory literature (see e.g. Diebold and Inoue 2001, p.133), this then implies that ϵ_t^2 behaves as if it were $I(d)$.

The largest eigenvalue different from unity of the transition matrix connects to the staying probabilities p_{ii} via

$$
\lambda_K \geq (\sum_{i=1}^{M} p_{ii} - 1)/(M - 1),
\tag{17}
$$

so $\lambda_K \longrightarrow 1$ whenever $p_{ii} \longrightarrow 1$.

4   Discussion

The argument above has taken the existence of $var(\epsilon_t^2)$ for granted. Francq and Zakoian (2005, section 2) give necessary and sufficient conditions for the

existence of higher moments of Markov-switching GARCH models which depend on the moments of the innovations η_t, on the GARCH-parameters and on the transition matrix of the Markov process. As the present paper is only concerned with conditions that lead to the appearance of long memory in otherwise standard situations, such subtleties are here ignored.

Another issue concerns the estimated persistence parameter $\hat{\alpha}+\hat{\beta}$ when Markov switching is ignored when estimating the model (1). It has long been known that $\hat{\alpha}+\hat{\beta}$ is then biased upwards towards unity, and from Krämer and Tameze (2006) we see that

$$\hat{\alpha} + \hat{\beta} \quad \xrightarrow{p} \quad 1 \quad as \quad T \to \infty$$

whenever $d \geq \frac{1}{2}$ and estimation is done with the Baillie-Chung (2001) minimum distance estimator. This is so because empirical autocorrelations of $I(d)$ - processes for $d \geq \frac{1}{2}$ tend to 1 in probability as $T \to \infty$ and the distance between theoretical and empirical autocorrelation of the ε_t^2 - process is then minimized whenever $\hat{\alpha} + \hat{\beta} = 1$.

References

Baillie, R.T. and Chung, H. (2001): "Estimation of GARCH models from the autocorrelation of the squares of a process." *Journal of Time Series Analysis* 22, 631 – 650.

Cai, J. (1994): "A Markov model of switching-regime ARCH." *Journal of Business and Economic Statistics* 12,309 – 316.

Diebold, F.X. and Inoue, A. (2001): "Long memory and regime switching." *Journal of Econometrics* 105, 131 – 159.

Francq, C., Roussignol, M. and Zakoian, J.-M. (2001): "Conditional heteroscedasticity driven by hidden Markov chains." *Journal of Time Series Analysis* 22, 197 – 220.

Francq, C. and Zakoian, J.-M. (2005): "L^2-structures of standard and switching regime GARCH models." *Stochastic Processes and their Applications* 115, 1557 – 1582.

Haas, M., Mittnik, S. and Paolella, M. S. (2004): "A new approach to Markov-switching GARCH models." *Journal of Financial Econometrics* 2, 493 – 530.

Hamilton, J. D., Susmel, R. (1994): "Autoregressive conditional heteroscedasticity and changes in regime." *Journal of Econometrics* 64, 307–333.

Hilebrand, E. (2005): "Neglecting parameter changes in GARCH models." *Journal of Econometrics* 129, 121–138.

Klaassen, F. (2002): "Improving GARCH volatility forecasts with regime-switching GARCH." *Empirical Economics* 27, 363 – 394.

Krämer, W. and Tameze, B. (2006): "Structural change and estimated persistence in the GARCH(1,1)-model." *SFB* 475 *discussion paper No. 33/06, Universität Dortmund.*

Mikosch, T. and Starica, C. (2004): "Nonstationarities in Financial Time Series, Long-Range Dependence, and the IGARCH Effect." *Review of Economics and Statistics* 86, 378 – 390.