



HAL
open science

The autocorrelation structure of the Markov-switching asymmetric power GARCH process

Markus Haas

► **To cite this version:**

Markus Haas. The autocorrelation structure of the Markov-switching asymmetric power GARCH process. *Statistics and Probability Letters*, 2010, 78 (12), pp.1480. 10.1016/j.spl.2007.12.025 . hal-00634357

HAL Id: hal-00634357

<https://hal.science/hal-00634357>

Submitted on 21 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

The autocorrelation structure of the Markov-switching asymmetric power GARCH process

Markus Haas

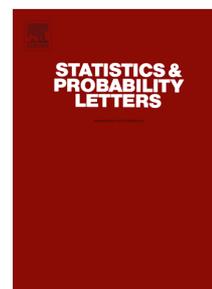
PII: S0167-7152(07)00447-6
DOI: [10.1016/j.spl.2007.12.025](https://doi.org/10.1016/j.spl.2007.12.025)
Reference: STAPRO 4865

To appear in: *Statistics and Probability Letters*

Received date: 15 August 2007
Revised date: 3 December 2007
Accepted date: 26 December 2007

Please cite this article as: Haas, M., The autocorrelation structure of the Markov-switching asymmetric power GARCH process. *Statistics and Probability Letters* (2008), doi:[10.1016/j.spl.2007.12.025](https://doi.org/10.1016/j.spl.2007.12.025)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



The Autocorrelation Structure of the Markov-switching Asymmetric Power GARCH Process

Markus Haas
Department of Statistics, University of Munich
Akademiestrasse 1
D-80799 Munich
Germany
Phone: +49 0 89 21 80 25 70
Fax: +49 0 89 21 80 50 44
E-mail: haas@stat.uni-muenchen.de

The author is grateful for constructive comments from an associate editor, which led to significant improvements of the paper. The research for this paper was supported by the German Research Foundation (DFG).

Abstract

In a recent paper, Liu (2007) proposed a generalization of the Markov-switching GARCH model of Haas et al. (2004) to allow both for a nonlinear relation between past shocks and future volatility as well as for the leverage effect, which refers to the observation that stock market volatility reacts differently to positive and negative shocks. For the new model, Liu (2007) derived conditions for stationarity and the existence of moments. This article complements Liu's (2007) results in two directions. First, a simple method for calculating the moments and the autocorrelation structure of the power-transformed absolute process is devised, which is of vital interest in applied GARCH modeling. Second, in an application to stock returns, the relevance of the extended Markov-switching GARCH process proposed by Liu (2007), as compared to simpler versions, is illustrated.

Keywords— Autocorrelations, Conditional volatility, GARCH, Leverage effect, Markov-switching, Stock market

1 Introduction

In a recent paper, building on work of Hwang and Basawa (2004), Liu (2007) provided a generalization of the Markov-switching GARCH model of Haas et al. (2004) to allow both for a nonlinear relation between past shocks and future volatility as well as for the leverage effect, which refers to the observation that stock market volatility reacts differently to positive and negative shocks. For the new model, Liu (2007) derived conditions for stationarity and the existence of moments.

The present article complements Liu's (2007) work in two directions. First, we suggest a representation of the process which admits straightforward computation of the unconditional moments and the dynamic autocorrelation structure of the power-transformed absolute returns (residuals), which are taken as a measure of volatility. Such results are of vital interest in the context of GARCH modeling, because it is the desire to capture the dependence structure of these measures of financial volatility which motivates the use of GARCH models in the first place, and, therefore, researchers will want to assess how well the dynamic properties of an estimated model mimic the corresponding quantities directly estimated from the data under consideration. Second, in an application to stock returns, we compare the fit of Liu's (2007) model with several simpler single-regime and Markov-switching GARCH specifications to illustrate the usefulness of the extended model. To the best of our knowledge, Liu's model has not been confronted with real data so far.

2 The Markov-switching Asymmetric Power GARCH Model (MS-APGARCH)

Time series $\{\epsilon_t, t \in \mathbb{Z}\}$ is generated by a k -regime Markov-switching asymmetric power GARCH process, or, in short, MS-APGARCH(k), if it can be described by

$$\epsilon_t = \eta_t \sigma_{\Delta_t, t}, \quad t \in \mathbb{Z}, \quad (1)$$

where $\{\eta_t, t \in \mathbb{Z}\}$ is an iid sequence of symmetric zero mean random variables, and $\{\Delta_t, t \in \mathbb{Z}\}$ is a Markov chain with finite state space $S = \{1, \dots, k\}$ and primitive (i.e., irreducible and aperiodic) transition matrix, P , with typical element $p_{ij} = p(\Delta_t = j | \Delta_{t-1} = i)$, that is,

$$P = [p_{ij}] = [p(\Delta_t = j | \Delta_{t-1} = i)], \quad i, j = 1, \dots, k. \quad (2)$$

In addition, $\{\eta_t, t \in \mathbb{Z}\}$ and $\{\Delta_t, t \in \mathbb{Z}\}$ are assumed to be independent. The stationary distribution of the Markov chain will be denoted by $\pi_\infty = (\pi_{1,\infty}, \dots, \pi_{k,\infty})'$. We note that

the assumption of η_t being symmetric is by no means crucial, but it holds for all candidate densities commonly employed in regime-switching GARCH models (in particular, Gaussian and Student's t), and it helps to simplify the notation below.

In the parametrization adopted in Liu (2007), the conditional variance of regime j , σ_{jt}^2 , is driven by a univariate asymmetric power GARCH (APGARCH) equation of the form

$$\sigma_{jt}^\delta = \omega_j + \theta_{1j}|\epsilon_{t-1}^+|^\delta + \theta_{2j}|\epsilon_{t-1}^-|^\delta + \beta_j\sigma_{j,t-1}^\delta, \quad \delta > 0, \quad (3)$$

where $\omega_j > 0$, $\theta_{1j}, \theta_{2j}, \beta_j \geq 0$, $j = 1, \dots, k$, $\epsilon_t^+ = \max\{0, \epsilon_t\}$, and $\epsilon_t^- = \min\{0, \epsilon_t\}$. Equation (3) nests the MS-GARCH(k) process of Haas et al. (2004) for $\delta = 2$ and $\theta_{1j} = \theta_{2j}$, $j = 1, \dots, k$. The introduction of the power parameter δ in (3), which dates back to Higgins and Bera (1992) and Ding et al. (1993), is rooted in the observation of Ding et al. (1993) that “there is no obvious reason why one should assume the conditional variance is a linear function of lagged squared returns (residuals)”, and, in fact, the time series dependencies of power-transformed residuals, $|\epsilon_t|^\delta$, are often found to be strongest for $\delta \approx 1$. The rationale behind the allowance for asymmetry in (3), i.e., $\theta_{1j} \neq \theta_{2j}$, is an empirical regularity known as the leverage effect, i.e., the “stylized fact” that, for stock returns, past negative shocks have a deeper impact on volatility than positive shocks of the same magnitude, which corresponds to $\theta_{1j} < \theta_{2j}$ in (3) (for a deeper discussion of the leverage effect and its economic interpretation, see, e.g., Bekaert and Wu, 2000; and Bouchaud et al., 2001). The idea of capturing the asymmetries in volatility as in (3) goes back to Glosten et al. (1993) and Zakoïan (1994).

In this article, we find a different parametrization of (3) more convenient. Namely, we use the APGARCH structure of Ding et al. (1993), which is given by

$$\sigma_{jt}^\delta = \omega_j + \alpha_j(|\epsilon_{t-1}| - \gamma_j\epsilon_{t-1})^\delta + \beta_j\sigma_{j,t-1}^\delta, \quad \delta > 0, \quad (4)$$

where $\omega_j > 0$, $\alpha_j, \beta_j \geq 0$, and $\gamma_j \in (-1, 1)$, $j = 1, \dots, k$. Clearly (3) and (4) are equivalent if $\theta_{1j} = \alpha_j(1 - \gamma_j)^\delta$ and $\theta_{2j} = \alpha_j(1 + \gamma_j)^\delta$, $j = 1, \dots, k$. For applications of the single-regime APGARCH model of Ding et al. (1993), where $k = 1$, see, e.g., Brooks et al. (2000), Giot and Laurent (2003), and Brooks (2007).

3 Autocorrelation Structure

To derive the autocorrelation structure of the power-transformed process, i.e.,

$$\varrho(\tau; \delta) := \text{Corr}(|\epsilon_{t-\tau}|^\delta, |\epsilon_t|^\delta) = \frac{\text{E}(|\epsilon_t|^\delta |\epsilon_{t-\tau}|^\delta) - \text{E}^2(|\epsilon_t|^\delta)}{\text{E}(|\epsilon_t|^{2\delta}) - \text{E}^2(|\epsilon_t|^\delta)}, \quad (5)$$

we introduce the following matrices:

$$X_t = \begin{pmatrix} \sigma_{1t}^\delta \\ \sigma_{2t}^\delta \\ \vdots \\ \sigma_{kt}^\delta \\ |\epsilon_{t-1}|^\delta \end{pmatrix}, \quad \tilde{\omega} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_k \\ 0 \end{pmatrix}, \quad \tilde{\alpha} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \\ 1 \end{pmatrix}, \quad \tilde{\eta}_t = \begin{pmatrix} (|\eta_t| - \gamma_1 \eta_t)^\delta \\ (|\eta_t| - \gamma_2 \eta_t)^\delta \\ \vdots \\ (|\eta_t| - \gamma_k \eta_t)^\delta \\ |\eta_t|^\delta \end{pmatrix},$$

$$A_t = \tilde{\alpha} \odot \tilde{\eta}_t, \quad \tilde{\gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_k \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} \beta_1 & 0 & \cdots & 0 & 0 \\ 0 & \beta_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & \beta_k & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix},$$

where \odot denotes the Hadamard product, i.e., the elementwise multiplication of conformable matrices. Moreover, let e_j be the j th unit vector in \mathbb{R}^{k+1} , $j = 1, \dots, k+1$. Then we can write the process as

$$X_t = \tilde{\omega} + (A_{t-1} e'_{\Delta_{t-1}} + B) X_{t-1}. \quad (6)$$

Representation (6) differs from that in Liu (2007) insofar as $|\epsilon_{t-1}|^\delta$ has been included into the state vector X_t . It turns out that this considerably simplifies the computation of the autocorrelation structure of the process.

Assume that $E(|\eta_t|^{2\delta}) < \infty$, and define $\kappa(\mu) = E(|\eta_t|^\mu)$, $\mu > 0$. For example, under Gaussianity of η_t (as assumed in the example in Section 5), we have

$$\kappa(\mu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |\eta|^\mu e^{-\eta^2/2} d\eta = \frac{2^{\mu/2}}{\sqrt{\pi}} \Gamma\left(\frac{\mu+1}{2}\right).$$

Note that, by symmetry,

$$\bar{A} := E(A_t) = \frac{\kappa(\delta)}{2} \left\{ \tilde{\alpha} \odot [(\mathbf{1}_{k+1} + \tilde{\gamma})^{(\delta)} + (\mathbf{1}_{k+1} - \tilde{\gamma})^{(\delta)}] \right\},$$

and, using the easily checked identity $(\tilde{\alpha} \odot \tilde{\eta}_t) \otimes (\tilde{\alpha} \odot \tilde{\eta}_t) = (\tilde{\alpha} \otimes \tilde{\alpha}) \odot (\tilde{\eta}_t \otimes \tilde{\eta}_t)$,

$$\begin{aligned} \overline{A \otimes A} &:= E(A_t \otimes A_t) = (\tilde{\alpha} \otimes \tilde{\alpha}) \odot E(\tilde{\eta}_t \otimes \tilde{\eta}_t) \\ &= \frac{\kappa(2\delta)}{2} (\tilde{\alpha} \otimes \tilde{\alpha}) \odot [(\mathbf{1}_{k+1} + \tilde{\gamma})^{(\delta)} \otimes (\mathbf{1}_{k+1} + \tilde{\gamma})^{(\delta)} + (\mathbf{1}_{k+1} - \tilde{\gamma})^{(\delta)} \otimes (\mathbf{1}_{k+1} - \tilde{\gamma})^{(\delta)}],
 \end{aligned}$$

where, for an $n \times 1$ vector x , $x^{(\delta)} := (x_1^\delta, \dots, x_n^\delta)'$, and $\mathbf{1}_n$ is an n -dimensional column of ones.

Representation (6) is a first-order vector difference equation with stochastic coefficient matrix. The (regime-specific) expectations of this matrix and its second-order Kronecker power are needed for determining whether the process has a stationary solution with a finite

(2δ)th-order moment, as required for the autocorrelation function (5) to be well-defined (see Proposition 3.1), as well as for the computation of the unconditional moments in Proposition 3.2. By the independence of $\{\eta_t\}$ and $\{\Delta_t\}$, we have

$$C_{11}(j) := E[(A_t e'_{\Delta_t} + B) | \Delta_t = j] = \bar{A} e'_j + B, \quad (7)$$

$$C_{21}(j) := \tilde{\omega} \otimes C_{11}(j) + C_{11}(j) \otimes \tilde{\omega},$$

$$\begin{aligned} C_{22}(j) &:= E[(A_t e'_{\Delta_t} + B) \otimes (A_t e'_{\Delta_t} + B) | \Delta_t = j] \\ &= (\bar{A} \otimes \bar{A})(e_j \otimes e_j)' + (\bar{A} e'_j) \otimes B + B \otimes (\bar{A} e'_j) + B \otimes B, \end{aligned} \quad (8)$$

$$j = 1, \dots, k,$$

and, borrowing notation from Francq and Zakoian (2005), we define the matrices

$$\mathbb{P}_{C_{\ell m}} = \begin{pmatrix} p_{11} C_{\ell m}(1) & p_{21} C_{\ell m}(1) & \cdots & p_{k1} C_{\ell m}(1) \\ p_{12} C_{\ell m}(2) & p_{22} C_{\ell m}(2) & \cdots & p_{k2} C_{\ell m}(2) \\ \vdots & \vdots & \ddots & \vdots \\ p_{1k} C_{\ell m}(k) & p_{2k} C_{\ell m}(k) & \cdots & p_{kk} C_{\ell m}(k) \end{pmatrix} \quad (9)$$

for $\ell = 1, 2$ and $m \leq \ell$.

Now let $\rho(A)$ denote the spectral radius of a square matrix A , i.e.,

$$\rho(A) := \max\{|z| : z \text{ is an eigenvalue of } A\}.$$

Assumption 3.1 $E(|\eta_t|^{2\delta}) < \infty$, $\rho(\mathbb{P}_{C_{11}}) < 1$, and $\rho(\mathbb{P}_{C_{22}}) < 1$, where matrices $\mathbb{P}_{C_{11}}$ and $\mathbb{P}_{C_{22}}$ are defined via (9) along with (7) and (8).

Assumption 3.1 ensures that the first two unconditional fractional moments of the process, i.e., $E(|\epsilon_t|^\delta)$ and $E(|\epsilon_t|^{2\delta})$, are finite, so that the autocorrelations (5) are well-defined. This is stated in the following proposition due to Liu (2007).

Proposition 3.1 (Liu, 2007, Corollary 2.1 and Theorem 3.2) *If Assumption 3.1 holds, then the MS-APGAR $\text{CH}(k)$ process defined by (1), (2), and (4) has a unique strictly stationary solution with finite (2δ)th-order moment.*

Assumption 3.1 therefore enables us to compute the autocorrelation structure of the power-transformed absolute process, where use will be made of the following lemma (cf. Francq and Zakoian, 2005, Lemma 3).

Lemma 3.1 *For $\ell \geq 1$, if the variable $Y_{t-\ell}$ belongs to the information set generated by $\Psi_{t-\ell} := \{\epsilon_s : s \leq t - \ell\}$, then*

$$\pi_{j,\infty} E(Y_{t-\ell} | \Delta_t = j) = \sum_{i=1}^k \pi_{i,\infty} p_{ij}^{(\ell)} E(Y_{t-\ell} | \Delta_{t-\ell} = i),$$

where the $p_{ij}^{(\ell)} := p(\Delta_t = j | \Delta_{t-\ell} = i)$, $i, j = 1, \dots, k$, denote the ℓ -step transition probabilities, as given by the elements of P^ℓ .

Lemma 3.1 is based on the basic relation

$$E(Y_{t-\ell} | \Delta_t = j) = \sum_{i=1}^k p(\Delta_{t-\ell} = i | \Delta_t = j) E(Y_{t-\ell} | \Delta_{t-\ell} = i \cap \Delta_t = j),$$

along with the identity $\pi_{j,\infty} p(\Delta_{t-\ell} = i | \Delta_t = j) = \pi_{i,\infty} p(\Delta_t = j | \Delta_{t-\ell} = i) = \pi_{i,\infty} p_{ij}^{(\ell)}$, and essentially says that “given the regime at time t , the expectation of any variable determined by the information up to time t does not depend on the future regime history”, i.e., $E(Y_{t-\ell} | \Delta_{t-\ell} = i \cap \Delta_t = j) = E(Y_{t-\ell} | \Delta_{t-\ell} = i)$.

Proposition 3.2 *If Assumption 3.1 holds, then the first two fractional moments of the stationary distribution of the MS-APGARCH(k) are given by*

$$E(|\epsilon_t|^\delta) = (\mathbf{1}'_k \otimes e'_{k+1})Q, \quad (10)$$

and

$$E(|\epsilon_t|^{2\delta}) = (\mathbf{1}'_k \otimes e'_{k+1} \otimes e'_{k+1})R, \quad (11)$$

where Q and R are defined by the equations

$$Q = \pi_\infty \otimes \tilde{\omega} + \mathbb{P}_{C_{11}}Q, \quad (12)$$

and

$$R = \pi_\infty \otimes \tilde{\omega} \otimes \tilde{\omega} + \mathbb{P}_{C_{21}}Q + \mathbb{P}_{C_{22}}R, \quad (13)$$

respectively. Furthermore, for $\tau \geq 1$,

$$E(|\epsilon_t|^\delta | \epsilon_{t-\tau} |^\delta) = (\mathbf{1}'_k \otimes e'_{k+1})S(\tau)e_{k+1}, \quad (14)$$

where the sequence of $k(k+1) \times (k+1)$ matrices $S(\tau)$ is defined by the recursion

$$S(\tau) = (P^\tau \otimes \tilde{\omega})\tilde{Q} + \mathbb{P}_{C_{11}}S(\tau-1), \quad \tau \geq 1, \quad (15)$$

and the $k \times (k+1)$ matrix \tilde{Q} and the $k(k+1) \times (k+1)$ matrix $S(0)$ are obtained by reshaping the vectors Q and R , respectively, as indicated in the proof.

Proof. Using the representation (6) and the fact that X_t belongs to the information set generated by $\Psi_{t-1} = \{\epsilon_s : s \leq t-1\}$, Lemma 3.1 gives

$$\pi_{j,\infty} E(X_t | \Delta_{t-1} = j) = \pi_{j,\infty} \tilde{\omega} + \sum_{i=1}^k p_{ij} C_{11}(j) \pi_{i,\infty} E(X_{t-1} | \Delta_{t-2} = i), \quad j = 1, \dots, k. \quad (16)$$

Now define the $k(k+1) \times 1$ vector Q as

$$Q = (\pi_{1,\infty}E(X_t|\Delta_{t-1} = 1)', \dots, \pi_{k,\infty}E(X_t|\Delta_{t-1} = k)')'$$

Then (16) implies (12), and (10) follows from $E(X_t) = \sum_j \pi_{j,\infty}E(X_t|\Delta_{t-1} = j) = (\mathbf{1}'_k \otimes I_{k+1})Q$, and $E(|\epsilon_t|^\delta) = e'_{k+1}E(X_t) = (1 \otimes e'_{k+1})(\mathbf{1}'_k \otimes I_{k+1})Q = (\mathbf{1}'_k \otimes e'_{k+1})Q$, using $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$ for conformable matrices A, B, C , and D , and where I_n denotes the identity matrix of dimension n .

Analogously, using the identities $\text{vec}(xy') = y \otimes x$ and $\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$, we obtain

$$\begin{aligned} \pi_{j,\infty}E[\text{vec}(X_t X'_t)|\Delta_{t-1} = j] &= \pi_{j,\infty}(\tilde{\omega} \otimes \tilde{\omega}) + \sum_{i=1}^k p_{ij}C_{21}(j)\pi_{i,\infty}E(X_{t-1}|\Delta_{t-2} = i) \\ &\quad + \sum_{i=1}^k p_{ij}C_{22}(j)\pi_{i,\infty}E[\text{vec}(X_{t-1}X'_{t-1})|\Delta_{t-2} = i], \\ j &= 1, \dots, k. \end{aligned} \quad (17)$$

Now let, analogously to Q , R be the $k(k+1)^2 \times 1$ vector with elements $\pi_{j,\infty}E[\text{vec}(X_t X'_t)|\Delta_{t-1} = j]$, $j = 1, \dots, k$. Then (17) establishes (13), and (11) follows from

$$\begin{aligned} E(|\epsilon_t|^{2\delta}) &= e'_{k+1}E(X_t X'_t)e_{k+1} = (e'_{k+1} \otimes e'_{k+1})E[\text{vec}(X_t X'_t)] \\ &= (e'_{k+1} \otimes e'_{k+1})(\mathbf{1}'_k \otimes I_{(k+1)^2})R = (\mathbf{1}'_k \otimes e'_{k+1} \otimes e'_{k+1})R. \end{aligned}$$

To calculate $E(|\epsilon_t|^\delta|\epsilon_{t-\tau}|^\delta)$, we observe from the representation (6) that, for $\tau \geq 1$, and $j = 1, \dots, k$,

$$\begin{aligned} \pi_{j,\infty}E(X_t X'_{t-\tau}|\Delta_{t-1} = j) &= \tilde{\omega}\pi_{j,\infty}E(X'_{t-\tau}|\Delta_{t-1} = j) \\ &\quad + C_{11}(j)\pi_{j,\infty}E(X_{t-1}X'_{t-\tau}|\Delta_{t-1} = j). \end{aligned} \quad (18)$$

From Lemma 3.1,

$$\pi_{j,\infty}E(X'_{t-\tau}|\Delta_{t-1} = j) = \sum_{i=1}^k \pi_{i,\infty}p_{ij}^{(\tau)}E(X'_{t-\tau}|\Delta_{t-\tau-1} = i), \quad j = 1, \dots, k, \quad (19)$$

and

$$\pi_{j,\infty}E(X_{t-1}X'_{t-\tau}|\Delta_{t-1}) = \sum_{i=1}^k \pi_{i,\infty}p_{ij}E(X_{t-1}X'_{t-\tau}|\Delta_{t-2} = i), \quad j = 1, \dots, k. \quad (20)$$

Let $S(\tau)$, $\tau = 0, 1, \dots$, be the $k(k+1) \times (k+1)$ matrix where the $(k+1) \times (k+1)$ matrices $\pi_{j,\infty}E(X_t X'_{t-\tau}|\Delta_{t-1} = j)$, $j = 1, \dots, k$, are arranged one underneath the other. Then $S(0)$ is obtained by reshaping R . Furthermore, let \tilde{Q} be the $k \times (k+1)$ matrix where $\pi_{j,\infty}E(X_t|\Delta_{t-1} =$

j) in Q is replaced with its transpose, i.e., $\pi_{j,\infty}E(X'_t|\Delta_{t-1} = j)$. Then (18)–(20) imply (15), and we have

$$E(|\epsilon_t|^\delta|\epsilon_{t-\tau}|^\delta) = e'_{k+1}E(X_t X'_{t-\tau})e_{k+1} = e'_{k+1}(\mathbf{1}'_k \otimes I_{k+1})S(\tau)e_{k+1} = (\mathbf{1}'_k \otimes e'_{k+1})S(\tau)e_{k+1},$$

showing (14). ■

4 Simple Cases

For some concreteness, assume $k = 2$, $\omega_1 \neq \omega_2$, $\alpha_1 = \alpha_2 =: \alpha$, $\beta_1 = \beta_2 = 0$, $\gamma_1 = \gamma_2 = 0$, $\delta = 2$, and normally distributed innovations η_t . This corresponds to a special case of a Markov-switching ARCH process, where only the intercept in the variance equation is subject to regime-switching, as considered in the classic paper of Cai (1994). Straightforward calculations show that, for this model, if $\alpha < 1$,

$$E(\epsilon_t^2) = \frac{\pi_{1,\infty}\omega_1 + \pi_{2,\infty}\omega_2}{1 - \alpha}, \quad (21)$$

which has already been calculated by Cai (1994). Moreover, if $\alpha < \sqrt{1/3} \approx 0.577$, the fourth moment of the unconditional distribution is

$$E(\epsilon_t^4) = \frac{3(\pi_{1,\infty}\omega_1 + \pi_{2,\infty}\omega_2)^2(1 + \alpha)}{(1 - \alpha)(1 - 3\alpha^2)} + \frac{3(1 + \delta\alpha)}{1 - \delta\alpha} \frac{\pi_{1,\infty}\pi_{2,\infty}(\omega_1 - \omega_2)^2}{1 - 3\alpha^2}, \quad (22)$$

where $\delta := p_{11} + p_{22} - 1$ is the degree of regime persistence. Expressions (21) and (22) can be compared with the moments of the single-regime ARCH(1) process of Engle (1982), i.e., $\epsilon_t \sim N(0, \sigma_t^2)$, $\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2$, which has second and fourth moments given by $\omega/(1 - \alpha)$ and $3\omega^2(1 + \alpha)(1 - \alpha)^{-1}(1 - 3\alpha^2)^{-1}$, respectively. Suppose that $\omega = \pi_{1,\infty}\omega_1 + \pi_{2,\infty}\omega_2$. Then both processes have the same unconditional variance, but we observe from (22) that the fourth moment of the regime-switching process, and hence its kurtosis, is greater than that of the single-regime ARCH model. Interestingly, the difference between the unconditional fourth moments of the models increases not only with the difference between ω_1 and ω_2 but also with the persistence of the regimes, δ . The covariance function of this Markov-switching ARCH process is

$$\text{Cov}(\epsilon_t^2, \epsilon_{t-\tau}^2) = \alpha^\tau [E(\epsilon_t^4) - E^2(\epsilon_t^2)] + \delta \frac{\delta^\tau - \alpha^\tau}{\delta - \alpha} \frac{\pi_{1,\infty}\pi_{2,\infty}(\omega_1 - \omega_2)^2}{1 - \delta\alpha},$$

which reduces to $\text{Cov}(\epsilon_t^2, \epsilon_{t-\tau}^2) = \delta^\tau \pi_{1,\infty}\pi_{2,\infty}(\omega_1 - \omega_2)^2$ for the basic Markov-switching process with constant regime-specific variances, where $\alpha = 0$, as calculated by Timmermann (2000). Also, for $\omega_1 = \omega_2$, we have the single-regime ARCH(1) process with $\text{Corr}(\epsilon_t^2, \epsilon_{t-\tau}^2) = \alpha^\tau$.

Another special case is the single-regime APGARCH process, where $k = 1$, which has been investigated in He and Teräsvirta (1999), Hwang and Basawa (2004), and Karanasos and Kim (2006). For this model, provided that $c_{22} < 1$,

$$\mathbb{E}(|\epsilon_t|^\delta) = \kappa(\delta)\mathbb{E}(\sigma_t^\delta) = \frac{\kappa(\delta)\omega_1}{1 - c_{11}}, \quad \mathbb{E}(|\epsilon_t|^{2\delta}) = \kappa(2\delta)\mathbb{E}(\sigma_t^{2\delta}) = \frac{\kappa(2\delta)(1 + c_{11})\omega_1^2}{(1 - c_{11})(1 - c_{22})},$$

where $c_{ii} = \mathbb{E}\{\alpha_1(|\eta_t| - \gamma_1\eta_t)^\delta + \beta_1\}^i$, $i = 1, 2$, and

$$\text{Cov}(|\epsilon_t|^\delta, |\epsilon_{t-\tau}|^\delta) = c_{11}^{\tau-1} \{ \kappa(\delta)[\kappa(\delta)\beta_1 + \tilde{\kappa}(\delta; \gamma_1)\alpha_1]\mathbb{E}(\sigma_t^{2\delta}) - c_{11}\mathbb{E}^2(|\epsilon_t|^\delta) \},$$

where $\tilde{\kappa}(\delta; \gamma_1) = \mathbb{E}[(|\eta_t| - \gamma_1\eta_t)^\delta |\eta_t|^\delta] = \frac{1}{2}[(1 + \gamma_1)^\delta + (1 - \gamma_1)^\delta]\kappa(2\delta)$. For the original GARCH model of Bollerslev (1986), where $\delta = 2$, $\gamma_1 = 0$, and η_t is standard normal, this reduces to the well-known formulas $\mathbb{E}(\epsilon_t^2) = \omega_1/(1 - \alpha_1 - \beta_1)$, $\mathbb{E}(\epsilon_t^4) = 3\omega_1^2(1 + \alpha_1 + \beta_1)(1 - \alpha_1 - \beta_1)^{-1}(1 - 3\alpha_1^2 - 2\alpha_1\beta_1 - \beta_1^2)^{-1}$, and $\text{Corr}(\epsilon_t^2, \epsilon_{t-\tau}^2) = (\alpha_1 + \beta_1)^{\tau-1}\alpha_1(1 - \alpha_1\beta_1 - \beta_1^2)/(1 - 2\alpha_1\beta_1 - \beta_1^2)$.

5 Illustrative Example

To illustrate the relevance of the generalization from MS-GARCH(k) to MS-APGARCH(k), we consider the time series of daily returns of the New York Stock Exchange (NYSE) composite index from January 1966 to August 2001, a sample of 8979 observations.¹ Continuously compounded percentage returns are used, i.e., returns r_t are defined as $r_t = 100 \times \log(I_t/I_{t-1})$, where I_t is the index level at time t . As the series does not exhibit any significant first-order dependencies, we model returns as $r_t = \mu + \epsilon_t$, where μ is a constant mean parameter, and ϵ_t is described by (1) with standard normal η_t .

The return series is displayed in the top left plot of Figure 1, revealing the well-established phenomenon of pronounced volatility clustering. The top right panel plots, for the largest 900 absolute (demeaned) return observations (i.e., approximately 10% of the sample), the empirical complementary distribution function of $|\epsilon_t|$, i.e., $\mathbb{P}(|\epsilon_t| > x)$, against x on a log-log scale, showing that the tail behavior of the data may be well approximated by a power law of the form $\mathbb{P}(|\epsilon_t| > x) \sim cx^{-\alpha}$ with a tail exponent of $\alpha \approx 3.3$. Along with the conditional heteroskedasticity observed in the left plot, this indicates the potential appropriateness of the GARCH class of models in general, because it is known that these processes give rise to asymptotic power laws even with light-tailed (e.g., Gaussian) innovation sequences (Mikosch and Stărică, 2000; Liu, 2006a,b).

The bottom panel of Figure 1 illustrates properties of the return series relevant for the specification of the volatility dynamics. The left plot shows, for $\tau = 1, 2, 5, 10$, the sample ACF

¹ Thanks are due to John Maheu for making the data available.

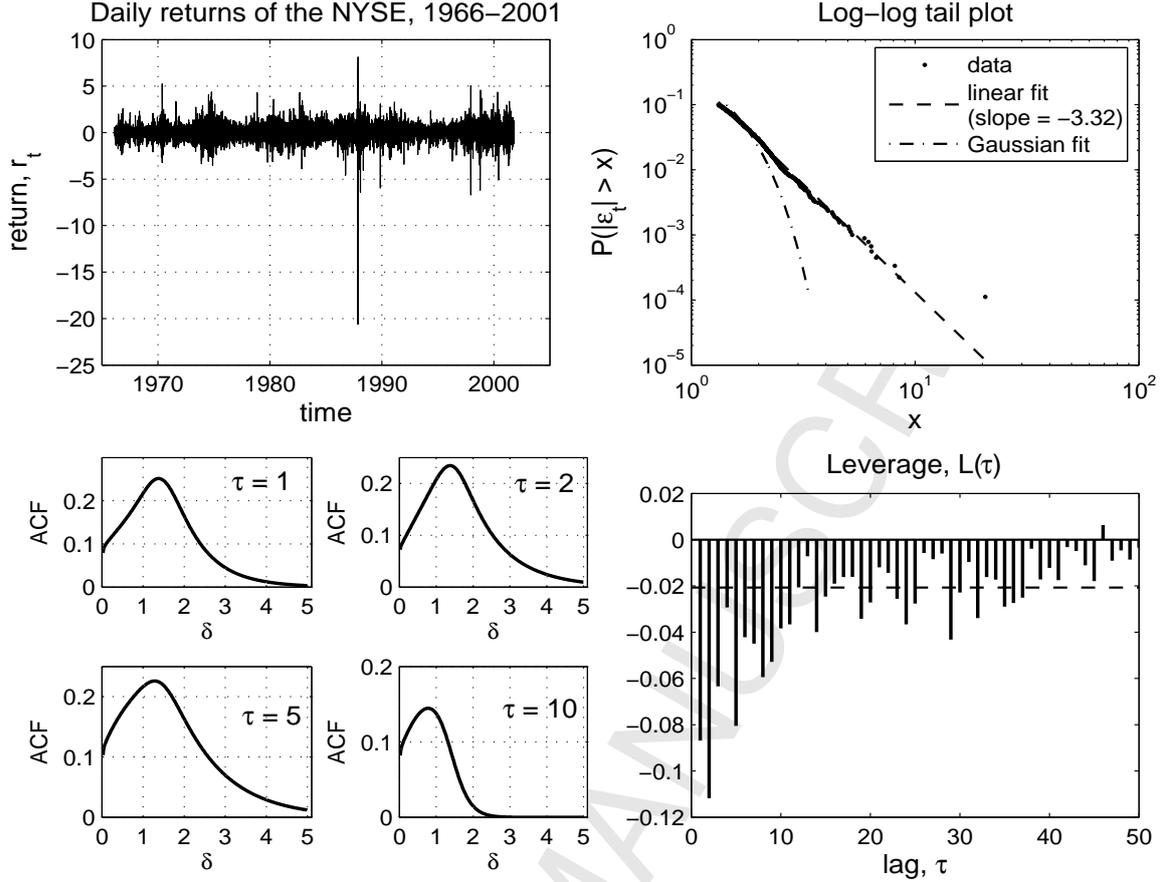


Figure 1: The top left plot shows the return data under study. The top right plot depicts, for the 900 largest absolute (demeaned) return observations, the log of the empirical complementary cdf, $P(|\epsilon_t| > x)$, versus $\log x$, where $\{\epsilon_t\}$ is the demeaned return series. Approximate linearity suggests an asymptotic power law, i.e., for large x , $P(|\epsilon_t| > x) \approx cx^{-\alpha}$ for some $\alpha > 0$. The bottom left plot shows, for various lags, τ , the sample ACF (5) of the power transformed residuals, $\varrho(\tau; \delta) = \text{Corr}(|\epsilon_{t-\tau}|^\delta, |\epsilon_t|^\delta)$, as a function of the power parameter δ . The right plot illustrates the importance of the leverage effect, as defined by (23).

$\varrho(\tau; \delta) = \text{Corr}(|\epsilon_{t-\tau}|^\delta, |\epsilon_t|^\delta)$ as a function of δ , revealing that the empirical autocorrelations tend to be strongest for δ -values considerably different from 2, suggesting potential modeling gains from relaxing the standard GARCH model's restriction that $\delta = 2$. The right plot depicts the correlation between past returns and future volatility, as measured by the absolute (demeaned) return observations, which may be taken as a measure for the leverage effect, $L(\tau)$, i.e.,

$$L(\tau) = \text{Corr}(r_{t-\tau}, |r_t|), \quad \tau > 0. \quad (23)$$

The leverage effect appears to be strong and highly significant for the first 10–15 lags, and, therefore, should be incorporated into a sound volatility model for the returns under study.

Table 1: Estimation results for the NYSE returns.

Model	k	$\hat{\delta}$	$\hat{\gamma}$	$\rho(\mathbb{P}_{C_{11}})$	$\rho(\mathbb{P}_{C_{22}})$	loglik	K	BIC	
								Value	Rank
GARCH	1	2	0	0.986	0.985	-10364	4	20765	8
	2	2	0	0.986	0.987	-10157	9	20395	6
	3	2	0	0.989	0.994	-10119	16	20384	4
PGARCH	1	1.742 (0.132)	0	0.986	0.982	-10363	5	20771	9
	2	1.402 (0.140)	0	0.988	0.980	-10148	10	20388	5
	3	1.345 (0.146)	0	0.991	0.983	-10108	17	20371	3
APGARCH	1	1.254 (0.099)	0.518 (0.053)	0.984	0.974	-10271	6	20598	7
	2	1.099 (0.102)	0.558 (0.060)	0.985	0.974	-10078	11	20257	2
	3	1.085 (0.101)	0.675 (0.068)	0.994	0.989	-10044	18	20252	1

Standard errors are given in parentheses. “loglik” is the value of the maximized log-likelihood function, “ K ” refers to the number of parameters of a given model (including the mean parameter μ), and “BIC” is the Bayesian information criterion, i.e., $\text{BIC} = -2 \times \text{loglik} + K \log T$, where T is the sample size. Smaller values of BIC are preferred. Both the criterion value and the corresponding ranking of the models are shown.

Maximum likelihood estimation results for three different kinds of GARCH models with $k = 1, 2, 3$ regimes are reported in Table 1, where $k = 1$ corresponds to the standard single-regime GARCH process. In Table 1, “GARCH” refers to the specification with $\delta = 2$ and without leverage effect. In the models labeled “PGARCH”, the power parameter δ is freely estimated, but there is still no leverage, while the full specification, with no restriction on δ and allowance for asymmetric volatility dynamics, is denoted as “APGARCH”. Preliminary analysis shows that, in the multi-regime models, there are no statistically significant differences between the regime-specific asymmetry parameters γ_j in (4), and, therefore, Table 1 only reports the results for the specifications with $\gamma_1 = \dots = \gamma_k =: \gamma$.

According to the Bayesian information criterion (BIC) of Schwarz (1978), as reported in the last two columns of Table 1, the multi-regime models are clearly preferred over the single-regime specification. Moreover, for fixed k , $k \geq 2$, PGARCH is always favored against GARCH but dominated by APGARCH. These results are supported by the parameter estimates of δ and γ , as reported in the third and fourth column of Table 1, respectively. These quantities are significantly different from 2 and 0, respectively, in all cases where they are freely estimated. This highlights the importance of both kinds of generalizations introduced by the MS-APGARCH(k) as compared to the MS-GARCH(k) model.

Within the class of multi-regime MS-APGARCH(k) models, the three-component specifi-

cation exhibits a superior fit according to the BIC. It may be worth mentioning that this may not have been expected, because the BIC is a rather conservative model selection criterion due to its strong penalty term. The superiority of the three-component model is also supported by inspection of Figures 2 and 3, which show, for the three APGARCH specifications, the empirical autocorrelations (ACFs) of the power-transformed residuals along with their theoretical counterparts implied by the estimated models. Note that Assumption 3.1 is satisfied for all estimated models, as the dominant eigenvalues of matrices $\mathbb{P}_{C_{11}}$ and $\mathbb{P}_{C_{22}}$, $\rho(\mathbb{P}_{C_{11}})$ and $\rho(\mathbb{P}_{C_{22}})$, as reported in the fifth and sixth column of Table 1, respectively, are smaller than unity. As often reported in the literature since Ding et al. (1993) and Ding and Granger (1993), the empirical autocorrelations decay rapidly at the beginning and then decrease rather slowly, displaying significant positive spikes over very long lags. The single-component APGARCH model is not capable of reproducing this pattern, as its theoretical ACF, shown in the top plot of Figure 2, decreases too slowly at the beginning and then too fast at higher lags. Model MS-APGARCH(2), the ACF of which is depicted in the bottom plot of Figure 2, is better able to capture the empirical ACF over the approximately first 100 lags, but its autocorrelations then also decay to zero too fast. On the other hand, the ACF of model MS-APGARCH(3), shown in Figure 3, does a good job in reproducing the empirical autocorrelation structure with its fast decay at the beginning and its slow decrease afterwards.

References

- Bekaert, G. and Wu, G. (2000). Asymmetric Volatility and Risk in Equity Markets. *The Review of Financial Studies*, 13:1–42.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31:307–327.
- Bouchaud, J.-P., Matacz, A., and Potters, M. (2001). Leverage Effects in Financial Markets: The Retarded Volatility Model. *Physical Review Letters*, 87:228701/1–228701/4.
- Brooks, R. (2007). Power ARCH Modelling of the Volatility of Emerging Equity Markets. *Emerging Markets Review*, 8:124–133.
- Brooks, R. D., Faff, R. W., McKenzie, M. D., and Mitchell, H. (2000). A Multi-Country Study of Power ARCH Models and National Stock Market Returns. *Journal of International Money and Finance*, 19:377–397.
- Cai, J. (1994). A Markov Model of Switching-Regime ARCH. *Journal of Business and Economic Statistics*, 12:309–316.
- Ding, Z. and Granger, C. W. J. (1996). Modeling Volatility Persistence of Speculative Returns: A New Approach. *Journal of Econometrics*, 73:185–215.

- Ding, Z., Granger, C. W. J., and Engle, R. F. (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, 1:83–106.
- Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity With Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50:987–1008.
- Franco, C. and Zakoian, J.-M. (2005). The L^2 -structures of Standard and Switching-regime GARCH Models. *Stochastic Processes and their Applications*, 115:1557–1582.
- Giot, P. and Laurent, S. (2003). Value-at-Risk for Long and Short Trading Positions. *Journal of Applied Econometrics*, 18:641–664.
- Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 48:1779–1801.
- Haas, M., Mittnik, S., and Paoletta, M. S. (2004). A New Approach to Markov-Switching GARCH Models. *Journal of Financial Econometrics*, 2:493–530.
- He, C. and Teräsvirta, T. (1999). Statistical Properties of the Asymmetric Power ARCH Process. In Engle, R. F. and White, H., editors, *Cointegration, Causality, and Forecasting*, pages 462–474. Oxford University Press, New York.
- Higgins, M. L. and Bera, A. K. (1992). A Class of Nonlinear ARCH Models. *International Economic Review*, 33:137–158.
- Hwang, S. Y. and Basawa, I. V. (2004). Stationarity and Moment Structure for Box-Cox Transformed Threshold GARCH(1,1) Processes. *Statistics and Probability Letters*, 68:209–220.
- Karanasos, M. and Kim, J. (2006). A Re-Examination of the Asymmetric Power ARCH Model. *Journal of Empirical Finance*, 13:113–128.
- Liu, J.-C. (2006a). On the Tail Behaviors of a Family of GARCH Processes. *Econometric Theory*, 22:852–862.
- Liu, J.-C. (2006b). On the Tail Behaviors of Box-Cox Transformed Threshold GARCH(1,1) Process. *Statistics and Probability Letters*, 76:1323–1330.
- Liu, J.-C. (2007). Stationarity for a Markov-switching Box-Cox Transformed Threshold GARCH Process. *Statistics and Probability Letters*, 77:1428–1438.
- Mikosch, T. and Stărică, C. (2000). Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process. *Annals of Statistics*, 28:1427–1451.
- Timmermann, A. (2000). Moments of Markov Switching Models. *Journal of Econometrics*, 96:75–111.
- Zakoian, J.-M. (1994). Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, 18:931–955.

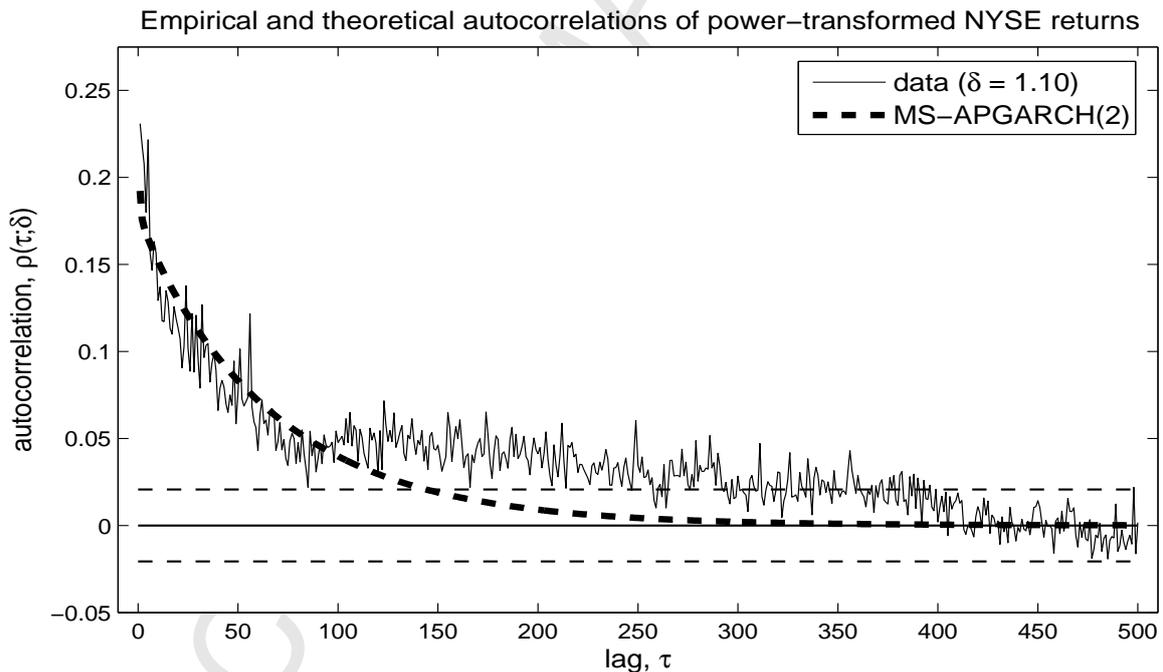
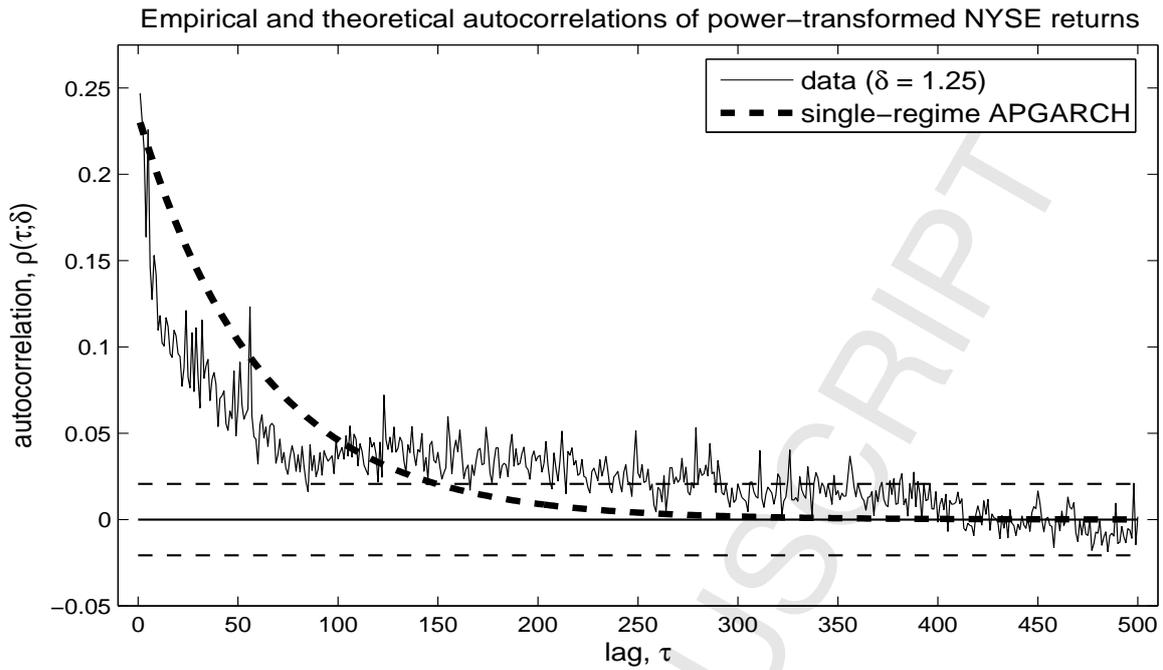


Figure 2: The top panel shows the empirical autocorrelations of the power-transformed absolute NYSE returns (solid line), as defined in (5), along with their theoretical counterparts implied by the fitted single-component APGARCH model (dashed line). The usual 95% asymptotic confidence intervals are also included. The bottom panel repeats this, but for the two-regime Markov-switching APGARCH process, i.e., model MS-APGARCH(2).

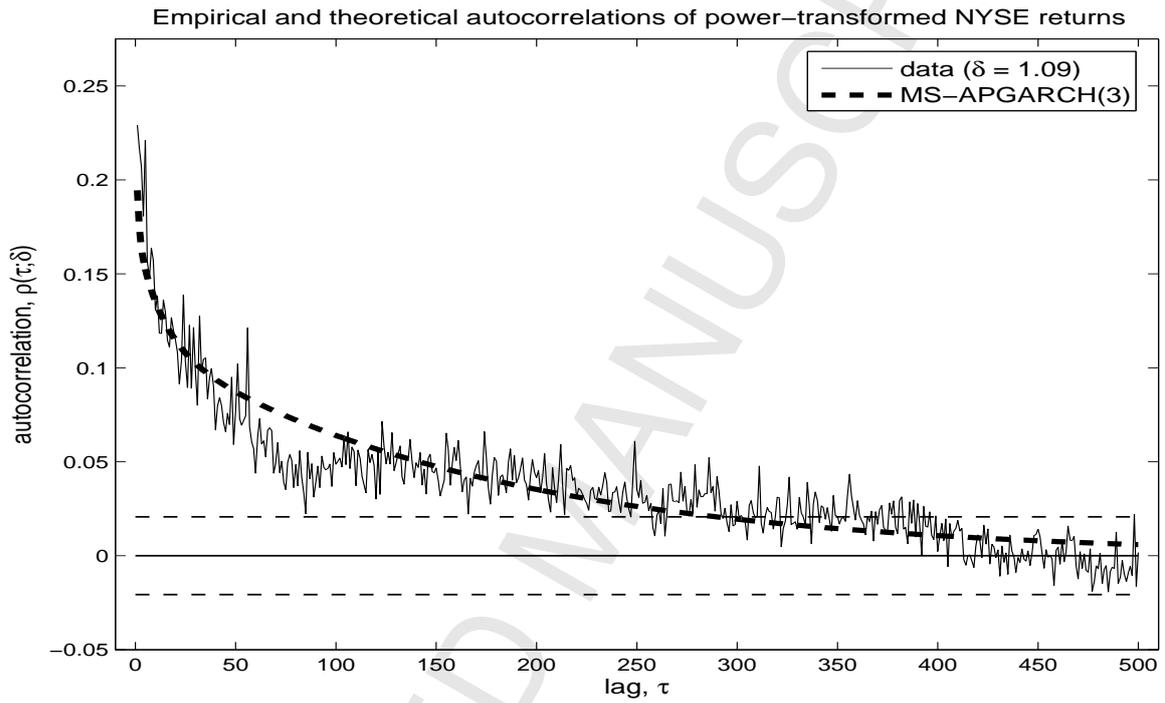


Figure 3: This figure displays the empirical autocorrelations of the power-transformed absolute NYSE returns (solid line), as defined in (5), along with their theoretical counterparts implied by the fitted three-component Markov-switching APGARCH process, i.e., model MS-APGARCH(3) (dashed line). 95% asymptotic one-at-a-time confidence intervals are also included.