This is a postprint version of the following published document:


© Elsevier
Estimation and inference in univariate and multivariate log-GARCH-X models when the conditional density is unknown

Genaro Sucarrat\textsuperscript{a,}\textsuperscript{*}, Steffen Grønneberg\textsuperscript{a}, Alvaro Escribano\textsuperscript{b}

\textsuperscript{a} Department of Economics, BI Norwegian Business School, Oslo, Norway
\textsuperscript{b} Department of Economics, Universidad Carlos III de Madrid, Madrid, Spain

\begin{abstract}
A general framework for the estimation and inference in univariate and multivariate Generalised log-ARCH-X (i.e. log-GARCH-X) models when the conditional density is unknown is proposed. The framework employs (V)ARMA-X representations and relies on a bias-adjustment in the log-volatility intercept. The bias is induced by (V)ARMA estimators, but the remaining parameters can be estimated in a consistent and asymptotically normal manner by usual (V)ARMA methods. An estimator of the bias and a closed-form expression for the asymptotic variance is derived. Adding covariates and/or increasing the dimension of the model does not change the structure of the problem, so the univariate bias-adjustment procedure is applicable not only in univariate log-GARCH-X models estimated by the ARMA-X representation, but also in multivariate log-GARCH-X models estimated by VARMA-X representations. Extensive simulations verify the properties of the log-moment estimator, and an empirical application illustrates the usefulness of the methods.

\textbf{Keywords:} Log-GARCH-X, ARMA-X, Multivariate log-GARCH-X, VARMA-X, Volatility
\end{abstract}

1. Introduction

The Autoregressive Conditional Heteroscedasticity (ARCH) class of models due to Engle (1982) is useful in a wide range of applications. In finance, in particular, it has been extensively used to model the clustering of large (in absolute value) financial returns. Engle himself, however, originally motivated the class as useful in modelling the time-varying conditional uncertainty (i.e. conditional variance) of economic variables in general, and of UK inflation in particular. Other areas of application include, among others, the uncertainty of electricity prices (e.g. Escribano et al., 2011, Koopman et al., 2007), the evolution of temperature data (e.g. Franses et al., 2001) and – more generally – positively valued variables, i.e. so-called Multiplicative Error Models (MEMs), see Brownlees et al. (2012) for a survey.

Within the ARCH class of models exponential versions are of special interest. This is because they enable richer autoregressive volatility dynamics (e.g. contrarian or cyclical) compared with non-exponential ARCH models, and because their fitted values of volatility are guaranteed to be positive. The latter is not necessarily the case for ordinary (i.e. non-exponential) ARCH models, particularly when covariates or other conditioning variables (“X”) are added to the volatility equation. In fact, the greater the dimension of X, the more restrictions are needed in order to ensure positivity. Another desirable property is that volatility forecasts are more robust to jumps and outliers. Robustness can be important in order to avoid volatility fore-
cast failure subsequent to jumps and outliers. The log-GARCH class was independently proposed by Pantula (1986); Geweke (1986) and Milhøj (1987). Engle and Bollerslev (1986) argued against log-ARCH models because of the possibility of applying the log-operator (in the log-ARCH terms) on zero-values, which occurs whenever the error term in a regression equals zero. A solution to this problem, however, is provided in Sucarrat and Escribano (2013) for the case where the zero-probability is zero (e.g. because zeros are due to discreteness or missing values). The solution is only available when estimation is via the (V)ARMA representation. Finally, two competing classes of exponential ARCH models are the EGARCH (Nelson, 1991) and the Beta-t-EGARCH model (Harvey, 2013). The former has proved to be more difficult theoretically (more on this below), and the latter is not – by its very nature – amenable to the assumption of an unknown conditional density (i.e. the conditional density must be known).

The assumption that the conditional density is unknown is particularly convenient from a practitioner’s point of view. This is because the user then does not need to worry about changing the conditional density from application to application, or alternatively to work with a sufficiently general density that will often make estimation and inference numerically more challenging. This explains the attraction of Quasi Maximum Likelihood Estimators (QMLEs). In the univariate case, consistency and asymptotic normality of QMLE for GARCH models under mild conditions were first established by Berkes et al. (2003), and Francq and Zakoïan (2004). In the exponential case, most of the attention has been directed at the EGARCH, whose asymptotic properties have turned out to be very difficult to establish, see e.g. Straumann and Mikosch (2006). Only recently was consistency and asymptotic normality proved (for the univariate EGARCH(1,1) only) under the complicated condition of continuous invertibility, see Wintenberger (2013). The log-GARCH model is much more tractable. Francq et al. (forthcoming) prove consistency and asymptotic normality of the Gaussian QMLE for an asymmetric log-GARCH(p, q) model under mild conditions. Their method does not employ ARMA representations, which means it is more efficient when the conditional error is normal or close to normal, but not when the conditional density is fat-tailed, see the asymptotic efficiency comparison in Francq and Sucarrat (2013). Moreover, the estimator of Francq et al. (forthcoming) cannot handle zero-errors or missing values in the manner suggested by Sucarrat and Escribano (2013). Finally, Francq and Sucarrat (2013) propose an estimator that achieves efficiency for conditional densities that are normal or close to normal, by combining the ARMA-approach with the Centred Exponential Chi-Squared as instrumental QML-density. In the multivariate case, QML results have been established for the BEKK model of Engle and Kroner (1995) by Comte and Lieberman (2003), for an ARMA–GARCH with constant conditional correlations (CCCs) by Ling and McAleer (2003), for a factor GARCH model by Hafner and Preminger (2009), for a multivariate GARCH with CCCs by Francq and Zakoïan (2010), and for a multivariate GARCH with stochastic correlations by Francq and Zakoïan (2015) under the assumption that the system is estimable equation-by-equation. For exponential ARCH models there are no multivariate results. Kawakatsu (2006) proposed a multivariate exponential ARCH model, the matrix exponential GARCH, which contains a multivariate version of the EGARCH model. But there are no proofs for the estimation and inference methods that he proposes.

This paper makes three contributions. It is well-known that all the coefficients apart from the log-volatility intercept in a univariate log-GARCH specification can be estimated consistently (under suitable assumptions) via an ARMA representation, see for example Psaradakis and Tzavalis (1999), and Francq and Zakoïan (2006). However, the estimate of the log-volatility intercept will be asymptotically biased, and the bias is made up of a log-moment expression that depends on the unknown density of the conditional error. A simple estimator of the log-moment expression made up of the empirical residuals of the ARMA regression (Section 2.2) is derived. The implication of this is that the log-volatility intercept can be estimated consistently, and hence that all the log-GARCH parameters can be estimated consistently via the ARMA representation. An expression for the asymptotic variance (Section 2.3) of the estimator of the log-moment expression is also derived.

In the second contribution of the paper (Section 2.4), it is shown that the addition of covariates, i.e. the log-GARCH-X model, does not alter the relation between the ARMA coefficients and the log-GARCH coefficients. In other words, consistent and asymptotically normal estimation of the ARMA-X representation will produce exactly the same bias as earlier, and so the bias correction procedure described above is applicable also for ARMA-X models. Next, a multivariate log-GARCH-X model that admits time-varying conditional correlations is proposed (Section 3). The model has a VARMA-X representation with a vector of error-terms. The vector is either IID, which corresponds to the Constant Conditional Correlation (CCC) case, or independent but non-identical (ID), which corresponds to the time-varying correlations case. In both cases, however, each entry in the vector of standardised errors is marginally IID. So the bias-correction from the univariate case can be used equation-by-equation – under suitable assumptions – subsequent to the consistent estimation of the VARMA-X representation.

In the third contribution (Section 4) the usefulness of the results is illustrated by means of an application to the modelling of the uncertainty of electricity prices. Electricity prices are characterised by autoregressive persistence, day-of-the-week effects, large spikes or jumps, ARCH and non-normal conditional errors that are possibly skewed. For robust (to jumps) forecasts of uncertainty (i.e. volatility) that accommodate all these characteristics, the log-GARCH-X model is particularly suited. The investigation shows that estimated volatility can be substantially biased if sufficient ARCH-lags and day-of-the-week effects are not included.

The rest of the paper is organised as follows. Section 2 presents the univariate log-GARCH model, the relation between the univariate log-GARCH model and its ARMA representation, and derives the log-moment estimator and its asymptotic variance. Also, it is shown that the addition of covariates does not alter the relationship between the log-GARCH and ARMA parameters. Section 3 shows how the ideas extend to the multivariate case. Section 4 contains our empirical application, and Section 5 concludes.
2. Univariate log-GARCH

The univariate log-GARCH\((p, q)\) model is given by
\[
\begin{align*}
\epsilon_t &= \sigma_t z_t, \quad z_t \sim \text{IID}(0, 1), \quad P(z_t = 0) = 0, \quad \sigma_t > 0, \quad (1) \\
\ln \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2, \quad t \in \mathbb{Z}, \quad (2)
\end{align*}
\]
where \(p\) is the ARCH order and \(q\) is the GARCH order. In finance, \(\epsilon_t\) is often interpreted as return or mean-corrected return, but more generally it is simply the error in a regression model. Throughout, it is assumed that \(\epsilon_t\) is observable and known. Of course, this is neither a realistic nor a desirable assumption, but simply reflects the current state of the theoretical literature. 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 and if \(|E(\ln z_t^2)| < \infty\), then \(\ln \sigma_t^2\) is stable. For common densities like the Student’s \(t\) with degrees of freedom greater than 2, and the Generalised Error Distribution (GED) with shape parameter greater than 1, then \(\sigma_t^2\) will generally be stable as well if \(\ln \sigma_t^2\) is stable. Practitioners are often interested in the dynamics of other powers, e.g. the conditional standard deviation. For that purpose, it should be noted that the \(d\)th power log-GARCH\((p, q)\) model can be written as
\[
\ln \sigma_t^{d} = \alpha_{0,d} + \sum_{i=1}^{p} \alpha_i \ln |\epsilon_{t-i}|^d + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^d, \quad d > 0, \quad (3)
\]
where \(\alpha_{0,d} = \alpha_0 d/2\). This means a complete analysis of the \(d\)th power log-GARCH model can be undertaken in terms of the \(d = 2\) representation.

2.1. The ARMA representation

If \(|E(\ln z_t^2)| < \infty\), then the log-GARCH\((p, q)\) model (1)-(2) admits the ARMA\((p, q)\) representation
\[
\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + u_t, \quad u_t = \ln z_t^2 - E(\ln z_t^2), \quad (4)
\]
where
\[
\phi_0 = \alpha_0 + \left(1 - \sum_{j=1}^{q} \beta_j\right) \cdot E(\ln z_t^2), \quad \phi_i = \alpha_i + \beta_i \quad \text{and} \quad \theta_j = -\beta_j. \quad (5)
\]
Consistent and asymptotically normal estimates of all the ARMA parameters – and hence all the log-GARCH parameters except the log-volatility intercept \(\alpha_0\) – are thus readily obtained via usual ARMA estimation methods subject to appropriate assumptions, see e.g. Brockwell and Davis (2006). To estimate \(\alpha_0\), the most common solutions have been to either impose restrictive assumptions regarding the distribution of \(z_t\) (say, normality, see e.g. Psaradakis and Tzavalis (1999)), or to use an ex post scale-adjustment (see e.g. Bauwens and Sucarrat, 2010, and Sucarrat and Escribano, 2012). What is shown below is that the ex post scale-adjustment (i.e. formula (6) and its modified version (8)) provides a consistent estimate of \(E(\ln z_t^2)\). Consequently, the final log-GARCH parameter \(\alpha_0\) can also be estimated consistently.

2.2. On consistency

The scale-adjustment employed by Bauwens and Sucarrat (2010), and Sucarrat and Escribano (2012), is essentially a smearing estimate (more on this below). Consider writing (1) as
\[
\epsilon_t = \sigma_t^z z_t^*, \quad z_t^* \sim \text{IID}(0, \sigma_t^2),
\]
where \(\sigma_t^z\) is a time-varying scale, not necessarily equal to the standard deviation, and where \(z_t^*\) does not necessarily have unit variance. Of course, by construction \(\sigma_t = \sigma_t^z \sigma_{z^*}\) and \(z_t = z_t^*/\sigma_{z^*}\). Next, suppose a log-scale specification (e.g. an ARMA specification contained in (4)) is fitted to \(\ln \epsilon_t^2\), with \(\ln \hat{\sigma}_{t^2}\) denoting the fitted value of the ARMA specification such that \(\hat{\sigma}_{t^2} = \exp(\ln \hat{\sigma}_{t^2})\), and with the ARMA residual defined as \(\hat{u}_t = \ln \epsilon_t^2 - \ln \hat{\sigma}_{t^2}\). In order to obtain an estimate of the time-varying conditional standard deviation, which is needed for comparison with other volatility models, then it is natural to consider adjusting \(\hat{\sigma}_{t^2}\) by multiplying it with an estimate of \(\sigma_{z^*}\), say, the sample standard deviation of the standardised residuals \(\hat{z}_t^*\). Although this argument is fine heuristically, it is not straightforwardly apparent what underlying magnitude the adjustment actually estimates. It transpires that, in the log-GARCH model, the log of the scale-adjustment provides an estimate of \(-E(\ln z_t^2)\). To see this, consider the scale adjustment and its approximation:
\[
\hat{\sigma}_{t^2}^2 = \frac{1}{T-1} \sum_{t=1}^{T} (\hat{z}_t^* - \hat{z}_t^*_t)^2 \approx \frac{1}{T} \sum_{t=1}^{T} (\hat{z}_t^*)^2 = \frac{1}{T} \sum_{t=1}^{T} \exp(\hat{u}_t).
\]
The population analogue of the final expression is \( E[\exp(u_t)] \). Taking the natural log of \( E[\exp(u_t)] \) gives \( \ln E[\exp(u_t)] = -E(\ln z_t^2) \) under the assumption that \( E(z_t^2) = 1 \), i.e. the identifiability assumption from (1). This suggests that

\[
- \ln \left[ \frac{1}{T} \sum_{t=1}^{T} \exp(u_t) \right]
\]

provides a consistent estimate of \( E(\ln z_t^2) \), due to the continuity of the logarithm function.

The expression in square brackets in (6), i.e. \( T^{-1} \sum_{t=1}^{T} \exp(u_t) \), is well-known as the “smearing estimate”, see Duan (1983). It provides an estimate of the adjustment needed for an unbiased estimate of \( E(y_t | x_t) \) when the left-hand side of the estimated model is \( \ln y_t \). The proof of Duan (1983), however, is for static models. In dynamic models, e.g. when the \( u_t \)'s are ARMA residuals, then a different proof strategy and additional assumptions are needed. Complete proofs under mild assumptions that hold under all the configurations covered in this paper, however, are beyond our scope. For simplicity and convenience, therefore, a set of minimal assumptions and conditions relied upon throughout is formulated, and a proof of the key condition (A2) is provided only in the log-ARCH(\( p \)) case (recently, Francq and Sucarrat (2015) prove A2 for an equation-by-equation least squares estimator of a multivariate log-GARCH-X model with Dynamic Conditional Correlations).

Formally, the following assumptions are relied upon:

A1: \( E(z_t^2) = 1 \) and \( |E(\ln z_t^2)| < \infty \).

A2: Let \( \hat{u}_t, t = 1, \ldots, T \) denote the ARMA-residuals resulting from estimating the ARMA representation (4). Then:

\[
\frac{1}{T} \sum_{t=1}^{T} \exp(\hat{u}_t) - \frac{1}{T} \sum_{t=1}^{T} \exp(u_t) = o_p(1).
\]

In A1, the first moment condition is simply the identifiability condition from (1), whereas the other moment condition \( |E(\ln z_t^2)| < \infty \) is required for the ARMA representation (4) to exist. For the two most commonly used densities of \( z_t \) in finance, i.e. \( N(0, 1) \) and \( t, E(\ln z_t^2) \) are finite. Regarding A2, it immediately implies that (6) is a consistent estimator of \( E(\ln z_t^2) \) due to the continuity of the logarithm function. As already noted, though, a complete proof of A2 under all the configurations covered by this paper is beyond our scope. In the log-ARCH(\( p \)) case, however, the proof is relatively straightforward.

**Theorem 1.** Suppose \( \alpha^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i, \ln e_{t-i}^2 \) in (1)-(2), that \( \ln e_t^2 \) is strictly stationary and that A1 holds. The mean-corrected AR(\( p \)) representation is then given by \( \ln e_t^2 - E(\ln e_t^2) = \sum_{i=1}^{p} \phi_i(\ln e_{t-i}^2 - E(\ln e_t^2)) + u_t \), where \( \phi_i = \alpha_i \) as in (5). Define \( \tilde{Y}_t = \ln e_t^2 - T^{-1} \sum_{t=1}^{T} \ln e_t^2 \). Let \( \hat{\phi}_1, \ldots, \hat{\phi}_p \) denote the OLS estimates of \( \phi_1, \ldots, \phi_p \) based on the \( \tilde{Y}_t \)'s, let \( \hat{u}_t = \tilde{Y}_t - \sum_{i=1}^{p} \hat{\phi}_i \tilde{Y}_{t-i} \) for \( t > p \) and let \( \hat{u}_t = 0 \) for \( 0 < t \leq p \). If \( E(z_t^4) < \infty \) and \( |E(\ln z_t^2)| < \infty \), then A2 holds.

**Proof.** See Appendix A. \( \square \)

The theorem states that A2 holds when the mean-corrected AR(\( p \)) representation of a log-ARCH(\( p \)) model is estimated by OLS, which then implies that (6) is a consistent estimator of \( E(\ln z_t^2) \). Next, it follows straightforwardly that all the log-ARCH(\( p \)) parameters can be estimated via the relationships in (5), since \( \hat{\phi}_0 = \left(1 - \sum_{i=1}^{p} \hat{\phi}_i \right) \cdot T^{-1} \sum_{t=1}^{T} \ln e_t^2 \) provides a consistent estimate of \( \phi_0 \) under the assumptions of the theorem. Strict stationarity of \( \ln e_t^2 \) follows if the roots of the AR-polynomial are all outside the unit-circle.

2.3. On normality

Our main interest is a consistent estimator of \( E(\ln z_t^2) \), so that the ARMA-estimates can be used to consistently estimate all the log-GARCH parameters via the relationships in (5). To this end, the limiting distribution of the estimator of \( E(\ln z_t^2) \) is of minor interest. In simulations, however, the limiting distribution and an expression for the asymptotic variance can be useful in verifying simulation results.

Let (6) be modified to

\[
\hat{T} = - \ln \left[ \frac{1}{T} \sum_{t=1}^{T} \exp(\hat{u}_t - \tilde{u}_T) \right],
\]

where \( \tilde{u}_T \) is the empirical mean of the ARMA-residuals. The mean-correction term \( \tilde{u}_T \) is needed, since asymptotic normality may not be achieved without it. See e.g. the related discussion in Yu (2007), where high moment partial sum processes of residuals in ARMA models are treated, and where a mean-correction term is needed for asymptotic normality. In some cases, e.g. when OLS is used to estimate the AR(\( p \)) representation of a log-ARCH(\( p \)) model, then \( \tilde{u}_T \) is zero by construction, and so (8) equals (6). The following two assumptions will give us asymptotic normality of (8):
A3: Let \( \{\tilde{u}_t\}_{t=1}^T \) denote the ARMA-residuals resulting from estimating the ARMA representation (4). Denoting \( \overline{u}_T \) and \( \overline{u}_T \) as the averages of \( \tilde{u}_t \) and \( u_t \), respectively:

\[
\sqrt{T} \left[ \frac{1}{T} \sum_{t=1}^T \exp(\tilde{u}_t - \overline{u}_T) - \frac{1}{T} \sum_{t=1}^T \exp(u_t - \overline{u}_T) \right] = o_p(1).
\]

A4: \( E(z_t^4) < \infty \) and \( |E([\ln z_t^2])^2| < \infty \).

Condition A3 is slightly stronger than A2, since A3 implies that (8) provides a consistent estimate of \( E(\ln z_t^2) \) as long as A1 holds. The moment conditions in A4 are needed for the asymptotic variance of (8) to be finite.

**Theorem 2.** Suppose (1)–(2), A1, A3 and A4 hold. Then

\[
\sqrt{T} \left[ \tilde{r}_T - E(\ln z_t^2) \right] \xrightarrow{D} N(0, \xi^2), \quad \text{where } \xi^2 = \text{Var} (z_t^2 - \ln z_t^2).
\]

**Proof.** See Appendix B. □

The key assumption for asymptotic normality to hold is A3, but a complete proof under all the configurations covered by this paper is beyond our scope. Just as for consistency in the log-ARCH\((p)\) case (see Theorem 1), however, a proof of asymptotic normality is relatively straightforward.

**Theorem 3.** Suppose the assumptions of Theorem 1 hold. If in addition \( E(u_t^4) < \infty \), then A3 holds.

**Proof.** See Appendix A. □

Assumption A4 holds under the assumptions of Theorem 1. The condition \( E(u_t^4) < \infty \) is, in fact, a very weak additional assumption, since it follows from \( E(e^{4u_t}) < \infty \). An extensive set of Monte Carlo simulations have been performed, of which a small subset is available as supplementary material from the Webpage of the first author (available as http://www.sucarrat.net/research/lgarchxsims.pdf). The simulations confirm that the usual ARMA-methods (e.g. Nonlinear Least Squares and Gaussian QML) provide consistent estimates, and that the empirical standard errors coincide with their asymptotic counterparts.

### 2.4. Log-GARCH-X

Additional covariates or conditioning variables (“X”) can be added linearly or nonlinearly to the log-volatility specification \( \ln \sigma_t^2 \) without affecting the relationship between the log-GARCH coefficients and the ARMA coefficients. Specifically, let the log-GARCH-X model be given by

\[
\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \ln \sigma_{t-j}^2 + g(\lambda, x_t),
\]

where \( g \) is a linear or nonlinear function of the conditioning variables \( x_t \) and a parameter vector \( \lambda \). The index \( t \) in \( x_t \) does not necessarily mean that all (or any) of its elements are contemporaneous. If \( |E(\ln z_t^2)| < \infty \), then (10) admits the ARMA-X representation

\[
\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^p \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^q \theta_j u_{t-j} + g(\lambda, x_t) + u_t, \quad u_t = \ln z_t^2 - E(\ln z_t^2),
\]

where the ARMA coefficients are defined as before, i.e. by (5). A complete proof of consistency and asymptotic normality would of course require precise assumptions on the behaviour of \( x_t \), see for example Francq and Sucarrat (2015), and Chapter 4 in Hannan and Deistler (2012).

One type of conditioning variable that is of special interest in financial applications is leverage or volatility asymmetry. Table 1 provides simulation results for a simple version of leverage, \( g(\lambda, x_t) = \lambda I_{z_{t-1} < 0}, \) where \( I_{z_{t-1} < 0} \) is an indicator function equal to 1 if \( z_{t-1} < 0 \) and 0 otherwise. Note that \( I_{z_{t-1} < 0} \) is observable, since \( I_{z_{t-1} < 0} = I_{\epsilon_{t-1} < 0} \). The simulations show that all the parameters are estimated consistently, and the second-to-last column shows that the finite sample empirical standard error of the estimate of \( E(\ln z_t^2) \) corresponds well to its asymptotic counterpart (last column), both for the normal and standardised \( t \) distributions (see Table 2 for additional simulations).
3. Multivariate log-GARCH

The $M$-dimensional log-GARCH model is given by

$$
e_t \sim \text{ID}(0, H_t), \quad t \in \mathbb{Z},$$

$$D_t^2 = \text{diag} \{ \sigma^2_{m,t} \}, \quad m = 1, \ldots, M,$$

$$z_t = D_t^{-1} \epsilon_t, \quad \forall m : z_{m,t} \sim \text{IID}(0, 1), \quad P(z_t = 0) = 0,$$

$$\ln \sigma^2_t = \alpha_0 + \sum_{i=1}^p \alpha_i \ln \sigma^2_{t-i-1} + \sum_{j=1}^q \beta_j \ln z_{t-j}^2, \quad p \geq q,$$

where $\epsilon_t, \sigma^2_t$ and $z_t$ are $M \times 1$ vectors, and where $H_t$ and $D_t$ are $M \times M$ matrices. In (15) we have that $\alpha_0 = (\alpha_{1,0}, \ldots, \alpha_{M,0})'$.

$$\alpha_t = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,M} \\ \vdots & \ddots & \vdots \\ \alpha_{M,1} & \cdots & \alpha_{M,M} \end{pmatrix} \quad \text{and} \quad \beta_j = \begin{pmatrix} \beta_{11,j} & \cdots & \beta_{1M,j} \\ \vdots & \ddots & \vdots \\ \beta_{M1,j} & \cdots & \beta_{MM,j} \end{pmatrix},$$

where $'$ is the transpose operator. Eq. (12) means that $\epsilon_t$ is independent with mean zero and a time-varying conditional covariance matrix $H_t$. The IID assumption in Eq. (14) states that each marginal series $\{z_{m,t}\}$ is IID(0, 1). Marginal identifiability is a key characteristic of the ARCH class of models, and is needed for (6) (or (8)) to be applicable after estimation via the VARMA representation. An implication of (14) is that $z_t \sim \text{ID}(0, R_t)$, where $R_t$ is both the conditional covariance and correlation matrix – possibly time-varying – of $z_t$. In other words, the vector $z_t$ is IID, but not necessarily IID, even though each marginal series $\{z_{m,t}\}$ is IID. In the special case where vector $z_t$ is IID, then $R_t$ is a Constant Conditional Correlation (CCC) model. Estimation of the volatilities $D_t^2$ does not require that the off-diagonals of $H_t$ (i.e. the covariances) are specified explicitly. Nor need we assume that $\epsilon_t$ is distributed according to a certain density, say, the normal.

3.1. The VARMA representation

If $|E(\ln z_t^2)| < \infty$, then the $M$-dimensional log-GARCH($p, q$) model (15) admits the VARMA($p, q$) representation

$$\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^p \phi_i \ln \epsilon_{t-i-1}^2 + \sum_{j=1}^q \theta_j u_{t-j} + u_t, \quad u_t = \ln z_t^2 - E(\ln z_t^2),$$

where

$$\phi_0 = \alpha_0 + \left( I_M - \sum_{j=1}^q \beta_j \right) : E(\ln z_t^2), \quad \phi_i = \alpha_i + \beta_i, \quad \text{and} \quad \theta_j = -\beta_j.$$ 

In the special case where vector $z_t$ is IID, which implies a CCC model for the correlations (assuming they exist), then vector $u_t$ is IID as well. In this case, it is well-known that the multivariate Gaussian QMLE provides consistent and asymptotically normal estimates of the VARMA coefficients under suitable assumptions, see e.g. Lütkepohl (2005). Accordingly, consistent estimation and asymptotically normal inference regarding all the log-GARCH coefficients – apart from the log-volatility
\[ \text{The estimated model is } \ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln e_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2 + \lambda x_t, \]  
\[ \text{(19)} \]

where \( x_t \) is an \( L \times 1 \) vector of covariates, and where \( \lambda \) is an \( M \times L \) matrix. For notational economy, the covariates \( x_t \) enter linearly, but in principle they can enter non-linearly in the univariate case, see (11). Similarly, index \( t \) in \( x_t \) does not necessarily mean that all (or any) of its elements are contemporaneous. The VARMA-X representation of (19) is then given by

\[ \ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^{p} \phi_i \ln e_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + \lambda x_t + u_t, \]

with the VARMA coefficients and \( u_t \) defined as before, i.e. by (18). In other words, the relation between the VARMA coefficients and the log-GARCH coefficients are not affected by adding \( \lambda x_t \) to (19). So VARMA-X methods can be used to estimate all the log-GARCH parameters (under suitable assumptions on \( x_t \)) except the log-volatility intercept \( \alpha_0 \) in a first step, and then in a second step equation-by-equation application of (6) can be used to estimate each element in \( E(\ln z_t^2) \) and, hence, the log-volatility intercept \( \alpha_0 \). Also here it is useful to distinguish between the CCC and time-varying correlations cases. If \( u_t \) is IID, i.e. the CCC case, then – under suitable assumptions – the multivariate Gaussian QMLE provides consistent estimates of the VARMA-X representation, see e.g. Hannan and Deistler (2012). If correlations are time-varying, and if the matrices \( \beta_i \) are diagonal, then each equation can be estimated separately in terms of their ARMA-X representations, see Francq and Sucarrat (2015).

### Table 2

Finite sample properties of equation-by-equation Gaussian QML of a 2-dimensional log-GARCH(1,1) w/diagonal matrix \( \beta_t \) when the correlations follow the DCC of Engle (2002).

| DGP | \( T \) | \( m(\hat{\alpha}_{10}) \) | \( m(\hat{\alpha}_{20}) \) | \( m(\hat{\alpha}_{11}) \) | \( m(\hat{\alpha}_{21}) \) | \( m(\hat{\beta}_{11}) \) | \( m(\hat{\beta}_{12}) \) | \( m(\hat{\beta}_{21}) \) | \( m(\hat{\beta}_{22}) \) | \( m(\hat{\tau}_{1}) \) | \( m(\hat{\tau}_{2}) \) | \( m(\hat{\tau}_{2}) \) | \( m(\hat{\tau}_{2}) \) | \( m(\hat{\tau}_{2}) \) | \( \text{s.e} (\hat{\tau}) \) |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| A:  | 1000 | -0.065 | -0.229 | 0.046 | 0.101 | 0.101 | 0.048 | 0.902 | 0.680 | -1.270 | 0.056 | -1.270 | 0.054 | -1.270 | 0.054 |
| B:  | 1000 | -0.029 | -0.026 | 0.098 | 0.053 | 0.053 | 0.097 | 0.791 | 0.792 | -1.270 | 0.055 | -1.270 | 0.053 | -1.270 | 0.054 |
| C:  | 1000 | -0.005 | -0.023 | 0.099 | 0.099 | 0.099 | 0.099 | 0.799 | 0.799 | -1.270 | 0.056 | -1.270 | 0.054 | -1.270 | 0.054 |

The estimated model is \( \ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln e_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2 + \lambda x_t \), where \( \alpha_0 = (\alpha_{10}, \alpha_{20})' \), \( \alpha_1 = (\alpha_{11}, \alpha_{21})' \) and \( \beta_1 = (\beta_{11}, \beta_{12}) \). The standardised errors (\( z_t, z_{2t} \)) are governed by the DCC of Engle (2002): (\( z_t, z_{2t} \) \( \sim N(0, \Sigma_t) \)). Table 2 contains simulation results of \( E(\ln z_t^2) \) needed. Since the process \( \{u_{mt}\} \) is marginally IID for each \( m \), equation-by-equation application of (6) (or of (8)) after estimation of the VARMA can be used to estimate each element in \( E(\ln z_t^2) \).
4. Application: Modelling the uncertainty of electricity prices

Daily electricity prices are characterised by autoregressive persistence, day-of-the-week effects, large spikes or jumps, ARCH and non-normal conditional errors that are possibly skewed. Koopman et al. (2007), Escribano et al. (2011), and Bauwens et al. (2013) have proposed univariate and multivariate models that contain some or several of these features. However, in none of these models is the volatility specification – a non-exponential GARCH – robust to the large spikes that are a common characteristic of electricity prices. Nor, are they flexible enough to accommodate a complex and rich heteroscedasticity dynamics similar to that of the mean specification without imposing very strong parameter restrictions (e.g. non-negativity). Finally, automated model selection with a large number of variables is infeasible in practice due to computational complexity and positivity constraints. The log-GARCH-X class of models, by contrast, remedies these deficiencies.

The data consist of the daily peak and off-peak spot electricity prices (in Euros per kWh) from 1 January 2010 to 20 May 2014 (i.e. 1601 observations before lag-adjustments) for the Oslo region in Norway. The source of the data is http://www.nordpoolspot.com/, and the sample was determined by availability: Observations prior to the sample period are not available, and the data were downloaded just after 20 May 2014. Electricity forwards for this region are traded at the Nord Pool Spot energy exchange, a leading European market for electrical energy. Factories, companies and other institutions with electricity consumption may want to shift part of their activity to and from peak hours for efficient cost management, since the difference between peak and off-peak prices can be very large at times, see the bottom graphs of Fig. 1. As an aid in the decision-making process, forecasts of future prices and of price uncertainty (i.e. volatility or risk) can, therefore, be of great usefulness. The daily peak spot price \( S_{1,t} \) is computed as the average of the spot prices during peak hours, i.e. \( S_{1,t} = (S_{t(8\, \text{am})} + \cdots + S_{t(9\, \text{pm})})/14 \), whereas the daily off-peak spot price \( S_{2,t} \) is computed as the average of the spot prices during off-peak hours, i.e. \( S_{2,t} = (S_{t(0\, \text{am})} + \cdots + S_{t(7\, \text{am})} + S_{t(10\, \text{pm})} + S_{t(11\, \text{pm})})/10 \). Note that \( S_{t(8\, \text{am})} \) should be interpreted as the electricity price from 8 am to 9 am, \( S_{t(9\, \text{am})} \) should be interpreted as the electricity price from 9 am to 10 am, and so on. Graphs of \( S_{1,t}, S_{2,t} \) and their log-returns \((r_t = \Delta \ln S_t)\) are contained in Fig. 1. The price and returns figures exhibit the usual characteristics of electricity prices, namely that the price variability is substantially larger than the variability of, say, stock prices, stock indices and exchange rates, and that big jumps occur relatively frequently.

The conditional mean is specified as a two-dimensional Vector Error Correction Model (VECM) augmented with day-of-the-week dummies in both equations (the R-squared of the two equations are 0.26 and 0.17, respectively; more details are available on request). The residuals or mean-corrected returns from the estimated model are then used for the estimation of the log-volatility specifications. The univariate models of the two mean-corrected returns are

\[
\text{log-GARCH-1} : \quad \ln \sigma_t^2 = a_0 + a_1 \ln \epsilon_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2, \tag{20}
\]

\[
\text{log-GARCH-2} : \quad \ln \sigma_t^2 = a_0 + \sum_{i=1}^{7} a_i \ln \epsilon_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2, \tag{21}
\]
Fig. 2. Fitted standard deviations (SDs) of the univariate log-GARCH-1 and log-GARCH-4 models, and then nominal differences and ratios between the SDs (computed as log-GARCH-4 minus log-GARCH-1 and log-GARCH-4 over log-GARCH-1, respectively).

\[
\text{log-GARCH-3: } \ln \sigma^2_t = \alpha_0 + \sum_{i=1}^{7} \alpha_i \ln \epsilon^2_{t-i} + \beta_1 \ln \sigma^2_{t-1} + \sum_{l=1}^{6} \lambda_l x_l, \tag{22}
\]

\[
\text{log-GARCH-4: } \ln \sigma^2_{1,t} = \alpha_0 + \sum_{i=1}^{7} \alpha_{1,i} \ln \epsilon^2_{1,t-i} + \sum_{i=1}^{7} \alpha_{2,i} \ln \epsilon^2_{2,t-i} + \beta_1 \ln \sigma^2_{1,t-1} + \sum_{l=1}^{6} \lambda_l x_l, \tag{23}
\]

\[
\text{log-GARCH-5: } \ln \sigma^2_{2,t} = \alpha_0 + \sum_{i=1}^{7} \alpha_{1,i} \ln \epsilon^2_{1,t-i} + \sum_{i=1}^{7} \alpha_{2,i} \ln \epsilon^2_{2,t-i} + \sum_{l=1}^{6} \lambda_l x_l, \tag{24}
\]

where \( \epsilon_t \) is the mean-corrected return in question, and where \( x_{1t}, \ldots, x_6 \) are six day-of-the-week dummies for Tuesday to Sunday. In the last two specifications, \( \epsilon_{2,t} \) is the mean-corrected off-peak return when \( \epsilon_{1,t} \) is the mean-corrected on-peak return, and vice-versa, \( \epsilon_{2,t} \) is the mean-corrected on-peak return when \( \epsilon_{1,t} \) is the mean-corrected off-peak return. Of course, this means that the last two equations could be considered as an equation-by-equation estimation scheme similar to that of Francq and Zakoïan (2015), except that we do not estimate the time-varying correlations. The last specification, i.e. log-GARCH-5, actually refers to a more parsimonious version than the one displayed. The parsimonious specification is obtained by automated General-to-Specific (GETS) modelling starting from (24), see Sucarrat and Escribano (2012). Arguably, the most important specifications are log-GARCH-4 and log-GARCH-1. The former since it nests all the others, the latter for benchmarking.

The upper part of Table 3 contains the estimation results of the peak models (an * to the right of the standard errors means the t-ratio is greater than 2 in absolute value). The first striking characteristic is that volatility is much more volatile (i.e. less persistent) than is usually the case for financial returns. The ARCH(1) estimate is large and about 0.2 in all models—in daily financial returns it is typically about 0.05 (or lower), and the GARCH(1) estimate falls from about 0.7 in log-GARCH-1 to an insignificant 0 in log-GARCH-4. Moreover, several additional ARCH-lags and day-of-the-week dummies are significant in log-GARCH-4. In particular, the results show that the most precise peak return forecasts are produced on Fridays, whereas the most uncertain ones are produced on Mondays. Also, in addition to several significant own ARCH-lags, four off-peak ARCH-lags are significant. This means that there is a feedback effect from off-peak volatility. Altogether, daily intra-week dynamics and day-of-the-week effects account for all the variation in volatility, as there is little – if any – persistence.

The lower part of Table 3 contains the estimation results of the off-peak models. These are much more in line with what one usually finds in other financial returns. In log-GARCH-4 the ARCH(1) and GARCH(1) estimates are 0.092 and 0.845, respectively, which compares with 0.137 and 0.792 in log-GARCH-1. In other words, the inclusion of lags and day-of-the-week dummies do not affect these estimates very much. However, just as for peak volatility, several ARCH-lags and day-of-the-week dummies are significant. In particular, the most precise forecasts of off-peak return are produced on Fridays—just as in the peak case, whereas the most uncertain ones are produced on Sundays. Also, just as in the peak case, there is volatility-feedback, since several (three) peak lags are significant.

Fig. 2 contains the fitted standard deviations of the log-GARCH-1 and log-GARCH-4 models, their nominal difference and their ratio. The bottom graphs, in particular, show that they can produce fundamentally different volatility forecasts. Specifically, they show that the log-GARCH-1 underestimates volatility on average, and that the log-GARCH-4 models can
Table 3
Estimation results of the models (20)–(24).

<table>
<thead>
<tr>
<th>Peak specifications:</th>
<th>LogL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1: ( \ln \hat{\sigma}^2_{1,t} )</strong></td>
<td>-0.434 + 0.202 ( f^2_{t-1} ) + 0.639 ( \hat{\sigma}^2_{1,t-1} )</td>
</tr>
<tr>
<td><strong>2: ( \ln \hat{\sigma}^2_{2,t} )</strong></td>
<td>-0.976 + 0.232 ( f^2_{t-1} ) + 0.124 ( \ln \hat{\sigma}^2_{2,t-2} ) + 0.053 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.010 ( \ln \hat{\sigma}^2_{1,t-4} )</td>
</tr>
<tr>
<td><strong>3: ( \ln \hat{\sigma}^2_{3,t} )</strong></td>
<td>-0.127 + 0.228 ( f^2_{t-1} ) + 0.119 ( \ln \hat{\sigma}^2_{2,t-2} ) + 0.059 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.002 ( \ln \hat{\sigma}^2_{1,t-4} )</td>
</tr>
<tr>
<td><strong>4: ( \ln \hat{\sigma}^2_{4,t} )</strong></td>
<td>0.224 + 0.203 ( f^2_{t-1} ) + 0.103 ( \ln \hat{\sigma}^2_{2,t-2} ) - 0.041 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.014 ( \ln \hat{\sigma}^2_{1,t-4} ) + 0.046 ( \ln \hat{\sigma}^2_{1,t-5} )</td>
</tr>
<tr>
<td><strong>5: ( \ln \hat{\sigma}^2_{5,t} )</strong></td>
<td>-0.071 + 0.209 ( f^2_{t-1} ) + 0.120 ( \ln \hat{\sigma}^2_{2,t-2} ) + 0.066 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.093 ( \ln \hat{\sigma}^2_{1,t-4} )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Off-peak specifications:</th>
<th>LogL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1: ( \ln \hat{\sigma}^2_{2,t} )</strong></td>
<td>-0.070 + 0.137 ( \ln \hat{\sigma}^2_{2,t-1} ) - 0.792 ( \ln \hat{\sigma}^2_{1,t-1} )</td>
</tr>
<tr>
<td><strong>2: ( \ln \hat{\sigma}^2_{2,t} )</strong></td>
<td>-0.548 + 0.202 ( \ln \hat{\sigma}^2_{2,t-1} ) + 0.065 ( \ln \hat{\sigma}^2_{2,t-2} ) + 0.083 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.064 ( \ln \hat{\sigma}^2_{1,t-4} )</td>
</tr>
<tr>
<td><strong>3: ( \ln \hat{\sigma}^2_{2,t} )</strong></td>
<td>-0.148 + 0.202 ( \ln \hat{\sigma}^2_{2,t-1} ) + 0.106 ( \ln \hat{\sigma}^2_{2,t-2} ) + 0.091 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.068 ( \ln \hat{\sigma}^2_{1,t-4} )</td>
</tr>
<tr>
<td><strong>4: ( \ln \hat{\sigma}^2_{4,t} )</strong></td>
<td>-0.603 + 0.092 ( \ln \hat{\sigma}^2_{1,t-1} ) - 0.004 ( \ln \hat{\sigma}^2_{2,t-2} ) - 0.060 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.022 ( \ln \hat{\sigma}^2_{1,t-4} ) - 0.013 ( \ln \hat{\sigma}^2_{1,t-5} )</td>
</tr>
<tr>
<td><strong>5: ( \ln \hat{\sigma}^2_{5,t} )</strong></td>
<td>-0.047 + 0.087 ( \ln \hat{\sigma}^2_{1,t-1} ) + 0.018 ( \ln \hat{\sigma}^2_{1,t-2} ) + 0.179 ( \ln \hat{\sigma}^2_{1,t-3} ) - 0.078 ( \ln \hat{\sigma}^2_{1,t-4} ) + 0.056 ( \ln \hat{\sigma}^2_{1,t-5} )</td>
</tr>
</tbody>
</table>

LogL, Gaussian log-likelihood computed as \( \sum_{t=1}^{T} \ln f(e_t; \hat{\theta}) \), where \( f(e_t; \hat{\theta}) \) is the univariate normal density, \( e_t \) is the mean-corrected return and \( \hat{\theta} \) is the fitted standard deviation (\( T = 1586 \) is the number of observations), \( k \) the number of log-GARCH parameters (\( r \) not included). Estimation of the ARMA representation is with the LSE without mean-correction. An asterisk \(*\) to the right of the standard error means \(| t | > 2 \), where \( t \) is the t-ratio. Computations in R (R Core Team, 2014) with the igarch (Sucarrat, 2014) and AutoSEARCH (Sucarrat, 2012) packages.

produce fitted standard deviations that are more than twice as big on specific days. In other words, volatility may be seriously underestimated if lag and day-of-the-week effects are not accommodated.

5. Conclusions

A general and flexible framework for the estimation of and inference in univariate and multivariate Generalised log-ARCH-X (i.e. log-GARCH-X) models when the conditional density is unknown is proposed. Estimation is via the (V)ARMA-X representation, which induces a bias in the log-volatility intercept made up of a log-moment expression that depends on the conditional density. An estimator of the log-moment expression, together with its asymptotic variance, is derived under mild assumptions. Due to the structure of the problem, the bias-correction procedure is likely to also hold for univariate log-GARCH-X models, and — equation-by-equation — in multivariate log-GARCH-X models with time-varying correlations. An extensive number of simulations support the conjecture. Finally, an empirical application shows that the methods are particularly useful when the volatility dynamics are complex and possibly affected by many factors.

An early version of this paper (Sucarrat and Escribano, 2010) initiated the larger research agenda of which it is a part. Sucarrat and Escribano (2012) relies explicitly on the results of this paper, whereas (Bauwens and Sucarrat, 2010) is a precursor. These led to the development of the R (R Core Team, 2014) software packages AutoSEARCH (Sucarrat, 2012) and getts (Sucarrat, 2014) for automated General-to-Specific (Gets) modelling of log-ARCH-X models. An early critique of the log-ARCH class of models was that the log-ARCH terms in the log-volatility specification may not exist, since the errors
of a regression in empirical practice can be zero. A solution to this problem, however, is proposed in Sucarrat and Escribano (2013). This solution is only available when estimation is via the (V)ARMA representation. Francq and Sucarrat (2013) propose another ARMA-based QMLE for log-GARCH models (with the centred exponential chi-squared as instrumental density) that is asymptotically more efficient when the conditional error is normal or close to normal. Finally, Francq and Sucarrat (2015) prove the consistency and asymptotic normality of a least squares equation-by-equation estimator of a multivariate log-GARCH-X model with Dynamic Conditional Correlations by using the VARMA-X representation.

Acknowledgements

We are grateful to the Editor, two anonymous referees, Jonas Andersson, Luc Bauwens, Christian Francq, Andrew Harvey, Emma Iglesias, Sebastien Laurent, Enrique Sentana and seminar and conference participants at Nuffield College (Oxford University), Université de Lille, Universidad Carlos III de Madrid, BI Norwegian Business School (Oslo), IHS (Vienna), 2nd. Rimini Time Series Workshop (2013), 21st SNDE Symposium (Milan, 2013), ESEM 2011 (Oslo), Interdisciplinary Workshop in Louvain-la-Neuve 2011, CFE conference 2010 (London), FIBE 2010 (Bergen), IWAP 2010 (Madrid) and Foro de Finanzas 2010 (Elche) for useful comments, suggestions and questions. Support from the Ministerio Economía y Competitividad (Spain), grant MDM 2014-0431, funding from the 6th European Community Framework Programme, MICIN ECO2009-08308, and funding from The Bank of Spain Excellence Program are gratefully acknowledged.

Appendix A. Proof of Theorems 1 and 3

Since estimation is by OLS, a mean correction is irrelevant in A3. Let \( Y_t = \ln e_i^2 \), and let \( \phi := (\hat{\phi}_1, \ldots, \hat{\phi}_p) \) be the least squares estimator of \( \phi := (\phi_1, \ldots, \phi_p) \) based on mean corrected observations \( (Y_t - \bar{Y}_T)_{1 \leq t \leq T} \) where \( \bar{Y}_T = T^{-1} \sum_{t=1}^T Y_t \). Note that \( \sqrt{T}(\hat{\phi} - \phi) = \mathcal{O}_p(1) \) by standard theory (Brockwell and Davis, 2006). Denote \( \gamma := Ee^{u_i^2} \) and \( \hat{\gamma} = \frac{1}{T} \sum_{t=1}^T \exp(\hat{u}_t) \).

A3 implies \( \text{Var}(u_i^2) = E[(\ln z_i^2)^2] - E[\ln z_i^2]^2 < \infty \) as well as \( E \exp(u_i) = \frac{1}{1/(E[\exp(\ln z_i^2)])} < \infty \), and A4 implies that \( \text{Var}[\exp(u_i)] = (Ez_i^4 - 1)/(E[\exp(\ln z_i^2)])^2 < \infty \), and so \( E[\exp(u_i)] < \infty \). To prove Theorem 1, we show what we call Case (i): If \( \tilde{U}_T < \infty \) and \( E^{u_i^2} < \infty < \infty \), then \( \hat{\gamma} = \gamma + o_p(1) \). To prove Theorem 3 we prove Case (ii): If \( \tilde{U}_T < \infty \) and \( E^{u_i^2} < \infty < \infty \), then \( \sqrt{T}(\hat{\gamma} - \gamma) = T^{-1/2} \sum_{t=1}^T (\hat{u}_t - \gamma) + o_p(1) \). Both cases use the following expansions. Let \( \delta_t := u_t - \bar{u}_T \), when \( t \leq p \). For \( t \geq p + 1 \), a Taylor expansion shows that

\[
e^{u_t + \delta_t} = e^{u_t} + \delta_t e^{u_t} + \delta_t^2 \int_0^1 (1-x)e^{u_t + \delta_t x} \, dx = e^{u_t} + \delta_t e^{u_t} + \delta_t^2 \int_0^1 (1-x)e^{u_t} \, dx.
\]

(A.1)

We follow the main argument in Theorem 1 of Lee (1997) to bound \( \delta_t \) uniformly in \( t \). Let \( \mu = EY_0 = \phi_0/(1 - \sum_{i=1}^p \phi_i) \), so that \( u_t = Y_t - \mu - \sum_{i=1}^p \phi_i (Y_{t-i} - \mu) \). The definition of \( \bar{u}_t \), as well as addition and subtraction show that for \( t \geq p + 1 \), we have that \( \bar{u}_t = u_t - (\hat{\phi} - \phi)'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - T^{-1} \sum_{i=1}^p (Y_{t-i} - \mu)(1 - \sum_{i=1}^p \phi_i) \) so that \( \delta_t = -(\hat{\phi} - \phi)'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - T^{-1} \sum_{i=1}^p (Y_{t-i} - \mu)(1 - \sum_{i=1}^p \phi_i) \) as in Lee (1997). Lee (1997) shows \( T^{-1/2} \sum_{i=1}^p (Y_{t-i} - \mu) = \sqrt{T} \bar{u}_T (1 - \sum_{i=1}^p \phi_i)^{-1} \) with \( \xi_t = o_p(1) \), see the proof of Theorem 1 in Lee (1997) immediately before his equation 2.6. For \( t \geq p + 1 \) we therefore get that

\[
\delta_t = -(\hat{\phi} - \phi)'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - \bar{u}_T + R_T
\]

(A.2)

where \( R_T = o_p(T^{-1/2}) \) does not depend on \( t \). To see this, note that \( \sqrt{T} R_T = (\sqrt{T} \bar{u}_T) - (\sqrt{T} \bar{u}_T) - (1 - \sum_{i=1}^p \phi_i)^{-1}(1 - \sum_{i=1}^p \hat{\phi}_i) + (1 - \sum_{i=1}^p \phi_i) + o_p(1) \) we have \( (1 - \sum_{i=1}^p \phi_i) + o_p(1) \). Hence, \( \sqrt{T} R_T = (\sqrt{T} \bar{u}_T) - (\sqrt{T} \bar{u}_T) = (1 - \sum_{i=1}^p \phi_i) + o_p(1) \).

We have \( \sup_{1 \leq t \leq T} |\hat{\phi} - \phi)'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) + |\bar{u}_T| + o_p(1) \) since \( (Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) \) is a strictly stationary linear process with exponentially decreasing coefficients, \( E^{u^2}_T < \infty \) implies \( E^{u^2}_T < \infty \) for any \( \kappa > 0 \). Suppose \( 0 \leq \kappa < 1/2 \). We have \( T^{-1/2} \sup_{1 \leq t \leq T} |Y_t - \mu| = o_p(1) \) if \( E^{u^2}_T < \infty \) as well as \( \leq E^{u^2}_T < \infty \), see e.g. Lemma 12.4 of Ibragimov and Phillips (2008). For case (i), we know that \( E^{u^2}_T < \infty < \infty \), i.e. \( \alpha = 1/4 \). For both cases, we have \( \sup_{1 \leq t \leq T} |\hat{\phi}_t - \phi)'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) + o_p(1) \) since \( \sqrt{T}(\hat{\phi} - \phi) = \mathcal{O}_p(1) \). Hence if \( E^{u^2}_T < \infty < \infty \), i.e. case (i), we conclude that \( M_T = \sup_{1 \leq t \leq T} |\hat{\phi} - \phi)'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) + |\bar{u}_T| + o_p(1) \) since \( u_T = o_p(1) \) by the LLN (Law of Large Numbers). If \( E^{u^2}_T < \infty \) we get \( T^{1/2} M_T = T^{1/2} \sup_{1 \leq t \leq T} |\hat{\phi} - \phi)'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) + T^{-1/2} |\bar{u}_T| + o_p(1) \). Therefore \( T^{1/2} \bar{u}_T = o_p(1) \) by the CLT, we see that \( T^{-1/4} |\bar{u}_T| = o_p(1) \). For \( \alpha = 1/4 \) we conclude that \( T^{1/4} M_T = o_p(1) \).
We now show consistency, i.e. case (i). Eq. (A.1) shows that \( \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} = \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} + \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} + \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} \). Clearly, \( \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} = f^{T}_{0}(1-x) e^{\phi \delta t} dx \). Hence, the triangle inequality implies that \( |\gamma - \gamma| \leq \frac{1}{\delta T} \sum_{t=0}^{T} |\delta_{t}| e^{\phi \delta t} + \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} \). Using \( |\delta_{t}| \leq M_{T} \) we get that \( |\gamma - \gamma| \leq M_{T} \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} + \frac{1}{\delta T} \sum_{t=0}^{T} \epsilon_{t}^{\phi} + \sigma(1) \), which is \( \sigma(1) \) because \( M_{T} = \sigma(1) \) and \( T^{-1} \sum_{t=0}^{T} \epsilon_{t}^{\phi} = E^{\sigma(1)} + \sigma(1) = \sigma(1) \).

Let us now show asymptotic Normality, i.e. case (ii). From Eq. (A.1), we see that \( \sqrt{T} (\gamma - \gamma) = \frac{1}{\sqrt{T}} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) + \frac{1}{\sqrt{T}} \sum_{t=0}^{T} \delta_{t} e^{\phi \delta t} + \frac{1}{\sqrt{T}} \sum_{t=0}^{T} \epsilon_{t}^{\phi} \). The last sum is \( \sigma(1) \). To see this, we again use \( f^{T}_{0}(1-x) e^{\phi \delta t} dx \). Combined with \( T^{1/4} M_{T} = \sigma(1) \) and see that \( \frac{1}{\sqrt{T}} \sum_{t=0}^{T} \epsilon_{t}^{\phi} \). which is \( \sigma(1) \) because \( (T^{1/4} M_{T})^{2} = \sigma(1)^{2} = \sigma(1) \) by continuity, that \( M_{T} = \sigma(1) \) and \( T^{-1} \sum_{t=0}^{T} \epsilon_{t}^{\phi} = \sigma(1) \).

We have therefore shown that \( \sqrt{T} (\gamma - \gamma) = T^{-1/2} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) + T^{-1/2} \sum_{t=0}^{T} \delta_{t} e^{\phi \delta t} + \sigma(1) = T^{-1/2} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) + T^{-1/2} \sum_{t=0}^{T} \delta_{t} e^{\phi \delta t} + \sigma(1) \). For the sum with \( \delta_{t} \), we apply Eq. (A.2), and get \( T^{-1/2} \sum_{t=0}^{T} \delta_{t} e^{\phi \delta t} = -(\phi - \phi') T^{-1/2} \sum_{t=0}^{T} (Y_{t} - 1 - \mu) e^{\phi \delta t} - \hat{T} (T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t}) + R_{T} T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t} \). Since \( Y_{t} \) and \( u_{t} \) are independent for \( j \geq 0 \), we get \( T^{-1} \sum_{t=0}^{T} (Y_{t} - 1 - \mu) e^{\phi \delta t} = E(Y_{t} - 1 - \mu) e^{\phi \delta t} = \sigma(1) \). Hence, \( (\phi - \phi') T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t} = -(\phi - \phi') \sigma(1) \). \( \sqrt{T} (\hat{T} - \gamma) \) implies that \( R_{T} T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t} = \sigma(1) \). We further have that \( \sqrt{T} (\hat{T} - \gamma) = \sigma(1) \).

In conclusion we get \( \sqrt{T} (\gamma - \gamma) = T^{-1/2} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) - \sqrt{T} \hat{T} e^{\phi \delta t} + \sigma(1) \). We now complete the proof by showing that \( T^{-1/2} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) \) and \( T^{-1/2} \sum_{t=0}^{T} \epsilon_{t}^{\phi} \) fulfills exactly the same expansion as we found for \( \sqrt{T} (\gamma - \gamma) \). We have \( T^{-1/2} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) = T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t} - \sqrt{T} \hat{T} e^{\phi \delta t} = e^{-\phi \delta t} (T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t}) - \sqrt{T} \hat{T} e^{\phi \delta t} = e^{-\phi \delta t} (T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t} - \sqrt{T} \hat{T} e^{\phi \delta t}) + e^{\phi \delta t} (T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t} - \sqrt{T} \hat{T} e^{\phi \delta t}) + e^{-\phi \delta t} (T^{-1/2} \sum_{t=0}^{T} e^{\phi \delta t} - \sqrt{T} \hat{T} e^{\phi \delta t}) + (1 + \sigma(1)) \). Since \( E \sigma(1) \), the CLT implies that \( T^{-1/2} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) = \sigma(1) \) and hence \( E \sigma(1) \). \( T^{-1/2} \sum_{t=0}^{T} (e^{\phi \delta t} - \gamma) = (1 + \sigma(1)) \). The delta method gives \( (e^{-\phi \delta t} - 1) \sqrt{T} \hat{T} e^{\phi \delta t} = -\sqrt{T} \hat{T} e^{\phi \delta t} + \sigma(1) \) and the conclusion follows.

Appendix B. Proof of Theorem 2

Let \( \tau_{T} = -\ln[T^{-1} \sum_{t=1}^{T} e^{u_{t} - u_{t-1}}] \). Assumption A4 and the smoothness of the logarithm function imply that \( \tau_{T} \) and \( \tau_{T} \) have the same behaviour up to \( \sigma(1) \). Denoting \( \tau = E \ln(z_{T}^{2}) = -E \sigma(1) \), this means \( \sqrt{T} (\tau_{T} - \tau) \) is asymptotically normal. We have that \( \tau_{T} = \ln T - \sum_{t=1}^{T} e^{u_{t} - u_{t-1}} \). By the smoothness of \( f \), the delta method implies that \( \sqrt{T} (\tau_{T} - \tau) \) is asymptotically normal. We have that \( \ln T - \sum_{t=1}^{T} e^{u_{t} - u_{t-1}} \rightarrow \ln T \sim N(0,0) \), Cov \( e^{u_{t} - u_{t-1}} \). Hence, \( \sqrt{T} (\tau_{T} - \tau) \) is asymptotically normal with variance equal to

\[
\tau^{2} = (f' (E e^{u_{t}}))^{2} + Var X + Var Y + 2f' (E e^{u_{t}}) \frac{Var [exp(u_{t})]}{[E exp(u_{t})]^{2}} + Var (u_{t}) - 2 \frac{E E [u_{t} exp(u_{t})]}{E exp(u_{t})}.
\]

Using that \( Var [exp(u_{t})] = (E^{2} - 1) / \{ exp [E \ln(z_{T}^{2})] \}^{2} \), \( E exp(u_{t}) = 1 / \{ exp [E \ln(z_{T}^{2})] \} \), \( Var (u_{t}) = E \{ \ln(z_{T}^{2}) \}^{2} - E \{ \ln(z_{T}^{2}) \} \) / \{ exp [E \ln(z_{T}^{2})] \}, \( E e^{u_{t} exp(u_{t})} \), we obtain

\[
\tau^{2} = E \{ \ln(z_{T}^{2}) \}^{2} - E \{ \ln(z_{T}^{2}) \}^{2} + (E^{2} - 1) - 2 E [E \ln(z_{T}^{2})^{2}] + 2 E \ln(z_{T}^{2}).
\]

From A4 we have that \( E(z_{T}^{2}) < \infty \) and \( E \{ \ln(z_{T}^{2}) \}^{2} < \infty \), and the Cauchy–Schwarz inequality implies that \( E \{ \ln(z_{T}^{2})^{2} \} \leq (E \{ \ln(z_{T}^{2}) \}^{2}) (E^{2}) \). So \( \tau^{2} \) is finite. Finally, the expression simplifies to \( \tau^{2} = Var (z_{T}^{2} - 1) \).
Appendix C. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.csda.2015.12.005.

References


Francq, C., Sucarrat, G., 2013. An exponential chi-squared QMLE for Log-GARCH models via the ARMA representation, URL: http://mpre.ub.uni-muenchen.de/51783/.


Sucarrat, G., 2014. gets: General-to-Specific (GETS) Model Selection, r package version 0.2. URL: http://cran.r-project.org/web/packages/gets/.


