UNIVERSIDAD CARLOS III DE MADRID working papers

Working Paper
Economic series 10 - 13
Junio 2010

Departamento de Economía
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 916249875

"The Power Log-GARCH Model*"

Genaro Sucarrat† Alvaro Escribano‡

Departamento de Economía
Universidad Carlos III

9 June 2010

Abstract

Exponential models of autoregressive conditional heteroscedasticity (ARCH) are attractive in empirical analysis because they guarantee the non-negativity of volatility, and because they enable richer autoregressive dynamics. However, the currently available models exhibit stability only for a limited number of conditional densities, and the available estimation and inference methods in the case where the conditional density is unknown hold only under very specific and restrictive assumptions. Here, we provide results and simple methods that readily enables consistent estimation and inference of univariate and multivariate power log-GARCH models under very general and non-restrictive assumptions when the power is fixed, via vector ARMA representations. Additionally, stability conditions are obtained under weak assumptions, and the power log-GARCH model can be viewed as nesting certain classes of stochastic volatility models, including the common ASV(1) specification. Finally, our simulations and empirical applications suggest the model class is very useful in practice.

JEL Classication: C22, C32, C51, C52
Keywords: Power ARCH, exponential GARCH, log-GARCH, Multivariate GARCH, Stochastic volatility

---

* We are grateful to Jonas Andersson, Luc Bauwens, Andrew Harvey, Emma Iglesias, SebastienLaurent and participants at the FIBE 2010 conference at NHH (Bergen) for useful comments, suggestions and questions. Funding from the 6th. European Community Framework Programme, MICIN ECO2009-08308 and from The Bank of Spain Excellence Program is gratefully acknowl-edged.
† Corresponding author. Department of Economics, Universidad Carlos III de Madrid (Spain). Email: gsucarra@eco.uc3m.es. Webpage: http://www.eco.uc3m.es/sucarrat/index.html.
‡ Department of Economics, Universidad Carlos III de Madrid (Spain).

1 Introduction

The Autoregressive Conditional Heteroscedasticity (ARCH) class of models due to Engle (1982) is widely used to model the clustering of large (in absolute value) financial returns. Within this class of models a type that is of special interest is exponential ARCH models, because their fitted values of volatility are guaranteed to be non-negative in empirical practice (this is not the case for ordinary ARCH models), and because they enable richer autoregressive volatility dynamics. For example, as an extreme case, all parameters can be negative while volatility is ensured positive. In contrast with ordinary ARCH models, however, in exponential ARCH models stability conditions in general and the existence of unconditional moments in particular depend to a greater extent on the conditional density. For example, the most common exponential ARCH model, Nelson's (1991) EGARCH, is generally not stable for t-distributed errors, see Nelson (1991, p. 365). This is a serious shortcoming since the t-distribution is the preferred choice by practitioners among the densities that are more fat-tailed than the normal, and it has prompted specific work on EGARCH models with t-distributed conditional densities, see for example Harvey and Chakravarty (2010). Furthermore, in contrast to ordinary ARCH models, fewer theoretical results exist that enable consistent estimation and valid asymptotic inference in exponential ARCH models when the density of the conditional density is unknown. For example, Straumann and Mikosch (2006, p. 2452) proves consistency of the Quasi Maximum Likelihood (QML) estimator for Nelson's (1991) univariate EGARCH(1,1). However, the result of Straumann and Mikosch is limited in that it does not apply to higher order EGARCH models, nor to models where the power differs from 2, nor to multivariate versions. Furthermore, their result does not enable ordinary inference strategies: "At the moment we cannot provide a proof of the asymptotic normality of the QMLE in the general EGARCH model.." (same place, p. 2490). Zaffaroni (2009) proves consistency and asymptotic normality of the Whittle estimator for Nelson's (1991) univariate EGARCH(P, Q) model of general orders P and Q. However, a number of restrictive regularity restrictions must be satisfied, including that the conditional density depends on a single parameter only (this is implied by assumption E; see the discussion on pp. 193-194). This effectively rules out skewed distributions like the skewed t and the skewed Generalised Error Distribution (GED), which depend on two parameters, one for shape and one for skewness. Again, this is a severe limitation in practice because the standardised errors of financial returns are often found to be skewed. Dahl and Iglesias (2008) prove consistency and asymptotic normality of QML for a univariate exponential GARCH(1,1) structure that nests the 2nd. power log-GARCH(1,1) with (non-logarithmic) asymmetry, but not the EGARCH of Nelson (1991). Again, their result is limited in the same way as Mikosch and Straumann's in that it does not apply to higher order models, nor to models where the power differs from 2, nor to multivariate versions. Also, many stability properties of their model is unknown. Finally, Kawakatsu (2006) has proposed a multivariate exponential ARCH model, the matrix exponential GARCH, which contains a multivariate version of Nelson's 1991 model. However, general conditions for the existence of its unconditional moments are not available, and a general estimation and inference theory for the case where the conditional density is unknown has yet to be provided.

In this paper we provide a result and methods that enables consistent estimation and ordinary inference methods for a general class of univariate and multivariate exponential ARCH models that we term the power log-GARCH model, via vector autoregressive moving average (VARMA) representations. This class of exponential ARCH models is stable for a much larger class of densities than the EGARCH of Nelson (1991), including the t-distribution. The univariate second power log-GARCH model can be viewed as a dynamic version of Harvey's (1976) multiplicative heteroscedasticity model, and the univariate second power log-GARCH model was first proposed by Pantula (1986), Geweke (1986) and Milhøj (1987). The main motivation was that it ensured non-negative variances. However, it does so at the cost of possibly applying the log-operator on zero-values of the squared residuals of the mean specification, which occurs whenever the residual is equal to zero. If the residuals are rarely equal to zero, then this is not a serious shortcoming in practice since an adequately small positive number may replace the zero value.[1] Nevertheless, this problem is not present in the EGARCH model of Nelson (1991), which might explain why so little work has been devoted to the log-GARCH model compared with the EGARCH model. Some theoretical results apply to structures that nest specific cases of the log-GARCH model, for example some of the results in He et al. (2002), Carrasco and Chen (2002), and Dahl and Iglesias (2008). But these works do not have the log-GARCH model as their main focus.

Another strand of literature that is of relevance for log-GARCH models is the stochastic volatility (SV) literature, since the power log-GARCH can be viewed as nesting certain classes of SV models, including the common autoregressive SV (ASV) model. Viewed in this way, it is well-known that all the coefficients apart from the volatility constant in a univariate second power log-GARCH specification can be estimated consistently (under suitable assumptions) via its autoregressive moving average (ARMA) representation, see for example Psaradakis and Tzavalis (1999), and Francq and Zakoïan (2006). However, the estimate of the volatility constant will generally be biased and the bias depends on the distribution of the standardised error. This is another reason that explains in part the hitherto unattractiveness of the log-GARCH model in empirical finance, since *ad hoc* assumptions and possibly tedious estimation procedures would be needed in order to obtain a valid estimate

[1] What "adequately small" is depends on the data. Financial prices are discrete in the sense that they are recorded with a finite number of digits, typically between 0 and 6. Accordingly, if the positive number is too small then this will induce a negative outlier (when applying the logarithm) that is likely to affect estimation and inference results. Another practical issue to contend with is that the discreteness of a price series can be time-varying. With these two considerations in mind, we use the following simple rule throughout. If $\{\hat{\epsilon}_t\}$ denote the residuals of the mean, then the zero-adjusting value is set equal to the 10% sample quantile of $\{\hat{\epsilon}_t^2\}$.

of the constant. For example, in the context of an SV model, Harvey et al. (1994, section 6) propose a method that can be adapted to the log-GARCH model. Specifically, they propose a way of estimating the bias under the assumption of Student's t distributed standardised errors. By contrast, the result we provide enables a consistent estimate of the bias by means of simple formulas made up of the residuals from the ARMA regression, without having to specify the density of any of the errors (only weak moment assumptions are needed). So a consistent estimate of the variance constant is readily available under very general assumptions on the errors, for any (fixed) power—integer or non-integer—greater than zero.[2] Moreover, when interpreted as an SV model, consistent estimation of the coefficients of the log-GARCH terms can be undertaken with unknown distribution on the SV term under very general assumptions. Our result also holds under very general assumptions when the mean specification differs from zero, and the generalisation to a flexible multivariate version of the power log-GARCH model is straightforward, since consistent estimation can be undertaken via the vector-ARMA (VARMA) representation. Finally, our simulations and our empirical applications suggest our results and methods are very useful for empirical practice.

The rest of the paper is organised as follows. The next section, section 2, presents the univariate power log-GARCH model. The key theoretical result of this paper, proposition 1 and its proof, is contained in subsection 2.2. Section 3 presents the multivariate power log-GARCH. Section 4 contains three empirical applications. Section 5 concludes, whereas the subsequent appendix contains various supporting information. Tables and figures are located at the end.

2 The univariate power log-GARCH model

2.1 Notation and specification

For each t the univariate δth. power log-GARCH(P,Q) model is given by

$$r_t = \mu(\phi, x_t) + \epsilon_t, \quad E(r_t|\mathcal{I}_t) = \mu(\phi, x_t), \tag{1}$$

$$\epsilon_t = \sigma_t z_t, \quad z_t \sim IID(0,1), \quad Prob(z_t = 0) = 0, \quad \sigma_t > 0, \tag{2}$$

$$\begin{aligned} \log \sigma_t^\delta &= h(\gamma, w_t) \\ &= \alpha_0 + \sum_{p=1}^{P} \alpha_p \log|\epsilon_{t-p}|^\delta + \sum_{q=1}^{Q} \beta_q \log \sigma_{t-q}^\delta, \quad \delta > 0, \end{aligned} \tag{3}$$

where $\mu(\phi, x_t) = E(r_t|\mathcal{I}_t)$ is the expectation of r_t conditional on the information set $\mathcal{I}_t$, $Var(r_t|\mathcal{I}_t) = \sigma_t^2$ is the conditional variance of r_t, δ is the power, P is the ARCH

[2] Our estimation methods assumes the power is fixed and known. However, in practice, grid search methods may readily be implemented to search for the power.

order, Q is the GARCH order, ϕ and γ are parameter vectors, and x_t and w_t are the vectors of variables at t in the mean and variance specifications, respectively. The mean $\mu(\phi, x_t)$ allows for a large class of possible specifications, linear or non-linear, and it may contain autoregressive (AR) and moving average (MA) terms. However, it cannot contain functions of σ_t, for example GARCH-in-mean terms,[3] since our methods essentially assume the mean error ϵ_t is determined by $r_t - \mu(\phi, x_t)$ only. But information in w_t may of course appear in x_t, that is, we allow for $x_t \cap w_t \neq \emptyset$. Finally, denoting $P^* = \max\{P, Q\}$, if the roots of the lag polynomial $1 - (\alpha_1 + \beta_1)L - \cdots - (\alpha_{P^*} + \beta_{P^*})L^{P^*}$ are all greater than 1 in modulus, then $\{\log \sigma_t^\delta\}$ is covariance stationary. For common densities like the GED with shape parameter greater than 1, and the Student's t with degrees of freedom greater than 2, $\{\epsilon_t\}$ will in general be covariance stationary, see subsections 2.3 and 3.1.

Table 1 contains the autocorrelations of $\{\epsilon_t^2\}$ for the 1st. and 2nd. power log-GARCH(1,1) specifications for empirically relevant parameter values similar to those of section 4.1. When α_0 is exactly equal to zero, then the autocorrelations of $\{\epsilon_t^2\}$ do not depend on the power. This explains presumably why there is virtually no difference between the autocorrelations in the simulations of the 1st. and 2nd. power specifications. In contrast to the GARCH(1,1) model the autocorrelations of the power log-GARCH(1,1) models depends on the distribution of z_t: The more fat-tailed, the weaker correlations. Nevertheless, the power log-GARCH(1,1) is capable of generating stronger autocorrelations than the GARCH(1,1), although not as persistent—or at least not for the parameter values used in the table (this is consistent with the findings of He et al. (2002)). This might suggest that the log-GARCH(1,1), as Nelson's (1991) EGARCH(1,1), may not be appropriate for some types of financial series, in particular high frequency versions.

2.2 ARMA representations

The error ϵ_t can be written as $\sigma_t z_t = \sigma_t^* z_t^*$, where

$$\sigma_t^* = \sigma_t (E|z_t|^\delta)^{1/\delta}, \quad z_t^* = \frac{z_t}{(E|z_t|^\delta)^{1/\delta}}, \quad E(|z_t^*|^\delta) = 1. \tag{4}$$

This decomposition is useful because it enables an ARMA representation of the power log-GARCH specification that is readily estimable by means of common estimation methods. For example, the δth. power log-ARCH(1) specification is given by $\log \sigma_t^\delta = \alpha_0 + \alpha_1 \log |\epsilon_{t-1}|^\delta$. Adding $\log E|z_t|^\delta + \log |z_t^*|^\delta$ to each side and then adding $E(\log |z_t|^\delta) - E(\log |z_t|^\delta)$ to the right-hand side, yields the AR(1) representation $\log |\epsilon_t|^\delta = \alpha_0^* + \alpha_1 \log |\epsilon_{t-1}|^\delta + u_t^*$, where $\alpha_0^* = \alpha_0 + \log E|z_t|^\delta + E(\log |z_t^*|^\delta)$, and where $u_t^* = \log |z_t^*|^\delta - E(\log |z_t^*|^\delta)$ is a zero-mean IID process. In other words, the power log-ARCH(1) model admits an AR(1) representation. For a given power

[3] This is not a serious drawback since proxies for financial price variability (say, functions of past squared returns, bid-ask spreads, functions of high-low values, etc.) are readily available and can be included as regressors instead.

$\delta > 0$, the parameters α_0^* and α_1 can thus be estimated consistently by means of ordinary estimation methods subject to usual assumptions. However, in order to recover α_0 we need estimates of $\log E|z_t|^\delta$ and $E(\log|z_t^*|^\delta)$, and the proposition we state below provides simple formulas for consistent estimation of $\log E|z_t|^\delta$ and $E(\log|z_t^*|^\delta)$ under very general assumptions. A useful aspect to point out in that regard is that we will in the process also obtain an estimate of $E(\log|z_t|^\delta)$, since $E(\log|z_t|^\delta) = \log E|z_t|^\delta + E(\log|z_t^*|^\delta)$.

More generally the power log-GARCH(P, Q) model with $P \geq Q$ admits the ARMA(P, Q) representation

$$\log|\epsilon_t|^\delta = \alpha_0^* + \sum_{p=1}^{P} \alpha_p^* \log|\epsilon_{t-p}|^\delta + \sum_{q=1}^{Q} \beta_q^* u_{t-q}^* + u_t^* \tag{5}$$

with probability 1, where

$$\begin{aligned}
\alpha_0^* &= \alpha_0 + (1 - \sum_{q=1}^{Q} \beta_q) \cdot \left[\log E|z_t|^\delta + E(\log|z_t^*|^\delta)\right] \\
\alpha_1^* &= \alpha_1 + \beta_1 \\
&\vdots \\
\alpha_P^* &= \alpha_P + \beta_P \\
\beta_1^* &= -\beta_1 \\
&\vdots \\
\beta_Q^* &= -\beta_Q,
\end{aligned}$$

and where $u_t^* = \log|z_t^*|^\delta - E(\log|z_t^*|^\delta) = \log|z_t|^\delta - E(\log|z_t|^\delta)$ as earlier. When $P > Q$, then $\beta_{Q+1} = \cdots = \beta_P = 0$ by assumption. Also, it should be noted that the equations are not affected by the (linear) inclusion of other variables in the log-variance specification (3). The consequence of all this is that consistent estimates of all the ARMA parameters—and hence all the log-GARCH parameters except α_0—can readily be obtained by means of common estimation procedures (least squares, QML in the errors $\{u_t^*\}$, etc.) subject to usual assumptions,[4] as long as the power δ is given, and as long as $P \geq Q$. If $P < Q$, then the ARMA representation may contain common factors. To see this consider for example a δth. power log-GARCH(0,1) specification whose ARMA representation is $\log|\epsilon_t|^\delta = \alpha_0^* + \beta_1 \log|\epsilon_{t-1}|^\delta - \beta_1 u_{t-1}^* + u_t^*$. That is, the AR parameter is equal to the negative of the MA parameter. It is also worth noting the ease with which some non-stationary specifications can be formulated and estimated. For example, an integrated power log-GARCH(1,1) with specification $\log\sigma_t^\delta = \alpha_0 + (1 - \beta_1)\log|\epsilon_{t-1}|^\delta + \beta_1 \log\sigma_{t-1}^\delta$

[4]For example, in the case of estimating an AR(P) representation by means of OLS, the most important assumptions for the current purposes are that the roots of $(1 - \alpha_1 c - \cdots - \alpha_P c^P) = 0$ are outside the unit circle, that $E(u_t^{*2}) < \infty$ and that $E(u_t^{*4}) < \infty$.

can be written as the MA(1) representation $\Delta \log|\epsilon_t|^\delta = \alpha_0^* + \beta_1^* u_{t-1}^* + u_t^*$. More generally, if $\log|\epsilon_t|^\delta$ is I(1), then the estimates of the stationary $AR(P)$ representation $\Delta \log|\epsilon_t|^\delta = \alpha_0^* + \sum_{p=1}^P \alpha_p \Delta \log|\epsilon_{t-p}|^\delta + u_t^*$ can in many cases be used to obtain estimates of the non-stationary representation, or at least as a reasonable approximation.

In order to recover α_0 we need estimates of $\log E(|z_t|^\delta)$ and $E(\log|z_t^*|^\delta)$, and the following proposition gives very general conditions under which they can be estimated consistently after estimation of the ARMA-representation (5).

Proposition 1. Suppose the power δ is known and that a consistent estimation procedure of the ARMA representation (5) of the power log-GARCH specification (3) exhibits the property $\hat{u}_t^* \xrightarrow{P} u_t^*$ for each t, where $\{\hat{u}_t^*\}$ are estimates of $\{u_t^*\}$. If $0 < E|z_t|^\delta < \infty$ and if $|E(\log|z_t|)| < \infty$, then

a)
$$-\log\left[\frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t^*)\right] \xrightarrow{P} E(\log|z_t^*|^\delta), \tag{6}$$

and

b)
$$-\frac{\delta}{2}\log\left[\frac{1}{T}\sum_{t=1}^T \hat{z}_t^{*2}\right] \xrightarrow{P} \log E(|z_t|^\delta), \tag{7}$$

where $\{\hat{z}_t^*\} = \{\epsilon_t/\sqrt[\delta]{\hat{\sigma}_t^{*\delta}}\}$, $\log\hat{\sigma}_t^{*\delta} = \widehat{\log|\epsilon_t|^\delta} - E(\widehat{\log|z_t^*|^\delta})$, and where $\widehat{\log|\epsilon_t|^\delta}$ is the fitted value of the ARMA representation (5).

Proof. In proving a), we first show that $\log E[\exp(u_t^*)] = -E(\log|z_t^*|^\delta)$, then that $\frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t^*) \xrightarrow{P} E[\exp(u_t^*)]$. Since $u_t^* = \log|z_t^*|^\delta - E(\log|z_t^*|^\delta)$ straightforward algebra yields

$$\begin{aligned}
\log E[\exp(u_t^*)] &= \log E\{\exp[\log|z_t^*|^\delta - E(\log|z_t^*|^\delta)]\} \\
&= \log E\left\{\frac{|z_t^*|^\delta}{\exp[E(\log|z_t^*|^\delta)]}\right\} \\
&= \log\left\{\frac{E|z_t^*|^\delta}{\exp[E(\log|z_t^*|^\delta)]}\right\} \\
&= \log E|z_t^*|^\delta - E(\log|z_t^*|^\delta) \\
&= -E(\log|z_t^*|^\delta),
\end{aligned}$$

since $E|z_t^*|^\delta = 1$ and since $|E(\log|z_t^*|^\delta)| < \infty$. The latter follows from the assumptions $0 < E|z_t|^\delta < \infty$ and $|E(\log|z_t|)| < \infty$. Accordingly, $(-1)\cdot\log E[\exp(u_t^*)] = E(\log|z_t^*|^\delta)$. We now turn to the proof of $\frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t^*) \xrightarrow{P} E[\exp(u_t^*)]$. We have that $\frac{1}{T}\sum_{t=1}^T \exp(u_t^*) \xrightarrow{P} E[\exp(u_t^*)]$ due to Khinshine's theorem (see for example

Davidson 1994, theorem 23.5) since $\{u_t^*\}$ is IID, and the properties $E|z_t^*|^\delta = 1$ and $|E(\log|z_t^*|^\delta)| < \infty$ ensure that $E[\exp(u_t^*)]$ exists. Consider $\frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t^*) - \frac{1}{T}\sum_{t=1}^T \exp(u_t^*)$, which can be rewritten as $\frac{1}{T}\sum_{t=1}^T[\exp(\hat{u}_t^*) - \exp(u_t^*)]$. Since $\hat{u}_t^* \xrightarrow{P} u_t^*$ for each t, we have that $\exp(\hat{u}_t^*) \xrightarrow{P} \exp(u_t^*)$ for each t due to the continuity of the $\exp(\cdot)$ function. Accordingly, $\frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t^*) \rightarrow \frac{1}{T}\sum_{t=1}^T \exp(u_t^*)$ as $T \rightarrow \infty$, and since $\frac{1}{T}\sum_{t=1}^T \exp(u_t^*) \rightarrow E[\exp(u_t^*)]$ as $T \rightarrow \infty$ it follows that $\frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t^*) \xrightarrow{P} E[\exp(u_t^*)]$.

We now prove b). Due to the continuity of the $\exp(\cdot)$ operator, the assumption of consistent estimation of the ARMA representation ensures that the fitted values $\{\hat{\sigma}_t^{*\delta}\}$ are consistent estimates of their true counterparts. Next, taking the δth. square root and dividing each ϵ_t by means of $\sqrt[\delta]{\hat{\sigma}_t^{*\delta}}$ implies that the $\{\hat{z}_t^*\}$ are consistent estimates of their true counterparts $\{z_t^*\}$. Finally, using a similar argument to the proof of a) yields that $\frac{1}{T}\sum_{t=1} \hat{z}_t^{*2} \xrightarrow{P} 1/E(|z_t|^\delta)^{2/\delta}$, and so $-\frac{\delta}{2}\log(\frac{1}{T}\sum_{t=1}^T \hat{z}_t^{*2}) \xrightarrow{P} \log E|z_t|^\delta$. ■

When the power δ is equal to 2, then $\log E|z_t|^\delta = 0$ and so the second correction b) is not needed. The a) can thus be viewed as a correction due to the application of the logarithm operator, and b) can be viewed as a "power correction". In the process we obtain estimates of $E(\log|z_t|^\delta)$ and $E|z_t|^\delta$, which are sometimes usefulness in practical applications.[5] Another feature of practical interest is that the corrections constitute a standardisation of the errors. In other words, the sample variance of the $\{\hat{z}_t\}$ will always be equal to or close to 1. The property $\hat{u}_t^* \xrightarrow{P} u_t^*$ is essentially a consequence of consistent estimation of the ARMA representation (5). For the two most common powers, $\delta = 1$ and $\delta = 2$, the proposition holds under very general assumptions. Specifically, the conditions $0 < E|z_t|^\delta < \infty$ and $|E(\log|z_t|)| < \infty$ are satisfied for the most commonly used densities in finance: The Normal, the Generalised Error Distribution (GED) and the Student's t for appropriate number of degrees of freedom. It should also be noted that the proposition is likely to hold in many cases if the $\{\epsilon_t\}$ are estimated in a previous step, as long as the estimation procedure exhibits $\hat{\epsilon}_t \xrightarrow{P} \epsilon_t$ for each t. In words, in sufficiently large samples the estimated residuals are distributed as the true errors, and so are the $\{\log|\hat{\epsilon}_t|^\delta\}$ with probability 1 due to continuity. An important example is the case where one fits a power log-ARCH(P) specification to $\log|\hat{\epsilon}_t|^\delta$ by means of OLS.

2.3 On stability

A serious shortcoming in Nelson's (1991) EGARCH model is that its unconditional variance (and other, higher order integer moments) may not exist for many com-

[5] The estimate of $E|z_t|^\delta$ is obtained by first noting that $E(\log|z_t^*|^\delta) = E(\log|z_t|^\delta) - \log E(|z_t|^\delta)$, and then by setting $E(|z_t|^\delta) = \exp[E(\log|z_t|^\delta) - E(\log|z_t^*|^\delta)]$ replacing the population values by the corresponding estimates.

mon distributions of the standardised errors z_t. For example, if $z_t \overset{IID}{\sim} t_\nu$ in an EGARCH(1,1) with log-variance specification equal to

$$\log \sigma_t^2 = \alpha_0 + \alpha_1[|z_{t-1}| - E|z_{t-1}|] + \theta z_{t-1} + \beta_1 \log \sigma_{t-1}^2,$$

and if the degrees of freedom $\nu > 2$, then the theoretically and empirically unreasonable assumption $\alpha_1 < 0$ is a necessary condition for the existence of the unconditional variance, see condition (A1.6) and the subsequent discussion in Nelson (1991, p. 365). Moreover, if $\theta \neq 0$, then α_1 has to be even more negative for the unconditional variance to exist. These are the shortcomings that prompted the work by Harvey and Chakravarty (2010) on the Beta-t-EGARCH model.

In the δth. power log-GARCH(1,1) with t_ν distributed standardised errors, the unconditional variance will generally exist for $\nu > 2$, regardless of the signs of the parameters α_1 and β_1. The following proposition is a special case of proposition 4 in section 3, and provides a set of exact sufficient conditions.

Proposition 2. Consider a univariate δth. power log-GARCH(1,1) specification with either $z_t \overset{IID}{\sim} GED(\tau), \tau > 1$ or $z_t \overset{IID}{\sim} t(\nu), \nu > 2$. If $|\alpha_1 + \beta_1| < 1$ and if $2\alpha_1(\alpha_1 + \beta_1)^{i-1} \in (-1, 2]$ for each $i = 1, 2, \ldots$, then $E(\epsilon_t^2) < \infty$ and is given by equation (22) (see appendix) with $s = 2$.

Proof. From equation (22) in the appendix with $s = 2$, it follows that $E\left(|z_{t-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}}\right)$ must be finite for each $i = 1, 2, \ldots$ for the expression $E(\epsilon_t^2)$ to exist. For $z_t \sim GED(\tau), \tau > 1$, then $E(|z_t|^c) < \infty$ for $c > -1$, see Zhu and Zinde-Walsh (2009, p. 94). For $z_t \sim t(\nu), \nu > 2$, then $E(|z_t|^c) < \infty$ for $-1 < c < \nu$, see Harvey and Shephard (1996, p. 434). So if $|\alpha_1+\beta_1| < 1$ and $2\alpha_1(\alpha_1+\beta_1)^{i-1} \in (-1, 2]$ for all i, then $E\left(|z_{t-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}}\right) < \infty$ for each $i = 1, 2, \ldots$ Finally, due to proposition 4, the infinite product converges and so $E(\epsilon_t^2) < \infty$. ■

In practice, the restrictions of proposition 2 are very weak and will generally be satisfied, since the typical estimates of α_1 and β_1 are about 0.05 and 0.90, respectively (see the empirical section). In particular, if $|\alpha_1 + \beta_1| < 1$ and if both α_1 and β_1 are equal to or greater than zero, then $2\alpha_1(\alpha_1 + \beta_1)^{i-1}$ takes values in $[0, 2]$ for all $i = 1, 2, \ldots$ Finally, a set of stability conditions for more general univariate δth. power log-GARCH specifications is provided in the following corollary, which follows from proposition 4 in section 3.

Corollary 1. Consider a univariate δth. power log-GARCH(P, Q) model with $P \geq Q$. Suppose the roots of $1 - (\alpha_1 + \beta_1)c - \cdots - (\alpha_P + \beta_P)c^P$ are all greater than 1 in modulus, such that $\log \sigma_t^\delta$ admits the representation $\alpha_0/[1 - (\alpha_1 + \beta_1) - \cdots - (\alpha_P + \beta_P)] + \sum_{i=1}^{\infty} \psi_i \log |z_{t-i}|^\delta$, where $\sum_{i=1}^{\infty} |\psi_i| < \infty$. Then the sth. unconditional moment $E(\epsilon_t^s)$, $s \in \{1, 2, \ldots\}$, exists if $|E(z_t^s)| < \infty$ and if $E|z_{t-i}|^{s\psi_i} < \infty$ for each $i = 1, 2, \ldots$

Proof. Set $M = 1$ in proposition 4 in section 4. ■

The conditions of proposition 1 provides a set of relatively mild restrictions for the sth. unconditional moment to exist. For example, for $E(\epsilon_t^s)$ to exist when $z_t \sim t(\nu), \nu > 2$, we need that $s < \nu$ and $-1 < s\psi_i < \nu$ for each $i = 1, 2, \ldots$ For $z_t \sim GED(\tau), \tau > 1$, $E(\epsilon_t^s)$ will exist as long as $s\psi_i > -1$ for each $i = 1, 2, \ldots$

2.4 On estimation efficiency

It is well known that GARCH models may be consistently estimated via ARMA representations. However, it is also well-known that such estimation methods do not have very good properties. By contrast, estimation of power log-GARCH models via ARMA representations has much better properties for several reasons. First, the error term in GARCH regressions is heteroscedastic. By contrast, the error term in power log-GARCH regressions is IID. Second, the distribution of the error term in the ARCH regression has an exponential-like shape, and takes on values in $[-1, \infty)$. In power log-GARCH regressions, by contrast, it is almost symmetric with the left-tail usually being "longer", and the error takes on values in $(-\infty, \infty)$. This means estimators and test-statistics in the power log-GARCH case are likely to correspond much closer to their asymptotic approximations in finite samples than in the GARCH case, since the convergence to their asymptotic counterparts will be much faster. Also, coefficient tests will exhibit greater power under the alternative, since the error is "smaller" due to the log-transformation. Finally, power log-GARCH regressions impose much weaker restrictions on the parameter space due to the exponential variance specification. In ARCH regressions, by contrast, strong parameter restrictions might be needed in order to ensure positive variance. For these reasons estimation of power log-GARCH models via ARMA representations is likely to work much better than for ordinary ARCH models.

Table 2 contains some simulations that shed light on the finite sample accuracy of some common estimation methods for selected specifications. The finite sample biases are acceptable for many purposes, and estimating the errors $\{\epsilon_t\}$ in a previous step does not seem to affect the estimation precision of α_0 and α_1 substantially in the second step, or at least not when the persistence in the mean specification is small. Tables 3 and 4 compare least squares estimation via ARMA representations with Gaussian QML estimation (in the standardised errors z_t). In table 3 the simulations suggest OLS compares favourably to QML in the estimation of a log-ARCH(1) model, when the standardised errors are more fat-tailed than the normal. When this is the case, then OLS exhibits smaller finite-sample estimation bias of the parameters, and the estimation variances are comparable to or smaller than those of QML. In table 4 the simulations suggest NLS compares well with QML in the estimation of a log-GARCH(1,1) model, when the standardised errors are more fat-tailed than the normal. In this case NLS is more efficient and generally the bias is smaller. The only exception is when $T = 200$. So all in all our simulations suggest

NLS via the ARMA representation compares well with QML in finite samples.[6]

2.5 Inference

In many practical finance applications the mean is either equal to zero or adequately treated as if equal to zero. Or, alternatively, the residuals from the mean specification are treated *as if* observable. When this is the case, and when the logarithmic variance specification does not contain log-GARCH terms, then inference regarding the parameters α—apart from the first element α_0—can be undertaken by means of the usual ordinary least squares theory. When log-GARCH terms enter the power log-variance specification, then a different approach is needed for both the log-ARCH and log-GARCH terms.

Suppose no log-GARCH terms enter the log-variance specification (3), which means (5) reduces to an AR(P) specification with $\alpha_p^* = \alpha_p$. In this case, if $\mathbf{W}$ is the matrix of observations on the regressors, that is, the first column consists of ones and each row of $\mathbf{W}$ is denoted by w_t, then the usual test statistic

$$\frac{\hat{\alpha}_p}{se(\hat{\alpha}_p)} \tag{8}$$

is approximately $N(0,1)$ in large samples for $p = 1, \ldots, P$ under the null of $\alpha_p = 0$, where $\hat{\alpha}_p$ is the OLS estimate of the pth. coefficient, and where $se(\hat{\alpha}_p)$ is the pth. element of the diagonal of the ordinary covariance matrix estimate $\hat{\sigma}^2_{u_t^*}(\mathbf{W}'\mathbf{W})^{-1}$. The $\hat{\sigma}^2_{u_t^*}$ is the standard error of u_t^* and equal to $\frac{1}{T-K}\sum_{t=1}^{T}\hat{u}_t^*$. In order to conduct asymptotic inference regarding α_0, we may proceed by means of a Wald parameter restriction test. In the case when the power $\delta = 2$ for example, OLS estimation provides us with the estimate $\hat{\alpha}_0^*$. Next, we may test $\alpha = 0$ by testing whether $\hat{\alpha}_0^*$ is equal to $-\log \hat{E}[\exp(\hat{u}_t^*)] = -\log[\frac{1}{T}\sum_{t=1}^{T}\exp(\hat{u}_t^*)]$, since $\alpha_0^* = E(\log z_t^2)$ under the null of $\alpha_0 = 0$. The Wald-statistic under the null of $\alpha = 0$ then becomes

$$\frac{\{\hat{\alpha}_0^* + \log \hat{E}[\exp(\hat{u}_t^*)]\}^2}{\widehat{Var}(\hat{\alpha}_0^*)} \overset{asy.}{\sim} \chi^2(1),$$

where $\widehat{Var}(\hat{\alpha}_0^*)$ is the ordinary coefficient variance estimate of α_0^*.

Table 5 contains the simulated finite sample size for two-sided tests of $\alpha_0 = 0$ and $\alpha_1 = 0$ using a nominal size of 5%, and when the power $\delta = 2$. The simulations suggest least squares inference is appropriately sized in finite samples for α_1, since the simulated rejection frequencies range between 4.4% and 5.4% across density shapes. For the test of $\alpha_0 = 0$, the simulations suggest the Wald test is undersized, since the simulated rejection frequencies are close to 0%. Deviations from

[6]Our simulations results depend of course on the exact structure of the numerical algorithms we use. Surely both the NLS and ML algorithms can be improved, so further exploration is needed for a more accurate comparison.

the normal brings the size closer to the nominal, but the discrepancy is nevertheless still notable although acceptable in many practical applications. The undersizedness might suggest that the test lacks power under reasonable departures from the null of $\alpha_0 = 0$. However, additional simulations (not reported) suggest this is not the case. Even though the Wald test is undersized under the null, the test carries reasonable power even when the departure from the null is small.

When the logarithmic variance specification contains log-GARCH terms, then one might consider using the usual theory for inference regarding the parameters of the ARMA representation. However, it is doubtful that this theory will be of value in practice, since the AR and MA coefficient estimates will typically be strongly correlated (recall: $\alpha_p^* = \alpha_p + \beta_p$ and $\beta_q^* = -\beta_q$). An alternative approach is to conduct inference by means of Wald parameter restriction tests. For example, in log-GARCH(1,1) specifications, one may test whether $\alpha_1 = 0$ by testing its implication, namely that $\alpha_t^* = (-1) \cdot \beta_1^*$, and so on. Another possibility is to use the property that a $\{\log \sigma_t^\delta\}$ stationary power log-GARCH specification is (in general) invertible in the ARMA specification. One may then approximate the log-GARCH part by means of a (possibly long) log-ARCH specification, and next conduct inference on each of the lags. A third approach is to use an information criterion to select between alternative specifications. Finally, one may include a regressor that acts as a local approximation to $\log \sigma_{t-1}^\delta$, a "volatility proxy", and subsequently undertake ordinary inference on the associated parameter.

2.6 Stochastic volatility

The simplicity of the estimation and inference methods described hitherto are not affected by "stochastic volatility" terms in the power log-variance specification. To see this define

$$\log \sigma_t^\delta = \alpha_0 + h(\gamma, w_t) + (\log \kappa^\delta) y_t, \qquad \kappa > 0, \qquad y_t \notin x_t, \tag{9}$$

where $h(\gamma, w_t)$ is an abbreviation for the sum of the log-ARCH and log-GARCH terms, that is, $h(\gamma, w_t) = \sum_{p=1}^P \alpha_p \log |\epsilon_{t-p}|^\delta + \sum_{q=1}^Q \beta_q \log \sigma_{t-q}^\delta$, $\{y_t\}$ is IID and independent with $\{z_t\}$, and where the requirement $y_t \notin x_t$ means y_t does not enter the mean equation. It should be noted that, without affecting our argument, y_t can be replaced by y_{t-1}. Indeed, this is needed for $\{\log \sigma_t^\delta\}$ to be a Martingale difference sequence.[7] The specification (9) nests many types of stochastic volatility models, including the autoregressive stochastic volatility (ASV) model of order 1, or ASV(1). Now, denote the information set that does not contain y_t for $\mathcal{I}_t^{sv}$, and denote the information set that includes y_t for $\mathcal{I}_t$. If we condition on $\mathcal{I}_t$, then consistent estimation via the ARMA representation identifies all the parameters, that is, α_0, γ and κ. By contrast, if we condition on $\mathcal{I}_t^{sv}$, then σ_t or volatility is stochastic. In this case consistent estimation via the ARMA representation does

[7]We are grateful to Andrew Harvey for pointing this out to us.

not identify two parameters, namely α_0 and κ, but the others (γ in the example) are identified. Moreover, the estimate of the conditional variance $Var(r_t|\mathcal{I}_t^{sv})$ will be consistent. To see this recall that

$$\begin{aligned}
\epsilon_t &= \sigma_t z_t \\
&= \exp[\alpha_0 + h(\gamma, w_t)]^{\frac{1}{\delta}} \kappa^{y_t} z_t \\
Var(r_t|\mathcal{I}_t) &= \sigma_t^2 \\
Var(r_t|\mathcal{I}_t^{sv}) &= \exp[\alpha_0 + h(\gamma, w_t)]^{\frac{2}{\delta}} E(\kappa^{2y_t}),
\end{aligned}$$

assuming the variances exist. Now, the last term can be rewritten as $Var(r_t|\mathcal{I}_t^{sv}) = \exp[\tilde{\alpha}_0 + h(\gamma, w_t)]^{\frac{2}{\delta}}$, where $\tilde{\alpha}_0 = \alpha_0 + \log E(\kappa^{2y_t})$. In other words, the estimation procedures described above can be used to estimate $\tilde{\alpha}_0$ and γ, while the standardised error will now be equal to $\tilde{z}_t = \frac{\kappa^{y_t} z_t}{E(\kappa^{2y_t})^{\frac{1}{2}}}$ instead of z_t.

2.7 Extensions

Several extensions of the power log-GARCH model suggest themselves. One is the multivariate extension that will be explored in the next section. Another extension, which we do not pursue here, is to specify $\log \sigma_t^\delta$ as a Fractionally Integrated EGARCH process (FIEGARCH) along the lines of Bollerslev and Mikkelsen (1996). A third type of extension consists simply of adding variables linearly to the $\log \sigma_t^\delta$ specification. This can in many case be done straightforwardly without compromising the applicability of the simple estimation and inference methods we have outlined above.

One type of variables that can be added linearly are asymmetry-terms, and in the current context we consider three different types. The first and simplest is of the indicator type $I_{\{z_{t-1}<0\}}$, which are equal to 1 when $z_{t-1} < 0$ and 0 otherwise.[8] In practice this type of asymmetry terms can in general be approximated by means of $I_{\epsilon_{t-1}<0}$, which means the log-GARCH model augmented with such asymmetry terms can be estimated via an ARMA-X representation. As for stability, the following proposition provides quite general sufficient conditions for the existence of the unconditional variance when the standardised errors are either distributed as a student t or as a GED.

Proposition 3. Consider the asymmetric δth. power log-GARCH(1,1) specification

$$\log \sigma_t^\delta = \alpha_0 + \alpha_1 \log|\epsilon_{t-1}|^\delta + \beta_1 \log \sigma_{t-1}^\delta + (\log \lambda^\delta) I_{\{z_{t-1}<0\}}, \quad 0 < \lambda < \infty$$

[8]The original economic justification for asymmetry variables is to capture socalled "leverage" effects in stock markets, see Nelson (1991). So the impact of the regressor is expected to be negative. In some markets, however, for example exchange rate markets, the impact may be either negative or positive depending on which currency is in the denominator of the exchange rate. So we prefer the more general term asymmetry rather than leverage.

with either $z_t \overset{IID}{\sim} GED(\tau), \tau > 1$ or $z_t \overset{IID}{\sim} t(\nu), \nu > 2$. If $|\alpha_1 + \beta_1| < 1$ and if $2\alpha_1(\alpha_1+\beta_1)^{i-1} \in (-1, 2]$ for each $i = 1, 2, \ldots$, then $E(\epsilon_t^2) < \infty$.

Proof. The assumption $|\alpha_1+\beta_1| < 1$ means $\log_t^\delta$ admits the representation $\alpha_0/(1-\alpha_1-\beta_1) + \sum_{i=1}^{\infty}(\alpha_1+\beta_1)^{i-1} \cdot [\alpha_1 \log|z_{t-i}|^\delta + (\log\lambda^\delta)I_{\{z_{t-i}<0\}}]$. This implies that $(\sigma_t^\delta)^{2/\delta} = \sigma_t^2 = \exp[\alpha_0\delta^{-1}/(1-\alpha_1-\beta_1)] \cdot \prod_{i=1}^{\infty}(|z_{t-i}|^{\alpha_1}\lambda^{I_{\{z_{t-i}<0\}}})^{2(\alpha_1+\beta_1)^{i-1}}$, and that $E(\epsilon_t^2) = \exp[\alpha_0\delta^{-1}/(1-\alpha_1-\beta_1)] \cdot \prod_{i=1}^{\infty} a_i$, where $a_i = E[(|z_{t-i}|^{\alpha_1}\lambda^{I_{\{z_{t-i}<0\}}})^{2(\alpha_1+\beta_1)^{i-1}}]$. When $\lambda^{2(\alpha_1+\beta_1)^{i-1}} \in (0,1)$, then $\lambda^{2(\alpha_1+\beta_1)^{i-1}} E[|z_{t-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}}] \leq a_i \leq \lambda^{2(\alpha_1+\beta_1)^{i-1}} E(|z_{t-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}})$, and when $\lambda^{2(\alpha_1+\beta_1)^{i-1}} > 1$, then $E[|z_{t-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}}] \leq a_i \leq \lambda^{2(\alpha_1+\beta_1)^{i-1}} E(|z_{t-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}})$. So each a_i will exist $|\alpha_1+\beta_1| < 1$ and if $2\alpha_1(\alpha_1+\beta_1)^{i-1} \in (-1, 2]$. Finally, since both the two upper bounds and the two lower bounds will tend to 1 as $i \to \infty$, then $a_i \to 1$ and so $E(\epsilon_t^2) < \infty$ by means of the same type of reasoning as in the proof of proposition 2. ■

Another type of asymmetry-term that can also straightforwardly be included and estimated via an ARMA-X representation, are asymmetry terms analogous to those of Glosten et al. (1993). In this case the specification of a δth. power log-GARCH(1,1) takes the form

$$\log\sigma_t^\delta = \alpha_0 + \alpha_1 \log|\epsilon_{t-1}|^\delta + \beta_1 \log\sigma_{t-1}^\delta + \lambda \log|\epsilon_{t-1}|^\delta I_{\{z_{t-1}<0\}}.$$

The exact stability conditions for this specifications are more difficult to derive. Nevertheless, in the case where $\alpha_1, \beta_1 \geq 0$ and $\lambda \in (-1, 0)$, then it follows straightforwardly from the results above that $\alpha_1+\beta_1 < 1$ is a sufficient condition for stability. In particular, the 2nd. moment will exist for student's $t_\nu, \nu > 2$ and GED(τ), $\tau > 1$ distributions. The third type of asymmetry-term that can also straightforwardly be included are analogous to those of Nelson (1991). In this case the specification of a δth. power log-GARCH(1,1) takes the form

$$\log\sigma_t^\delta = \alpha_0 + \alpha_1 \log|\epsilon_{t-1}|^\delta + \beta_1 \log\sigma_{t-1}^\delta + \lambda\epsilon_{t-1}.$$

However, the stability conditions for this type of specification has not been studied (but see Dahl and Iglesias (2008) where $\{\epsilon_t\}$ is assumed strictly stationary and ergodic). Also, it is not clear that the results and methods above are applicable, since least squares and maximum likelihood methods may not provide consistent estimates of an ARMA-X representation.

A second type of variables of special interest that can be added linearly are volatility proxies. For example, if V_t^δ is a volatility proxy in the δth. power, then a diagnostic tool of the volatility proxy that naturally suggests itself is a logarithmic version of Mincer and Zarnowitz (1969) regressions. In logarithmic versions of Mincer-Zarnowitz regressions the log-variance $\log\sigma_t^\delta$ is equal to $\gamma_0 + \gamma_1 \log V_t^\delta$, and the joint test $\gamma_0 = 0$ and $\gamma_1 = 1$ is a test of whether V_t^δ is an "unbiased" estimate of σ_t^δ. Moreover, adding variables to the Mincer-Zarnowitz specification readily permits encompassing tests of V_t^δ. For example, suppose one would like to investigate whether V_t^δ parsimoniously encompasses the other candidate variables (log-ARCH terms, log-GARCH terms, volume variables, etc.). Then this can simply be done in terms of a joint hypothesis test framework of a general specification that nest the variables. A volatility proxy of particular interest is the lag of the equally weighted moving average (EWMA) of past squared errors ($EWMA_{t-1}$). The EWMA is very simple to compute and is always available since it does not require the acquisition of high-frequency data like (say) realised volatility (RV).[9] Also, the EWMA often performs well in practice when compared with many of its technically more sophisticated competitors.

3 A multivariate power log-GARCH model

Financial markets tend to move together, and the extent to which they do so varies over time. This is the main motivation behind multivariate ARCH models, and the implications for asset pricing was the original context in which Bollerslev, Engle and Wooldridge (1988) first proposed a multivariate ARCH moel, see Bauwens et al. (2006) for a recent surveys. For the power log-GARCH class of models, there exists a straightforward multivariate generalisation of the univariate class that can be estimated by means of common methods via its vector ARMA (VARMA) representation. This multivariate version is *not* simply a collection of univariate power log-GARCH models. Indeed, the model is truly multivariate in that P log-ARCH terms of each of the M variables enter each of the M equations, and in that Q log-GARCH terms enter in each of the M equations.

3.1 Notation and specification

Suppose $\{\epsilon_t\}$ is a sequence of $(M \times 1)$ vectors of mean errors. Then the M-dimensional power log-GARCH(P, Q) model is given by

$$\epsilon_t = \text{diag}(\sigma_t) z_t, \qquad z_t | \mathcal{I}_t \sim IID(0, Cov(z_t)), \qquad Var(z_t | \mathcal{I}_t) = I_M, \tag{10}$$

where σ_t is the $(M \times 1)$ vector of conditional standard deviations, $\text{diag}(\sigma_t)$ is an $(M \times M)$ diagonal matrix with σ_t on the diagonal and zeros elsewhere, z_t is the $(M \times 1)$ vector of standardised errors, $Cov(z_t)$ is the variance-covariance matrix of $\{z_t\}$, and $\mathcal{I}_t$ is the conditioning set in question. Here, $\mathcal{I}_t = \{z_{t-1}, z_{t-2}, \ldots\}$. At this point it is worth noting that we do not impose any restrictions on the off-diagonal entries of $Cov(z_t)$. In other words, the covariances of z_t may not be positive definite (we will return to this issue below in subection 3.3). The M-dimensional log-variance

[9] The P period $EWMA_{t-1}$ is equivalent to an integrated ARCH(P) model with the variance constant α_0 being equal to zero, and the ARCH parameters $\alpha_1 = \cdots = \alpha_P$ all equal to $\frac{1}{P}$.

specification is given by

$$\log \sigma_t^{\delta} = \alpha_0 + \sum_{p=1}^{P} \alpha_p \log |\epsilon_{t-p}|^{\delta} + \sum_{q=1}^{Q} \beta_q \log \sigma_{t-q}^{\delta}, \qquad P \geq Q, \tag{11}$$

where

$$\log \sigma_t^{\delta} = \begin{pmatrix} \log \sigma_{1,t}^{\delta} \\ \vdots \\ \log \sigma_{m,t}^{\delta} \\ \vdots \\ \log \sigma_{M,t}^{\delta} \end{pmatrix}, \quad \alpha_0 = \begin{pmatrix} \alpha_{1,0} \\ \vdots \\ \alpha_{m,0} \\ \vdots \\ \alpha_{M,0} \end{pmatrix}, \quad \alpha_p = \begin{pmatrix} \alpha_{11.p} & \cdots & \alpha_{1m.p} & \cdots & \alpha_{1M.p} \\ \vdots & \ddots & \vdots & & \vdots \\ \alpha_{m1.p} & \cdots & \alpha_{mm.p} & \cdots & \alpha_{mM.p} \\ \vdots & & \vdots & \ddots & \vdots \\ \alpha_{11.p} & \cdots & \alpha_{1m.p} & \cdots & \alpha_{1M.p} \end{pmatrix},$$

$$\log |\epsilon_{t-p}|^{\delta} = \begin{pmatrix} \log |\epsilon_{1,t-p}|^{\delta} \\ \vdots \\ \log |\epsilon_{m,t-p}|^{\delta} \\ \vdots \\ \log |\epsilon_{M,t-p}|^{\delta} \end{pmatrix}, \quad \beta_q = \begin{pmatrix} \beta_{11.q} & \cdots & \beta_{1m.q} & \cdots & \beta_{1M.q} \\ \vdots & \ddots & \vdots & & \vdots \\ \beta_{m1.q} & \cdots & \beta_{mm.q} & \cdots & \beta_{mM.q} \\ \vdots & & \vdots & \ddots & \vdots \\ \beta_{M1.q} & \cdots & \beta_{Mm.q} & \cdots & \beta_{MM.q} \end{pmatrix}.$$

For example, the specification of a two-dimensional δth. power log-ARCH(1) model is

$$\log \sigma_{1,t}^{\delta} = \alpha_{1,0} + \alpha_{11.1} \log |\epsilon_{1,t-1}|^{\delta} + \alpha_{12.1} \log |\epsilon_{2,t-1}|^{\delta}$$

$$\log \sigma_{2,t}^{\delta} = \alpha_{2,0} + \alpha_{21.1} \log |\epsilon_{1,t-1}|^{\delta} + \alpha_{22.1} \log |\epsilon_{2,t-1}|^{\delta},$$

whereas the specification of a two-dimensional δth. power log-GARCH(2,1) is

$$\begin{aligned} \log \sigma_{1,t}^{\delta} = {} & \alpha_{1,0} + \alpha_{11.1} \log |\epsilon_{1,t-1}|^{\delta} + \alpha_{12.1} \log |\epsilon_{2,t-1}|^{\delta} + \alpha_{11.2} \log |\epsilon_{2,t-2}|^{\delta} \\ & + \alpha_{12.2} \log |\epsilon_{2,t-2}|^{\delta} + \beta_{11,1} \log \sigma_{1,t-1}^{\delta} + \beta_{12,1} \log \sigma_{2,t-1}^{\delta} \end{aligned}$$

$$\begin{aligned} \log \sigma_{2,t}^{\delta} = {} & \alpha_{2,0} + \alpha_{21.1} \log |\epsilon_{1,t-1}|^{\delta} + \alpha_{22.1} \log |\epsilon_{2,t-1}|^{\delta} + \alpha_{21.2} \log |\epsilon_{2,t-2}|^{\delta} \\ & + \alpha_{22.2} \log |\epsilon_{2,t-1}|^{\delta} + \beta_{21,1} \log \sigma_{1,t-1}^{\delta} + \beta_{22,1} \log \sigma_{2,t-1}^{\delta}, \end{aligned}$$

and so on.

The following proposition provides a general set of non-restrictive sufficient conditions for the existence of the unconditional moments.

Proposition 4. Consider an M-dimensional δth. power log-GARCH(P, Q) model with $P \geq Q$ that admits the representation $\log \sigma_t^{\delta} = \Psi_0 + \sum_{i=1}^{\infty} \Psi_i \log |z_{t-i}|^{\delta}$ with $\{\Psi_i\}$ being an absolutely summable sequence of $(M \times M)$ matrices. Then the sth. unconditional moment $E(\epsilon_{m,t}^s) = \exp(s\delta^{-1}\psi_{m,0}) \cdot \prod_{i=1}^{\infty} E\left[|z_{1,t-i}|^{s\psi_{i,m1}} |z_{2,t-i}|^{s\psi_{i,m2}} \cdots |z_{M,t-i}|^{s\psi_{i,mM}}\right]$, $s \in \{1, 2, \ldots\}$, of variable $m \in \{1, \ldots, M\}$ exists if $E\left[|z_{1,t-i}|^{s\psi_{i,m1}} |z_{2,t-i}|^{s\psi_{i,m2}} \cdots |z_{M,t-i}|^{s\psi_{i,mM}}\right] < \infty$ for each i

Proof. By definition, absolute summability of the matrix sequence $\{\Psi_i\}$ means $\sum_{i=1}^{\infty} |\psi_{i,mn}| < \infty$ for each $m, n \in \{1, 2, \ldots, M\}$. Next, a sufficient condition for an infinite product $\prod_{i=1}^{\infty} a_i$ to converge to a finite, nonzero number is that the series $\sum_{i=1}^{\infty} |a_i - 1|$ converges (Gradshteyn and Ryzhik (2007, section 0.25)). Since $E\left[|z_{1,t-i}|^{s\psi_{i,m1}}|z_{2,t-i}|^{s\psi_{i,m2}} \cdots |z_{M,t-i}|^{s\psi_{i,mM}}\right] \rightarrow 1$ as $i \rightarrow \infty$ due to absolute summability, it follows that $|a_i - 1| \rightarrow 0$ as $i \rightarrow \infty$. Accordingly, if $a_i = E\left[|z_{1,t-i}|^{s\psi_{i,m1}}|z_{2,t-i}|^{s\psi_{i,m2}} \cdots |z_{M,t-i}|^{s\psi_{i,mM}}\right] < \infty$ for each i, it follows that $E(\epsilon_{mt}^s)$ exists. ■

In practice, the natural condition to check is whether all the eigenvalues of the $(M \times M)$ matrix $\sum_{p=1}^{P^*}(\alpha_p + \beta_p)$ are smaller than 1 in modulus. If this is the case, then $\{\Psi_i\}$ is absolutely summable. Whether the second condition is satisfied or not, that is, $E\left[|z_{1,t-i}|^{s\psi_{i,m1}}|z_{2,t-i}|^{s\psi_{i,m2}} \cdots |z_{M,t-i}|^{s\psi_{i,mM}}\right] < \infty$ for each i, will depend on the distribution of z_t.

3.2 VAR and VARMA representations

The parameters of the power log-GARCH(P, Q) model can be consistently estimated by means of common methods via its VARMA representation subject to appropriate assumptions. Specifically, the VAR(P) representation of an M-dimensional power log-ARCH(P) model is given by

$$\log|\epsilon_t|^\delta = \alpha_0^* + \sum_{p=1}^{P} \alpha_p \log|\epsilon_{t-p}|^\delta + u_t^*, \tag{12}$$

where α_p is defined as above, and where

$$\log|\epsilon_t|^\delta = \begin{pmatrix} \log|\epsilon_{1,t}|^\delta \\ \vdots \\ \log|\epsilon_{m,t}|^\delta \\ \vdots \\ \log|\epsilon_{M,t}|^\delta \end{pmatrix}, \quad \alpha_0^* = \begin{pmatrix} \alpha_{1,0} + \log E|z_{1,t}|^\delta + E(\log|z_{1,t}^*|^\delta) \\ \vdots \\ \alpha_{m,0} + \log E|z_{m,t}|^\delta + E(\log|z_{m,t}^*|^\delta) \\ \vdots \\ \alpha_{M,0} + \log E|z_{M,t}|^\delta + E(\log|z_{M,t}^*|^\delta) \end{pmatrix}$$

$$u_t^* = \begin{pmatrix} \log|z_{1,t}^*|^\delta - E(\log|z_{1,t}^*|^\delta) \\ \vdots \\ \log|z_{m,t}^*|^\delta - E(\log|z_{m,t}^*|^\delta) \\ \vdots \\ \log|z_{M,t}^*|^\delta - E(\log|z_{M,t}^*|^\delta) \end{pmatrix}.$$

In other words, $\{u_t^*\}$ is now a vector zero-mean IID process. The VARMA(P, Q) representation of an M-dimensional power log-GARCH(P, Q) model is given by

$$\log|\epsilon_t|^\delta = \alpha_0^* + \sum_{p=1}^{P} \alpha_p^* \log|\epsilon_{t-p}|^\delta + \sum_{q=1}^{Q} \beta_q^* u_{t-q}^* + u_t^*, \tag{13}$$

where

$$\alpha_p^* = \alpha_p + \beta_p, \qquad \beta_q^* = -\beta_q, \qquad \alpha_0^* = \alpha_0 + (I_M - \sum_{q=1}^{Q} \text{diag}(\beta_q))[\log E|z_t|^\delta + E(\log |z_t^*|^\delta)],$$

$$\alpha_0 = \begin{pmatrix} \alpha_{1,0} \\ \vdots \\ \alpha_{m,0} \\ \vdots \\ \alpha_{M,0} \end{pmatrix}, \qquad \log E|z_t|^\delta + E(\log |z_t^*|^\delta) = \begin{pmatrix} \log E|z_{1,t}|^\delta + E(\log |z_{1,t}^*|^\delta) \\ \vdots \\ \log E|z_{m,t}|^\delta + E(\log |z_{m,t}^*|^\delta) \\ \vdots \\ \log E|z_{M,t}|^\delta + E(\log |z_{M,t}^*|^\delta) \end{pmatrix}.$$

As in the univariate case, if $P > Q$ then $\beta_{Q+1} = \cdots = \beta_P = 0$ by assumption, and the formulas in proposition 1 can be used to estimate $\log E|z_t|^\delta$ and $E(\log |z_t^*|^\delta)$ once the VARMA representation has been estimated.

In theory, multivariate δth. power log-GARCH models can be consistently estimated by means of common estimation methods (say, least squares or QML) via its VARMA representation. However, it is well known that, in practice, VARMA models may not be readily estimated due to numerical issues. The question of how well the available estimation estimation algorithms actually work for the multivariate log-GARCH we leave for future research.

3.3 Modelling conditional covariances

A key motivation for multivariate GARCH models is that they can be used in the computation of portfolio variances. However, unless restrictions are imposed on the off-diagonals of the covariance matrix $H_t = Cov(\epsilon_t)$, then one cannot be ensured that such portfolio variances will be positive. This is the motivation for the socalled positive definiteness requirement of H_t. In the power log-GARCH model this amounts to positive definiteness of $Cov(z_t)$. The methods we have outlined so far do not rely on any specific property on the covariance-matrix of $\{z_t\}$. Indeed, the only assumption we rely upon is that $Var(z_t) = I_M$. Our estimation methods are thus compatible with $Cov(z_t)$ being positive definite or not.

4 Empirical applications

One type of problems that specifications in the power log-GARCH class of models might be particularly suited for are those that involves many explanatory variables. Indeed, specifications contained in the 2nd. power log-GARCH-X class have proved useful in explanatory exchange rate modelling and forecasting, in stock price volatility proxy evaluation, and in value-at-risk portfolio forecasting, see Bauwens et al. (2006), Rime and Sucarrat (2007), Bauwens and Sucarrat (2008), and Sucarrat and

Escribano (2009). Here we explore further the usefulness of power log-GARCH models in three empirical applications. The first two are devoted to the univariate and multivariate analysis of stock market variability. This choice is motivated by the fact that the modelling and forecasting of stock market variability is one of the more frequent use of GARCH models. The third empirical application applies the power log-GARCH model to the complex problem of modelling daily electricity prices. The relative change in daily electricity prices differ from most other financial returns in that they exhibit strong and complex patterns of autoregressive persistence and periodicity, both in the mean and volatility specifications.

4.1 Log-ARCH models vs. the GARCH(1,1)

In the ARCH class of models introduced by Engle (1982), the GARCH(1,1) specification of Bollerslev (1986) is possibly the most common model of financial variability. In this subsection we therefore compare the GARCH(1,1) with four simple models of the power log-ARCH class in modelling SP500 variability both in-sample and out-of-sample. We choose the SP500 stock market index because it is widely used for the purpose of comparison.

Our estimation sample is 1 January 2001 - 31 December 2005 (1305 observations), whereas our out-of-sample evaluation period is 1 January 2006 - 30 October 2009 (997 observations).[10] The ARCH models are all fitted to the "demeaned" log-returns in percent (see figure 1), where the mean is an OLS estimated AR(1) specification equal to $r_t = \phi_0 + \phi_1 r_{t-1} + \epsilon_t$. The same mean specification is used out-of-sample to demean the returns. The estimation results of the GARCH(1,1) specification, a 2nd. power log-GARCH(1,1) specification, a 1st. power log-GARCH(1,1) specification, a 2nd. power log-ARCH(1) specification augmented with the log of a 20-day[11] equally weighted moving average (EWMA$(20)_{t-1}$) of past squared residuals as regressor, that is, a volatility proxy, and a 2nd. power log-ARCH(0) specification with $\log EWMA(20)_{t-1}$ as only regressor, are[12]

$$\hat{\sigma}_t^2 = 0.006 + 0.061\hat{\epsilon}_{t-1}^2 + 0.934\hat{\sigma}_{t-1}^2 \tag{14}$$

[10] The source of the raw series is Reuters-EcoWin Pro, and the series identifier is ew:usa15510200.

[11] 20 trading days corresponds in general to 4 weeks or approximately one calendar month.

[12] AR_1 is the 1st. order serial correlation of $\{\hat{z}_t\}$, whereas $ARCH_1$, $ARCH_2$ and $ARCH_5$ are the 1th., 2nd. and 5th. order serial correlations of $\{\hat{z}_t^2\}$. The values in square brackets are p-values from Ljung and Box (1979) tests of no serial correlation up to the lag order in question. JB is the Jarque and Bera (1980) test statistic with the associated p-value in square brackets. Var is the sample variance of the standardised residuals $\{\hat{z}_t\}$. The GARCH(1,1) model is estimated by means of QML using the `garch()` function in the *tseries* R package, see Trapletti and Hornik (2009). The log-GARCH(1,1) model is estimated by means of NLS via the ARMA representation using the `arma()` function, which is also part of the *tseries* R package. The volatility proxy model is estimated by means of OLS.

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var
$\{\hat{z}_t\}$ in-sample:	-0.04 [0.17]	-0.04 [0.12]	0.02 [0.21]	0.02 [0.30]	23 [0.00]	1.00
$\{\hat{z}_t\}$ out-of-sample:	-0.10 [0.00]	-0.04 [0.16]	-0.01 [0.35]	0.02 [0.73]	355 [0.00]	1.10

$$\log \hat{\sigma}_t^2 = 0.091 + 0.0533 \log \hat{\epsilon}_{t-1}^2 + 0.9242 \log \hat{\sigma}_{t-1}^2 \tag{15}$$

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var
$\{\hat{z}_t\}$ in-sample:	0.00 [0.99]	0.01 [0.68]	0.06 [0.05]	0.12 [0.00]	$8K$ [0.00]	0.99
$\{\hat{z}_t\}$ out-of-sample:	-0.11 [0.00]	-0.01 [0.84]	0.04 [0.42]	0.10 [0.01]	237 [0.00]	1.17

$$\log \hat{\sigma}_t = 0.045 + 0.0532 \log |\hat{\epsilon}_{t-1}| + 0.9243 \log \hat{\sigma}_{t-1} \tag{16}$$

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var
$\{\hat{z}_t\}$ in-sample:	0.00 [0.99]	0.01 [0.68]	0.07 [0.05]	0.12 [0.00]	$8K$ [0.00]	0.99
$\{\hat{z}_t\}$ out-of-sample:	-0.11 [0.00]	-0.01 [0.84]	0.04 [0.42]	0.10 [0.01]	237 [0.00]	1.17

$$\log \hat{\sigma}_t^2 = 0.100 + 0.005 \log \hat{\epsilon}_{t-1}^2 + 0.901 \log \widehat{EWMA}(20)_{t-1} \tag{17}$$

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var
$\{\hat{z}_t\}$ in-sample:	-0.04 [0.14]	-0.04 [0.12]	0.03 [0.20]	0.02 [0.20]	18 [0.00]	1.00
$\{\hat{z}_t\}$ out-of-sample:	-0.10 [0.00]	-0.04 [0.23]	-0.02 [0.44]	0.01 [0.84]	430 [0.00]	1.12

$$\log \hat{\sigma}_t^2 = 0.089 + 0.906 \log \widehat{EWMA}(20)_{t-1} \tag{18}$$

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var
$\{\hat{z}_t\}$ in-sample:	-0.04 [0.13]	-0.04 [0.18]	0.03 [0.25]	0.02 [0.21]	15 [0.00]	1.00
$\{\hat{z}_t\}$ out-of-sample:	-0.10 [0.00]	-0.04 [0.26]	-0.02 [0.47]	0.01 [0.86]	425 [0.00]	1.12

Figure 1 graphs the demeaned residuals, whereas figures 2-3 contain graphs of the associated in-sample and out-of-sample standard deviations and standardised residuals, respectively. It should be noted that the standard deviations and the standardised residuals of the 1st. and 2nd. power log-GARCH(1,1) models are indistinguishable from each other graphically, and similarly for the 4th. and 5th. models. The reason why the 1st. and 2nd. power log-GARCH(1,1) models' are so similar is presumably that the variance constant α_0 in both log-GARCH(1,1) specifications are close to zero. When α_0 is exactly equal to zero, then all δth. power log-GARCH(1,1) models are equivalent. This can be seen by studying the effect of changing δ in the ARMA representation (5), and in the autocorrelation structure of $\{\epsilon_t^2\}$ (see the appendix). The similarity between the 4th. and 5th. models suggest the log-ARCH(1) term explain very little of the time-varying volatility compared with the volatility proxy.

Additionally, there are at least five more features worth noticing from the estimation results of equations (14)-(18), and from the figures. First, the estimates $\hat{\alpha}_0, \hat{\alpha}_1$ and $\hat{\beta}_1$ of the log-GARCH(1,1) specifications are relatively similar to those of the GARCH(1,1), and the sum $\hat{\alpha}_1 + \hat{\beta}_1$ is close to (but below) 1 in all three cases. Second, from the figures it is clear that the fitted conditional standard deviations are relatively similar. The greatest difference is that the two log-GARCH(1,1) specifications (standard deviations in red) seem to generate lower standard deviations in high variability periods (towards the end of 2008). A third feature of interest is that the estimation results above suggest all the models appear to be relatively (parameter) stable across the two periods. This interpretation is due to the sample variances $Var(\hat{z}_t)$ being relatively close to 1 out-of-sample. Fourth, the JB-statistics in the results above, and the figure with the standardised residuals, suggest the log-GARCH(1,1) specifications generate standardised residuals that are more fat-tailed than those of the first and last two models. Finally, the ARCH diagnostic tests in the results above do not suggest the log-GARCH(1,1) specifications depict the ARCH in SP500 variability as well as the other models both in and out-of-sample.

4.2 Multivariate ARCH models

In this subsection we compare four joint models of the daily S&P500 and the FTSE Euro 100 (EUR100) stock market index returns from 1 January 2001 to 30 October 2009.[13] The S&P500 return series is the same demeaned series as in the previous subsection, and we use the same approach in demeaning the EUR100 returns. Both demeaned series are displayed in figure 1. Also, we repeat the exercise of using the data up to and including 2005 in estimating the models, and then generating out-of-sample conditional variances, standardised residuals, etc., from the beginning of 2006 until 30 October 2009.

The first of the three models we fit is a diagonal BEKK(1,1,1) model estimated by means of multivariate Gaussian ML.[14] The second model is a two dimensional 2nd. power log-GARCH(1,1) model, which we estimate by means of two-stage equation-by-equation OLS via the VARMA representation. (In the first OLS step we set the common lag-length in the VAR representation as equal to the natural logarithm of the sample size, see Kascha (2007) for a comparison of various VARMA estimation methods, including ours.) The third is a four dimensional 2nd. power log-ARCH(1) model augmented with volatility proxy dynamics (we use the same method as in the

[13]The source of the EUR100 series is also Reuters-EcoWin Pro, and its series identifier is ew:emu15555.

[14]This model we estimate with the OxMetrics package G@RCH 5.1, see Laurent (2007). We only report the estimation results of the diagonal of the covariance matrix H_t, and the estimates are reported in their GARCH(1,1) form. For example, in the first equation, the estimate 0.01 is equal to c_{11}^2, 0.06 is equal to a_{11}^2, and so on.

previous subsection in generating a volatility proxy), which we estimate equation-by-equation by means of OLS. Finally, the fourth model is a parsimonious version of the third model. The estimation results and diagnostics of the models are:[15]

$$\begin{pmatrix}\hat{\sigma}^2_{1t}\\ \hat{\sigma}^2_{2t}\end{pmatrix} = \begin{pmatrix}0.01\\ 0.01\end{pmatrix} + \begin{pmatrix}0.06 & 0\\ 0 & 0.09\end{pmatrix}\cdot\begin{pmatrix}\hat{\epsilon}^2_{1t-1}\\ \hat{\epsilon}^2_{2t-1}\end{pmatrix} + \begin{pmatrix}0.93 & 0\\ 0 & 0.90\end{pmatrix}\cdot\begin{pmatrix}\hat{\sigma}^2_{1t-1}\\ \hat{\sigma}^2_{2t-1}\end{pmatrix}$$

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var	$Corr$
In-sample $\{\hat{z}_{1t}\}$:	−0.05 [0.07]	−0.00 [0.88]	0.03 [0.51]	0.01 [0.26]	90 [0.00]	1.02	0.00
$\{\hat{z}_{2t}\}$:	−0.06 [0.03]	−0.04 [0.14]	0.01 [0.31]	−0.00 [0.02]	62 [0.00]	1.00	0.00
Out-of-sample $\{\hat{z}_{1t}\}$:	−0.01 [0.68]	−0.05 [0.09]	0.01 [0.24]	0.03 [0.43]	243 [0.00]	1.08	0.61
$\{\hat{z}_{2t}\}$:	−0.00 [0.93]	−0.02 [0.46]	0.02 [0.61]	−0.01 [0.04]	210 [0.00]	1.10	0.61

$$\begin{pmatrix}\log\hat{\sigma}^2_{1t}\\ \log\hat{\sigma}^2_{2t}\end{pmatrix} = \begin{pmatrix}-0.06\\ 0.40\end{pmatrix} + \begin{pmatrix}0.05 & 0.02\\ 0.11 & 0.02\end{pmatrix}\cdot\begin{pmatrix}\log\hat{\epsilon}^2_{1t-1}\\ \log\hat{\epsilon}^2_{2t-1}\end{pmatrix} + \begin{pmatrix}0.32 & 0.35\\ 0.32 & 0.62\end{pmatrix}\cdot\begin{pmatrix}\log\hat{\sigma}^2_{1t-1}\\ \log\hat{\sigma}^2_{2t-1}\end{pmatrix}$$

AR eigenvalues: 0.92, 0.08 MA eigenvalues: -0.84, -0.10

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var	$Corr$
In-sample $\{\hat{z}_{1t}\}$:	0.08 [0.01]	0.01 [0.66]	0.06 [0.07]	0.06 [0.00]	1107 [0.00]	1.00	0.51
$\{\hat{z}_{2t}\}$:	−0.02 [0.43]	0.06 [0.02]	0.02 [0.06]	0.03 [0.00]	12840 [0.00]	1.00	0.51

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var	$Corr$
Out-of-sample $\{\hat{z}_{1t}\}$:	−0.03 [0.35]	0.02 [0.56]	0.11 [0.00]	0.17 [0.00]	50617 [0.00]	1.30	0.60
$\{\hat{z}_{2t}\}$:	−0.01 [0.84]	0.01 [0.67]	0.04 [0.46]	−0.01 [0.06]	66 [0.00]	0.76	0.60

$$\begin{pmatrix}\log\hat{\sigma}^2_{1t}\\ \log\hat{\sigma}^2_{2t}\\ \log\widehat{EWMA}_{1t}\\ \log\widehat{EWMA}_{2t}\end{pmatrix} = \begin{pmatrix}0.04\\ 0.33\\ -0.01^{*}\\ 0.01^{*}\end{pmatrix} + \begin{pmatrix}0.01 & -0.01 & 0.77 & 0.12\\ 0.08 & -0.01 & 0.20 & 0.73\\ -0.00 & 0.00 & 0.97 & 0.02\\ 0.00 & -0.00 & 0.03 & 0.97\end{pmatrix}\cdot\begin{pmatrix}\log\hat{\epsilon}^2_{1t-1}\\ \log\hat{\epsilon}^2_{2t-1}\\ \log\widehat{EWMA}_{1t-1}\\ \log\widehat{EWMA}_{2t-1}\end{pmatrix}$$

AR eigenvalues: 0.996+0i, 0.95+0i, -0.00+0i, -0.00+0i

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var	$Corr$
In-sample $\{\hat{z}_{1t}\}$:	0.05 [0.07]	−0.04 [0.19]	0.03 [0.28]	0.01 [0.33]	19 [0.00]	1.00	0.52
$\{\hat{z}_{2t}\}$:	0.00 [0.97]	−0.03 [0.35]	0.00 [0.64]	0.00 [0.00]	73 [0.00]	1.00	0.52
Out-of-sample $\{\hat{z}_{1t}\}$:	−0.02 [0.61]	−0.02 [0.45]	−0.01 [0.68]	0.02 [0.84]	399 [0.00]	1.16	0.61
$\{\hat{z}_{2t}\}$:	0.01 [0.68]	0.09 [0.00]	0.04 [0.01]	−0.02 [0.00]	207 [0.00]	1.03	0.61

[15]Diagnostic tests are univariate and the values in square brackets are p-values. The AR eigenvalues are of the $\alpha_1 + \beta_1$ matrix, whereas the MA eigenvalues are of the $-\beta_1$ matrix.

$$\begin{pmatrix} \log \hat{\sigma}^2_{1t} \\ \log \hat{\sigma}^2_{2t} \end{pmatrix} = \begin{pmatrix} 0.04 \\ 0.23 \end{pmatrix} + \begin{pmatrix} 0.78 & 0.11 \\ 0.32 & 0.69 \end{pmatrix} \cdot \begin{pmatrix} \log \widehat{EWMA}_{1t-1} \\ \log \widehat{EWMA}_{2t-1} \end{pmatrix}$$

	AR_1	$ARCH_1$	$ARCH_2$	$ARCH_5$	JB	Var	$Corr$
In-sample $\{\hat{z}_{1t}\}$:	0.05 [0.07]	−0.03 [0.27]	0.03 [0.34]	0.01 [0.38]	15 [0.00]	1.00	0.52
$\{\hat{z}_{2t}\}$:	0.02 [0.57]	−0.02 [0.39]	0.04 [0.27]	0.01 [0.20]	41 [0.00]	1.00	0.52
Out-of-sample $\{\hat{z}_{1t}\}$:	0.02 [0.61]	−0.02 [0.44]	−0.01 [0.70]	0.02 [0.86]	412 [0.00]	1.16	0.61
$\{\hat{z}_{2t}\}$:	0.01 [0.72]	0.08 [0.01]	0.04 [0.02]	−0.02 [0.00]	207 [0.00]	1.03	0.61

It should be noted that the variance constants with an asterisk "*" in the log-ARCH(1) model with volatility dynamics (the third model) are not adjusted by means of the formulas in proposition 1.

Several of the features from the previous subsection are reproduced in the multivariate exercise: The diagnostic tests above suggest the log-GARCH(1,1) is not capable of depicting ARCH variability (for neither series) as well as the other models, and from figures 4 and 5 it is clear that the fitted conditional standard deviations are relatively similar across models, and that the log-GARCH(1,1) yields standardised residuals that are more fat-tailed than those of the two other models. A property that was less apparent in the previous subsection is the instability of the log-GARCH(1,1). This is suggested by the fact that the out-of-sample variances of its standardised residuals in the results above are much further away from 1. Finally, the two models with volatility proxy dynamics suggest that SP500 and EUR100 volatility can be parsimoniously modelled by means of the volatility proxies. The results of the firs model with volatility proxy dynamics suggests the log-ARCH(1) terms have little or no impact, and that the volatility proxies are endogenously determined. The fourth model, which do not contain the log-ARCH(1) terms nor equations for volatility proxy dynamics, yields parameters estimates similar to the third model, and diagnostics that are almost identical. Indeed, graphically the both the conditional standard deviations and the standardised residuals are indistinguishable, so only those of the first model with volatility proxy dynamics are contained in figures 4 and 5.

4.3 Modelling daily electricity prices

Daily electricity prices are often characterised by strong autoregressive persistence and ARCH, and by day-of-the week and seasonal effects in both the mean and variance specifications, see for example Escribano et al. (2009), and Koopman et al. (2007). The power log-GARCH model augmented with explanatory variables (the "power log-GARCH-X" model) permits a flexible and rich characterisation of all these effects in a single model that can readily be estimated by means of OLS. As an illustration we revisit the Spanish daily electricity price data in Escribano et al. (2009), which spans the period 1 January 1998 to 31 December 2003 ($T = 2191$ observations), see the upper two graphs of figure 6.

If $r_t = \Delta \log S_t$ denotes the return of the daily Spanish electricity price S_t, then we start from the general model

$$\begin{aligned} r_t &= \phi_0 + \sum_{m \in M} \phi_m r_{t-m} + \sum_{n=1}^{34} \eta_n x_{nt} + \epsilon_t, \\ \epsilon_t &= \sigma_t z_t, \quad z_t \sim IID(0,1), \quad Prob(z_t = 0) = 0, \quad \sigma_t > 0, \\ \log \sigma_t^2 &= \alpha_0 + \sum_{p \in P} \alpha_p \log \epsilon_{t-p}^2 + \sum_{p \in P} \lambda_p \log \epsilon_{t-p}^2 I_{\epsilon_{t-p}<0} + \omega_0 \log EWMA(7)_{t-1} \\ &\quad + \sum_{d=1}^{33} \omega_d y_{dt}, \end{aligned}$$

where $M = \{1, \ldots, 14, 21, 28, 35\}$, and where the 34 x_{nt} variables comprise 12 variables $I_{r_t>0}, \ldots, I_{r_{t-7}>0}, I_{r_{t-14}>0}, I_{r_{t-21}>0}, I_{r_{t-28}>0}, I_{r_{t-35}>0}$ intended to capture asymmetries in the constant, a GARCH-in-mean proxy $(r_{t-1}^2 - 1)$, 4 threshold variables $I_{r_{t-1}<-0.5}, I_{r_{t-1}>0.5}, I_{r_{t-2}<-0.5}, I_{r_{t-2}>0.5}$ that seek to capture the (possibly differing) impact of large negative and large positive price changes, respectively, 6 day-of-the-week dummies (Tuesday to Sunday) and 11 month-of-the-year dummies (February to December). This means the general unrestricted mean specification contains a total of 51 deletable regressors, and one regressor (the constant) that is restricted from deletion in the specification search. In the log-variance specification $P = \{1, \ldots, 7, 14, 21, 28, 35\}$, $EWMA(7)_{t-1}$ is a rolling average of $\epsilon_{t-1}^2, \ldots, \epsilon_{t-7}^2$, and the 33 y_{dt} variables are the same as the 34 x_{nt} variables except for the GARCH-in-mean proxy which is not included among the y_{dt} variables. This means the general unrestricted log-variance specification contains a total of 56 deletable regressors, and one regressor (the constant) that is restricted from deletion in the specification search. Automated General-to-Specific (GETS) multi-path model selection with AutoSEARCH (Sucarrat 2010) yields a parsimonious model, which we further simplify by imposing economically meaningful parameter restrictions among the regressors.

The end result is ($|t|$-statistics in parentheses and p-values in square brackets)

$$\begin{aligned}
\hat{r}_t &= \underset{(4.28)}{0.114} - \underset{(22.70)}{0.045} \cdot (8r_{t-1} + 4r_{t-2} + 3r_{t-3} + 2r_{t-4} + 2r_{t-5} + 2r_{t-6}) \\
&\quad + \underset{(15.42)}{0.056} \cdot (2r_{t-7} + 2r_{t-14} + r_{t-21} + r_{t-28} + r_{t-35}) + \underset{(7.05)}{0.186}(r_{t-1}^2 - 1) \\
&\quad + \underset{(32.05)}{0.018} \cdot (10I_{r_t>0} + I_{r_{t-1}>0} + I_{r_{t-6}>0}) - \underset{(3.68)}{0.016} \cdot (I_{r_{t-7}>0} + I_{r_{t-14}>0}) \\
&\quad - \underset{(9.00)}{0.040}(Sat_t + 2Sun_t) - \underset{(2.72)}{0.023}Dec_t + \underset{(2.16)}{0.062}I_{r_{t-2}<-0.5} \qquad (19)
\end{aligned}$$

$$\begin{aligned}
\log \hat{\sigma}_t^2 &= -2.517 - \underset{(2.40)}{0.054}\log \hat{\epsilon}_{t-3}^2 + \underset{(6.67)}{0.393}\log EWMA(7)_{t-1} - \underset{(3.60)}{0.247}(I_{r_t>0} + I_{r_{t-1}>0}) \\
&\quad - \underset{(3.47)}{0.329}(Wed_t + Fri_t + Sat_t) - \underset{(3.09)}{0.392}(Apr_t + Jul_t) + \underset{(2.19)}{1.218}I_{r_{t-1}<-0.5} \qquad (20)
\end{aligned}$$

$$\hat{z}_t \sim SGED(\hat{\tau}_{shape} = 1.31, \hat{\tau}_{skew} = 0.82) \qquad (21)$$

AR_1	AR_6	AR_7	AR_{14}	$ARCH_1$	$ARCH_6$	$ARCH_7$	$ARCH_{14}$	R^2
−0.00 [0.95]	0.01 [0.98]	−0.00 [0.99]	−0.02 [0.21]	0.02 [0.48]	−0.02 [0.58]	0.03 [0.47]	0.03 [0.77]	0.70

The model is well-specified in the sense that the AR and ARCH tests exhibit little or no signs of autocorrelation in the standardised residuals (see also the bottom graph of figure 6), and in the squared standardised residuals. In the mean specification, the lag structure suggests a negative but declining effect of the previous 6 days, whereas the effect of lag-multiples of 7—a day-of-the-week effect—is positive albeit also declining. The GARCH-in-mean proxy $(r_{t-1}^2 - 1)$, is positive which means that very large returns in absolute value (of 100% or more) has a positive impact of about 0.19 on next day's returns. The next two terms suggests there is a asymmetry in the size of return, and that the effect depends on the day-of-the-week. The retention of the Saturday and Sunday dummies suggests prices tend to fall in the weekends compared with the price level of the rest of the week (the effect is the double for Sunday compared with Saturday), and similarly the effect of December is negative. Finally, the last term suggests there is a large positive effect—a "return-reversal" effect—from large falls of more than 0.5 in the log-price on the previous day.

The log-variance specification is a measure of the time-varying accuracy of the mean specification: The greater $\log \sigma_t^2$ is, the more inaccurate is the mean specification. Only the 3rd. log-ARCH term is retained in the specification search, which suggests that there is little ARCH and that the one there is negative (cyclical). By contrast, the lagged impact of the log of the volatility proxy $EWMA(7)_t$ is positive and about 0.4. The retention of the asymmetry terms suggests positive returns affect the precision of the mean equation negatively both contemporaneously and tomorrow, whereas the day-of-the-week and month-of-the-year dummies means there are some periodicity and seasonality effects on the precision. By contrast, a drop in the log-price larger than 0.5 increases the precision of the mean specification. Finally, the fitted Skew GED distribution of the standardised residuals suggests that they are fat-tailed with shape parameter equal to 1.31 ($\tau_{shape} = 2$ corresponds to the normal and $\tau_{shape} \in (1, 2)$ means the tails are fatter), and that they are negatively skewed with skewness parameter equal to 0.82 ($\tau_{skew} = 1$ corresponds to symmetry and $\tau_{skew} \in (0, 1)$ means the density is negatively skewed).[16]

5 Conclusions

We have provided results and generic methods that readily enables consistent estimation and inference of a general class of univariate and multivariate exponential ARCH models, even when the conditional density is not known. Specifically, consistent estimation and inference of univariate and multivariate power log-GARCH models can be undertaken subject to very weak assumptions for a given power, via vector ARMA representations. The unconditional moments of the power log-GARCH model exist subject to restrictions that are much weaker than for other exponential ARCH models, say, Nelson's (1991) EGARCH model, including when the power log-GARCH model is augmented with asymmetry terms. Furthermore, our simulations suggest least squares estimation and inference via ARMA representations fares well in finite samples, and the power log-GARCH model can be viewed as nesting certain classes of stochastic volatility models, for example the common ASV(1) specification. In the empirical section we estimate and evaluate several models contained in the power log-GARCH-X class. In the case of US and Euro-zone stock market index returs, the results suggest certain specifications in the power log-GARCH class of models compares well with more ordinary GARCH models, both in-sample and out-of-sample. Finally, our third empirical application shows that the complex problem of modelling daily electricity prices with rich persistence and periodicity structures in both the mean and volatility specifications can be readily resolved by means of the methods we propose.

References

Bauwens, L., S. Laurent, and J. Rombouts (2006). Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics 21*, 79–109.

Bauwens, L., D. Rime, and G. Sucarrat (2006). Exchange Rate Volatility and the Mixture of Distribution Hypothesis. *Empirical Economics 30*, 889–911.

Bauwens, L. and G. Sucarrat (2008). General to Specific Modelling of Exchange Rate Volatility: A Forecast Evaluation. Forthcoming in the *International Journal of*

[16] We use the method proposed by Fernández and Steel (1998) to skew Nelson's (1991) parametrisation of the standardised GED/exponential power distribution.

Forecasting. UC3M Working Paper version: WP 08-18 in the Economic Series (`http://hdl.handle.net/10016/2591`).

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics 31*, 307–327.

Bollerslev, T., R. F. Engle, and J. Wooldridge (1988). A Capital Asset Pricing Model with Time-Varying Covariances. *Econometric Reviews 11*, 143–172.

Bollerslev, T. and H. O. Mikkelsen (1996). Modelling and pricing long memory in stock market volatility. *Journal of Econometrics 73*, 151–184.

Carrasco, M. and X. Chen (2002). Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory 18*, pp. 17–39.

Dahl, C. M. and E. M. Iglesias (2008). The limiting properties of the QML in a general class of asymmetric volatility models. Unpublished manuscript, currently available as `https://www.msu.edu/~iglesia5/DahlIglesias2009.pdf`.

Davidson, J. (1994). *Stochastic Limit Theory*. Oxford: Oxford University Press.

Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflations. *Econometrica 50*, 987–1008.

Escribano, Á., J. I. Peña, and P. Villaplana (2009). Modelling Electricity Prices: International Evidence. Unpublished manuscript.

Fernández, C. and M. Steel (1998). On Bayesian Modelling of Fat Tails and Skewness. *Journal of the American Statistical Association 93*, 359–371.

Francq, C. and J.-M. Zakoïan (2006). Linear-representation Based Estimation of Stochastic Volatility Models. *Scandinavian Journal of Statistics 33*, 785–806.

Geweke, J. (1986). Modelling the Persistence of Conditional Variance: A Comment. *Econometric Reviews 5*, 57–61.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, 1779–1801.

Gradshteyn, I. S. and I. M. Ryzhik (2007). *Table of Integrals, Series and Products*. New York: Academic Press. Seventh edition. Available via `http://books.google.com/`.

Harvey, A. C. (1976). Estimating Regression Models with Multiplicative Heteroscedasticity. *Econometrica 44*, 461–465.

Harvey, A. C. and T. Chakravarty (2010). Beta-t-EGARCH. Unpublished working paper.

Harvey, A. C., E. Ruiz, and N. Shephard (1994). Multivariate Stochastic Variance Models. *Review of Economic Studies 61*, 247–264.

Harvey, A. C. and N. Shephard (1996). Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns. *Journal of Business and Economic Statistics 14*, 429–434.

He, C., T. Teräsvirta, and H. Malmsten (2002). Moment Structure of a Family of First-Order Exponential GARCH Models. *Econometric Theory 18*, 868–885.

Jarque, C. and A. Bera (1980). Efficient Tests for Normality, Homoskedasticity, and Serial Independence. *Economics Letters 6*, pp. 255–259.

Kascha, C. (2007). A Comparison of Estimation Methods for Vector Autoregressive Moving-Average Models. Department of Econmics, European University Institute: EUI Working Paper ECO 2007/12. Available as `http://cadmus.iue.it/dspace/bitstream/1814/6921/1/ECO-2007-12.pdf`.

Kawakatsu, H. (2006). Matrix exponential GARCH. *Journal of Econometrics 134*, 95–128.

Koopman, S. J., M. Ooms, and M. A. Carnero (2007). Periodic Seasonal REG-ARFIMA-GARCH Models for Daily Electricity Spot Prices. *Journal of the American Statistical Association 102*, 16–27.

Laurent, S. (2007). *Estimating and Forecasting ARCH Models Using G@RCH 5*. London: Timberlake Consultants Ltd.

Ljung, G. and G. Box (1979). On a Measure of Lack of Fit in Time Series Models. *Biometrika 66*, 265–270.

Milhøj, A. (1987). A Multiplicative Parametrization of ARCH Models. Research Report 101, University of Copenhagen: Institute of Statistics.

Mincer, J. and V. Zarnowitz (1969). The Evaluation of Economic Forecasts. In J. Zarnowitz (Ed.), *Economic Forecasts and Expectations*. New York: National Bureau of Economic Research.

Nelson, D. B. (1991). Conditional Heteroscedasticity in Asset Returns: A New Approach. *Econometrica 51*, 485–505.

Pantula, S. (1986). Modelling the Persistence of Conditional Variance: A Comment. *Econometric Reviews 5*, 71–73.

Psaradakis, Z. and E. Tzavalis (1999). On regression-based tests for persistence in logarithmic volatility models. *Econometric Reviews 18*, 441–448.

Rime, D. and G. Sucarrat (2007). Exchange Rate Variability, Market Activity and Heterogeneity. Universidad Carlos III de Madrid: UC3M Working Paper 07-70 in the Economic Series. Url: `http://e-archivo.uc3m.es:8080/dspace/bitstream/10016/984/1/we077039.pdf`.

Straumann, D. and T. Mikosch (2006). Quasi-Maximum-Likelihood Estimation in Conditionally Heteroscedastic Time Series: A Stochastic Recurrence Equations Approach. *The Annals of Statistics 34*, 2449–2495.

Sucarrat, G. (2010). AutoSEARCH: An R Package for Automated Financial GETS Modelling. Available via `http://www.eco.uc3m.es/sucarrat/autosearch/autosearch.htm`.

Sucarrat, G. and Á. Escribano (2009). Automated Model Selection in Finance: General-to-Specific Modelling of the Mean, Variance and Density. Available as: `http://www.eco.uc3m.es/sucarrat/research/autofim.pdf`.

Trapletti, A. and K. Hornik (2009). *tseries: Time Series Analysis and Computational Finance*. R package available via `http://CRAN.R-project.org/package=tseries`.

Zaffaroni, P. (2009). Whittle estimation of EGARCH and other exponential volatility models. *Journal of Econometrics 151*, 190–200.

Zhu, D. and V. Zinde-Walsh (2009). Properties and estimation of asymmetric exponential power distribution. *Journal of Econometrics 148*, 86–99.

Appendix: $E(\epsilon_t^s)$ and $E(\epsilon_t^2\epsilon_{t-j}^2)$ for the δth. power log-GARCH(1,1) model

For the δth. power log-GARCH(1,1) model the unconditional variance of $\{\epsilon_t\}$, and the autocovariances and autocorrelations of $\{\epsilon_t^2\}$, are all made up of $E(\epsilon_t^2)$, $E(\epsilon_t^2\epsilon_{t-j}^2)$ and $E(\epsilon_t^4)$. Assuming the terms exist and that $|\alpha_1 + \beta_1| < 1$, then the terms can readily be estimated through Bootstrapping and/or simulations if the relevant expressions are not available in closed form. The expression for the sth. unconditional moment $E(\epsilon_t^s)$, $s \in \{1, 2, \ldots\}$, is

$$E(\epsilon_t^s) \quad = \quad E(z_t^s) \cdot \exp\left(\frac{s\alpha_0}{\delta \cdot (1 - \alpha_1 - \beta_1)}\right) \cdot \prod_{i=1}^{\infty} E\left(|z_{t-i}|^{s\alpha_1(\alpha_1+\beta_1)^{i-1}}\right), \quad (22)$$

whereas for $j = 1, 2, \ldots$ the formula of $E(\epsilon_t^2 \epsilon_{t-j}^2)$ is

$$
\begin{aligned}
E(\epsilon_t^2 \epsilon_{t-j}^2) \quad = \quad & \exp\left[\frac{2\alpha_0}{\delta}\left(\sum_{i=1}^{j}(\alpha_1+\beta_1)^{i-1} + \frac{1+(\alpha_1+\beta_1)^j}{(1-\alpha_1-\beta_1)}\right)\right] \\
& \cdot \prod_{i=1}^{j} E\left(|z_{t-i-1}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}+2I_{(i=j)}}\right) \\
& \cdot \prod_{i=1}^{\infty} E\left(|z_{t-j-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}\cdot[1+(\alpha_1+\beta_1)^j]}\right), \qquad (23)
\end{aligned}
$$

where $I_{(i=j)}$ is an indicator function equal to 1 when $i = j$ and zero otherwise.

Table 1: Autocorrelations of $\{\epsilon_t^2\}$ for the 1st. and 2nd. power log-GARCH(1,1) model (with $z_t \sim GED(\tau)$) and for the GARCH(1,1) model, with $\alpha_0 = 0.005$, $\alpha_1 = 0.05$, $\beta_1 \in \{0.9, 0.94\}$ and $\tau \in \{1.1, 2\}$

	Log-GARCH(1,1,$\delta = 1$)				Log-GARCH(1,1,$\delta = 2$)				GARCH(1,1)	
	$\beta_1 = 0.90$		$\beta_1 = 0.94$		$\beta_1 = 0.90$		$\beta_1 = 0.94$		$\beta_1 = 0.9$	$\beta_1 = 0.94$
Lag	$\tau = 2$	$\tau = 1.1$	$\tau = 2$	$\tau = 1.1$	$\tau = 2$	$\tau = 1.1$	$\tau = 2$	$\tau = 1.1$		
1	0.10	0.06	0.33	0.21	0.09	0.06	0.33	0.21	0.07	0.15
2	0.09	0.06	0.31	0.18	0.09	0.06	0.32	0.19	0.07	0.15
3	0.09	0.05	0.30	0.17	0.09	0.05	0.31	0.17	0.07	0.15
4	0.08	0.05	0.30	0.16	0.08	0.05	0.29	0.16	0.06	0.15
5	0.08	0.05	0.28	0.15	0.08	0.05	0.28	0.16	0.06	0.15
6	0.07	0.05	0.27	0.14	0.07	0.05	0.27	0.15	0.06	0.15
7	0.07	0.05	0.26	0.14	0.07	0.04	0.26	0.14	0.05	0.15
8	0.07	0.04	0.25	0.13	0.07	0.04	0.24	0.13	0.05	0.14
9	0.06	0.04	0.24	0.12	0.06	0.04	0.23	0.12	0.05	0.14
10	0.06	0.04	0.23	0.11	0.06	0.04	0.22	0.11	0.05	0.14
11	0.05	0.04	0.22	0.10	0.06	0.04	0.22	0.11	0.04	0.14
12	0.05	0.03	0.21	0.10	0.05	0.03	0.21	0.10	0.04	0.14
13	0.05	0.03	0.20	0.10	0.05	0.03	0.20	0.09	0.04	0.14
14	0.05	0.03	0.19	0.09	0.05	0.03	0.19	0.09	0.04	0.14
15	0.04	0.02	0.19	0.09	0.05	0.03	0.19	0.09	0.04	0.13
16	0.04	0.03	0.18	0.08	0.04	0.03	0.18	0.08	0.03	0.13
17	0.04	0.03	0.18	0.07	0.04	0.03	0.17	0.07	0.03	0.13
18	0.04	0.02	0.17	0.07	0.04	0.03	0.17	0.07	0.03	0.13
19	0.04	0.02	0.16	0.07	0.03	0.02	0.16	0.07	0.03	0.13
20	0.03	0.02	0.16	0.06	0.03	0.02	0.16	0.06	0.03	0.13
21	0.03	0.02	0.15	0.06	0.03	0.02	0.15	0.06	0.03	0.13
22	0.03	0.02	0.15	0.05	0.03	0.02	0.15	0.05	0.02	0.13
23	0.03	0.02	0.14	0.05	0.03	0.02	0.14	0.05	0.02	0.12
24	0.03	0.02	0.14	0.05	0.03	0.02	0.13	0.05	0.02	0.12
25	0.03	0.02	0.13	0.05	0.03	0.02	0.13	0.05	0.02	0.12

The simulations of the expectations in the log-GARCH autocorrelations are in R with 100 000 draws of $z_t \sim GED(\tau)$. $GED(\tau)$ is short for Generalised Error Distribution with shape parameter τ. The standard normal is obtained when $\tau = 2$, and $\tau \in (1, 2)$ yields standardised densities that are more fat-tailed than the normal.

Table 2: Finite sample precision of estimation methods

Model: ϕ_1	α_0	α_1	β_1	Meth.	T	$M(\hat{\phi}_1)$	$V(\hat{\phi}_1)$	$M(\hat{\alpha}_0)$	$V(\hat{\alpha}_0)$	$M(\hat{\alpha}_1)$	$V(\hat{\alpha}_1)$	$M(\hat{\beta}_1)$	$V(\hat{\beta}_1)$	$M[\hat{E}(\log z_t^2)]$	$V[\hat{E}(\log z_t^2)]$
	0			OLS	200			-0.004	0.009					-1.263	0.014
					500			-0.003	0.004					-1.270	0.006
					1000			-0.001	0.002					-1.271	0.003
	0	0.10		OLS	200			-0.013	0.020	0.093	0.005			-1.268	0.015
					500			-0.008	0.008	0.096	0.002			-1.269	0.006
					1000			-0.005	0.004	0.098	0.001			-1.272	0.003
	0	0.10	0.8	TS-OLS	200			-0.396	0.825	0.100	0.006	0.478	0.504	-1.297	0.018
					500			-0.131	0.114	0.108	0.002	0.673	0.084	-1.297	0.007
					1000			-0.077	0.030	0.109	0.001	0.715	0.025	-1.293	0.003
				NLS	200			-0.432	0.454	0.104	0.004	0.443	0.244	-1.294	0.047
					500			-0.103	0.059	0.105	0.001	0.705	0.041	-1.276	0.016
					1000			-0.030	0.009	0.102	0.001	0.771	0.008	-1.275	0.005
	0	0.05	0.9	TS-OLS	200			-0.822	1.864	0.049	0.005	0.255	1.009	-1.293	0.019
					500			-0.396	0.549	0.061	0.002	0.559	0.382	-1.295	0.007
					1000			-0.216	0.213	0.057	0.001	0.711	0.146	-1.297	0.003
				NLS	200			-0.794	0.692	0.048	0.007	0.272	0.356	-1.292	0.090
					500			-0.405	0.399	0.057	0.001	0.557	0.247	-1.283	0.007
					1000			-0.128	0.127	0.057	0.001	0.783	0.076	-1.288	0.016
0.1	0	0.10		OLS	200	0.097	0.006	-0.023	0.022	0.082	0.005			-1.262	0.014
					500	0.101	0.002	-0.013	0.008	0.092	0.002			-1.274	0.006
					1000	0.102	0.001	-0.004	0.004	0.097	0.001			-1.271	0.003

The simulation DGP is $r_t = \phi_1 r_{t-1} + \epsilon_t$, $\epsilon_t = \sigma_t z_t$, $z_t \overset{IID}{\sim} N(0,1)$, $\log\sigma_t^2 = \alpha_0 + \alpha_1 \log\epsilon_{t-1}^2 + \beta_1 \log\sigma_{t-1}^2$ for $t = 1, \ldots, T$. $M(\cdot)$ and $V(\cdot)$ are the sample mean and sample variance of the estimates, respectively. The TS-OLS method consists of estimating an (assumed invertible) ARMA representation in two steps. First, use OLS to estimate the residuals by means of a (possibly) long AR-regression that approximates the ARMA representation. Second, estimate by means of OLS the ARMA representation using the estimated residuals from the first step as MA regressors. NLS is the non-linear least squares estimation procedure and is implemented using the `arma` function in the R package `tseries` (Trapletti and Hornik 2009). Simulations in R with 1000 replications, and a prior burn-in sample of 100 observations was discarded at each replication in order to avoid initial value issues.

Table 3: Finite sample precision of estimation methods: OLS *vs.* ML for a log-ARCH(1)

Method	$f(z_t)$	T	$M(\hat{\alpha}_0)$	$V(\hat{\alpha}_0)$	$M(\hat{\alpha}_1)$	$V(\hat{\alpha}_1)$
OLS	$N(0,1)$	200	-0.013	0.020	0.093	0.005
		500	-0.008	0.008	0.096	0.002
		1000	-0.005	0.004	0.098	0.001
	$GED(1.1)$	200	-0.025	0.039	0.094	0.005
		500	-0.012	0.017	0.097	0.002
		1000	-0.001	0.008	0.099	0.001
	$st(4.1)$	200	-0.040	0.065	0.093	0.005
		500	-0.019	0.024	0.099	0.002
		1000	-0.015	0.014	0.097	0.001
ML	$N(0,1)$	200	0.013	0.014	0.107	0.003
		500	0.000	0.005	0.104	0.001
		1000	0.001	0.003	0.101	0.000
	$GED(1.1)$	200	0.038	0.039	0.114	0.004
		500	0.035	0.015	0.112	0.002
		1000	0.027	0.009	0.109	0.001
	$st(4.1)$	200	0.023	0.078	0.116	0.005
		500	0.035	0.039	0.115	0.002
		1000	0.027	0.026	0.111	0.001

The DGP of the simulation is $r_t = \sigma_t z_t, z_t \stackrel{IID}{\sim} f(z_t), \log \sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2$ for $t = 1, \ldots, T$ with $\alpha_0 = 0$ and $\alpha_1 = 0.1$. $N(0,1)$ is short for standard normal, $GED(1.1)$ is short for Generalised Error Distribution with shape parameter $\tau = 1.1$ (the standard normal is obtained when $\tau = 2$) and $st(4.1)$ is short for standardised t-distribution with 4.1 degrees of freedom. ML estimation consists of Gaussian maximum likelihood estimation with initial parameter values provided by OLS. ML estimation is implemented as a Newton-Raphson algorithm with analytical gradient and Hessian, unit step-size and 0.0001 as convergence criterion in the log-likelihood. $M(\cdot)$ and $V(\cdot)$ are the sample mean and variances of the estimates, respectively. Simulations in R with 1000 replications, and a prior burn-in sample of 100 observations was discarded at each replication in order to avoid initial value issues.

Table 4: Finite sample precision of estimation methods: NLS *vs.* ML for a log-GARCH(1,1)

Meth.	$f(z_t)$	T	$M(\hat{\alpha}_0)$	$V(\hat{\alpha}_0)$	$M(\hat{\alpha}_1)$	$V(\hat{\alpha}_1)$	$M(\hat{\beta}_1)$	$V(\hat{\beta}_1)$
NLS	$N(0,1)$	200	-0.432	0.454	0.104	0.004	0.443	0.244
		500	-0.103	0.059	0.105	0.001	0.705	0.041
		1000	-0.030	0.009	0.102	0.001	0.771	0.008
	$GED(1.1)$	200	-0.530	0.746	0.108	0.004	0.472	0.224
		500	-0.156	0.155	0.105	0.001	0.698	0.051
		1000	-0.043	0.010	0.104	0.001	0.766	0.006
	$st(4.1)$	200	-0.568	0.803	0.106	0.005	0.455	0.225
		500	-0.129	0.100	0.105	0.001	0.714	0.040
		1000	-0.043	0.010	0.103	0.001	0.769	0.006
ML	$N(0,1)$	200	-0.290	0.739	0.092	0.004	0.592	0.512
		500	-0.128	0.095	0.103	0.002	0.691	0.050
		1000	-0.097	0.031	0.104	0.001	0.719	0.016
	$GED(1.1)$	200	-0.425	2.076	0.097	0.004	0.558	0.659
		500	-0.189	0.190	0.106	0.002	0.678	0.068
		1000	-0.115	0.078	0.107	0.001	0.723	0.024
	$st(4.1)$	200	-0.430	1.563	0.096	0.005	0.550	0.541
		500	-0.208	0.235	0.106	0.002	0.663	0.089
		1000	-0.144	0.106	0.109	0.001	0.699	0.040

The simulation DGP is $r_t = \sigma_t z_t$, $z_t \overset{IID}{\sim} f(z_t)$, $\log\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta \log\sigma_{t-1}^2$ for $t = 1, \ldots, T$ with $\alpha_0 = 0, \alpha_1 = 0.1$ and $\beta = 0.8$. $N(0,1)$ is short for standard normal, $GED(1.1)$ is short for Generalised Error Distribution with shape parameter $\tau = 1.1$ (the standard normal is obtained when $\tau = 2$) and $st(4.1)$ is short for standardised t-distribution with 4.1 degrees of freedom. NLS estimation is undertaken using the `arma` function in the R package `tseries` (Trapletti and Hornik 2009). ML is short for Gaussian maximum likelihood estimation with initial parameter values provided by consistent two-stage OLS. Numerically, ML is implemented as a Newton-Raphson algorithm with analytical gradient and Hessian, unit step-size and 0.0001 as convergence criterion in the log-likelihood. $M(\cdot)$ and $V(\cdot)$ are the sample mean and variances of the estimates, respectively. Simulations in R with 1000 replications, and a prior burn-in sample of 100 observations was discarded at each replication in order to avoid initial value issues.

Table 5: Finite sample size in the logarithmic variance specification, using a nominal level of 5%

H_0	H_1	Fitted specification	T	$\tau = 1.1$	$\tau = 2$	$\tau = 3$
$\alpha_1 = 0$	$\alpha_1 \neq 0$	$\alpha_0 + \alpha_1 \log \epsilon_{t-1}^2$	10	0.054	0.049	0.047
			100	0.047	0.046	0.044
			1000	0.052	0.049	0.051
			10000	0.048	0.049	0.048
$\alpha_0 = 0$	$\alpha_0 \neq 0$	α_0	10	0.070	0.044	0.027
			100	0.027	0.004	0.001
			1000	0.015	0.001	0.000
			10000	0.020	0.001	0.002

The simulation DGP is $r_t = \epsilon_t$, $\epsilon_t = \sigma_t z_t$, $z_t \overset{IID}{\sim} GED(\tau)$, $\log \sigma_t^2 = 0$, for $t = 1, \ldots, T$. $GED(\tau)$ is short for Generalised error distribution with shape parameter τ (= 2 yields the normal, $\tau \in (1, 2)$ yields distributions that are more fat-tailed compared with the normal and $\tau > 2$ yields distributions that are less fat-tailed) and tests are two-sided. Simulations in R with 10 000 replications.

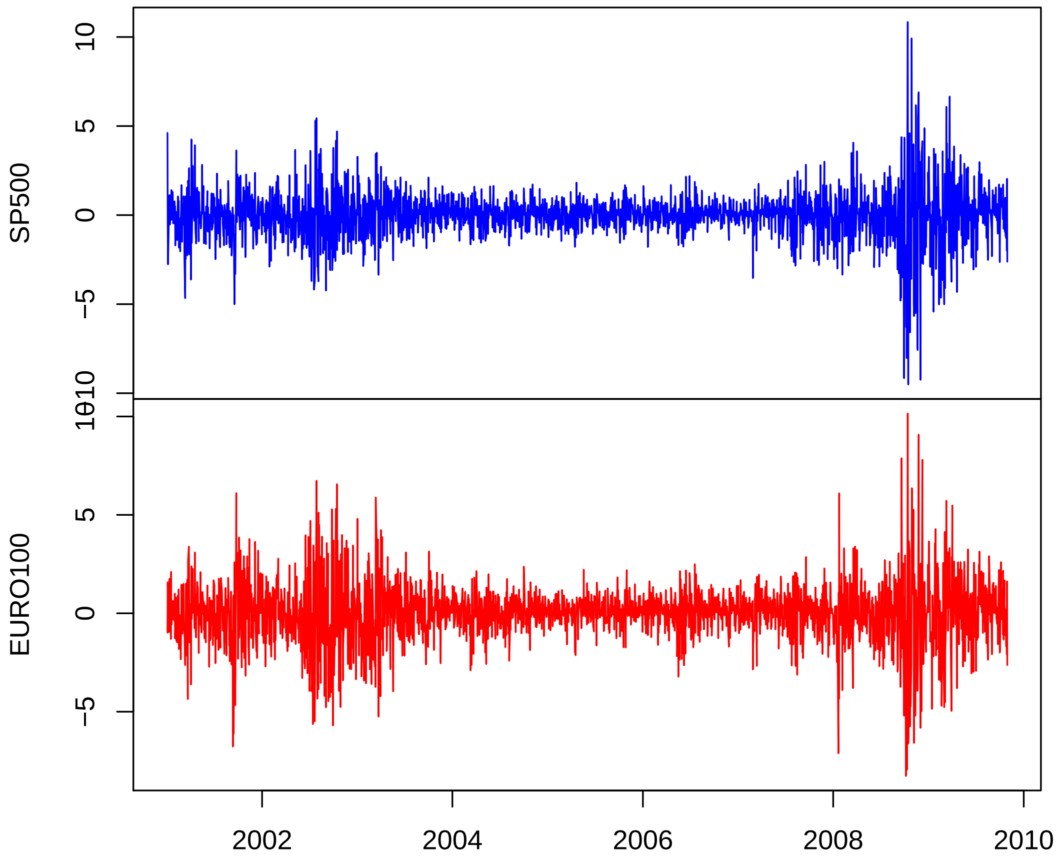

Figure 1: Daily demeaned log-returns (in percent) of the SP500 and EURO100 stock market indices 1 January 2001 - 30 October 2009 (2302 observations)

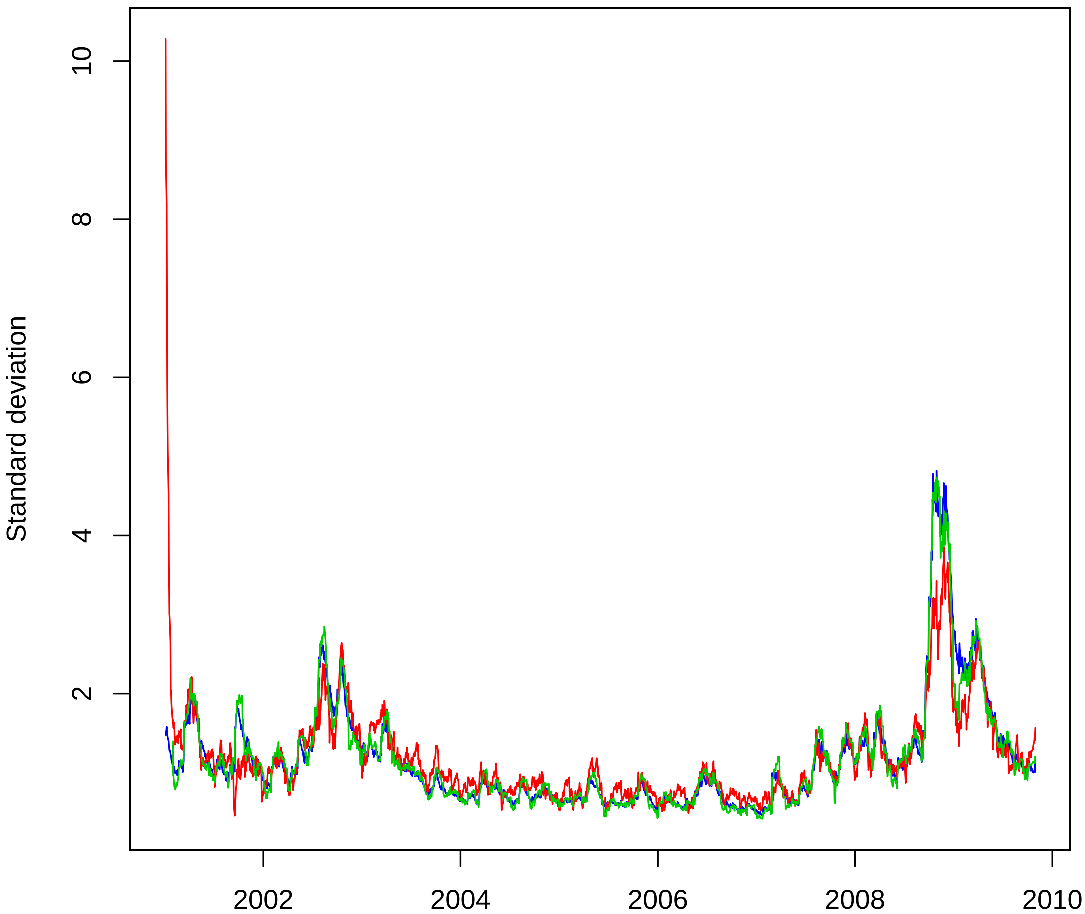

Figure 2: In-sample (1 January 2001 - 31 December 2005) and out-of-sample (1 January 2006 - 30 October 2009) conditional standard deviations of univariate models. Blue line: GARCH(1,1); Red line: 1st and 2nd power Log-GARCH(1,1); Green line: Log-ARCH(1) and log-ARCH(0) models with volatility proxy

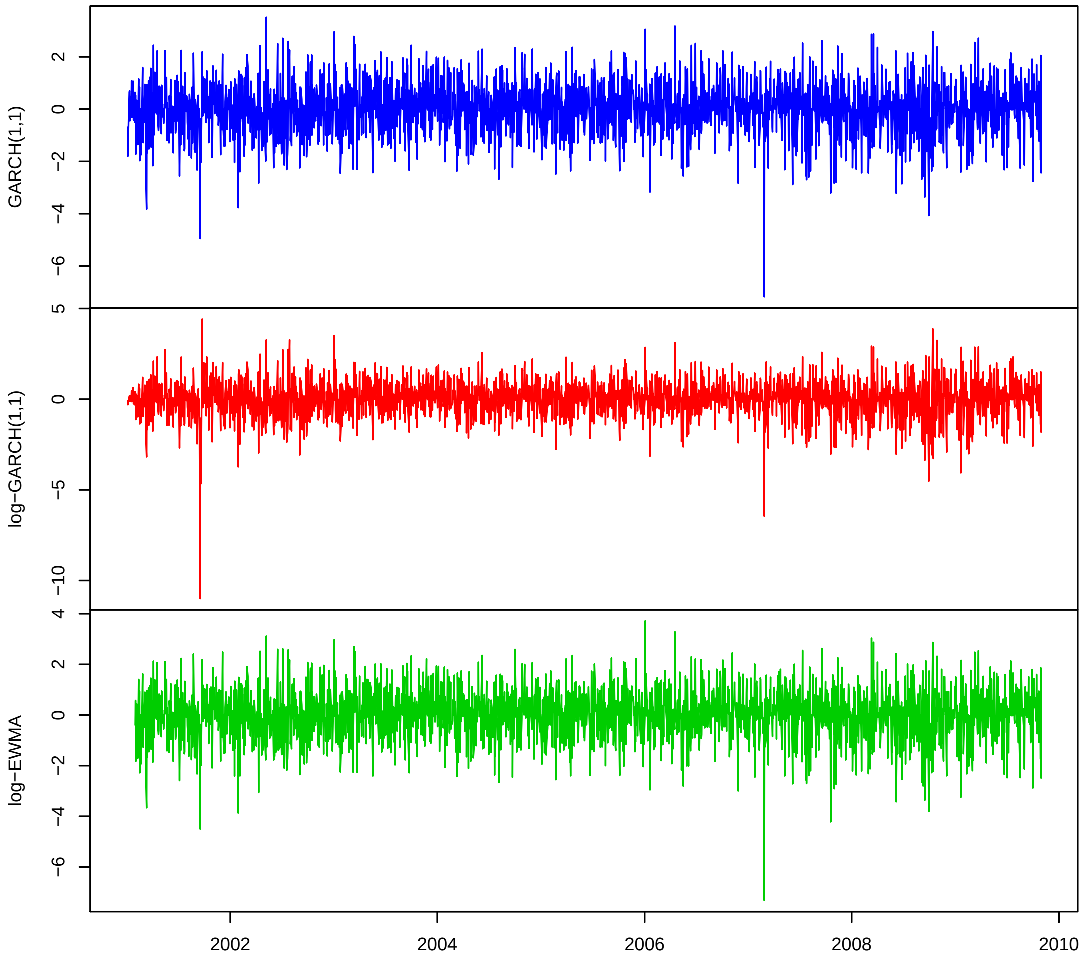

Figure 3: In-sample (1 January 2001 - 31 December 2005) and out-of-sample (1 January 2006 - 30 October 2009) univariate standardised residuals of the GARCH(1,1) specification (upper graph), of the 1st. and 2nd. power log-GARCH(1,1) specifications (middle graph), and of the log-ARCH(1) and log-ARCH(0) specifications with volatility proxy (bottom graph)

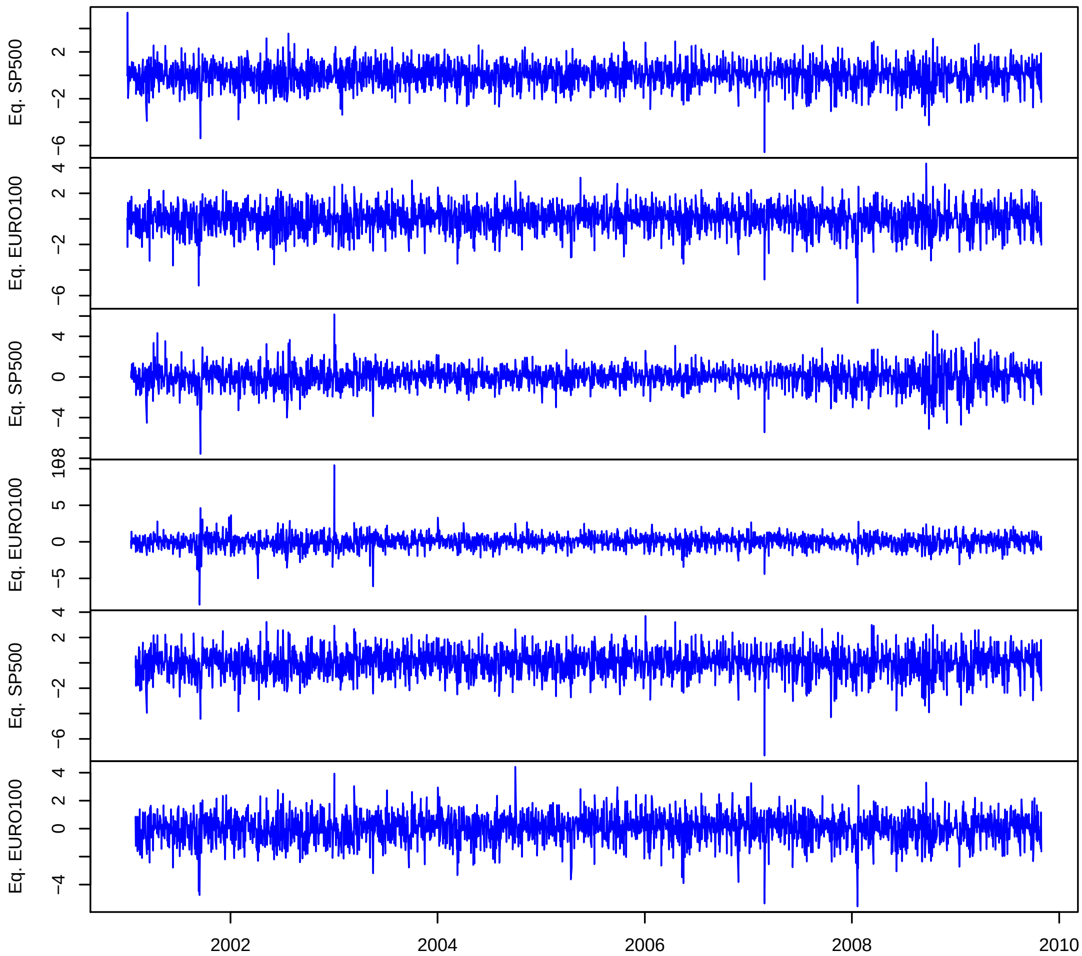

Figure 4: In-sample and out-of-sample standardised residuals of the multivariate models. Diagonal BEKK(1,1,1): The two upper graphs; Log-GARCH(1,1): The two middle graphs; The two models with volatility proxy dynamics: The two lower graphs

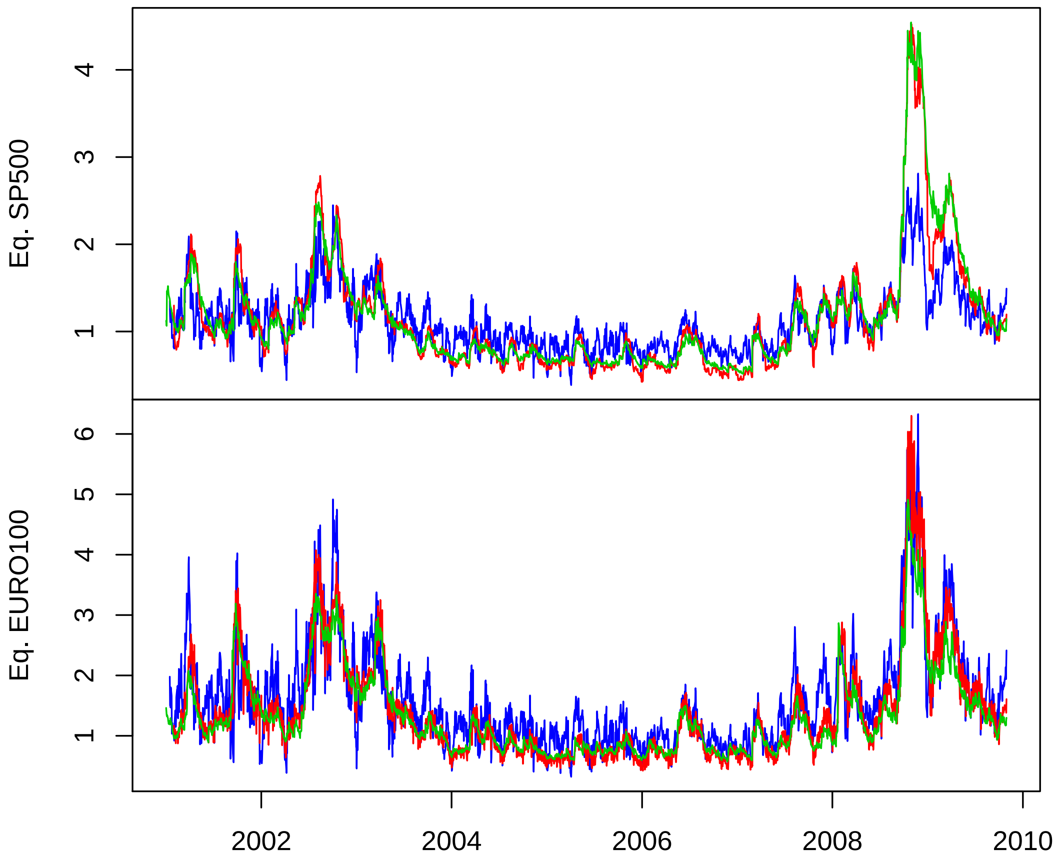

Figure 5: In-sample and out-of-sample conditional standard deviations of the multivariate models. Blue line: Log-GARCH(1,1); Red line: Log-ARCH(1) w/volatility proxy dynamics; Green line: Diagonal BEKK(1,1,1)

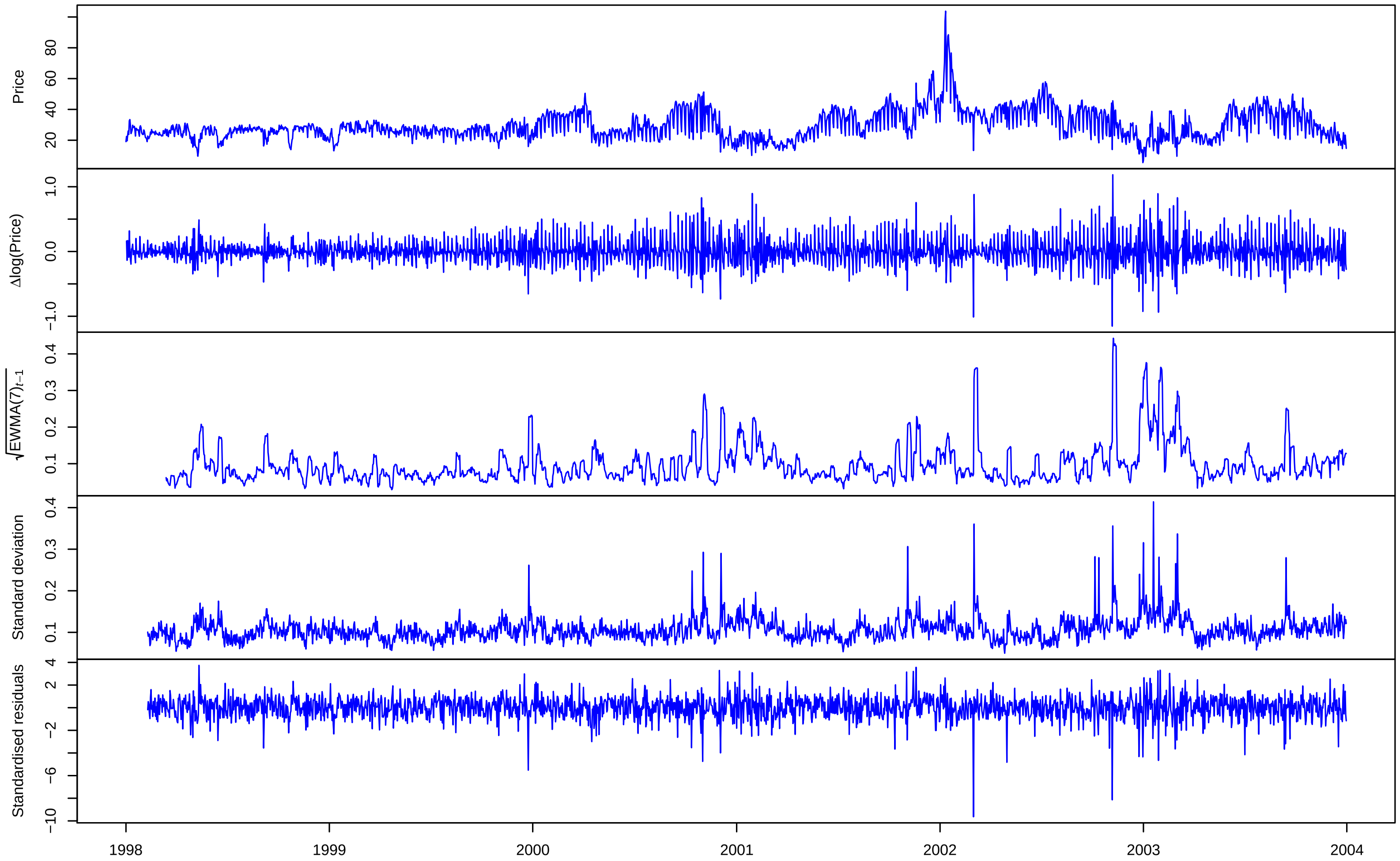

Figure 6: Daily electricity price (upper graph), log-returns (second graph), $\sqrt{\widehat{EWMA}(7)_{t-1}}$ (third graph), fitted standard deviation (fourth graph) and the standardised residuals (bottom graph) of the (19) - (21) model for Spain (1 January 1998 - 31 December 2003).

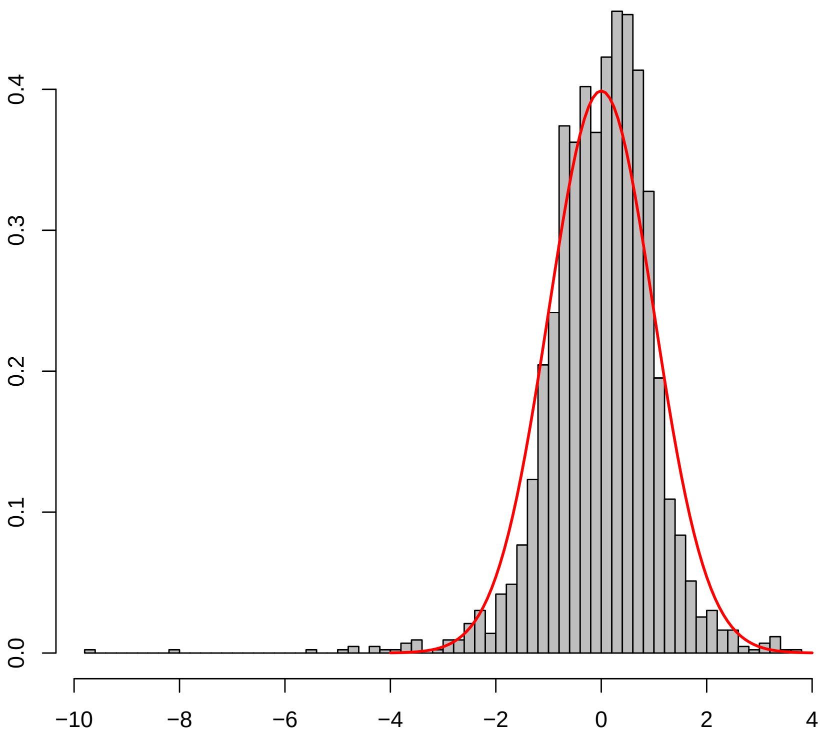

Figure 7: Empirical relative frequency distribution *vs.* the standard normal (red graph) of the standardised residuals of the specific model (19)-(21)