The Power Log-GARCH Model*

Genaro Sucarrat† and Álvaro Escribano ‡

24 December 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 are valid only under very restrictive assumptions. Here, we provide results and simple methods that readily enables consistent estimation and inference of univariate and multivariate power log-GARCH(P, Q) models with time-varying correlations under very general and non-restrictive assumptions, via vector ARMA(P, Q) representations. Augmented by explanatory or exogenous regressors in the volatility specification(s), our empirical applications show that the models are particularly suited for complex modelling problems where many series and/or variables are involved.

JEL Classification: C22, C32, C51, C52
Keywords: Power ARCH, exponential GARCH, log-GARCH, Multivariate GARCH, Time-varying correlations

---

*We are grateful to Jonas Andersson, Luc Bauwens, Andrew Harvey, Emma Iglesias, Sebastien Laurent, Enrique Sentana and participants at the 4th. CSDA International Conference (London), FIBE 2010 (Bergen), IWAP 2010 (Madrid) and Foro de Finanzas 2010 (Elche) conferences 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 acknowledged.

†Corresponding author. Department of Economics, BI Norwegian School of Management. Email: genaro.sucarrat@bi.no. Webpage: http://www.sucarrat.net/

‡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, exponential ARCH models are of special interest because their fitted values of volatility are guaranteed to be non-negative (this is not the case for ordinary ARCH models), and because they enable richer autoregressive volatility dynamics. As an extreme case, in exponential ARCH models the volatility remains positive even when all parameters are negative.

In contrast to ordinary ARCH models, however, stability conditions and the existence of unconditional moments depend to a greater extent on the conditional density in exponential ARCH models. For example, the best known 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.[1] Furthermore, in contrast to ordinary ARCH models, fewer theoretical results exist that enable consistent estimation and valid asymptotic inference when the conditional density is unspecified. 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. Also, 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. But a number of restrictive assumptions must be satisfied for the result to be applicable, 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 (standardised) skewed t and the (standardised) 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. Also, a set of identifying assumptions (see p. 194) are needed which effectively requires that a conditional density is specified. 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

[1] Indeed, the shortcoming has prompted specific work on exponential ARCH models with t-distributed conditional densities, see Harvey and Chakravarty (2010).

multivariate versions. Also, many stability properties of their model is unknown. Kristensen and Rahbek (2009) proves that the QML is consistent for a class of univariate non-linear ARCH models that includes the second power log-ARCH(P) family. Nevertheless, their result does not apply to powers different from 2, nor to models that includes log-GARCH terms, nor to multivariate models. 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 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, and the multivariate version admits time-varying correlations. 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.[2] 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,[3] and why its presence in applied work is limited.

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. 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

[2]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 $\{\hat{\epsilon}_t^2\}$.

[3]Examples of 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.

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 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.[4] 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.

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, is contained in subsection 2.2. Section 3 presents the multivariate power log-GARCH. Section 4 contains two empirical applications that shows that the model class is particularly suited for complex modelling problems that involves many variables, including exogenous conditioning information. 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

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

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

$$\ln \sigma_t^\delta = \alpha_0 + \sum_{p=1}^{P} \alpha_p \ln |\epsilon_{t-p}|^\delta + \sum_{q=1}^{Q} \beta_q \ln \sigma_{t-q}^\delta, \quad \delta > 0, \tag{3}$$

where $E(r_t|\mathcal{I}_t)$ is the expectation of r_t conditional on the information set $\mathcal{I}_t$, δ is the power, P is the ARCH order and Q is the GARCH order. The contemporaneity of

[4]Our estimation methods assumes the power is fixed and known. However, in practice, grid search methods can be used to search for the power.

the time index t of the information set $\mathcal{I}_t$ means it may contain contemporaneous conditioning information. The mean μ_t allows for a large range of possible specifications, linear or non-linear, and may contain both autoregressive (AR) and/or moving average (MA) terms, and additional regressors. 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 $\{\ln \sigma_t^\delta\}$ 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 below. Conditions for the existence of unconditional moments are explored in section 2.3 below. If we divide through by δ in the log-volatility specification, then δ does not matter in the sense that it only appears in the constant. In particular, the unconditional moments of ϵ_t depend negatively on δ (everything else equal), see appendix B.

In order to obtain a feel of how the empirical estimates of log-GARCH models compare with those of the ordinary GARCH(1,1) of Bollerslev (1986), Table 1 contains the estimates of four models fitted to (demeaned) daily SP500 returns (in %). The GARCH(1,1) model is estimated by Gaussian QML in the $\{z_t\}$, whereas the log-ARCH models are fitted by Gaussian QML in the residuals of the ARMA representation, see sections 2.2 and 2.4 below. The volatility proxy is an equally weighted 20-period moving average (EqWMA) of the past squared residuals.[5] Both the parameter estimates and fitted values of the conditional standard deviations (see figure 2) are very similar. Table 2 contains the autocorrelations of $\{\epsilon_t^2\}$ for the 1st. and 2nd. power log-GARCH(1,1) specifications for empirically relevant parameter values on the ARCH and GARCH parameters. It should be noted that although the moments and autocovariances depend on the power parameter δ, the autocorrelations of the log-GARCH(1,1) do not, see Appendix B. Moreover, in contrast to the ordinary GARCH(1,1) model, the autocorrelations of the power log-GARCH(1,1) models depend on the distribution of z_t: The more fat-tailed, the weaker correlations. Also, the power log-GARCH(1,1) is capable of generating stronger autocorrelations than the GARCH(1,1), although maybe not as persistent—or at least not for the parameter values in table 2 (this is consistent with the findings of He et al. (2002)).

2.2 ARMA representations

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

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

[5]In fact, the lagged N period EqWMA of the past squared residuals, $\text{EqWMA}(N)_{t-1}$, is equivalent to the conditional variance of an integrated ARCH(N) model with the variance constant α_0 equal to zero, and the ARCH parameters $\alpha_1 = \cdots = \alpha_N$ all equal to $\frac{1}{N}$.

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 any power δ. For example, the δth. power log-ARCH(1) specification is given by $\ln \sigma_t^\delta = \alpha_0 + \alpha_1 \ln |\epsilon_{t-1}|^\delta$. Adding $\ln E|z_t|^\delta + \ln |z_t^*|^\delta$ to each side and then adding $E(\ln |z_t|^\delta) - E(\ln |z_t|^\delta)$ to the right-hand side, yields the AR(1) representation $\ln |\epsilon_t|^\delta = \alpha_0^* + \alpha_1 \ln |\epsilon_{t-1}|^\delta + u_t^*$, where $\alpha_0^* = \alpha_0 + \ln E|z_t|^\delta + E(\ln |z_t^*|^\delta)$, and where $u_t^* = \ln |z_t^*|^\delta - E(\ln |z_t^*|^\delta)$ is a zero-mean IID process. In other words, the power log-ARCH(1) model admits an AR(1) representation. For a given power $\delta > 0$, the parameters α_0^* and α_1 can thus be estimated consistently by means of ordinary estimation methods subject to usual assumptions. In order to recover α_0 we need estimates of $\ln E|z_t|^\delta$ and $E(\ln |z_t^*|^\delta)$, and the proposition we state below provides simple formulas for consistent estimation of $\ln E|z_t|^\delta$ and $E(\ln |z_t^*|^\delta)$ under very general assumptions. Moreover, in the process we also obtain an estimate of $E(\ln |z_t|^\delta)$, since $E(\ln |z_t|^\delta) = \ln E|z_t|^\delta + E(\ln |z_t^*|^\delta)$.

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

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

with probability 1, where

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

and where $u_t^* = \ln |z_t^*|^\delta - E(\ln |z_t^*|^\delta) = \ln |z_t|^\delta - E(\ln |z_t|^\delta)$ as earlier. When $P > Q$, then $\beta_{Q+1} = \cdots = \beta_P = 0$ implicitly. Also, it should be noted that the equations are not affected by the (linear) inclusion of other variables in the log-volatility 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,[6] as long as the power δ is given, and

[6] 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 $E(u_t^{*4}) < \infty$.

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 $\ln|\epsilon_t|^\delta = \alpha_0^* + \beta_1 \ln|\epsilon_{t-1}|^\delta - \beta_1 u_{t-1}^* + u_t^*$. 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 $\ln\sigma_t^\delta = \alpha_0 + (1-\beta_1)\ln|\epsilon_{t-1}|^\delta + \beta_1 \ln\sigma_{t-1}^\delta$ can be written as the MA(1) representation $\Delta \ln|\epsilon_t|^\delta = \alpha_0^* + \beta_1^* u_{t-1}^* + u_t^*$. More generally, if $\ln|\epsilon_t|^\delta$ is I(1), then the estimates of the stationary AR(P) representation $\Delta\ln|\epsilon_t|^\delta = \alpha_0^* + \sum_{p=1}^{P} \alpha_p \Delta \ln|\epsilon_{t-p}|^\delta + u_t^*$ 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 $\ln E(|z_t|^\delta)$ and $E(\ln|z_t^*|^\delta)$, 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 $\hat{u}_t^* - u_t^* \xrightarrow{P} 0$ for each t, where $\{\hat{u}_t^*\}$ are estimates of $\{u_t^*\}$. If $0 < E|z_t|^\delta < \infty$ and if $|E(\ln|z_t|)| < \infty$, then

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

and

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

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

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

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

since $E|z_t^*|^\delta = 1$ and since $|E(\ln|z_t^*|^\delta)| < \infty$. The latter follows from the assumptions $0 < E|z_t|^\delta < \infty$ and $|E(\ln|z_t|)| < \infty$. Accordingly, $(-1) \cdot \ln E[\exp(u_t^*)] = E(\ln|z_t^*|^\delta)$. 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 $|E(\ln|z_t^*|^\delta)| < \infty$ 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 $\hat{u}_t^* - u_t^* \xrightarrow{P} 0$ for each t, we have that $\exp(\hat{u}_t^*) - \exp(u_t^*) \xrightarrow{P} 0$ for each t due to the continuity of the $\exp(\cdot)$ function. Accordingly, $\frac{1}{T}\sum_{t=1}^T \exp(\hat{u}_t^*) \to \frac{1}{T}\sum_{t=1}^T \exp(u_t^*)$ as $T \to \infty$, and since $\frac{1}{T}\sum_{t=1}^T \exp(u_t^*) \to E[\exp(u_t^*)]$ as $T \to \infty$ 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 $-\frac{\delta}{2}\ln(\frac{1}{T}\sum_{t=1}^T \hat{z}_t^{*2}) \xrightarrow{P} \ln E|z_t|^\delta$. ■

When the power δ is equal to 2, then $\ln E|z_t|^\delta = 0$ 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 $E(\ln|z_t|^\delta)$ and $E|z_t|^\delta$, which are sometimes useful in empirical applications.[7] 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 $\hat{u}_t^* - u_t^* \xrightarrow{P} 0$ 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 $|E(\ln|z_t|)| < \infty$ are satisfied for the most commonly used densities in finance: The Normal, the 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 $\hat{\epsilon}_t - \epsilon_t \xrightarrow{P} 0$ for each t. In words, in sufficiently large samples the estimated residuals are distributed as the true errors, and so are the $\{\ln|\hat{\epsilon}_t|^\delta\}$ due to continuity.

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

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

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

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 (20) (see appendix) with $s = 2$.

Proof. From equation (20) 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 Appendix A. For $z_t \sim t(\nu), \nu > 2$, then $E(|z_t|^c) < \infty$ for $-1 < c < \nu$, see Appendix A. 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 $\ln \sigma_t^\delta$ admits the representation $\alpha_0/[1 - (\alpha_1 + \beta_1) - \cdots -$

$(\alpha_P + \beta_P)] + \sum_{i=1}^{\infty} \psi_i \ln |z_{t-i}|^{\delta}$, 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, see Francq and Zakoian (2010). 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 conditionally heteroscedastic. By contrast, the error term in the ARMA representation of the 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 $[-1, \infty)$. The error u_t^* in the ARMA representation, by contrast, is much closer to symmetry (in comparison), with the left-tail usually being "longer". Moreover, the error takes on values in $(-\infty, \infty)$. This means QML-type 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.

Tables 3 and 4 contain simulations that shed light on the finite sample accuracy of QML methods for selected specifications. Table 3 suggests the finite sample biases associated with QML via the AR representation in the estimation of log-ARCH(1) processes are acceptable for many purposes, and that estimating the errors $\{\epsilon_t\}$ in a previous step does not 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. In table 4 the simulations suggest QML in $\{u_t^*\}$ via the ARMA representation compares well with QML in $\{z_t\}$ in the estimation of a log-ARCH(1) model, when the standardised errors are more fat-tailed than the normal.[8]

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 volatility 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. 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 $-\ln \hat{E}[\exp(\hat{u}_t^*)] = -\ln[\frac{1}{T}\sum_{t=1}^{T}\exp(\hat{u}_t^*)]$, since $\alpha_0^* = E(\ln z_t^2)$ under the null of $\alpha_0 = 0$. The Wald-statistic under the null of $\alpha = 0$ then becomes

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

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 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 volatility specification contains log-GARCH terms, then one might consider using the usual theory for inference regarding the parameters of the ARMA representation. For the GARCH coefficients this will work fine, but it is not equally straightforward for the ARCH parameters, since the AR and MA coefficient estimates will typically be strongly correlated (recall: $\alpha_p^* = \alpha_p + \beta_p$). 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

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

is to use the property that a $\{\ln \sigma_t^\delta\}$ 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 the lags.

2.6 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 $\ln \sigma_t^\delta$ 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 $\ln \sigma_t^\delta$ specification, that is, a "power log-GARCH-X" model. This can in many cases 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.[9] In practice this type of asymmetry terms can very often be adequately approximated by $I_{\epsilon_{t-1}<0}$, 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_ν or GED(τ).

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

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

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 $\ln_t^\delta$ admits the representation $\alpha_0/(1 - \alpha_1 - \beta_1) + \sum_{i=1}^\infty (\alpha_1 + \beta_1)^{i-1} \cdot [\alpha_1 \ln |z_{t-i}|^\delta + (\ln \lambda^\delta) I_{\{z_{t-i}<0\}}]$. 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 if $|\alpha_1 + \beta_1| < 1$ and if $2\alpha_1(\alpha_1 +$

[9] 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.

$\beta_1)^{i-1} \in (-1, 2]$. 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 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

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

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

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

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).

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-volatility $\ln \sigma_t^\delta$ is equal to $\gamma_0 + \gamma_1 \ln V_t^\delta$, 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.

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 model, see Bauwens et al. (2006) for a survey. 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. In particular, whether the volatility of one variable Granger-causes another may readily be tested via the VAR-representation.

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

$$\epsilon_t = \text{diag}(\sigma_t) z_t, \qquad z_t|\mathcal{I}_t \sim ID(0, Cov(z_t|\mathcal{I}_t)), \qquad Var(z_t|\mathcal{I}_t) = I_M, \tag{8}$$

$$\forall m : z_{mt}|\mathcal{I}_t \sim IID(0, 1), \tag{9}$$

where σ_t is the $(M \times 1)$ vector of conditional standard deviations, diag(σ_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, $Cov(z_t|\mathcal{I}_t)$ is the variance-covariance matrix of $\{z_t\}$, possibly time-varying, and $\mathcal{I}_t$ is the conditioning set in question. ID means the vector sequence $\{z_t\}$ is independent but not necessarily identical over time, so as to allow for time-varying contemporaneous correlation between the elements of z_t. The assumption $Var(z_t|\mathcal{I}_t) = I_M$, however, means each diagonal entry of $Cov(z_t|\mathcal{I}_t)$ is constant and equal to 1. Also, for the formulas in proposition 1 to be applicable (they are needed in order to correct for the bias induced by the estimation of the VARMA representation) we need to assume that each z_{mt} is IID, hence the assumption of independence but not necessarily identicalness of $\{z_t\}$. We do not impose further restrictions on the off-diagonal entries of $Cov(z_t|\mathcal{I}_t)$, which means it may not be positive definite (we will return to this issue below in subection 3.3).

The M-dimensional log-volatility specification is given by

$$\ln \sigma_t^\delta \;=\; \alpha_0 + \sum_{p=1}^{P} \alpha_p \ln |\epsilon_{t-p}|^\delta + \sum_{q=1}^{Q} \beta_q \ln \sigma_{t-q}^\delta, \qquad P \geq Q, \tag{10}$$

where

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

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

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

$$\begin{aligned} \ln\sigma_{1,t}^\delta &= \alpha_{1,0} + \alpha_{11.1}\ln|\epsilon_{1,t-1}|^\delta + \alpha_{12.1}\ln|\epsilon_{2,t-1}|^\delta \\ \\ \ln\sigma_{2,t}^\delta &= \alpha_{2,0} + \alpha_{21.1}\ln|\epsilon_{1,t-1}|^\delta + \alpha_{22.1}\ln|\epsilon_{2,t-1}|^\delta, \end{aligned}$$

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

$$\begin{aligned} \ln\sigma_{1,t}^\delta &= \alpha_{1,0} + \alpha_{11.1}\ln|\epsilon_{1,t-1}|^\delta + \alpha_{12.1}\ln|\epsilon_{2,t-1}|^\delta + \alpha_{11.2}\ln|\epsilon_{2,t-2}|^\delta \\ &\quad + \alpha_{12.2}\ln|\epsilon_{2,t-2}|^\delta + \beta_{11,1}\ln\sigma_{1,t-1}^\delta + \beta_{12,1}\ln\sigma_{2,t-1}^\delta \\ \\ \ln\sigma_{2,t}^\delta &= \alpha_{2,0} + \alpha_{21.1}\ln|\epsilon_{1,t-1}|^\delta + \alpha_{22.1}\ln|\epsilon_{2,t-1}|^\delta + \alpha_{21.2}\ln|\epsilon_{2,t-2}|^\delta \\ &\quad + \alpha_{22.2}\ln|\epsilon_{2,t-1}|^\delta + \beta_{21,1}\ln\sigma_{1,t-1}^\delta + \beta_{22,1}\ln\sigma_{2,t-1}^\delta, \end{aligned}$$

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 $\ln\sigma_t^\delta = \Psi_0 + \sum_{i=1}^\infty \Psi_i \ln|z_{t-i}|^\delta$ 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 $|E(z_{m,t}^s)| < \infty$ and 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 $E\left[|z_{1,t-i}|^{s\psi_{i,m1}}|z_{2,t-i}|^{s\psi_{i,m2}} \cdots |z_{M,t-i}|^{s\psi_{i,mM}}\right] \to 1$ as $i \to \infty$ due to absolute summability, it follows that $|a_i - 1| \to 0$ as $i \to \infty$. 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, and if $|E(z_{m,t}^s)| < \infty$, 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 set of conditions is satisfied or not, that is, $|E(z_{m,t}^s)| < \infty$ and $E\left[|z_{1,t-i}|^{s\psi_{i,m1}}|z_{2,t-i}|^{s\psi_{i,m2}}\cdots|z_{M,t-i}|^{s\psi_{i,mM}}\right] < \infty$ 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

$$\ln|\epsilon_t|^\delta \quad = \quad \alpha_0^* + \sum_{p=1}^{P} \alpha_p \ln|\epsilon_{t-p}|^\delta + u_t^*, \tag{11}$$

where α_p is defined as above, and where

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

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

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

$$\ln|\epsilon_t|^\delta \quad = \quad \alpha_0^* + \sum_{p=1}^{P} \alpha_p^* \ln|\epsilon_{t-p}|^\delta + \sum_{q=1}^{Q} \beta_q^* u_{t-q}^* + u_t^*, \tag{12}$$

where

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

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

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 $\ln E|z_t|^\delta$ and $E(\ln|z_t^*|^\delta)$ 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. When there are no log-GARCH terms in the log-volatility specification, then the corresponding VAR representation is straightforwardly estimated by OLS/QML. When log-GARCH terms are present, however, then it is well known that the corresponding VARMA representation may not be readily estimated in practice due to numerical issues. The question of how well the available estimation algorithms actually work for multivariate log-GARCH models that contain GARCH terms, we leave for future research.

3.3 Modelling conditional correlations

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 $H_t = Cov(\epsilon_t|\mathcal{I}_t)$, then one cannot be ensured that such portfolio variances will be positive. This is why H_t is often required to be positive definite. In the power log-GARCH model this amounts to positive definiteness of $Cov(z_t|\mathcal{I}_t)$. The conditions (8)-9) are compatible with both constant and time-varying conditional covariances, $Cov(z_t|\mathcal{I}_t)$ being positive definite or not. So once the estimates of $\{z_t\}$ have been obtained, then adapted versions of several of the methods surveyed in Bauwens et al. (2006) or the method by Kawakatsu (2006) can be considered in order to fit appropriate models of the off-diagonals of $Cov(z_t|\mathcal{I}_t)$. In the empirical application below we use the specification suggested by Engle (2002).

4 Empirical applications

Specifications contained in the 2nd. power log-GARCH-X class have proved particularly useful in situations that involves many explanatory variables, see for example Bauwens et al. (2006), Rime and Sucarrat (2007), Bauwens and Sucarrat (2008), and Sucarrat and Escribano (2010). Here we explore further the usefulness of power log-GARCH-X models in two empirical applications. The first considers the complex problem of modelling daily electricity prices, and the second undertakes a multivariate analysis of stock market variability.

4.1 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 volatility specifications, see for example Escribano et al. (2009), and Koopman et al. (2007). The power log-GARCH 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. (2010), which spans the period 1 January 1998 to 31 December 2003 ($T = 2191$ observations), see the upper two graphs of figure 3.

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

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

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-volatility specification $P = \{1, \ldots, 7, 14, 21, 28, 35\}$, $EqWMA(7)_{t-1}$ 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-volatility 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)

$$\begin{aligned}
\hat{r}_t &= \underset{(4.28)}{0.114} - \underset{(22.70)}{0.045} \cdot (8r_{t-1} + 4r_{t-2} + 3r_{t-3} + 2r_{t-4} + 2r_{t-5} + 2r_{t-6}) \\
&\quad + \underset{(15.42)}{0.056} \cdot (2r_{t-7} + 2r_{t-14} + r_{t-21} + r_{t-28} + r_{t-35}) + \underset{(7.05)}{0.186}(r_{t-1}^2 - 1) \\
&\quad + \underset{(32.05)}{0.018} \cdot (10I_{r_t>0} + I_{r_{t-1}>0} + I_{r_{t-6}>0}) - \underset{(3.68)}{0.016} \cdot (I_{r_{t-7}>0} + I_{r_{t-14}>0}) \\
&\quad - \underset{(9.00)}{0.040}(Sat_t + 2Sun_t) - \underset{(2.72)}{0.023}Dec_t + \underset{(2.16)}{0.062}I_{r_{t-2}<-0.5} \qquad (13)
\end{aligned}$$

$$\begin{aligned}
\ln \hat{\sigma}_t^2 &= -2.517 - \underset{(2.40)}{0.054} \ln \hat{\epsilon}_{t-3}^2 + \underset{(6.67)}{0.393} \ln EqWMA(7)_{t-1} - \underset{(3.60)}{0.247}(I_{r_t>0} + I_{r_{t-1}>0}) \\
&\quad - \underset{(3.47)}{0.329}(Wed_t + Fri_t + Sat_t) - \underset{(3.09)}{0.392}(Apr_t + Jul_t) + \underset{(2.19)}{1.218}I_{r_{t-1}<-0.5} \qquad (14)
\end{aligned}$$

$$\hat{z}_t \sim SGED(\hat{\tau}_{shape} = 1.31, \hat{\tau}_{skew} = 0.82) \qquad (15)$$

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 3), 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.

Since R^2 is substantially different from zero the log-volatility specification should *not* be interpreted as a measure of price variability, but rather as a measure of the time-varying accuracy of the mean specification: The greater $\ln \sigma_t^2$ 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 is negative (cyclical). By contrast, the lagged impact of the log of the volatility proxy EqWMA$(7)_t$ 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).[10]

4.2 Multivariate ARCH modelling

The use of multivariate ARCH models is plagued by the curse of dimensionality due to the number of parameters that has to be estimated. Here, we show that a rich, multivariate log-ARCH model that is straightforwardly estimated equation by equation by means of OLS/QML—thus avoiding numerical estimation issues—can readily be simplified by means of the same automated GETS model selection methods as those of the previous subsection. The end result is a parsimonious model and, in the process, a test for Granger-causality in the volatilities. Finally, we fit a DCC structure of the Engle (2002) type on the time-varying correlations of the standardised residuals. For the purpose of comparison we also fit a DCC model of the Engle (2002) type to the same data, with a diagonal structure on the volatilities made up of ordinary GARCH(1,1) models. Engle's (2002) model is arguably the most common DCC model in empirical practice due to its simplicity, and the diagonality assumption means Granger non-causality is assumed from the outset in Engle's DCC. Both models are fitted to the daily, demeaned S&P500 and the FTSE Euro 100 (EUR100) stock market index log-returns (in %) from 1 January 2001 to 30 October 2009 ($T = 2302$).[11] The S&P500 log-return series is the same demeaned series as in the previous section, and we demean the EUR100 returns the same way. Both demeaned series are displayed in figure 1.

The estimates of the diagonal DCC model of Engle (2002) are (standard errors in parentheses)

$$\begin{aligned}
\hat{\sigma}^2_{1,t} &= \underset{(0.004)}{0.010} + \underset{(0.012)}{0.052}\hat{\epsilon}^2_{1,t-2} + \underset{(0.010)}{0.930}\hat{\sigma}^2_{1,t-1} \\
\hat{\sigma}^2_{2,t} &= \underset{(0.013)}{0.005} + \underset{(0.004)}{0.064}\hat{\epsilon}^2_{1,t-2} + \underset{(0.009)}{0.944}\hat{\sigma}^2_{1,t-1} \\
\hat{q}^2_{1,2,t} &= \hat{\bar{\rho}}_{1,2} + 0.050(\hat{z}_{1,t-1}\hat{z}_{2,t-1} - \hat{\bar{\rho}}_{1,2}) + 0.483(\hat{q}_{1,2,t-1} - \hat{\bar{\rho}}_{1,2}),
\end{aligned}$$

where $\bar{\rho}$ is the full sample sample correlation between $\hat{z}_{1,t}$ and $\hat{z}_{2,t}$, and where the fitted correlations $\hat{\rho}_t$ are computed as $\hat{q}_{1,2,t}/\sqrt{\hat{q}_{1,1,t}\hat{q}_{2,2,t}}$. The model is estimated with the R package `ccgarch`, see Nakatani (2010).

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

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

The initial, general log-ARCH model is given by

$$\begin{aligned}
\ln \sigma^2_{1,t} &= \alpha_{10} + \sum_{p=1}^{5} \alpha_{11.p} \ln \epsilon^2_{1,t-p} + \sum_{p=1}^{5} \alpha_{12.p} \ln \epsilon^2_{2,t-p} + \omega_{11} \ln EqWMA(20)_{1,t-1} \\
&\quad + \omega_{12} \ln EqWMA(20)_{2,t-1} + \sum_{i=1}^{4} c_{1i} day_{i,t} \\
\ln \sigma^2_{2,t} &= \alpha_{20} + \sum_{p=1}^{5} \alpha_{21.p} \ln \epsilon^2_{1,t-p} + \sum_{p=1}^{5} \alpha_{22.p} \ln \epsilon^2_{2,t-p} + \omega_{21} \ln EqWMA(20)_{1,t-1} \\
&\quad + \omega_{22} \ln EqWMA(20)_{2,t-1} + \sum_{i=1}^{4} c_{2i} day_{i,t},
\end{aligned}$$

where $day_1, \ldots, day_4$ are day-of-the-week dummies for Monday, Tuesday, Thursday and Friday. GETS model selection yields (standard errors in parentheses and p-values in square brackets)

$$\begin{aligned}
\ln \hat{\sigma}^2_{1,t} &= \underset{[0.00]}{0.400} + \underset{(0.02)}{0.075} \ln \hat{\epsilon}^2_{1,t-2} + \underset{(0.02)}{0.063} \ln \hat{\epsilon}^2_{2,t-1} + \underset{(0.06)}{0.734} \ln EqWMA(20)_{1,t-1} \\
\ln \hat{\sigma}^2_{2,t} &= \underset{[0.00]}{0.303} + \underset{(0.05)}{0.920} \ln EqWMA(20)_{2,t-1} - \underset{(0.13)}{0.495} Mon_t - \underset{(0.13)}{0.296} Fri_t,
\end{aligned}$$

and fitting an Engle (2002) like DCC model to the standardised residuals gives

$$\hat{q}_{1,2,t} = \bar{\rho} + 0.062(\hat{z}_{1,t-1}\hat{z}_{2,t-1} - \bar{\rho}) + 0.478\hat{q}_{1,2,t-1}.$$

The correlation parameters are estimated by ML using a standardised bivariate Gaussian density.

The GETS model selection does not suggest that the diagonality assumption in Engle's DCC is fully justified, since $\ln \hat{\epsilon}_{2,t-2}$ is retained in the EU100 log-volatility equation. Also, in the SP500 equation the final model contains a Monday and a Friday effect. Nevertheless, Figure 5 compares the fitted standard deviations and the fitted time-varying correlations of the two models, and graphically they are almost identical. So even though the diagonality assumption and the lack of day-of-the-week effects in Engle's DCC might not be justified statistically, the assumptions do not matter much graphically for the final fitted values.

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 with time-varying correlations, even when the conditional density is not known. Specifically, consistent QML 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. In the empirical section we show that complex modelling problems with rich persistence and periodicity structures in the volatility specifications, both univariate and multivariate with time-varying conditional correlations, can readily be resolved by the methods.

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 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.

Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics 20*, 339–350.

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

Escribano, Á., J. I. Peña, and P. Villaplana (2010). Modelling Electricity Prices: International Evidence. Forthcoming in the Oxford Bulletin of Economics and Statistics.

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.

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.

Kristensen, D. and A. Rahbek (2009). Asymptotics of the QMLE for Non-Linear ARCH Models. *Journal of Time Series Econometrics 1*, Issue 1, Article 2. Available at: `http://www.bepress.com/jtse/vol1/iss1/art2`.

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.

Nakatani, T. (2010). *ccgarch*. R package available via `http://CRAN.R-project.org/package=ccgarch`.

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.sucarrat.net/`.

Sucarrat, G. and Á. Escribano (2010). Automated Model Selection in Finance: General-to-Specific Modelling of the Mean and Volatility Specifications. Available via `http://www.sucarrat.net/`.

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

Appendices

Appendix A: Closed form expressions for $E|z_t|^c$ for the GED and t distributions

The expectation of the absolute value of a GED variate ε raised to the power c is readily available, since it can be showed that $|\varepsilon|^c$ is $Gamma(1/2, \tau)$ distributed where τ is the GED shape parameter ($\tau = 2$ yields the standard normal), see Harvey and Chakravarty (2010). Accordingly:

$$E|\varepsilon|^c = \frac{2^{c/\tau}\Gamma[(c+1)/\tau]}{\Gamma(1/\tau)}, \quad c > -1, \quad \tau > 0. \tag{16}$$

In particular, $Var(\varepsilon) = E|\varepsilon|^2 = 2^{2/\tau}\Gamma(3/\tau)/\Gamma(1/\tau)$, and so for the standardised (zero-mean, unit variance) GED variate $z = \varepsilon/\sqrt{Var(\varepsilon)}$ we obtain:

$$E|z|^c = \frac{\Gamma(1/\tau)^{c/2}\Gamma[(c+1)/\tau]}{\Gamma(3/\tau)^{c/2}\Gamma(1/\tau)}, \quad c > -1, \quad \tau > 0. \tag{17}$$

Using the property that a t-variate with $\nu > -1$ degrees of freedom can be written as $X\nu^{1/2}/Y_\nu^{1/2}$ where X is a standard normal and Y_ν is a Chi-squared with ν degrees of freedom, and where X and Y are independent, then the expectation of the absolute value of a t variate ε is:

$$E|\varepsilon|^c = \frac{\nu^{c/2}\Gamma(c/2 + 1/2)\Gamma(-c/2 + \nu/2)}{\Gamma(1/2)\Gamma(\nu/2)}, \quad -1 < c < \nu, \tag{18}$$

see Harvey and Shephard (1996, p. 434). Next, since $Var(\varepsilon) = \nu/(\nu - 2)$ we obtain (by setting $z = \varepsilon/\sqrt{Var(\varepsilon)}$):

$$E|z|^c = \frac{(\nu-2)^{c/2}\Gamma(c/2 + 1/2)\Gamma(-c/2 + \nu/2)}{\Gamma(1/2)\Gamma(\nu/2)}, \quad -1 < c < \nu, \quad \nu \neq 2. \tag{19}$$

Appendix B: $E(\epsilon_t^s)$ 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 sth. unconditional moment $E(\epsilon_t^s)$, $s \in \{1, 2, \ldots\}$, is

$$E(\epsilon_t^s) \quad = \quad E(z_t^s) \cdot \exp\left(\frac{s\alpha_0}{\delta \cdot (1 - \alpha_1 - \beta_1)}\right) \cdot \prod_{i=1}^{\infty} E\left(|z_{t-i}|^{s\alpha_1(\alpha_1+\beta_1)^{i-1}}\right), \tag{20}$$

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

$$\begin{aligned}
E(\epsilon_t^2 \epsilon_{t-j}^2) &= \exp\left[\frac{2\alpha_0}{\delta}\left(\frac{1+(\alpha_1+\beta_1)^j}{(1-\alpha_1-\beta_1)} + \sum_{i=1}^{j}(\alpha_1+\beta_1)^{i-1}\right)\right] \\
&\quad \cdot \prod_{i=1}^{j} E\left(|z_{t-i-1}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}+2I_{(i=j)}}\right) \\
&\quad \cdot \prod_{i=1}^{\infty} E\left(|z_{t-j-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}\cdot[1+(\alpha_1+\beta_1)^j]}\right) \\
&= \exp\left(\frac{4\alpha_0}{\delta(1-\alpha_1-\beta_1)}\right) \\
&\quad \cdot \prod_{i=1}^{j} E\left(|z_{t-i-1}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}+2I_{(i=j)}}\right) \\
&\quad \cdot \prod_{i=1}^{\infty} E\left(|z_{t-j-i}|^{2\alpha_1(\alpha_1+\beta_1)^{i-1}\cdot[1+(\alpha_1+\beta_1)^j]}\right), \qquad (21)
\end{aligned}$$

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

Although the integer moments $E(\epsilon_t^s)$, $E(\epsilon_t^2 \epsilon_{t-k}^2)$ and the autocovariances $E(\epsilon_t^2 \epsilon_{t-k}^2) - E(\epsilon_t^2)^2$ all depend on both α_0 and δ, the autocorrelations do in fact not depend on neither α_0 nor on δ. The reason is that in the expression for the autocorrelation $[E(\epsilon_t^2 \epsilon_{t-j}^2) - E(\epsilon_t^2)^2]/[E(\epsilon_t^2) - E(\epsilon_t^2)^2]$, the expression $\exp(4\alpha_0/[\delta(1-\alpha_1-\beta_1)])$ can be factored out both in the numerator and in the denominator.

Table 1: Estimates of first order ARCH models fitted to de-meaned SP500 returns 1 January 2001 - 30 October 2009

Model	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\omega}$
GARCH(1,1)	0.0092	0.0663	0.9274	
Log-GARCH(1,1), $\delta = 2$	0.0857	0.0498	0.9401	
Log-GARCH(1,1), $\delta = 1$	0.0428	0.0498	0.9401	
Log-ARCH(0) w/volatility proxy, $\delta = 2$	0.1391			0.9299

Log-returns in % de-meaned by an AR(1) with constant. GARCH(1,1) model: $\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$. Log-GARCH(1,1) models: $\ln \sigma_t^\delta = \alpha_0 + \alpha_1 \ln |\epsilon_{t-1}|^\delta + \beta_1 \ln \sigma_{t-1}^\delta$. Log-ARCH(0) w/volatility proxy: $\ln \sigma_t^2 = \alpha_0 + \omega \ln \text{EqWMA}(20)_{t-1}$, where $\text{EqWMA}(20)_t = (1/20) \sum_{n=1}^{20} \epsilon_{t-n}^2$.

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

	Log-GARCH(1,1), $\beta_1 = 0.9$			Log-GARCH(1,1), $\beta_1 = 0.94$			GARCH(1,1)	
Lag	N	$GED(1.1)$	$t(5)$	N	$GED(1.1)$	$t(5)$	$\beta_1 = 0.9$	$\beta_1 = 0.94$
1	0.096	0.062	0.031	0.215	0.135	0.070	0.073	0.155
2	0.091	0.058	0.029	0.212	0.133	0.069	0.069	0.153
3	0.086	0.055	0.028	0.209	0.132	0.068	0.065	0.152
4	0.081	0.052	0.026	0.207	0.130	0.067	0.062	0.150
5	0.077	0.049	0.025	0.204	0.128	0.066	0.059	0.149
6	0.073	0.047	0.023	0.201	0.126	0.065	0.056	0.147
7	0.069	0.044	0.022	0.199	0.124	0.065	0.053	0.146
8	0.066	0.042	0.021	0.196	0.123	0.064	0.051	0.144
9	0.062	0.040	0.020	0.194	0.121	0.063	0.048	0.143
10	0.059	0.038	0.019	0.191	0.119	0.062	0.046	0.142
11	0.056	0.036	0.018	0.189	0.118	0.061	0.043	0.140
12	0.053	0.034	0.017	0.187	0.116	0.060	0.041	0.139
13	0.050	0.032	0.016	0.184	0.114	0.060	0.039	0.137
14	0.048	0.030	0.015	0.182	0.113	0.059	0.037	0.136
15	0.045	0.029	0.014	0.180	0.111	0.058	0.035	0.135
16	0.043	0.027	0.014	0.177	0.110	0.057	0.034	0.133
17	0.041	0.026	0.013	0.175	0.108	0.057	0.032	0.132
18	0.039	0.024	0.012	0.173	0.107	0.056	0.030	0.131
19	0.037	0.023	0.012	0.171	0.105	0.055	0.029	0.129
20	0.035	0.022	0.011	0.169	0.104	0.054	0.027	0.128
21	0.033	0.021	0.010	0.167	0.103	0.054	0.026	0.127
22	0.031	0.020	0.010	0.165	0.101	0.053	0.025	0.125
23	0.030	0.019	0.009	0.163	0.100	0.052	0.023	0.124
24	0.028	0.018	0.009	0.161	0.099	0.052	0.022	0.123
25	0.027	0.017	0.008	0.159	0.097	0.051	0.021	0.122

The label N means $z_t \sim N(0,1)$, $GED(1,1)$ means $z_t \sim GED$ with shape parameter $\tau = 1.1$ ($\tau = 2$ gives the normal, $\tau \in (1,2)$ gives densities that are more fat-tailed than the normal), whereas $t(5)$ means z_t is Student's t with 5 degrees of freedom.

Table 3: Finite sample precision of QML methods

ϕ_1	α_0	α_1	$f(z_t)$	T	$M(\hat{\phi}_1)$	$V(\hat{\phi}_1)$	$M(\hat{\alpha}_0)$	$V(\hat{\alpha}_0)$	$M(\hat{\alpha}_1)$	$V(\hat{\alpha}_1)$	$M[\hat{E}(\ln z_t^2)]$	$V[\hat{E}(\ln z_t^2)]$
	0		$N(0,1)$	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	0.1	$N(0,1)$	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
			$st(5)$	200			-0.040	0.065	0.093	0.005	-1.658	0.053
				500			-0.019	0.024	0.099	0.002	-1.663	0.024
				1000			-0.015	0.014	0.097	0.001	-1.669	0.011
0.1	0	0.1	$N(0,1)$	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 DGPs are all nested in $r_t = \phi_1 r_{t-1} + \epsilon_t$, $\epsilon_t = \sigma_t z_t$, $z_t \overset{IID}{\sim} f(z_t)$, $\ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln \epsilon_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2$. $M(\cdot)$ and $V(\cdot)$ are the sample mean and variance 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: Estimation precision of log-GARCH(1,1): A comparison of QML methods

Method	$f(z_t)$	T	$M(\hat{\alpha}_0)$	$V(\hat{\alpha}_0)$	$M(\hat{\alpha}_1)$	$V(\hat{\alpha}_1)$	$M(\hat{\beta}_1)$	$V(\hat{\beta}_1)$	$M[\hat{E}(\ln z_t^2)]$	$V[\hat{E}(\ln z_t^2)]$
QML in $\{u_t^*\}$	$N(0,1)$	200	-0.066	0.137	0.104	0.014	0.731	0.049	-1.329	0.218
		500	-0.045	0.013	0.100	0.001	0.763	0.014	-1.274	0.007
		1000	-0.022	0.004	0.100	0.001	0.782	0.005	-1.273	0.003
	$t(5)$	200	-0.083	0.149	0.103	0.006	0.737	0.051	-1.612	0.260
		500	-0.055	0.034	0.102	0.001	0.762	0.016	-1.565	0.023
		1000	-0.027	0.006	0.101	0.000	0.781	0.004	-1.562	0.007
(Q)ML in $\{z_t\}$	$N(0,1)$	200	-0.059	0.031	0.103	0.001	0.750	0.025	–	–
		500	-0.015	0.003	0.100	0.000	0.788	0.004	–	–
		1000	-0.007	0.001	0.100	0.000	0.795	0.001	–	–
	$t(5)$	200	-0.110	0.100	0.105	0.002	0.728	0.051	–	–
		500	-0.036	0.015	0.104	0.001	0.773	0.012	–	–
		1000	-0.014	0.006	0.103	0.001	0.787	0.005	–	–

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. The simulation DGP is $r_t = \sigma_t z_t$, $z_t \overset{IID}{\sim} f(z_t)$, $\ln\sigma_t^2 = \alpha_0 + \alpha_1\epsilon_{t-1}^2 + \beta\ln\sigma_{t-1}^2$ with $\alpha_0 = 0, \alpha_1 = 0.1$ and $\beta = 0.8$. QML in $\{z_t\}$ consists of Gaussian maximum likelihood estimation with z_t as if Gaussian, with initial parameter values $(\alpha_0, \alpha_1, \beta_1) = (0, 0.05, 0.9)$. QML in $\{u_t^*\}$ consists of Gaussian maximum likelihood estimation with u_t^* as if Gaussian, with initial parameter values $(\alpha_1, \beta_1) = (0.05, 0.9)$. Estimation is undertaken using the `optim` function in the R `base` package. $M(\cdot)$ and $V(\cdot)$ are the sample mean and variances of the estimates, respectively.

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

H_0	H_1	Fitted specification	T	$\tau = 1.1$	$\tau = 2$	$\tau = 3$
$\alpha_1 = 0$	$\alpha_1 \neq 0$	$\alpha_0 + \alpha_1 \ln \epsilon_{t-1}^2$	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 $r_t = \epsilon_t$, $\epsilon_t = \sigma_t z_t$, $z_t \overset{IID}{\sim} GED(\tau)$, $\ln \sigma_t^2 = 0$, for $t = 1, \ldots, T$. Tests are two-sided. Simulations in R with 10 000 replications.

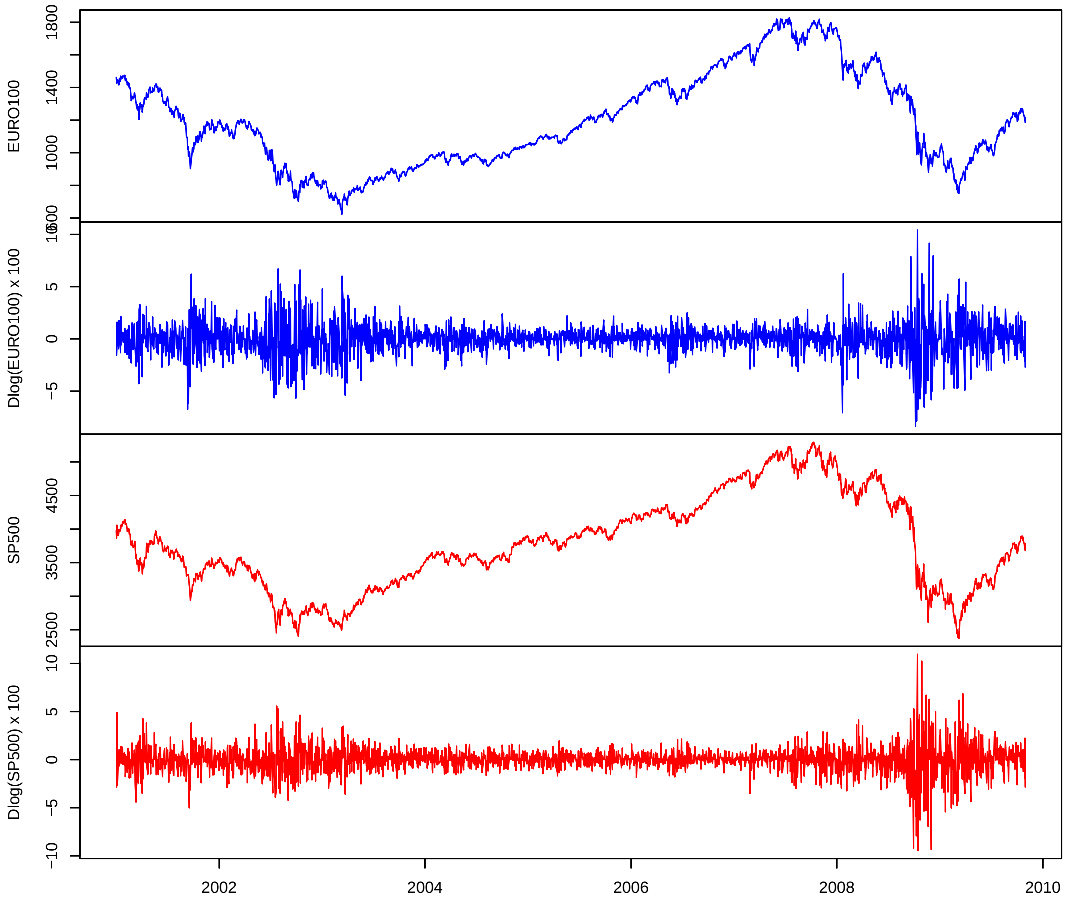

Figure 1: Daily demeaned log-returns (in percent) of the EURO100 and SP500 stock market indices 1 January 2001 - 30 October 2009 (2302 observations)

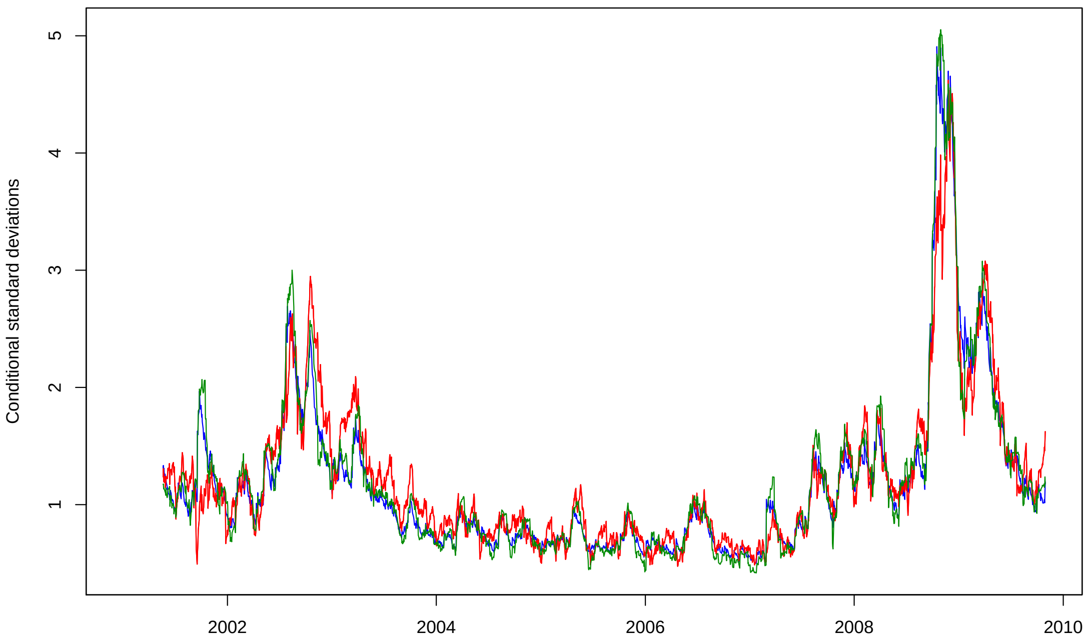

Figure 2: Fitted values of conditional standard deviations of univariate models. Blue line: GARCH(1,1); Red line: 1st and 2nd power Log-GARCH(1,1); Green line: Log-ARCH(0) with $EqWMA(20)_{t-1}$ as volatility proxy

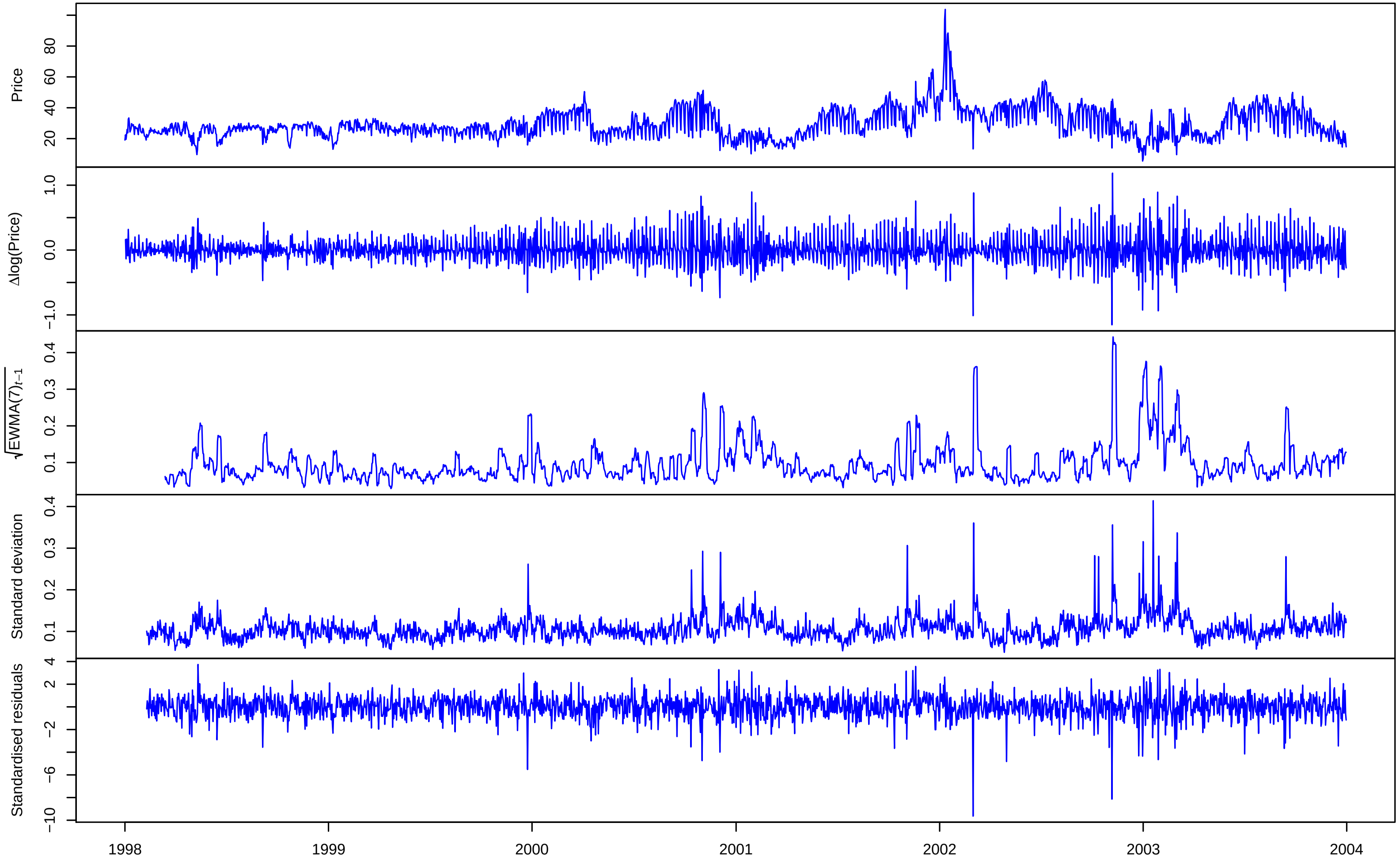

Figure 3: Daily electricity price (upper graph), log-returns (second graph), $\sqrt{\widehat{EqWMA}(7)_{t-1}}$ (third graph), fitted standard deviation (fourth graph) and the standardised residuals (bottom graph) of the (13) - (15) model for Spain (1 January 1998 - 31 December 2003).

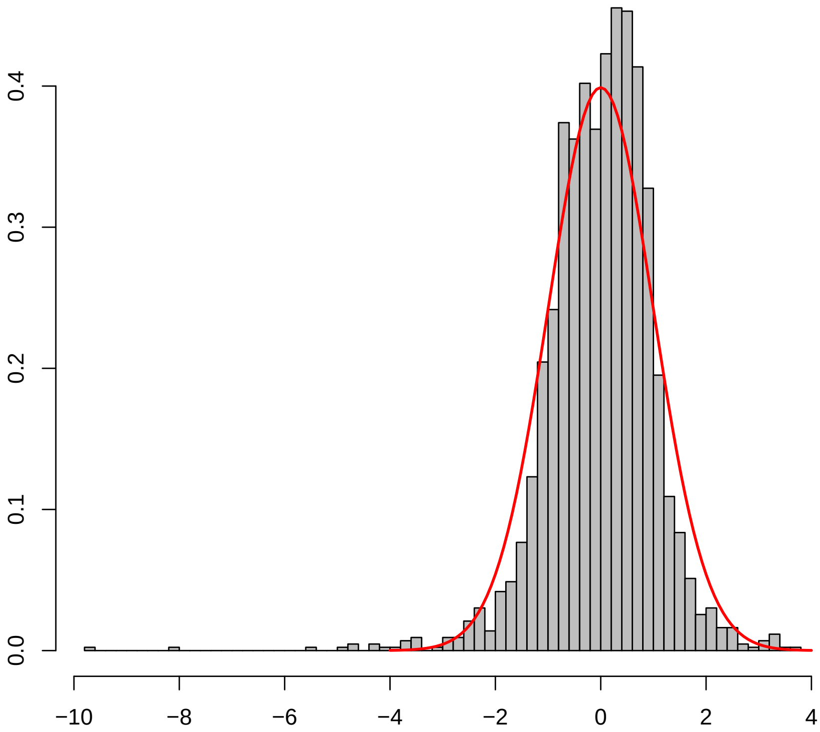

Figure 4: Empirical relative frequency distribution *vs.* the standard normal (red graph) of the standardised residuals of the specific model (13)-(15)

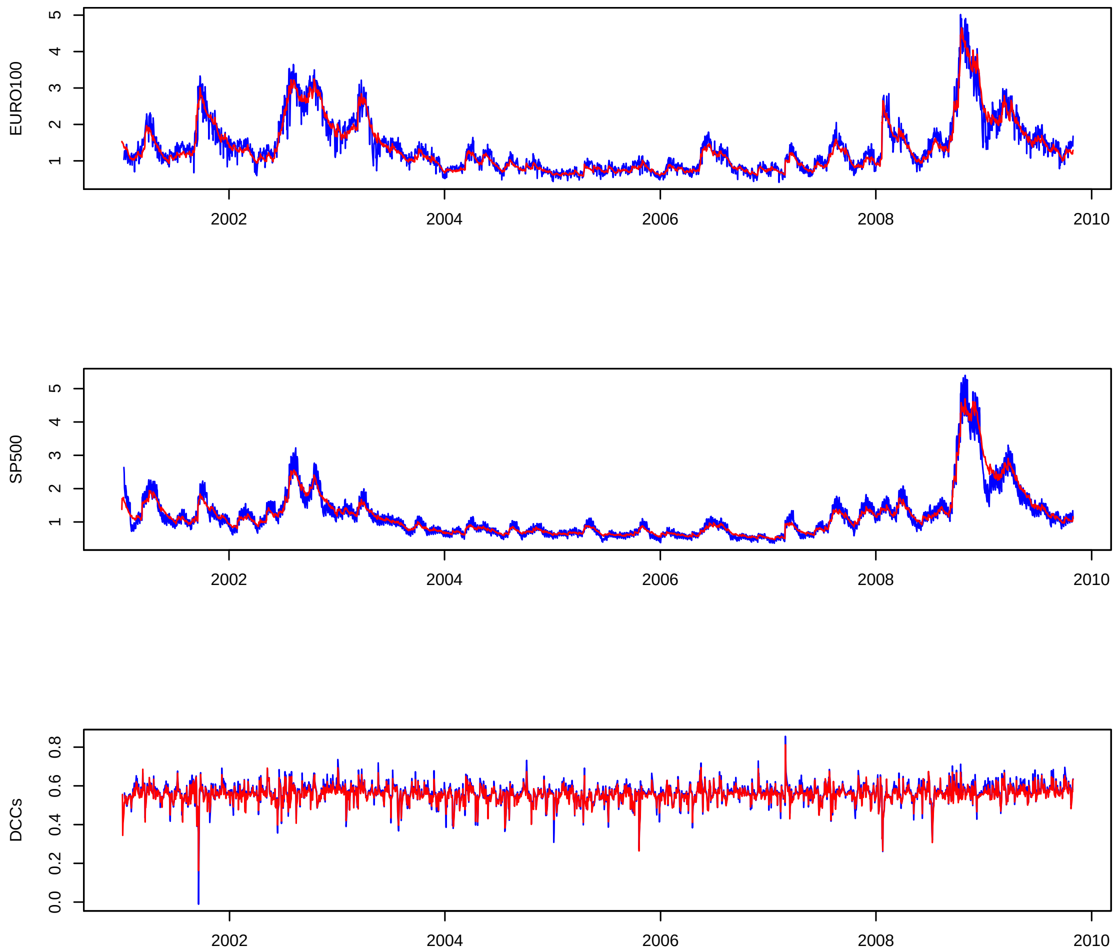

Figure 5: Conditional standard deviations (upper graphs) and time-varying correlations (bottom graph) of the bivariate log-ARCH model (blue lines) and the bivariate (diagonal) DCC model (red lines) of Engle (2002)