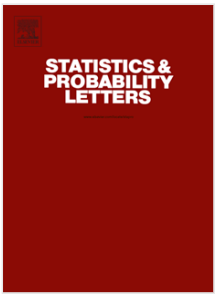
Persistence in volatility, conditional kurtosis, and the Taylor property in absolute value GARCH processes

Markus Haas



Persistence in volatility, conditional kurtosis, and the Taylor property in absolute value GARCH processes

Markus Haas

PII: S0167-7152(09)00154-0
DOI: 10.1016/j.spl.2009.04.017
Reference: STAPRO 5401

To appear in: *Statistics and Probability Letters*

Received date: 16 January 2009
Revised date: 8 April 2009
Accepted date: 20 April 2009

Please cite this article as: Haas, M., Persistence in volatility, conditional kurtosis, and the Taylor property in absolute value GARCH processes. *Statistics and Probability Letters* (2009), doi:10.1016/j.spl.2009.04.017

Persistence in volatility, conditional kurtosis, and the Taylor property in absolute value GARCH processes

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 two anonymous referees, which led to significant improvements of the paper. The research for this paper was supported by the German Research Foundation (DFG).

Abstract

Many authors have observed what is known as the *Taylor property*, namely that the time series dependencies of financial volatility as measured by the autocorrelation function of power–transformed absolute returns are stronger for absolute stock returns than for the squares. In this note, we devise a simple method to detect the Taylor property at any lag in a class of GARCH(1,1) models and fully characterize the relevant parameter space for several popular conditional distributions. It turns out that (i) very generally a first–order Taylor property implies the Taylor property at any lag, and (ii) the degree of conditional kurtosis is crucial for the appearance of the effect. This generalizes earlier findings in He and Teräsvirta (1999) and Gonçalves et al. (2009) which focus on first–order autocorrelations and/or pure ARCH processes only. An application to the S&P500 index illustrates the results.

Keywords—Autocorrelations, conditional volatility, GARCH, kurtosis

1 Introduction

The time–series dependencies of financial volatility are frequently studied by means of the autocorrelation function (ACF) of power–transformed absolute (demeaned) returns, r_t, i.e., $\rho_\delta(\tau) := \text{Corr}(|r_{t-\tau}|^\delta, |r_t|^\delta)$, $\delta > 0$. In this regard, it was observed by Taylor (1986) that, for stock returns, the ACF of the absolute returns tends to be higher than that of the squares. A generalized version of this observation has been identified by Ding et al. (1993) and Ding and Granger (1996) and termed the *Taylor effect* by Granger and Ding (1995), namely that one "almost invariably" (Franses and van Dijk, 2000, p. 30) finds that, for given τ, $\rho_\delta(\tau)$ is maximized for δ close to unity. Granger (2005) adds this effect to the list of *stylized facts* that characterize stock return dynamics.[1]

In view of this evidence, several authors have investigated whether popular volatility models are capable of reproducing the Taylor effect. He and Teräsvirta (1999) consider the absolute value GARCH(1,1) (AVGARCH(1,1)) process driven by Gaussian innovations and, for analytical tractability, concentrate on what they dub the *Taylor property*, namely $\rho_1(1) > \rho_2(1)$. They find that the model incorporates this property only for extremely large values of unconditional kurtosis, and even then the difference $\rho_1(1) - \rho_2(1)$ is very small. Recently, in the context of pure AVARCH(1) processes, Gonçalves et al. (2009) extended this finding by allowing for fat–tailed innovations and find that "high values of kurtosis of the generating white noise favor the appearance of the Taylor property".[2] However, apart from the limitation to the empirically less important ARCH specification, their focus is still on first–order autocorrelations. This is an undesirable restriction as the behavior of the autocorrelations at higher lags characterizes the persistence of the different measures of volatility. In this note, we propose a simple methodology for identifying the Taylor property in AVGARCH(1,1) models at all lags and for a wide variety of conditional distributions. It turns out that rather generally the "first–order" Taylor property $\rho_1(1) > \rho_2(1)$ implies $\rho_1(\tau) > \rho_2(\tau)$ for *all* τ. The role of conditional kurtosis appears to be even more prominent in the AVGARCH(1,1) than in the AVARCH(1) model.

[1] For further discussion and evidence, see, for example, Fornari and Mele (1994), Ryden et al. (1998), Brooks et al. (2000), Cont (2001), Kilic (2004), Malmsten and Teräsvirta (2004), Bulla and Bulla (2006), and Yoon (2008).

[2] For stochastic volatility models, Mora–Galán et al. (2004), Ruiz and Veiga (2008), and Veiga (2009) come to similar conclusions.

2 The absolute value GARCH(1,1) (AVGARCH(1,1)) Process

Consider the absolute value GARCH(1,1) process introduced by Taylor (1986),

$$\epsilon_t = \sigma_t \eta_t, \tag{1}$$

where $\{\eta_t\}$ is iid with zero mean and unit variance, and

$$\sigma_t = \omega + \alpha|\epsilon_{t-1}| + \beta\sigma_{t-1} = \omega + (\alpha|\eta_{t-1}| + \beta)\sigma_{t-1}, \quad \omega > 0, \quad \alpha, \beta \geq 0. \tag{2}$$

Following Gonçalves et al. (2009), we concentrate on innovation distributions satisfying

Assumption 2.1 *The innovations η_t have a symmetric density with $\kappa_1 > \kappa_4^{-1/4}$, where*

$$\kappa_r := E(|\eta_t|^r)$$

is the rth absolute moment of η_t.

In the AVARCH(1) process, where $\beta = 0$ in (2), the restriction $\kappa_1 > \kappa_4^{-1/4}$ was shown by Gonçalves et al. (2009) to guarantee the first–order Taylor effect for a subset of the admissible α–values. In subsequent calculations, we consider two popular distributions for η_t, both satisfying Assumption 2.1, namely Student's t with ν degrees of freedom and density

$$f(\eta_t; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma(\nu/2)\sqrt{(\nu-2)\pi}} \left(1 + \frac{\eta_t^2}{\nu-2}\right)^{-(\nu+1)/2}, \quad \nu > 2, \tag{3}$$

and the generalized exponential distribution (GED) with density

$$f(\eta_t; p) = \frac{\lambda p}{2^{1/p+1}\Gamma(1/p)} \exp\left\{-\frac{|\lambda\eta_t|^p}{2}\right\}, \quad \lambda = 2^{1/p}\sqrt{\frac{\Gamma(3/p)}{\Gamma(1/p)}}, \quad p > 0, \tag{4}$$

for which we have

$$\kappa_r := \mathrm{E}(|\eta_t|^r) = \begin{cases} \frac{\Gamma\left(\frac{r+1}{p}\right)}{\Gamma(1/p)} \left[\frac{\Gamma(1/p)}{\Gamma(3/p)}\right]^{r/2}, & \text{if } \eta_t \sim \mathrm{GED}(p) \\ (\nu-2)^{r/2} \frac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(\frac{\nu-r}{2}\right)}{\sqrt{\pi}\Gamma(\nu/2)}, \quad r < \nu, & \text{if } \eta_t \sim t(\nu). \end{cases} \tag{5}$$

The distributions have unit variance, so $\kappa_2 = 1$, and κ_4 is the conditional kurtosis.

In view of the evidence concerning the role of conditional kurtosis reported in Gonçalves et al. (2009), density (4) is particularly useful as it nests both leptokurtic ($p < 2$) as well as platykurtic ($p > 2$) distributions, and $p = 2$ gives normality. Moreover, as $p \to \infty$, the GED converges to the zero–mean and unit–variance uniform distribution on $(-\sqrt{3}, \sqrt{3})$ with $\kappa_r = (r+1)^{-1}3^{r/2}$ (Box, 1953). From the Gauss–Winckler inequality (cf. von Mises, 1931; Beesack, 1984; and Avkhadiev, 2005), stating that, for (zero–mean) symmetric unimodal

densities, $[(r+1)\kappa_r]^{1/r} \leq [(s+1)\kappa_s]^{1/s}$, $0 < r \leq s$, no such density can have kurtosis lower than that of the GED for $p \to \infty$, given by $\kappa_4 = 9/5$. But still (although tightly) $\kappa_1 = \sqrt{3}/2 \approx 0.866 > 0.863 \approx \sqrt[4]{5/9} = \kappa_4^{-1/4}$. This suggests that $\kappa_1 > \kappa_4^{-1/4}$ will be satisfied for practically any distribution one can anticipate in the context of GARCH models.[3]

The moment structure of model (1)–(2) has been investigated by He and Teräsvirta (1999), Ling and McAleer (2002), and Hwang and Basawa (2004).[4] To summarize their results relevant for the subsequent discussion, we define, for $m \in \mathbb{N}$,

$$c_{mm} = \mathrm{E}[(\alpha|\eta_{t-1}| + \beta)^m] = \sum_{i=0}^{m} \binom{m}{i} \kappa_{m-i} \alpha^{m-i} \beta^i. \tag{6}$$

Proposition 1 *(He and Teräsvirta, 1999, Theorem 1; Ling and McAleer, 2002, Theorem 2.2) The AVGARCH(1,1) process (1)–(2) has a strictly stationary solution with finite mth–order absolute moment, $E(|\epsilon_t|^m)$, $m \in \mathbb{N}$, if and only if $c_{mm} < 1$, and, in this case, the moment can be calculated recursively via*

$$E(|\epsilon_t|^m) = \frac{\kappa_m}{1 - c_{mm}} \sum_{i=0}^{m-1} \binom{m}{i} \omega^{m-i} (c_{ii}/\kappa_i) E(|\epsilon_t|^i). \tag{7}$$

We shall thus assume in the following that $c_{44} < 1$, i.e., $\mathrm{E}(\epsilon_t^4) < \infty$ and the ACFs of $|\epsilon_t|$ and ϵ_t^2 exist.

3 Autocorrelation Structure

The ACF of the absolute values, $\rho_1(\tau) = \mathrm{Corr}(|\epsilon_t|, |\epsilon_{t-\tau}|)$, is (cf. He and Teräsvirta, 1999, Theorem 5)

$$\rho_1(\tau) = c_{11}^{\tau-1} a_1, \quad a_1 = \frac{\kappa_1 \alpha (1 - \kappa_1 \alpha \beta - \beta^2)}{1 - 2\kappa_1 \alpha \beta - \beta^2}, \quad \tau \geq 1, \tag{8}$$

and exhibits an exponential decay at rate c_{11}. He and Teräsvirta (1999, Theorem 4) also derive $\rho_2(\tau) = \mathrm{Corr}(\epsilon_t^2, \epsilon_{t-\tau}^2)$. However, the following simple lemma provides a more straightforward calculation which directly leads to a more explicit expression for the ACF than the recursive relation given in He and Teräsvirta (1999) and is thus more suitable for our purposes.

[3] However, unit–variance distributions exist where $\kappa_1 \kappa_4^{1/4} < 1$. For example, for the two–sided gamma distribution with density $f(x) = \lambda^\theta |x|^{\theta-1} \exp\{-\lambda|x|\}/(2\Gamma(\theta))$, $-\infty < x < \infty$, where $\theta > 0$ and $\lambda = \sqrt{\theta(\theta+1)}$ (to have $\kappa_2 = 1$), which is unimodal for $\theta < 1$, we have $\kappa_r = \lambda^{-r}\Gamma(\theta+r)/\Gamma(\theta)$, and calculations show that $\kappa_1 < \kappa_4^{-1/4}$ for $\theta < (-3 + \sqrt{17})/4 \approx 0.281$. In fact, two–sided gamma densities have been sporadically used in GARCH models (Shaun and Satchell, 2006a,b). It is also possible to construct Gaussian mixtures with $\kappa_1 \kappa_4^{1/4} < 1$. In general, it appears that distributions exhibiting this property are characterized by an extreme (and unrealistic) degree of peakedness.

[4] See also Storti and Vitale (2003) and Liu (2006) for further results.

Lemma 2 *Assume* $c_{11} \neq c_{22}$. *Then*

$$\rho_2(\tau) = \tilde{a}_1 c_{11}^{\tau-1} + (\tilde{a}_2 - \tilde{a}_1) c_{22}^{\tau-1}, \quad \tau \geq 1, \tag{9}$$

where

$$\begin{aligned} \tilde{a}_1 &= \frac{c_{21}\left[E(\sigma_t^2)(\omega - E(\sigma_t)) + E(\sigma_t^3)(\alpha\kappa_3 + \beta)\right]}{(c_{11} - c_{22})(E(\epsilon_t^4) - E^2(\epsilon_t^2))}, \quad \textit{and} \\ \tilde{a}_2 &= \frac{E(\sigma_t^2)(\omega^2 - E(\sigma_t^2)) + 2\omega E(\sigma_t^3)(\alpha\kappa_3 + \beta) + E(\sigma_t^4)(\alpha^2\kappa_4 + \beta^2 + 2\alpha\beta\kappa_3)}{E(\epsilon_t^4) - E^2(\epsilon_t^2)}. \end{aligned} \tag{10}$$

Observe that, just as in a_1 in (8), ω cancels out both in $\tilde{a}_1$ and $\tilde{a}_2$. Also, in the AVARCH(1,1) case, where $\beta = 0$, $\mathrm{E}(\epsilon_t^4) \to \infty$ as $\alpha \to \kappa_4^{-1/4}$, and so $\lim_{\alpha \to \kappa_4^{-1/4}} \rho_2(1) = \lim_{\alpha \to \kappa_4^{-1/4}} \tilde{a}_2 = \kappa_4^{-1/2}$.

Proof. Let Ψ_t be the σ–algebra generated by $\{\eta_s : s \leq t\}$. We have

$$\mathrm{E}(\epsilon_{t-\tau}^2 \epsilon_t^2) = \mathrm{E}[\eta_{t-\tau}^2 \sigma_{t-\tau}^2 \mathrm{E}(\eta_t^2 \sigma_t^2 | \Psi_{t-\tau})] = \mathrm{E}[\eta_{t-\tau}^2 \sigma_{t-\tau}^2 \mathrm{E}(\sigma_t^2 | \Psi_{t-\tau})]. \tag{11}$$

To conveniently calculate $\mathrm{E}(\sigma_t^2 | \Psi_{t-\tau})$, define

$$X_t = \begin{pmatrix} \sigma_t \\ \sigma_t^2 \end{pmatrix}, \quad \tilde{\omega} = \begin{pmatrix} \omega \\ \omega^2 \end{pmatrix}, \quad C = \begin{pmatrix} c_{11} & 0 \\ c_{21} & c_{22} \end{pmatrix},$$

where $c_{21} = 2\omega c_{11}$. This gives

$$\mathrm{E}(X_t | \Psi_{t-2}) = \tilde{\omega} + C X_{t-1}. \tag{12}$$

Recursively substituting in (12) we obtain

$$\begin{aligned} \mathrm{E}(X_t | \Psi_{t-\tau}) &= \sum_{i=0}^{\tau-2} C^i \tilde{\omega} + C^{\tau-1} X_{t-\tau+1} = \mathrm{E}(X_t) + C^{\tau-1}(X_{t-\tau+1} - \mathrm{E}(X_t)) \\ &= \mathrm{E}(X_t) + C^{\tau-1}(\tilde{\omega} + C_{t-\tau} X_{t-\tau} - \mathrm{E}(X_t)), \end{aligned} \tag{13}$$

where $\mathrm{E}(X_t) = (I_2 - C)^{-1}\tilde{\omega}$, I_2 is the two–dimensional identity matrix, and

$$C_t = \begin{pmatrix} c_t & 0 \\ 2\omega c_t & c_t^2 \end{pmatrix}, \quad c_t = \alpha|\eta_t| + \beta. \tag{14}$$

We note that, provided $c_{11} \neq c_{22}$,

$$C^\tau = \begin{pmatrix} c_{11}^\tau & 0 \\ c_{21} \sum_{i=0}^{\tau-1} c_{11}^{\tau-1-i} c_{22}^i & c_{22}^\tau \end{pmatrix} = \begin{pmatrix} c_{11}^\tau & 0 \\ c_{21} \frac{c_{11}^\tau - c_{22}^\tau}{c_{11} - c_{22}} & c_{22}^\tau \end{pmatrix}. \tag{15}$$

Substituting (14) and (15) into (13), we obtain an expression for $\mathrm{E}(\sigma_t^2 | \Psi_{t-\tau})$ in terms of $\sigma_{t-\tau}$ and $\sigma_{t-\tau}^2$, which can in turn be plugged into (11), and after some algebra the final result is (9). ■

Equation (9) shows that the decay of $\rho_2(\tau)$ is described by a mixture of two exponentials with rates c_{11} and c_{22}. As $\mathrm{E}(\sigma_t) = \omega(1-c_{11})^{-1}$, the factor in brackets in the numerator of $\tilde{a}_1$ can be written as $(\alpha\kappa_3+\beta)\mathrm{E}(\sigma_t^3) - (\alpha\kappa_1+\beta)\mathrm{E}(\sigma_t^2)\mathrm{E}(\sigma_t)$ and is positive by Lyapunov's inequality. Thus $\tilde{a}_1 > 0$ if $c_{11} > c_{22}$, and a comparison with (8) shows that the Taylor property will *not* materialize in the sense that $\rho_1(\tau)$ decreases to zero slower than $\rho_2(\tau)$. But may it happen that both $c_{44} < 1$ and $c_{22} > c_{11}$, implying a reverse Taylor property (at least) at higher lags? On the other hand, suppose that $c_{44} < 1$ implies $c_{11} > c_{22}$, which will turn out to be the "regular" situation.

Corollary 3 *If $c_{11} > c_{22}$, then*

$$\rho_1(\tau) \gtreqless \rho_2(\tau)\ \forall\tau \Leftrightarrow a_1 \begin{cases} > \max\{\tilde{a}_1, \tilde{a}_2\} \\ = \tilde{a}_1 = \tilde{a}_2 \\ < \min\{\tilde{a}_1, \tilde{a}_2\} \end{cases}$$

Proof. Inspection of (9). ■

Let us first consider the AVARCH(1) process, where $\beta = 0$ in (2).

3.1 The AVARCH(1) process

Proposition 1 shows that, for $\mathrm{E}(\epsilon_t^4) < \infty$, we require $\alpha < \kappa_4^{-1/4}$, and hence $c_{44} < 1$ and Assumption 2.1 are sufficient for $c_{11} > c_{22}$. This gives rise to the following result.

Proposition 4 *In the AVARCH(1) process with finite fourth moment and an innovation distribution satisfying Assumption 2.1, we have*

$$\rho_1(1) \gtreqless \rho_2(1) \Rightarrow \rho_1(\tau) \gtreqless \rho_2(\tau) \quad \textit{for all } \tau. \tag{16}$$

Proof. Calculations show that

$$a_1 \gtreqless \tilde{a}_2 \Leftrightarrow a_1 \gtreqless \tilde{a}_1 \Leftrightarrow P(\alpha) \gtreqless 0, \quad \alpha \in (0, \kappa_4^{-1/4}), \tag{17}$$

where $P(\alpha)$ is the sixth–order polynomial given by

$$\begin{aligned} P(\alpha) \;=\; & (2\kappa_3\kappa_4 - 4\kappa_1^2\kappa_3\kappa_4 + 2\kappa_1\kappa_4)\alpha^6 + (4\kappa_4 - 4\kappa_1^2\kappa_4 - \kappa_1\kappa_3\kappa_4 + \kappa_1\kappa_3)\alpha^5 \\ & + (3\kappa_3 - 6\kappa_1\kappa_4 - \kappa_3\kappa_4 - \kappa_1^2\kappa_3 + 5\kappa_1^2\kappa_3\kappa_4)\alpha^4 + (3\kappa_1\kappa_3\kappa_4 + \kappa_1\kappa_3 - 4\kappa_4)\alpha^3 \\ & + (\kappa_1\kappa_4 + \kappa_1 - 2\kappa_3)\alpha^2 + (1 - 4\kappa_1\kappa_3 + \kappa_1^2 + 3\kappa_1^2\kappa_4 - \kappa_4)\alpha + \kappa_1 + \kappa_1\kappa_4 - 2\kappa_3. \end{aligned} \tag{18}$$

■

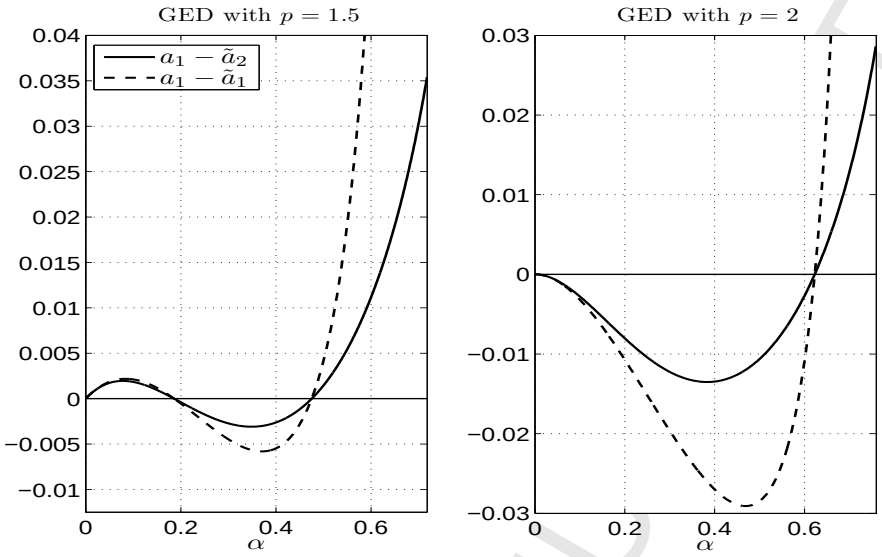


Figure 1: For the GED distribution (4) with $p = 1.5$, the left plot shows $a_1 - \tilde{a}_2 = \rho_1(1) - \rho_2(1)$ (solid line) and $a_1 - \tilde{a}_1$ (dashed line) as a function of parameter α for the absolute value ARCH(1) process, i.e., (1)–(2) with $\beta = 0$. Quantities a_1, $\tilde{a}_1$, and $\tilde{a}_2$ are defined in (8) and (10). The right plot repeats this, but for $p = 2$ (i.e., the normal distribution).

Remark 1 *Gonçalves et al. (2009) and (implicitly) He and Teräsvirta (1999) observed that* $\kappa_1 > \kappa_4^{-1/4}$ *guarantees the existence of* $\tilde{\alpha} < \kappa_4^{-1/4}$ *such that* $\rho_1(1) > \rho_2(1)$ *for* $\alpha \in (\tilde{\alpha}, \kappa_4^{-1/4})$. *This can also be arrived at by means of (18). In fact, putting* $\alpha = \kappa_4^{-1/4}$ *and bearing in mind that* $\kappa_4 > 1$, *we obtain*

$$\begin{aligned} P(\kappa_4^{-1/4}) &= (\kappa_1 - \kappa_4^{-1/4}) \left\{ 3(\kappa_4^{3/4} - \kappa_4^{-1/4}) \left(\kappa_1 - \frac{10 - 2\kappa_4 - 8\kappa_4^{1/2}}{6(\kappa_4^{3/4} - \kappa_4^{-1/4})} \right) \right. \\ &\quad \left. + \kappa_3(5 - 4\kappa_4^{-1/2} - \kappa_4^{-1}) \left(\kappa_1 - \frac{6(\kappa_4^{-3/4} - \kappa_4^{1/4})}{10 - 8\kappa_4^{-1/2} - 2\kappa_4^{-1}} \right) \right\} \\ &\gtreqless 0 \Leftrightarrow \kappa_1 \gtreqless \kappa_4^{-1/4}. \end{aligned}$$

(18) also shows that, if $\kappa_1 + \kappa_1\kappa_4 - 2\kappa_3 > 0$, *there will be* $\breve{\alpha}$ *such that* $\rho_1(\tau) > \rho_2(\tau)$ *for* $\alpha \in (0, \breve{\alpha})$. *The GED distribution satisfies this condition for* $p < 2$. *For the normal,* $\kappa_1 + \kappa_1\kappa_4 - 2\kappa_3 = 0$, *and then* $1 - 4\kappa_1\kappa_3 + \kappa_1^2 + 3\kappa_1^2\kappa_4 - \kappa_4 = 2(2/\pi - 1) < 0$ *implies that such an* $\breve{\alpha}$ *does not exist. Moreover, as illustrated in the left panel of Figure 1, even for* $p < 2$ *the Taylor property is quantitatively negligible as long as* α *is small.*

3.2 The AVGARCH(1,1) process

For the full model (1)–(2), solving $c_{22} - c_{11} = \alpha^2 + 2\kappa_1\alpha\beta + \beta^2 - \kappa_1\alpha - \beta = 0$ shows that, in (α, β)–space, $c_{22} > c_{11}$ for all parameter combinations to the right of the line described by[5]

$$\tilde{\beta} = \begin{cases} \frac{1-2\alpha\kappa_1+\sqrt{1-4\alpha^2(1-\kappa_1^2)}}{2}, & 0 < \alpha \leq (<)\kappa_1, \qquad \text{if } \kappa_1 \geq (<)\frac{1}{\sqrt{2}} \\ \frac{1-2\alpha\kappa_1\pm\sqrt{1-4\alpha^2(1-\kappa_1^2)}}{2}, & \kappa_1 \leq \alpha \leq \frac{1}{2\sqrt{1-\kappa_1^2}}, \quad \text{if } \kappa_1 < \frac{1}{\sqrt{2}}, \end{cases} \tag{19}$$

whereas parameter constellations admitting a finite fourth moment, i.e., $c_{44} < 1$, are between the axes and the line $\breve{\beta} = f(\alpha)$, where $\breve{\beta}$ is the unique positive solution of

$$P(\beta;\alpha) = \beta^4 + 4\kappa_1\alpha\beta^3 + 6\alpha^2\beta^2 + 4\kappa_3\alpha^3\beta - (1-\kappa_4\alpha^4) = 0, \quad 0 < \alpha < \frac{1}{\sqrt[4]{\kappa_4}}. \tag{20}$$

The region where the variance is finite ($c_{22} < 1$) is also of interest, and is between the axes and the line $\beta = -\kappa_1\alpha + \sqrt{1-\alpha^2(1-\kappa_1^2)}$, $0 < \alpha < 1$.

Now $c_{11} > c_{22}$ for any parameter combination such that $c_{44} < 1$ requires $\breve{\beta} < \tilde{\beta}$ for $\alpha < \kappa_4^{-1/4}$. In a neighborhood of $\alpha = \kappa_4^{-1/4}$, $\breve{\beta} < \tilde{\beta}$ is a consequence of Assumption 2.1, whereas for α–values in a neighborhood of zero it can be deduced by differentiating both relations twice (implicitly so in case of (20)),

$$\begin{aligned} \left.\frac{d\tilde{\beta}}{d\alpha}\right|_{\alpha=0} &= \left.\frac{d\breve{\beta}}{d\alpha}\right|_{\alpha=0} = -\kappa_1, \\ \left.\frac{d^2\tilde{\beta}}{d\alpha^2}\right|_{\alpha=0} &= -2(1-\kappa_1^2) > -3(1-\kappa_1^2) = \left.\frac{d^2\breve{\beta}}{d\alpha^2}\right|_{\alpha=0}. \end{aligned}$$

Similar to the approach in Gonçalves et al. (2009), for a given distribution, the entire shape of the lines defined by (19) and (20) can be elucidated graphically. This is done in Figure 2 for four different members of (3) and (4), namely Student's t with $\nu = 4.5$ (with kurtosis $\kappa_4 = 3(\nu-2)/(\nu-4) = 15$) and the GED with $p = 1, 2$, and ∞, corresponding to the Laplace ($\kappa_4 = 6$), Gaussian ($\kappa_4 = 3$) and uniform ($\kappa_4 = 9/5$), respectively. The fourth–moment condition is satisfied by all parameter constellations to the left of the dashed line, whereas $c_{22} > c_{11}$ for all those to the right of the dash–dotted line. Although these lines converge somewhat as the kurtosis decreases, the region where $1 > c_{22} > c_{11}$ (Region R_3) is, in all cases, entirely to the right of that where $\mathrm{E}(\epsilon_t^4)$ exists (Region R_1), so there is no pair (α, β) where both $c_{22} > c_{11}$ and $c_{44} < 1$. As the distributions in Figure 2 range from rather leptokurtic to extremely platykurtic, we expect that this is true for most of the densities that one would reasonably use in GARCH models. Thus, in case of existence, both (8) and (9) will be dominated by an exponential decay at rate c_{11}.

[5] Note that the case $\kappa_1 < 1/\sqrt{2}$ is of little practical relevance. For the t and GED distributions, $\kappa_1 < 1/\sqrt{2}$ if $\nu < 4$ (so that $\rho_2(\tau)$ is not defined) and $p < 1$, respectively.

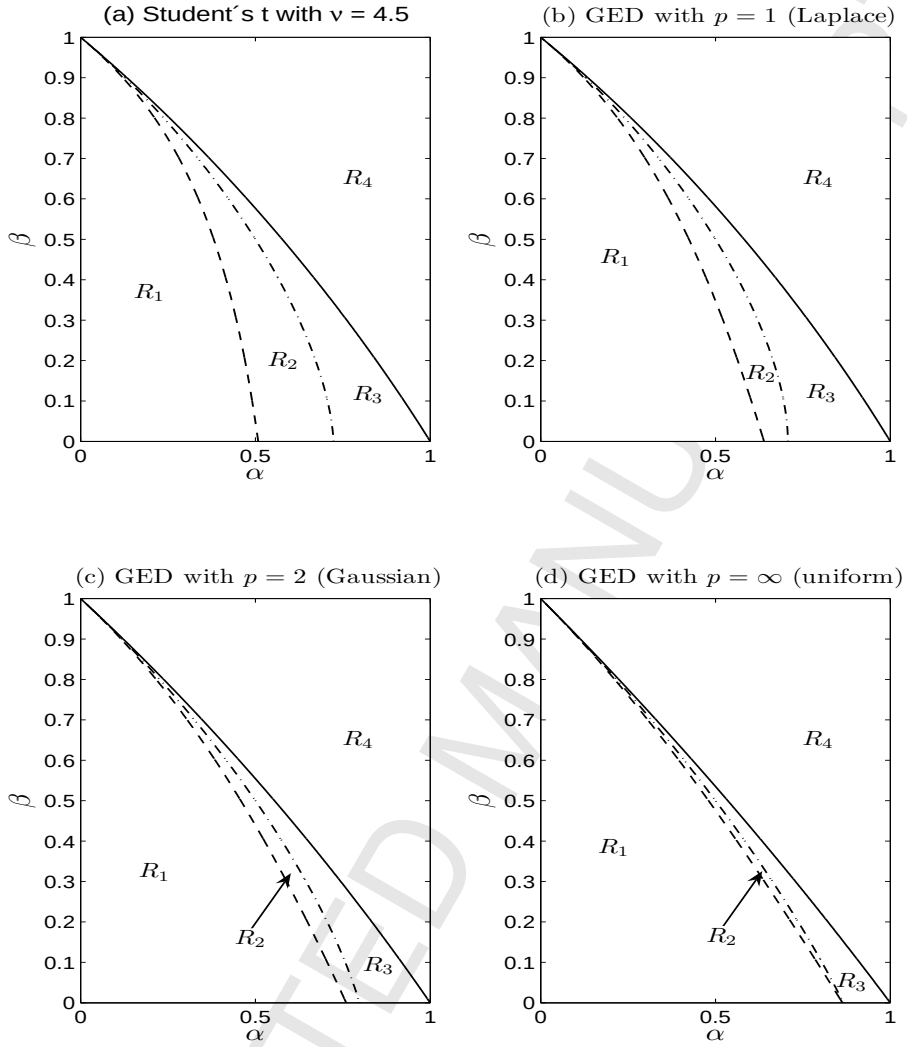


Figure 2: For each distribution, the dash–dotted and dashed lines represent relationships (19) and (20), respectively, and the solid line indicates the region of covariance stationarity, where $\mathrm{E}(\epsilon_t^2) < \infty$. Thus, the regions R_i, $i = 1, 2, 3, 4$, are characterized as follows:

$$
\begin{aligned}
R_1 &= \{(\alpha, \beta) : \alpha, \beta \geq 0, c_{22} < c_{11} < 1, c_{44} < 1\}, \\
R_2 &= \{(\alpha, \beta) : \alpha, \beta \geq 0, c_{22} < c_{11} < 1 < c_{44}\}, \\
R_3 &= \{(\alpha, \beta) : \alpha, \beta \geq 0, c_{11} < c_{22} < 1 < c_{44}\}, \\
R_4 &= \{(\alpha, \beta) : \alpha, \beta \geq 0, 1 < c_{22}\},
\end{aligned}
$$

where c_{mm}, $m \in \mathbb{N}$, is defined in (6).

In view of these results, we shall now, for a variety of distributions, classify all parameter combinations in (α, β)–space such that $c_{44} < 1$ according to their implied ordering of the coefficients a_1, $\tilde{a}_1$, and $\tilde{a}_2$, as indicated in Corollary 3. We do so in Figure 3, where for illustration we consider the GED with $p \in \{1, 1.25, 1.5, 2, \infty\}$. First observe that $a_1 > \tilde{a}_2$ implies $a_1 > \tilde{a}_1$ in all cases, that is, the first–order Taylor property implies the Taylor property for all τ (Region $\tilde{R}_1$). Obviously, conditional kurtosis is crucial for this to show up. It is thoroughly present for the Laplace (Panel (a)), whereas the relevant area shrinks (in a complex manner) as conditional kurtosis decreases. For the uniform (Panels (e) and (f)), there is still a tiny region where the Taylor property holds (as revealed in Panel (f)), but numerical calculations show it being quantitatively negligible. Figure 3, of course, tells us nothing about the magnitude of the effect, and even for leptokurtic densities it tends to be rather small as long as α is small. We also mention a fact that is not easily discernible in Panel (d) (Gaussian) of Figure 3, namely that even for higher values of β there is a very narrow strip between Regions $\tilde{R}_2$ and $\tilde{R}_4$ where the Taylor property holds; e.g., if $\beta = 0.9$, it is there for $0.1132 < \alpha < 0.1158$, with unconditional kurtosis greater than 18.

In Region $\tilde{R}_2$ we have $\tilde{a}_1 < a_1 < \tilde{a}_2$ (this cannot occur in the AVARCH(1) process due to Proposition 4), so $\rho_1(1) < \rho_2(1)$ but $\rho_1(\tau) > \rho_2(\tau)$ for

$$\tau > \tau^\star := \frac{\log\left(\frac{\tilde{a}_2 - \tilde{a}_1}{a_1 - \tilde{a}_1}\right)}{\log\left(\frac{c_{11}}{c_{22}}\right)} + 1. \tag{21}$$

Finally, in Region $\tilde{R}_3$, where $a_1 < \min\{\tilde{a}_1, \tilde{a}_2\}$, $\rho_2(\tau) > \rho_1(\tau)$ for all τ. There is no region where $\tilde{a}_1 > a_1 > \tilde{a}_2$. Thus, $\rho_1(1) > \rho_2(1) \Rightarrow \rho_1(\tau) > \rho_2(\tau)\ \forall \tau$ for these distributions.

4 Empirical Example

We consider daily log–returns, r_t, of the S&P500 price index (obtained from Datastream) over the period from January 2000 to December 2007 ($T = 1998$ observations), i.e., $r_t = 100 \times \log(I_t/I_{t-1})$, where I_t is the index level at time t. Maximum likelihood estimation results for three different AVGARCH(1,1) processes are reported in Table 1. As $c_{44} < 1$ for all models, the ACFs are well–defined. Interestingly, the models with GED and Student's t innovations achieve approximately the same log–likelihood, $\log L$, but both clearly dominate the Gaussian model. This is also reflected in the values of the estimated shape parameters (p/ν) and the associated conditional kurtosis (κ_4).

To illustrate the results of Section 3, we note that the models based on the GED or Student's t both feature the Taylor property ($a_1 > \tilde{a}_2$), but it is rather moderate for the GED.

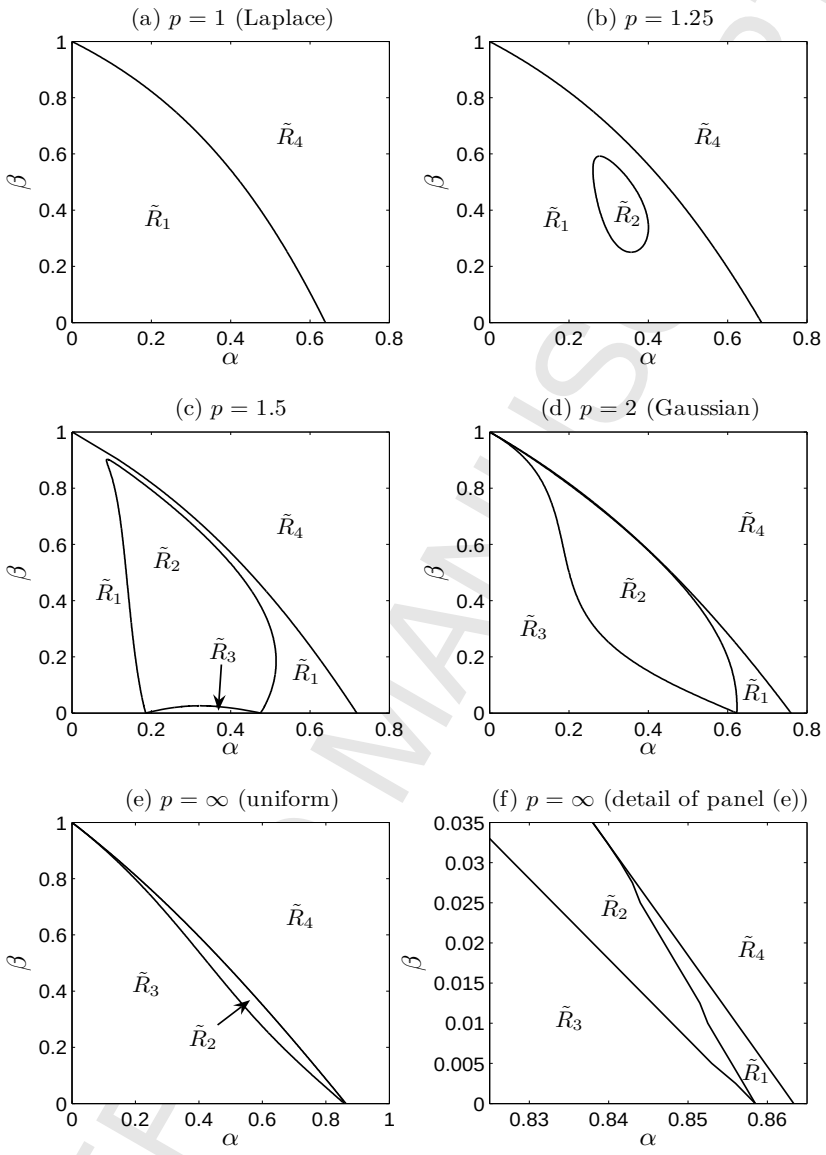


Figure 3: Parameter p is the shape parameter of the GED density (4). The regions $\tilde{R}_i$, $i = 1, 2, 3, 4$, are characterized as follows:

$$\begin{aligned}
\tilde{R}_1 &= \{(\alpha,\beta) : \alpha,\beta \geq 0, c_{44} < 1, a_1 > \max\{\tilde{a}_1, \tilde{a}_2\}\}, \\
\tilde{R}_2 &= \{(\alpha,\beta) : \alpha,\beta \geq 0, c_{44} < 1, \tilde{a}_1 < a_1 < \tilde{a}_2\}, \\
\tilde{R}_3 &= \{(\alpha,\beta) : \alpha,\beta \geq 0, c_{44} < 1, a_1 < \min\{\tilde{a}_1, \tilde{a}_2\}\}, \\
\tilde{R}_4 &= \{(\alpha,\beta) : \alpha,\beta \geq 0, c_{44} > 1\},
\end{aligned}$$

where $a_1, \tilde{a}_1$, and $\tilde{a}_2$ are defined in (8) and (10).

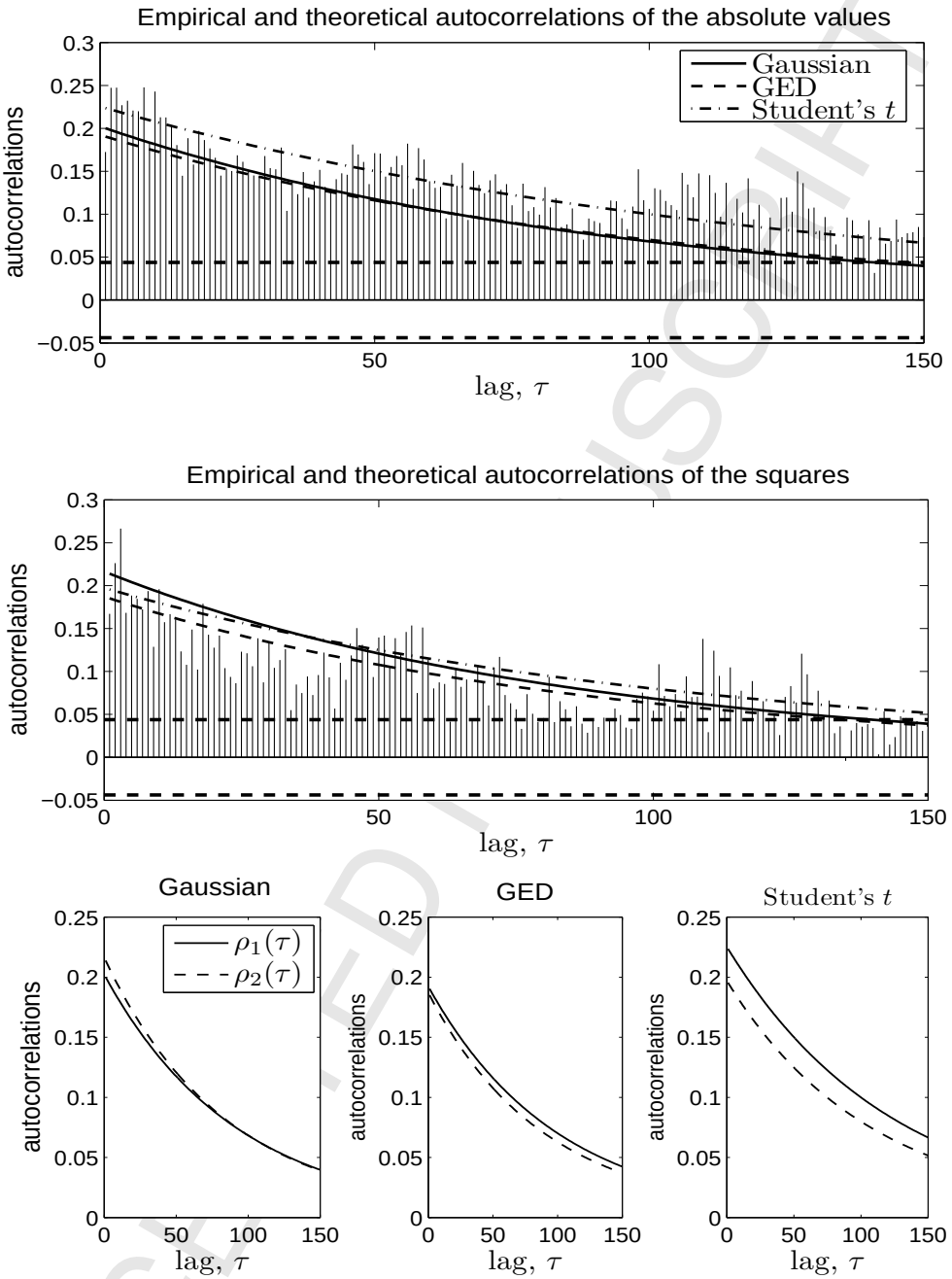


Figure 4: The top and center plots show the empirical autocorrelations of absolute and squared (demeaned) S&P500 returns, respectively, along with their theoretical counterparts implied by the fitted AVGARCH(1,1) models. For each model, the bottom panel compares the autocorrelations of the absolute and squared values in order to highlight the presence/magnitude or absence of the Taylor property.

Table 1: Estimation results for the S&P 500 returns

distribution	c_{11}	c_{22}	c_{44}	a_1	$\tilde{a}_1$	$\tilde{a}_2$	$\lceil\tau^\star\rceil$	κ_4	p/ν	$\log L$
Gaussian	0.989	0.981	0.971	0.200	0.190	0.214	103	3.00	$p=2$	–2781.0
GED	0.990	0.982	0.975	0.190	0.158	0.185	–	3.83	$p=1.47$ (0.07)	–2756.9
Student's t	0.992	0.986	0.983	0.224	0.157	0.195	–	4.08	$\nu=9.54$ (1.86)	–2756.6

[a]Reported are estimation results for AVGARCH(1,1) models (1)–(2) fitted to the S&P500 returns. c_{ii}, $i \in \{1,2,4\}$, is defined in (6). a_1, $\tilde{a}_1$ and $\tilde{a}_2$ are as in (8) and (10), and $\tau^\star$ is defined in (21). κ_i denotes the ith absolute moment of the innovations defined in (5), and p and ν are the estimated shape parameters of the GED and Student's t, respectively, with standard errors in parentheses. $\log L$ is the value of the maximized log–likelihood function.

On the other hand, for the Gaussian AVGARCH(1,1) process, $\tilde{a}_2 > a_1 > \tilde{a}_1$ (Region $\tilde{R}_2$ in Figure 3), so there is no Taylor property at lower lags.

The top and center panels of Figure 4 show the sample autocorrelations of the S&P500 returns along with their theoretical counterparts implied by the fitted models. The Taylor property is clearly visible in the empirical autocorrelations. The most pronounced difference between the models is the higher ACF of the absolute values implied by the Student's t process, whereas the GED is close to the Gaussian. The fact that the Taylor property is rather weak for the GED is also reflected in the bottom panel of Figure 4, where for each model both $\rho_1(\tau)$ and $\rho_2(\tau)$ are pictured in the same graph.

5 The Taylor Property and "Outliers"

Recently, Teräsvirta and Zhao (2007) argue that the Taylor effect may be due to "outliers" and show that it "vanishes when standard estimates of autocorrelation are replaced with robust ones", where the robust measures applied by these authors attach a lower weight to observations relatively far from the mean. These results are in in accordance with those of the present paper, as both indicate that the Taylor effect is actually an accompaniment of conditional leptokurtosis. The question is whether robust measures of autocorrelation are the appropriate tool to deal with this issue. This is in no way to be taken for granted, as it may be argued that "extraordinary price changes [...] are something that should be expected to occur occasionally in a speculative market, and such events are merely an outcome of the generating mechanism and not a break–down of the usual mechanism" (Granger and Ding, 1995), see also Stanley (2003) for discussion.

We finally note that the present analysis could be extended to compare the ACFs of power–transformed absolute returns, i.e., $\rho_\delta(\tau) = \mathrm{Corr}(|r_{t-\tau}|^\delta, |r_t|^\delta)$, for different $\delta > 0$ in the more general power GARCH(1,1) processes as considered in Ding et al. (1993), Hwang and

Basawa (2004), and Liu (2006). The extension to AVGARCH(p, q) models would be much more cumbersome due to the complicated nature of the associated moments, which are less amenable to analytical investigation. However, as the first–order GARCH model is by far the most commonly employed in empirical studies, a rather large part of the situations of practical interest is covered by an analysis of this specification.

References

Avkhadiev, F. G. (2005). A simple proof of the Gauss-Winckler inequality. *American Mathematical Monthly*, 112:459–462.

Beesack, P. R. (1984). Inequalities for absolute moments of a distribution: From Laplace to Von Mises. *Journal of Mathematical Analysis and Applications*, 98:435–457.

Box, G. E. P. (1953). A Note on regions for tests of kurtosis. *Biometrika*, 40:465–468.

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.

Bulla, J. and Bulla, I. (2006). Stylized facts of financial time series and hidden semi-Markov models. *Computational Statistics & Data Analysis*, 51:2192–2209.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.

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.

Fornari, F. and Mele, A. (1994). A stochastic variance model for absolute returns. *Economics Letters*, 46:211–214.

Franses, P. H. and van Dijk, D. (2000). *Non–Linear Time Series Models in Empirical Finance*. Cambridge University Press, Cambridge.

Gonçalves, E., Leite, J., and Mendes-Lopes, N. (2009). A mathematical approach to detect the Taylor property in TARCH processes. *Statistics and Probability Letters*, 79:602–610.

Granger, C. W. J. (2005). The past and future of empirical finance: some personal comments. *Journal of Econometrics*, 129:35–40.

Granger, C. W. J. and Ding, Z. (1995). Some properties of absolute return, an alternative measure of risk. *Annales D'économie et de Statistique*, 40:67–91.

He, C. and Teräsvirta, T. (1999). Properties of moments of a family of GARCH processes. *Journal of Econometrics*, 92:173–192.

Hwang, S. Y. and Basawa, I. V. (2004). Stationarity and moment structure for Box-Cox transformed threshold GARCH(1,1) processes. *Statistics & Probability Letters*, 68:209–220.

Kilic, R. (2004). On the long memory properties of emerging capital markets: evidence from Istanbul stock exchange. *Applied Financial Economics*, 14:915–922.

Ling, S. and McAleer, M. (2002). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, 106:109–117.

Liu, J.-C. (2006). On the tail behaviors of Box-Cox transformed threshold GARCH(1,1) process. *Statistics & Probability Letters*, 76:1323–1330.

Malmsten, H. and Teräsvirta, T. (2004). Stylized facts of financial time series and three popular models of volatility. SSE/EFI Working Paper Series in Economics and Finance 563, Stockholm School of Economics.

Mora-Galán, A., Pérez, A., and Ruiz, E. (2004). Stochastic volatility models and the Taylor effect. Statistics and Econometrics Working Paper 04-63, Department of Economics, Universidad Carlos III de Madrid.

Ruiz, E. and Veiga, H. (2008). Modelling long-memory volatilities with leverage effect: A-LMSV versus FIEGARCH. *Computational Statistics and Data Analysis*, 52:2846–2862.

Ryden, T., Teräsvirta, T., and Åsbrink, S. (1998). Stylized facts of daily return series and the hidden Markov model. *Journal of Applied Econometrics*, 13:217–244.

Shaun, A. B. and Satchell, S. E. (2006a). Asymmetry and downside risk in foreign exchange rates. *European Journal of Finance*, 12:313–332.

Shaun, A. B. and Satchell, S. E. (2006b). Asymmetry, loss aversion, and forecasting. *Journal of Business*, 79:1809–1830.

Stanley, H. E. (2003). Statistical physics and economic fluctuations: do outliers exist? *Physica A*, 318:279–292.

Storti, G. and Vitale, C. (2003). BL-GARCH models and asymmetries in volatility. *Statistical Methods & Applications*, 12:19–39.

Taylor, S. J. (1986). *Modelling Financial Time Series*. John Wiley & Sons, Chichester.

Teräsvirta, T. and Zhao, Z. (2007). Stylized facts of return series, robust estimates, and three popular models of volatility. SSE/EFI Working Paper Series in Economics and Finance 662, Stockholm School of Economics.

Veiga, H. (2009). Financial stylized facts and the Taylor-effect in stochastic volatility models. *Economics Bulletin*, 29:265–276.

von Mises, R. (1931). Über einige Abschätzungen von Erwartungswerten. *Journal für die Reine und Angewandte Mathematik*, 165:184–193.

Yoon, G. (2008). Long memory in volatility. Forthcoming in: *Applied Economics Letters*.