Estimation and Inference in Univariate and Multivariate Log-GARCH-X Models When the Conditional Density is Unknown

Genaro Sucarrat[a,*], Steffen Grønneberg[a], Alvaro Escribano[b]

[a]Department of Economics, BI Norwegian Business School, Oslo, Norway
[b]Department of Economics, Universidad Carlos III de Madrid, Madrid, Spain

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.

Keywords: Log-GARCH-X, ARMA-X, multivariate log-GARCH-X, VARMA-X, volatility

1. Introduction

The Autoregressive Conditional Heteroscedasticity (ARCH) class of models due to [1] 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. [2], [3]), the evolution of temperature data (e.g. [4]) and – more generally – positively valued variables, i.e. so-called Multiplicative Error Models (MEMs), see [5] 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

*Correspondence to: Nydalsveien 37, 0484 Oslo, Norway. Tel.: +47 46 41 07 79. Fax: +47 23 26 47 88.
URL: www.sucarrat.net (Genaro Sucarrat)

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 forecast failure subsequent to jumps and outliers. The log-GARCH class was independently proposed by [6], [7] and [8]. [9] 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 [10] 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 ([11]) and the Beta-t-EGARCH model ([12]). The former has proved to be much 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 [13], and [14]. 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. [15]. Only recently was consistency and asymptotic normality proved (for the univariate EGARCH(1,1) only) under the complicated condition of continuous invertibility, see [16]. The log-GARCH model is much more tractable. [17] 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 [18]). Moreover, the estimator of [17] cannot handle zero-errors or missing values in the manner suggested by [10]. Finally, [18] 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 [19] by [20], for an ARMA-GARCH with constant conditional correlations (CCCs) by [21], for a factor GARCH model by [22], for a multivariate GARCH with CCCs by [23], and for a multivariate GARCH with stochastic correlations by [24] under the assumption that the system is estimable equation-by-equation. For exponential ARCH models there are

no multivariate results. [25] 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 [26], and [27]. 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

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

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2, \quad t \in \mathbb{Z}, \tag{2}$$

where p is the ARCH order and q is the GARCH order. In finance, ϵ_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 ϵ_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 σ_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 dth power log-GARCH(p, q) model can be written as

$$\ln \sigma_t^d = \alpha_{0,d} + \sum_{i=1}^{p} \alpha_i \ln |\epsilon_{t-i}|^d + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^d, \quad d > 0, \tag{3}$$

where $\alpha_{0,d} = \alpha_0 d/2$. This means a complete analysis of the dth 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

$$\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + u_t, \quad u_t = \ln z_t^2 - E(\ln z_t^2), \tag{4}$$

where

$$\phi_0 = \alpha_0 + (1 - \sum_{j=1}^{q} \beta_j) \cdot E(\ln z_t^2), \quad \phi_i = \alpha_i + \beta_i \quad \text{and} \quad \theta_j = -\beta_j. \tag{5}$$

Consistent and asymptotically normal estimates of all the ARMA parameters – and hence all the log-GARCH parameters except the log-volatility intercept α_0 – are thus readily obtained via usual ARMA estimation methods subject to appropriate assumptions, see e.g. [28]. To estimate α_0, the most common solutions have been to either impose restrictive assumptions regarding the distribution of z_t (say, normality, see e.g. [26]), or to use an *ex post* scale-adjustment (see e.g. [29], and [30]). 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 α_0 can also be estimated consistently.

2.2. On consistency

The scale-adjustment employed by [29], and [30], is essentially a smearing estimate (more on this below). Consider writing (1) as

$$\epsilon_t = \sigma_t^* z_t^*, \quad z_t^* \sim IID(0, \sigma_{z^*}^2),$$

where σ_t^* 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 $\sigma_t = \sigma_t^* \sigma_{z^*}$ 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 $\ln \widehat{\sigma}_t^{*2}$ denoting the fitted value of the ARMA specification such that $\widehat{\sigma}_t^* = \exp(\ln \widehat{\sigma}_t^*)$, and with the ARMA residual defined as $\widehat{u}_t = \ln \epsilon_t^2 - \ln \widehat{\sigma}_t^{*2}$. 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 $\widehat{\sigma}_t^*$ by multiplying it with an estimate of σ_{z^*}, say, the sample standard deviation of the standardised residuals $\widehat{z}_t^*$. 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:

$$\widehat{\sigma}_{z^*}^2 = \frac{1}{T-1}\sum_{t=1}^{T}(\widehat{z}_t^* - \overline{\widehat{z}}_t^*)^2 \quad \approx \quad \frac{1}{T}\sum_{t=1}^{T}(\widehat{z}_t^*)^2 = \frac{1}{T}\sum_{t=1}^{T}\exp(\widehat{u}_t).$$

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

$$-\ln\left[\frac{1}{T}\sum_{t=1}^{T}\exp(\widehat{u}_t)\right] \tag{6}$$

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. $T^{-1}\sum_{t=T}\exp(\widehat{u}_t)$, is well-known as the "smearing estimate", see [31]. 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 [31], however, is for static models. In dynamic models, e.g. when the $\widehat{u}_t$'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, [32] 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 $\widehat{u}_t, t = 1, \ldots, T$ denote the ARMA-residuals resulting from estimating the ARMA representation (4). Then:

$$\frac{1}{T}\sum_{t=1}^{T}\exp(\widehat{u}_t) - \frac{1}{T}\sum_{t=1}^{T}\exp(u_t) = o_P(1). \tag{7}$$

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

Proof. See Appendix A. □

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 $\widehat{\phi}_0 = (1 - \sum_{i=1}^{p}\widehat{\phi}_i)\cdot T^{-1}\sum_{t=1}^{T}\ln\epsilon_t^2$ provides a consistent estimate of ϕ_0 under the assumptions of the Theorem. Strict stationarity of $\ln\epsilon_t^2$ 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

$$\widehat{\tau}_T = -\ln\left[\frac{1}{T}\sum_{t=1}^{T}\exp(\widehat{u}_t - \overline{\widehat{u}}_T)\right], \tag{8}$$

where $\overline{\widehat{u}}_T$ is the empirical mean of the ARMA-residuals. The mean-correction term $\overline{\widehat{u}}_T$ is needed, since asymptotic normality may not be achieved without it. See e.g. the related discussion in [33], 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 $\overline{\widehat{u}}_T$ is zero by construction, and so (8) equals (6). The following two assumptions will give us asymptotic normality of (8):

A3: Let $\{\widehat{u}_t\}_{t=1}^T$ denote the ARMA-residuals resulting from estimating the ARMA representation (4). Denoting $\overline{\widehat{u}}_T$ and $\overline{u}_T$ as the averages of $\widehat{u}_t$ and u_t, respectively:

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

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

$$\sqrt{T}\left[\widehat{\tau}_T-E(\ln z_t^2)\right]\xrightarrow{D}N(0,\zeta^2),\quad\text{where}\quad\zeta^2=\mathrm{Var}\,(z_t^2-\ln z_t^2). \tag{9}$$

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 $E(e^{|u_t|})<\infty$. 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

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2 + g(\lambda, x_t), \tag{10}$$

where g is a linear or nonlinear function of the conditioning variables x_t and a parameter vector λ. 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

$$\ln \epsilon_t^2 = \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + g(\lambda, x_t) + u_t, \quad u_t = \ln z_t^2 - E(\ln z_t^2), \tag{11}$$

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 [32], and Chapter 4 in [34].

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, $g(\lambda, x_t) = \lambda_1 I_{\{z_{t-1}<0\}}$, 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

$$\begin{aligned}
\epsilon_t &\sim ID(0, H_t), \quad t \in \mathbb{Z}, && (12)\\
D_t^2 &= \operatorname{diag}\left\{\sigma_{m,t}^2\right\}, \quad m = 1, \ldots, M, && (13)\\
z_t &= D_t^{-1}\epsilon_t, \quad \forall m : z_{m,t} \sim IID(0,1), \quad P(z_t = 0) = 0, && (14)\\
\ln \sigma_t^2 &= \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2, \quad p \geq q, && (15)
\end{aligned}$$

Table 1: Finite sample properties of the LSE via the ARMA representation for a log-GARCH(1,1) with leverage

DGP	T	$m(\widehat{\alpha}_0)$	$se(\widehat{\alpha}_0)$	$m(\widehat{\alpha}_1)$	$se(\widehat{\alpha}_1)$	$m(\widehat{\beta}_1)$	$se(\widehat{\beta}_1)$	$m(\widehat{\lambda}_1)$	$se(\widehat{\lambda}_1)$	$m(\widehat{\tau})$	$se(\widehat{\tau})$	$ase(\widehat{\tau})$
A:	500	-0.056	0.146	0.098	0.034	0.765	0.124	0.001	0.138	-1.272	0.078	0.067
	1000	-0.021	0.079	0.099	0.023	0.785	0.065	-0.011	0.088	-1.271	0.054	0.054
	2000	-0.011	0.048	0.099	0.016	0.795	0.041	-0.008	0.063	-1.270	0.039	0.038
B:	500	-0.101	0.271	0.050	0.032	0.830	0.198	-0.019	0.131	-1.271	0.078	0.067
	1000	-0.035	0.101	0.050	0.019	0.877	0.073	-0.016	0.079	-1.273	0.054	0.054
	2000	-0.013	0.045	0.050	0.012	0.891	0.039	-0.021	0.044	-1.270	0.038	0.038
C:	500	-0.064	0.157	0.099	0.035	0.756	0.132	-0.013	0.142	-1.384	0.086	0.085
	1000	-0.023	0.079	0.100	0.023	0.784	0.064	-0.010	0.094	-1.392	0.060	0.061
	2000	-0.009	0.050	0.100	0.016	0.793	0.039	-0.010	0.065	-1.391	0.043	0.043
D:	500	-0.099	0.251	0.049	0.030	0.841	0.173	-0.020	0.137	-1.394	0.092	0.085
	1000	-0.038	0.119	0.050	0.018	0.874	0.090	-0.027	0.078	-1.392	0.061	0.061
	2000	-0.016	0.051	0.050	0.013	0.889	0.040	-0.022	0.049	-1.390	0.045	0.043

The estimated model is $\ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln\epsilon_{t-1}^2 + \beta_1 \ln\sigma_{t-1}^2 + \lambda_1 I_{\{z_{t-1}<0\}}$. DGP A: $(\alpha_0, \alpha_1, \beta_1, \lambda_1, \tau) = (0, 0.1, 0.8, -0.01, -1.27)$ with $z_t \sim N(0,1)$. DGP B: $(\alpha_0, \alpha_1, \beta_1, \lambda_1, \tau) = (0, 0.05, 0.9, -0.02, -1.27)$ with $z_t \sim N(0,1)$. DGP C: $(\alpha_0, \alpha_1, \beta_1, \lambda_1, \tau) = (0, 0.1, 0.8, -0.01, -1.39)$ with $z_t \sim$ *standardised* $t(10)$. DGP D: $(\alpha_0, \alpha_1, \beta_1, \lambda_1, \tau) = (0, 0.05, 0.9, -0.02, -1.39)$ with $z_t \sim$ *standardised* $t(10)$. $m(x)$, sample mean of the estimate x. $se(x)$, sample standard deviation (division by R instead of $R-1$, where $R = 1000$ is the number of replications). $ase(\widehat{\tau})$, asymptotic standard error of $\widehat{\tau}$ (computed as $\sqrt{\zeta^2}/\sqrt{T}$, where ζ^2 is given in Theorem 2). Computations in *R* ([35]) with the `lgarch` package version 0.2 ([36]).

where ϵ_t, σ_t^2 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 $\alpha_0 = (\alpha_{1.0}, \ldots, \alpha_{M.0})'$,

$$\alpha_i = \begin{pmatrix} \alpha_{11.i} & \cdots & \alpha_{1M.i} \\ \vdots & \ddots & \vdots \\ \alpha_{M1.i} & \cdots & \alpha_{MM.i} \end{pmatrix} \quad \text{and} \quad \beta_j = \begin{pmatrix} \beta_{11.j} & \cdots & \beta_{1M.j} \\ \vdots & \ddots & \vdots \\ \beta_{M1.j} & \cdots & \beta_{MM.j} \end{pmatrix}, \tag{16}$$

where $'$ is the transpose operator. Equation (12) means that ϵ_t is independent with mean zero and a time-varying conditional covariance matrix H_t. The IID assumption in equation (14) states that each marginal series $\{z_{m,t}\}$ is $IID(0,1)$. Marginal identicality 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 $z_t \sim ID(0, R_t)$, 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 ID, but not necessarily IID, even though each marginal series $\{z_{mt}\}$ 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 ϵ_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

$$\ln \epsilon_t^2 \quad = \quad \phi_0 + \sum_{i=1}^{p} \phi_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \theta_j u_{t-j} + u_t, \quad u_t = \ln z_t^2 - E(\ln z_t^2), \tag{17}$$

where

$$\phi_0 = \alpha_0 + (I_M - \sum_{j=1}^{q} \beta_j) \cdot E(\ln z_t^2), \quad \phi_i = \alpha_i + \beta_i, \quad \text{and} \quad \theta_j = -\beta_j. \tag{18}$$

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. [37]. Accordingly, consistent estimation and asymptotically normal inference regarding all the log-GARCH coefficients – apart from the log-volatility intercept α_0 – are available as well. In order to obtain a consistent estimate of α_0, then an estimate of the $M \times 1$ vector $E(\ln z_t^2)$ is needed. Since the process $\{u_{m,t}\}$ 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 $E(\ln z_t^2)$.

In the case where vector z_t is only ID, which is implied by time-varying correlations, then vector u_t is only ID as well. This corresponds to a VARMA model with heteroscedastic error u_t. Fewer QML results are available in this case, e.g. [38]. However, in the special case where the β_j matrices are diagonal, then the M-dimensional VARMA model can be estimated equation-by-equation (along the lines of [24]) by univariate ARMA-X methods, since – equation-by-equation – each error term $u_{m,t}$ is IID. Next, equation-by-equation application of (6) provides an estimate of each element in $E(\ln z_t^2)$, and hence of the log-volatility intercept α_0. Table 2 contains simulation results of the case where the correlations are governed by the Dynamic Conditional Correlations (DCC) model of [39]. The finite sample bias tends to zero as T increases, and the last two columns show that the empirical sample standard errors coincide with their asymptotic counterparts as implied by (9). Additional simulations are available from the Webpage of the first author (available as `http://www.sucarrat.net/research/lgarchxsims.pdf`).

3.2. Multivariate log-GARCH-X

Just as in the univariate case, the multivariate log-GARCH model permits additional covariates or conditioning variables in each of the M equations. Specifically, write the multivariate log-GARCH-X specification as

$$\ln \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \ln \epsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \ln \sigma_{t-j}^2 + \lambda x_t, \tag{19}$$

Table 2: Finite sample properties of equation-by-equation Gaussian QML of a 2-dimensional log-GARCH(1,1) w/diagonal matrix β_1 when the correlations follow the DCC of [39]

DGP	T	$m(\widehat{\alpha}_{10})$	$m(\widehat{\alpha}_{20})$	$m(\widehat{\alpha}_{11})$	$m(\widehat{\alpha}_{21})$	$m(\widehat{\alpha}_{12})$	$m(\widehat{\alpha}_{22})$	$m(\widehat{\beta}_{11})$	$m(\widehat{\beta}_{22})$	$m(\widehat{\tau}_1)$	$se(\widehat{\tau}_1)$	$m(\widehat{\tau}_2)$	$se(\widehat{\tau}_2)$	$ase(\widehat{\tau})$
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	0.054
	5000	-0.013	-0.042	0.049	0.100	0.100	0.049	0.900	0.697	-1.271	0.024	-1.270	0.023	0.024
	10000	-0.005	-0.023	0.049	0.100	0.100	0.050	0.900	0.698	-1.271	0.017	-1.270	0.017	0.017
B:	1000	-0.029	-0.026	0.098	0.053	0.053	0.097	0.791	0.792	-1.270	0.055	-1.268	0.053	0.054
	5000	-0.005	-0.004	0.100	0.051	0.050	0.099	0.799	0.799	-1.270	0.024	-1.270	0.024	0.024
	10000	-0.003	-0.002	0.100	0.050	0.050	0.100	0.799	0.799	-1.269	0.017	-1.271	0.017	0.017

The estimated model is $\ln \sigma_t^2 = \alpha_0 + \alpha_1 \ln \epsilon_{t-1}^2 + \beta_1 \ln \sigma_{t-1}^2$, where $\alpha_0 = (\alpha_{10}, \alpha_{20})'$, $\alpha_1 = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$ and $\beta_1 = \text{diag}(\beta_{11}, \beta_{22})$. The standardised errors $(z_{1t}, z_{2t})'$ are governed by the DCC of [39]: $(z_{1t}, z_{2t})' \sim N(0, \Sigma_t)$, $\Sigma_t = \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix}$, $\rho_t = q_{12,t}/\sqrt{q_{1,t}q_{2,t}}$, $q_{12,t} = \overline{\rho} + a(z_{1,t-1}z_{2,t-1} - \overline{\rho}) + b(q_{12,t} - \overline{\rho})$, $q_{1,t} = 1 + a(z_{1,t-1}^2 - 1) + b(q_{1,t} - 1)$, $q_{2,t} = 1 + a(z_{2,t-1}^2 - 1) + b(q_{2,t} - 1)$ with $a = 0.05$ and $b = 0.9$. Estimation proceeds in three steps (the DCC is not estimated). First, a univariate ARMA-X specification is fitted to each of the two equations with the Gaussian QMLE. Second, the ARMA-X residuals $\widehat{u}_{1t}$ and $\widehat{u}_{2t}$, respectively, are used equation-by-equation to estimate τ_1 and τ_2, respectively, with formula (8). Finally, the ARMA-X estimates and $\widehat{\tau}_1$ and $\widehat{\tau}_2$ are combined using the relationships in (18) to obtain the log-GARCH estimates. $m(x)$, sample mean of the estimate x. $se(x)$, sample standard deviation (division by R instead of $R-1$, where $R = 1000$ is the number of replications). $ase(x)$, asymptotic standard error of x (computed as $\sqrt{av(x)}/\sqrt{T}$, where $av(\widehat{\tau}_1) = av(\widehat{\tau}_2) = \zeta^2$, see Theorem 2). In DGP A: $\alpha_1 = c(0,0)'$, $\alpha_1 = \begin{pmatrix} 0.05 & 0.10 \\ 0.10 & 0.05 \end{pmatrix}$, $\beta_1 = \text{diag}(0.90, 0.70)$ and $\overline{\rho} = -0.2$. In DGP B: $\alpha_1 = c(0,0)'$, $\alpha_1 = \begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.10 \end{pmatrix}$, $\beta_1 = \text{diag}(0.80, 0.80)$ and $\overline{\rho} = 0.4$. Computations in R ([35]) with the `lgarch` package version 0.2 ([36]).

where x_t is an $L \times 1$ vector of covariates, and where λ is an $M \times L$ matrix. For notational economy, the covariates x_t enter linearly, but in principle they can enter non-linearly as 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

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

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 λx_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 α_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 α_0. Also here it is useful to distinguish between 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. [34]. If correlations are time-varying, and if the matrices β_j are diagonal, then each equation can be estimated separately in terms of their ARMA-X representations, see [32].

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. [3], [2], and [40] 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 kw/h) 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 Figure 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. $S_{1,t} = (S_{t(8am)} + \cdots + S_{t(9pm)})/14$, 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. $S_{2,t} = (S_{t(0am)} + \cdots + S_{t(7am)} + S_{t(10pm)} + S_{t(11pm)})/10$. Note that $S_{t(8am)}$ should be interpreted as the electricity price from 8am to 9am, $S_{t(9am)}$ should be interpreted as the electricity price from 9am to 10am, and so on. Graphs of $S_{1,t}, S_{2,t}$ and their log-returns ($r_t = \Delta \ln S_t$) are contained in Figure 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

$$\text{log-GARCH-1:}\quad \ln\sigma_t^2 = \alpha_0 + \alpha_1 \ln\epsilon_{t-1}^2 + \beta_1 \ln\sigma_{t-1}^2, \tag{20}$$

$$\text{log-GARCH-2:}\quad \ln\sigma_t^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_i \ln\epsilon_{t-i}^2 + \beta_1 \ln\sigma_{t-1}^2, \tag{21}$$

$$\text{log-GARCH-3:}\quad \ln\sigma_t^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_i \ln\epsilon_{t-i}^2 + \beta_1 \ln\sigma_{t-1}^2 + \sum_{l=1}^{6} \lambda_l x_{lt}, \tag{22}$$

$$\text{log-GARCH-4:}\quad \ln\sigma_{1t}^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_{1.i} \ln\epsilon_{1,t-i}^2 + \sum_{i=1}^{7} \alpha_{2.i} \ln\epsilon_{2,t-i}^2 + \beta_1 \ln\sigma_{1,t-1}^2 + \sum_{l=1}^{6} \lambda_l x_{lt}, \tag{23}$$

$$\text{log-GARCH-5:}\quad \ln\sigma_{1t}^2 = \alpha_0 + \sum_{i=1}^{7} \alpha_{1.i} \ln\epsilon_{1,t-i}^2 + \sum_{i=1}^{7} \alpha_{2.i} \ln\epsilon_{2,t-i}^2 + \sum_{l=1}^{6} \lambda_l x_{lt}, \tag{24}$$

where ϵ_t is the mean-corrected return in question, and where $x_{1t}, \ldots, x_{6t}$ 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 [24], 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 [30]. 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 error 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

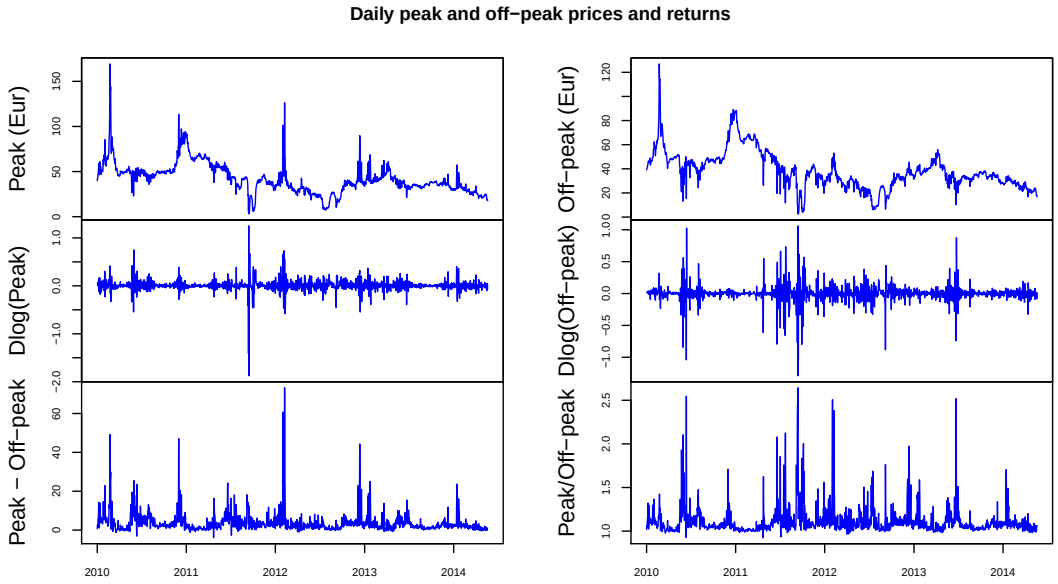

Figure 1: Daily peak and off-peak spot electricity prices (and their nominal and relative differences) in Euros per Mw/h, and log-returns for the Oslo area in Norway, 1 January 2010 - 20 May 2014 (1601 observations before lag-adjustments)

estimates very much. However, just as for peak volatiliy, 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.

Figure 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 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

Table 3: Estimation results of the models (20)-(24)

Peak specifications:		*LogL*
1: $\ln\widehat{\sigma}^2_{1,t} =$	$-0.434 + \underset{(0.008)^*}{0.202}\ln\epsilon^2_{1,t-1} + \underset{(0.018)^*}{0.639}\ln\widehat{\sigma}^2_{1,t-1}$	1890.3 [k=3]
2: $\ln\widehat{\sigma}^2_{1,t} =$	$-0.976 + \underset{(0.007)^*}{0.232}\ln\epsilon^2_{1,t-1} + \underset{(0.016)^*}{0.124}\ln\epsilon^2_{1,t-2} + \underset{(0.010)^*}{0.053}\ln\epsilon^2_{1,t-3} - \underset{(0.008)}{0.010}\ln\epsilon^2_{1,t-4}$ $+ \underset{(0.008)^*}{0.063}\ln\epsilon^2_{1,t-5} + \underset{(0.009)^*}{0.029}\ln\epsilon^2_{1,t-6} + \underset{(0.008)^*}{0.121}\ln\epsilon^2_{1,t-7} - \underset{(0.060)}{0.039}\ln\widehat{\sigma}^2_{1,t-1}$	1841.9 [k=9]
3: $\ln\widehat{\sigma}^2_{1,t} =$	$-0.127 + \underset{(0.007)^*}{0.228}\ln\epsilon^2_{1,t-1} + \underset{(0.021)^*}{0.119}\ln\epsilon^2_{1,t-2} + \underset{(0.013)^*}{0.059}\ln\epsilon^2_{1,t-3} - \underset{(0.009)}{0.002}\ln\epsilon^2_{1,t-4}$ $+ \underset{(0.008)^*}{0.064}\ln\epsilon^2_{1,t-5} + \underset{(0.010)}{0.018}\ln\epsilon^2_{1,t-6} + \underset{(0.008)^*}{0.093}\ln\epsilon^2_{1,t-7} + \underset{(0.086)}{0.014}\ln\widehat{\sigma}^2_{1,t-1} - \underset{(0.112)^*}{0.749}\,Tue_t$ $- \underset{(0.071)^*}{1.194}\,Wed_t - \underset{(0.071)^*}{1.066}\,Thu_t - \underset{(0.068)^*}{1.268}\,Fri_t - \underset{(0.074)^*}{0.923}\,Sat_t - \underset{(0.068)^*}{0.940}\,Sun_t$	1896.3 [k=15]
4: $\ln\widehat{\sigma}^2_{1,t} =$	$0.224 + \underset{(0.008)^*}{0.203}\ln\epsilon^2_{1,t-1} + \underset{(0.020)^*}{0.103}\ln\epsilon^2_{1,t-2} + \underset{(0.012)^*}{0.041}\ln\epsilon^2_{1,t-3} - \underset{(0.009)}{0.014}\ln\epsilon^2_{1,t-4} + \underset{(0.008)^*}{0.046}\ln\epsilon^2_{1,t-5}$ $+ \underset{(0.009)}{0.013}\ln\epsilon^2_{1,t-6} + \underset{(0.008)^*}{0.091}\ln\epsilon^2_{1,t-7} + \underset{(0.008)^*}{0.057}\ln\epsilon^2_{2,t-1} + \underset{(0.010)^*}{0.043}\ln\epsilon^2_{2,t-2} + \underset{(0.009)^*}{0.038}\ln\epsilon^2_{2,t-3}$ $+ \underset{(0.009)}{0.013}\ln\epsilon^2_{2,t-4} + \underset{(0.008)^*}{0.037}\ln\epsilon^2_{2,t-5} + \underset{(0.009)}{0.008}\ln\epsilon^2_{2,t-6} - \underset{(0.008)}{0.015}\ln\epsilon^2_{2,t-7} - \underset{(0.090)}{0.013}\ln\widehat{\sigma}^2_{t-1}$ $- \underset{(0.111)^*}{0.773}\,Tue_t - \underset{(0.073)^*}{1.208}\,Wed_t - \underset{(0.076)^*}{1.019}\,Thu_t - \underset{(0.072)^*}{1.295}\,Fri_t - \underset{(0.081)^*}{0.905}\,Sat_t - \underset{(0.070)^*}{0.863}\,Sun_t$	1963.0 [k=22]
5: $\ln\widehat{\sigma}^2_{1,t} =$	$-0.071 + \underset{(0.008)^*}{0.209}\ln\epsilon^2_{1,t-1} + \underset{(0.008)^*}{0.120}\ln\epsilon^2_{1,t-2} + \underset{(0.007)^*}{0.066}\ln\epsilon^2_{1,t-5} + \underset{(0.007)^*}{0.093}\ln\epsilon^2_{1,t-7}$ $+ \underset{(0.008)^*}{0.070}\ln\epsilon^2_{2,t-1} + \underset{(0.008)^*}{0.063}\ln\epsilon^2_{2,t-3} - \underset{(0.067)^*}{0.681}\,Tue_t - \underset{(0.070)^*}{1.194}\,Wed_t - \underset{(0.070)^*}{0.957}\,Thu_t$ $- \underset{(0.068)^*}{1.197}\,Fri_t - \underset{(0.069)^*}{0.801}\,Sat_t - \underset{(0.069)^*}{0.857}\,Sun_t$	1955.5 [k=13]

Off-peak specifications:		*LogL*
1: $\ln\widehat{\sigma}^2_{2,t} =$	$-0.070 + \underset{(0.006)^*}{0.137}\ln\epsilon^2_{2,t-1} + \underset{(0.010)^*}{0.792}\ln\widehat{\sigma}_{2,t-1}$	1676.0 [k=3]
2: $\ln\widehat{\sigma}^2_{2,t} =$	$-0.548 + \underset{(0.008)^*}{0.202}\ln\epsilon^2_{2,t-1} + \underset{(0.012)^*}{0.065}\ln\epsilon^2_{2,t-2} + \underset{(0.008)^*}{0.083}\ln\epsilon^2_{2,t-3} + \underset{(0.008)^*}{0.064}\ln\epsilon^2_{2,t-4}$ $+ \underset{(0.008)}{0.012}\ln\epsilon^2_{2,t-5} + \underset{(0.007)^*}{0.107}\ln\epsilon^2_{2,t-6} + \underset{(0.009)^*}{0.170}\ln\epsilon^2_{2,t-7} - \underset{(0.041)^*}{0.103}\ln\widehat{\sigma}^2_{2,t-1}$	1665.9 [k=9]
3: $\ln\widehat{\sigma}^2_{2,t} =$	$-0.148 + \underset{(0.008)^*}{0.202}\ln\epsilon^2_{2,t-1} + \underset{(0.024)^*}{0.106}\ln\epsilon^2_{2,t-2} + \underset{(0.012)^*}{0.091}\ln\epsilon^2_{2,t-3} + \underset{(0.012)^*}{0.068}\ln\epsilon^2_{2,t-4}$ $+ \underset{(0.010)^*}{0.048}\ln\epsilon^2_{2,t-5} + \underset{(0.009)^*}{0.111}\ln\epsilon^2_{2,t-6} + \underset{(0.014)^*}{0.094}\ln\epsilon^2_{2,t-7} - \underset{(0.108)}{0.135}\ln\widehat{\sigma}^2_{2,t-1} - \underset{(0.110)^*}{0.805}\,Tue_t$ $- \underset{(0.184)^*}{1.341}\,Wed_t - \underset{(0.231)}{0.401}\,Thu_t - \underset{(0.125)^*}{1.255}\,Fri_t - \underset{(0.232)^*}{1.091}\,Sat_t + \underset{(0.196)^*}{0.540}\,Sun_t$	1801.3 [k=15]
4: $\ln\widehat{\sigma}^2_{2,t} =$	$-0.603 + \underset{(0.008)^*}{0.092}\ln\epsilon^2_{1,t-1} - \underset{(0.012)}{0.004}\ln\epsilon^2_{1,t-2} - \underset{(0.011)^*}{0.060}\ln\epsilon^2_{1,t-3} - \underset{(0.011)}{0.022}\ln\epsilon^2_{1,t-4} - \underset{(0.011)}{0.013}\ln\epsilon^2_{1,t-5}$ $+ \underset{(0.011)^*}{0.034}\ln\epsilon^2_{1,t-6} - \underset{(0.008)}{0.014}\ln\epsilon^2_{1,t-7} + \underset{(0.009)^*}{0.155}\ln\epsilon^2_{2,t-1} - \underset{(0.011)^*}{0.085}\ln\epsilon^2_{2,t-2} + \underset{(0.012)^*}{0.030}\ln\epsilon^2_{2,t-3}$ $- \underset{(0.012)}{0.005}\ln\epsilon^2_{2,t-4} - \underset{(0.012)}{0.004}\ln\epsilon^2_{2,t-5} + \underset{(0.012)^*}{0.065}\ln\epsilon^2_{2,t-6} - \underset{(0.008)^*}{0.058}\ln\epsilon^2_{2,t-7} + \underset{(0.009)^*}{0.845}\ln\widehat{\sigma}^2_{2,t-1}$ $- \underset{(0.113)}{0.152}\,Tue_t + \underset{(0.092)}{0.030}\,Wed_t + \underset{(0.090)^*}{1.567}\,Thu_t - \underset{(0.091)^*}{0.356}\,Fri_t + \underset{(0.093)^*}{0.858}\,Sat_t + \underset{(0.110)^*}{2.168}\,Sun_t$	1849.2 [k=22]
5: $\ln\widehat{\sigma}^2_{2,t} =$	$-0.047 + \underset{(0.008)^*}{0.087}\ln\epsilon^2_{1,t-1} + \underset{(0.007)^*}{0.081}\ln\epsilon^2_{1,t-2} + \underset{(0.008)^*}{0.179}\ln\epsilon^2_{2,t-1} + \underset{(0.008)^*}{0.078}\ln\epsilon^2_{2,t-3}$ $+ \underset{(0.008)^*}{0.056}\ln\epsilon^2_{2,t-4} + \underset{(0.008)^*}{0.111}\ln\epsilon^2_{2,t-6} + \underset{(0.008)^*}{0.080}\ln\epsilon^2_{2,t-7} - \underset{(0.059)^*}{0.547}\,Tue_t - \underset{(0.061)^*}{1.119}\,Wed_t$ $- \underset{(0.058)^*}{0.997}\,Fri_t - \underset{(0.061)^*}{0.651}\,Sat_t + \underset{(0.060)^*}{0.839}\,Sun_t$	1812.6 [k=13]

LogL, Gaussian log-likelihood computed as $\sum_{t=1}^{T}\ln f_\epsilon(\epsilon_t;\widehat{\sigma}_t)$, where $f_\epsilon(\epsilon_t;\widehat{\sigma}_t)$ is the univariate normal density, ϵ_t is the mean-corrected return and $\widehat{\sigma}_t$ is the fitted standard deviation ($T = 1586$ is the number of observations). k, the number of of log-GARCH parameters (τ 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* ([35]) with the `lgarch` ([36]) and `AutoSEARCH` ([41]) packages.

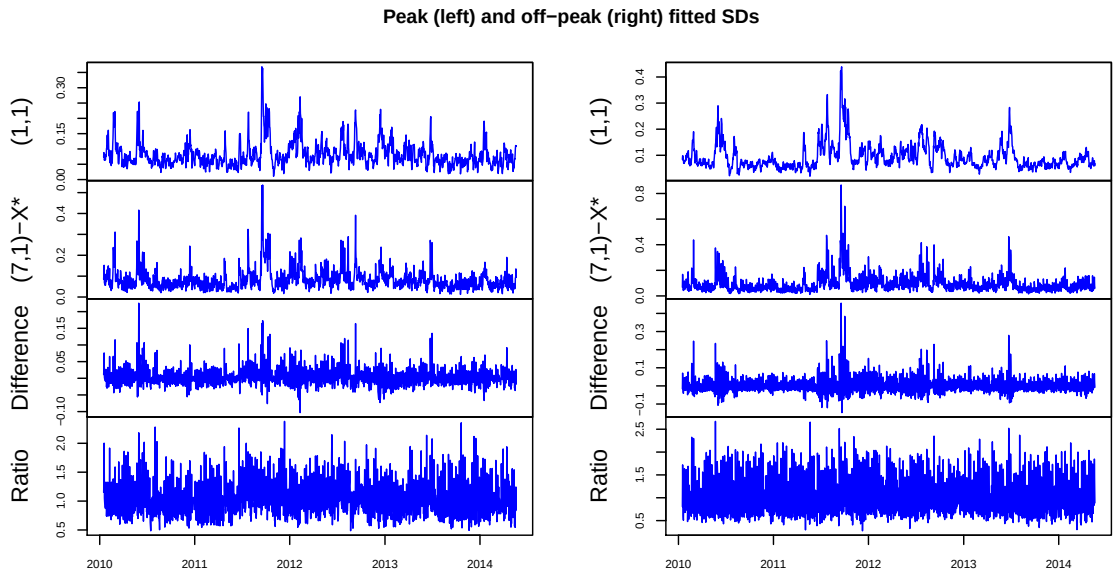

Figure 2: Fitted standard deviations (SDs) of the univariate log-GARCH-1 and log-GARCH-4 models, and the nominal differences and ratios between the SDs (computed as log-GARCH-4 minus log-GARCH-1 and log-GARCH-4 over log-GARCH-1, respectively)

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 ([42]) initiated the larger research agenda of which it is a part. [30] relies explicitly on the results of this paper, whereas [29] is a precursor. These led to the development of the *R* ([35]) software packages `AutoSEARCH` ([41]) and `gets` ([43]) 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 [10]. This solution is only available when estimation is via the (V)ARMA representation. [18] 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, [32] 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 Economia 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.

References

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

[2] A. Escribano, J. I. Peña, P. Villaplana, Modelling Electricity Prices: International Evidence, Oxford Bulletin of Economics and Statistics 73 (2011) 622–650, working paper version (2002): UC3M Working Paper 02-27 in the Economic Series.

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

[4] P. H. Franses, J. Neele, D. Van Dijk, Modelling asymmetric volatility in weekly Dutch temperature data, Environmental Modeling and Software 16 (2001) 131–137.

[5] C. Brownlees, F. Cipollini, G. Gallo, Multiplicative Error Models, in: L. Bauwens, C. Hafner, S. Laurent (Eds.), Handbook of Volatility Models and Their Applications, Wiley, New Jersey, 2012, pp. 223–247.

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

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

[8] A. Milhøj, A Multiplicative Parametrization of ARCH Models, research Report 101, University of Copenhagen: Institute of Statistics (1987).

[9] R. F. Engle, T. Bollerslev, Modelling the persistence of conditional variances, Econometric Reviews 5 (1986) 1–50.

[10] G. Sucarrat, Á. Escribano, Unbiased QML Estimation of Log-GARCH Models in the Presence of Zero Returns, mPRA Paper No. 50699. Online at `http://mpra.ub.uni-muenchen.de/50699/` (2013).

[11] D. B. Nelson, Conditional Heteroskedasticity in Asset Returns: A New Approach, Econometrica 59 (1991) 347–370.

[12] A. C. Harvey, Dynamic Models for Volatility and Heavy Tails, Cambridge University Press, New York, 2013.

[13] I. Berkes, L. Horvath, P. Kokoszka, GARCH processes: structure and estimation, Bernoulli 9 (2003) 201–227.

[14] C. Francq, J.-M. Zakoïan, Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes, Bernoulli 10 (2004) 605–637.

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

[16] O. Wintenberger, Continuous Invertibility and Stable QML Estimation of the EGARCH(1,1) model, Scandinavian Journal of Statistics 40 (2013) 846–867.

[17] C. Francq, O. Wintenberger, J.-M. Zakoïan, GARCH Models Without Positivity Constraints: Exponential or Log-GARCH?, forthcoming in *Journal of Econometrics*, `http//dx.doi.org/10.1016/j.jeconom.2013.05.004` (2013).

[18] C. Francq, G. Sucarrat, An Exponential Chi-Squared QMLE for Log-GARCH Models Via the ARMA Representation, `http://mpra.ub.uni-muenchen.de/51783/` (2013).

[19] R. F. Engle, K. F. Kroner, Multivariate simultaneous generalized ARCH, Econometric Theory 11 (1995) 122–150.

[20] F. Comte, O. Lieberman, Asymptotic Theory for Multivariate GARCH Processes, Journal of Multivariate Analysis 84 (2003) 61–84.

[21] S. Ling, M. McAleer, Asymptotic theory for a vector ARMA-GARCH model, Econometric Theory 19 (2003) 280–310.

[22] C. Hafner, A. Preminger, Asymptotic theory for a factor GARCH model, Econometric Theory 25 (2009) 336–363.

[23] C. Francq, J.-M. Zakoïan, QML estimation of a class of multivariate GARCH models without moment conditions on the observed process, unpublished working paper (2010).

[24] C. Francq, J.-M. Zakoïan, Estimating multivariate GARCH and stochastic correlation models equation by equation, mPRA Paper No. 54250. Online at http://mpra.ub.uni-muenchen.de/54250/ (2014).

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

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

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

[28] P. J. Brockwell, R. A. Davis, Time Series: Theory and Methods, Springer, New York, 2006, 2nd. Edition, first published in 1991.

[29] L. Bauwens, G. Sucarrat, General to Specific Modelling of Exchange Rate Volatility: A Forecast Evaluation, International Journal of Forecasting 26 (2010) 885–907.

[30] G. Sucarrat, Á. Escribano, Automated Model Selection in Finance: General-to-Specific Modelling of the Mean and Volatility Specifications, Oxford Bulletin of Economics and Statistics 74 (2012) 716–735.

[31] N. Duan, Smearing Estimate: A Nonparametric Retransformation Method, Journal of the Americal Statistical Association 78 (1983) pp. 605–610.

[32] C. Francq, G. Sucarrat, An Equation-by-Equation Estimator of a Multivariate Log-GARCH-X Model of Financial Returns, https://mpra.ub.uni-muenchen.de/67140/ (2015).

[33] H. Yu, High moment partial sum processes of residuals in ARMA models and their applications, Journal of Time Series Analysis 28 (2007) 72–91.

[34] E. Hannan, M. Deistler, The statistical theory of linear systems, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2012, originally published in 1988 by Wiley, New York.

[35] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria (2014).
URL http://www.R-project.org/

[36] G. Sucarrat, lgarch: Simulation and estimation of log-GARCH models, http://cran.r-project.org/web/packages/lgarch/ (2014).

[37] H. Lütkepohl, New Introduction to Multiple Time Series Analysis, Springer-Verlag, Berlin, 2005.

[38] J.-M. Bardet, O. Wintenberger, Asymptotic normality of the quasi maximum likelihood estimator for multidimensional causal processes, unpublished working paper (2009).

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

[40] L. Bauwens, C. Hafner, D. Pierret, Multivariate Volatility Modelling of Electricity Futures, Journal of Applied Econometrics 28 (2013) 743–761.

[41] G. Sucarrat, AutoSEARCH: General-to-Specific (GETS) Model Selection, r package version 1.2 (2012).
URL `http://CRAN.R-project.org/package=AutoSEARCH`

[42] G. Sucarrat, Á. Escribano, The Power Log-GARCH Model, universidad Carlos III de Madrid Working Paper 10-13 in the Economic Series, June 2010. `http://e-archivo.uc3m.es/bitstream/10016/8793/1/we1013.pdf` (2010).

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

[44] S. Lee, A note on the residual empirical process in autoregressive models, Statistics and Probability Letters 32 (4) (1997) 405–411.

[45] R. Ibragimov, P. C. Phillips, Regression asymptotics using martingale convergence methods, Econometric Theory 24 (4) (2008) 888–947.

Appendix A. Proof of Theorems 1 and 3

Since estimation is by OLS, a mean correction is irrelevant in **A3**. Let $Y_t = \ln \epsilon_t^2$, and let $\hat{\boldsymbol{\phi}} := (\hat{\phi}_1, \ldots, \hat{\phi}_p)$ be the least squares estimator of $\boldsymbol{\phi} := (\phi_1, \ldots, \phi_p)$ 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 $\sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}') = O_P(1)$ by standard theory [28]). Denote $\gamma := Ee^{u_0}$ and $\hat{\gamma} = \frac{1}{T} \sum_{t=1}^T \exp(\hat{u}_t)$. **A3** implies $\mathrm{Var}\,(u_0^2) = E[(\ln z_1^2)^2] - [E \ln(z_1^2)]^2 < \infty$ as well as $E \exp(u_1) = 1/(\exp[E \ln(z_1^2)]) < \infty$, and **A4** implies that $\mathrm{Var}\,[\exp(u_1)] = (Ez_1^4 - 1)/(\{\exp[E \ln(z_1^2)]\}^2) < \infty$, and so $Ee^{2u_0} < \infty$. To prove Theorem 1, we show what we call *Case (i)*: If $Eu_0^2 < \infty$ and $Ee^{u_0} < \infty$, then $\hat{\gamma} = \gamma + o_P(1)$. To prove Theorem 3 we prove *Case (ii)*: If $Eu_0^4 < \infty$ and $Ee^{2u_0} < \infty$, then $\sqrt{T}(\hat{\gamma} - \gamma) = T^{-1/2} \sum_{t=1}^T (e^{u_t - \bar{u}_T} - \gamma) + o_P(1)$. Both cases use the following expansions. Let

$\delta_{t,T} := \hat{u}_{t,T} - u_t$. Note that $\delta_{t,T} = -u_t$ when $t \leq p$. For $t \leq p$ we have $e^{\hat{u}_t} = e^0 = 1$. Note that the initial values are asymptotically insignificant. For $t \geq p+1$, a Taylor expansion shows that

$$e^{u_t+\delta_t} = e^{u_t} + \delta_t e^{u_t} + \delta_t^2 \int_0^1 (1-x) e^{u_t + x\delta_t} dx = e^{u_t} + \delta_t e^{u_t} + e^{u_t} \delta_t^2 \int_0^1 (1-x) e^{x\delta_t} dx. \tag{A.1}$$

We follow the main argument in Theorem 1 of [44] to bound δ_t uniformly in t. Let $\mu = EY_0 = \phi_0/(1 - \sum_{i=1}^p \phi_j)$, so that that $u_t = Y_t - \mu - \sum_{i=1}^p \phi_i (Y_{t-i} - \mu)$. The definition of $\hat{u}_t$, as well as addition and subtraction shows that for $t \geq p+1$, we have that $\hat{u}_t = u_t - (\hat{\boldsymbol{\phi}} - \boldsymbol{\phi})'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - T^{-1} \sum_{s=1}^T (Y_s - \mu)(1 - \sum_{i=1}^p \hat{\phi}_i)$ so that $\delta_{t,T} = -(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi})'(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - T^{-1} \sum_{s=1}^T (Y_s - \mu)(1 - \sum_{i=1}^p \hat{\phi}_i)$ as in [44]. [44] shows $T^{-1/2} \sum_{s=1}^T (Y_s - \mu) = \sqrt{T} \bar{u}_T (1 - \sum_{i=1}^p \phi_i)^{-1} + \xi_T$ with $\xi_T = o_P(1)$, see the proof of Theorem 1 in [44] immediately before his equation (2.6). For $t \geq p+1$ we therefore get that

$$\delta_{t,T} = -(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu) - \bar{u}_T + R_T \tag{A.2}$$

where $R_T = o_P(T^{-1/2})$ does not depend on t. To see this, note that $\sqrt{T} R_T = (\sqrt{T}\bar{u}_T) - (\sqrt{T}\bar{u}_T)(1 - \sum_{i=1}^p \phi_i)^{-1}(1 - \sum_{i=1}^p \hat{\phi}_i) + \xi_T(1 - \sum_{i=1}^p \hat{\phi}_i)$. Since $\sqrt{T}(\hat{\boldsymbol{\phi}} - \boldsymbol{\phi})' = O_P(1)$ we have $(1 - \sum_{i=1}^p \hat{\phi}_i) = (1 - \sum_{i=1}^p \phi_i) + o_P(1)$. Hence, $\sqrt{T} R_T = (\sqrt{T}\bar{u}_T) - (\sqrt{T}\bar{u}_T)(1 - \sum_{i=1}^p \phi_i)^{-1}(1 - \sum_{i=1}^p \hat{\phi}_i) + \xi_T(1 - \sum_{i=1}^p \hat{\phi}_i) = (\sqrt{T}\bar{u}_T) - (\sqrt{T}\bar{u}_T)(1 + o_P(1)) + o_P(1) = o_P(1)$, where the last equality follows by the CLT (Central Limit Theorem), i.e. that $\sqrt{T}\bar{u}_T = O_P(1)$. Hence, $M_T := \sup_{p+1 \leq t \leq T} |\delta_{t,T}| \leq \sup_{p+1 \leq t \leq T} |(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu)| + |\bar{u}_T| + |R_T| = \sup_{p+1 \leq t \leq T} |(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu)| + |\bar{u}_T| + o_P(T^{-1/2})$. We have $T^\alpha \sup_{p+1 \leq t \leq T} |(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu)| \leq T^\alpha \sup_{1 \leq j \leq p} |\hat{\phi}_j - \phi_j| \sup_{p+1 \leq t \leq T} |Y_t - \mu| = \sqrt{T} \sup_{1 \leq j \leq p} |\hat{\phi}_j - \phi_j| T^{\alpha - 1/2} \sup_{p+1 \leq t \leq T} |Y_t - \mu|$. Since (Y_t) is a strictly stationary linear process with exponentially decreasing coefficients, $Eu_0^\kappa < \infty$ implies $EY_0^\kappa < \infty$ for any $\kappa > 0$. Suppose $0 \leq \alpha < 1/2$. We have $T^{\alpha - 1/2} \sup_{p+1 \leq t \leq T} |Y_t - \mu| = o_P(1)$ if $Eu_0^{-1/(\alpha - 1/2)} < \infty$, see e.g. Lemma 12.4 of [45]. For case (i), we know that $Eu_0^2 < \infty$, i.e. $\alpha = 0$. For case (ii), we know $Eu_0^4 < \infty$, i.e. $\alpha = 1/4$. For both cases, we have $T^\alpha \sup_{p+1 \leq t \leq T} |(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu)| = o_P(1)$ since $\sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}') = O_P(1)$. Hence if $Eu_0^2 < \infty$, i.e. case (i), we conclude that $M_T = \sup_{p+1 \leq t \leq T} |(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu)| + |\bar{u}_T| + o_P(T^{-1/2}) = o_P(1)$ since $\bar{u}_T = o_P(1)$ by the LLN (Law of Large Numbers). If $Eu_0^4 < \infty$ we get $T^{1/4} M_T = T^{1/4} \sup_{p+1 \leq t \leq T} |(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')(Y_{t-1} - \mu, \ldots, Y_{t-p} - \mu)| + T^{-1/4} |T^{1/2} \bar{u}_T| + o_P(T^{-1/4})$. Because $T^{1/2}\bar{u}_T = O_P(1)$ by the CLT, we see that $T^{-1/4}|T^{1/2}\bar{u}_T| = o_P(1)$. 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 $\frac{1}{T}\sum_{t=1}^T e^{\hat{u}_t} = \frac{1}{T}\sum_{t=1}^p e^{\hat{u}_t} + \frac{1}{T}\sum_{t=q+1}^T e^{u_t} + \frac{1}{T}\sum_{t=q+1}^T \delta_t e^{u_t} + \frac{1}{T}\sum_{t=q+1}^T e^{u_t}\delta_t^2 \int_0^1 (1-x) e^{x\delta_t} dx$. Clearly, $\frac{1}{T}\sum_{t=1}^p e^{\hat{u}_t} = p/T = o_P(1)$. We have that $\int_0^1 (1-x) e^{x\delta_t} dx \leq e^{|\delta_t|}$ because for $0 \leq x \leq 1$ we have $(1-x) \leq 1$ and $e^{x\delta_t} \leq e^{|x\delta_t|} = e^{x|\delta_t|} \leq e^{|\delta_t|}$ so that $\int_0^1 (1-x) e^{x\delta_t} dx \leq \int_0^1 e^{|\delta_t|} dx = e^{|\delta_t|}$. By $Ee^{u_0} < \infty$, the LLN implies that $\frac{1}{T}\sum_{t=p+1}^T e^{u_t} = Ee^{u_0} + o_P(1)$. Hence, the triangle inequality implies that $|\hat{\gamma} - \gamma| \leq \frac{1}{T}\sum_{t=p+1}^T |\delta_t| e^{u_t} + \frac{1}{T}\sum_{t=p+1}^T e^{u_t}\delta_t^2 e^{|\delta_t|} + o_P(1)$. Using $|\delta_t| \leq M_T$ we get that $|\hat{\gamma} - \gamma| \leq M_T \frac{1}{T}\sum_{t=p+1}^T e^{u_t} + M_T^2 e^{M_t} \frac{1}{T}\sum_{t=p+1}^T e^{u_t} + o_P(1)$, which is $o_P(1)$

because $M_T = o_P(1)$ and $T^{-1}\sum_{t=p+1}^{T} e^{u_t} = Ee^{u_0} + o_P(1) = O_P(1)$.

Let us now show asymptotic Normality, i.e. case (ii). From eq. (A.1), we see that $\sqrt{T}(\hat{\gamma}-\gamma) = \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}(e^{u_t}-\gamma) + \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T}\delta_t e^{u_t} + \frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t}\delta_t^2\int_0^1(1-x)e^{x\delta_t}dx + o_P(1)$. The last sum is $o_P(1)$. To see this, we again use $\int_0^1(1-x)e^{x\delta_t}dx \leq e^{|\delta_t|}$ combined with $T^{1/4}M_T = o_P(1)$ and see that $\left|\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t}\delta_t^2\int_0^1(1-x)e^{x\delta_t}dx\right| \leq M_T^2 e^{M_T}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t} = \left(\frac{T^{1/4}}{T^{1/4}}M_T\right)^2 e^{M_T}\frac{1}{\sqrt{T}}\sum_{t=p+1}^{T} e^{u_t} = (T^{1/4}M_T)^2 e^{M_T}\frac{1}{T}\sum_{t=p+1}^{T} e^{u_t}$, which is $o_P(1)$ because $(T^{1/4}M_T)^2 = [o_P(1)]^2 = o_P(1)$ by continuity, that $e^{M_T} = e^{o_P(1)} = e^0 + o_P(1) = 1 + o_P(1) = O_P(1)$, and by the LLN we have $T^{-1}\sum_{t=p+1}^{T} e^{u_t} = O_P(1)$.

We have therefore shown that $\sqrt{T}(\hat{\gamma}-\gamma) = T^{-1/2}\sum_{t=p+1}^{T}(e^{u_t}-\gamma) + T^{-1/2}\sum_{t=p+1}^{T}\delta_t e^{u_t} + o_P(1) = T^{-1/2}\sum_{t=1}^{T}(e^{u_t}-\gamma) + T^{-1/2}\sum_{t=p+1}^{T}\delta_t e^{u_t} + o_P(1)$. For the sum with δ_t, we apply eq. (A.2), and get $T^{-1/2}\sum_{t=p+1}^{T}\delta_t e^{u_t} = -(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')T^{-1/2}\sum_{t=p+1}^{T}(Y_{t-1}-\mu,\ldots,Y_{t-p}-\mu)e^{u_t} - \bar{u}_T\left(T^{-1/2}\sum_{t=p+1}^{T} e^{u_t}\right) + R_T T^{-1/2}\sum_{t=p+1}^{T} e^{u_t}$. Since Y_{t-j} and u_t are independent for $j \geq 0$, we get $T^{-1}\sum_{t=p+1}^{T}(Y_{t-1}-\mu,\ldots,Y_{t-p}-\mu)e^{u_t} = E[(Y_{-1}-\mu,\ldots,Y_{-p}-\mu)e^{u_0}] + o_P(1) = (0,0,\ldots,0)Ee^{u_0} + o_P(1) = (0,0,\ldots,0) + o_P(1)$. Hence, $(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')T^{-1/2}\sum_{t=p+1}^{T}(Y_{t-1}-\mu,\ldots,Y_{t-p}-\mu)e^{u_t} = \sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}')[(0,0,\ldots,0) + o_P(1)]$, which is $o_P(1)$ because $\sqrt{T}(\hat{\boldsymbol{\phi}}' - \boldsymbol{\phi}') = O_P(1)$. Recalling $R_T = o_P(T^{-1/2})$ implies that $R_T T^{-1/2}\sum_{t=p+1}^{T} e^{u_t} = (T^{1/2}R_T)T^{-1}\sum_{t=p+1}^{T} e^{u_t} = o_P(1)[Ee^{u_0} + o_P(1)] = o_P(1)$. We further have that $\bar{u}_T\left(T^{-1/2}\sum_{t=p+1}^{T} e^{u_t}\right) = \sqrt{T}\bar{u}_T[Ee^{u_0} + o_P(1)] = \sqrt{T}\bar{u}_T Ee^{u_0} + o_P(1)\sqrt{T}\bar{u}_T = \sqrt{T}\bar{u}_T Ee^{u_0} + o_P(1)$ since $\sqrt{T}\bar{u}_T = O_P(1)$. In conclusion we get $\sqrt{T}(\hat{\gamma}-\gamma) = T^{-1/2}\sum_{t=1}^{T}(e^{u_t}-\gamma) - \sqrt{T}\bar{u}_T Ee^{u_1} + o_P(1)$. We now complete the proof by showing that $T^{-1/2}\sum_{t=1}^{T}(e^{u_t-\bar{u}_T} - Ee^{u_0})$ fulfils exactly the same expansion as we found for $\sqrt{T}(\hat{\gamma}-\gamma)$. We have $T^{-1/2}\sum_{t=1}^{T}(e^{u_t-\bar{u}_T} - Ee^{u_0}) = T^{-1/2}\sum_{t=1}^{T} e^{-\bar{u}_T}e^{u_t} - \sqrt{T}Ee^{u_0} = e^{-\bar{u}_T}(T^{-1/2}\sum_{t=1}^{T} e^{u_t}) - \sqrt{T}Ee^{u_0} = e^{-\bar{u}_T}(T^{-1/2}\sum_{t=1}^{T} e^{u_t} - Ee^{u_0} + Ee^{u_0}) - \sqrt{T}Ee^{u_0} = e^{-\bar{u}_T}(T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}]) + e^{-\bar{u}_T}\sqrt{T}Ee^{u_0} - \sqrt{T}Ee^{u_0} = e^{o_P(1)}(T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}]) + [e^{-\bar{u}_T} - 1]\sqrt{T}Ee^{u_0}$. Since $Ee^{2u_0} < \infty$, the CLT implies that $T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}] = O_P(1)$ and hence $e^{o_P(1)}T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}] = (1+o_P(1))T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}] = T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}] + o_P(1)T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}] = T^{-1/2}\sum_{t=1}^{T}[e^{u_t} - Ee^{u_0}] + o_P(1)$. The delta method gives $[e^{-\bar{u}_T} - 1]\sqrt{T}Ee^{u_0} = \sqrt{T}[e^{-\bar{u}_T} - e^0]Ee^{u_0} = -\sqrt{T}\bar{u}_T Ee^{u_0} + o_P(1)$ and the conclusion follows.

Appendix B. Proof of Theorem 2

Let $\widetilde{\tau}_T = -\ln[T^{-1}\sum_{t=1}^{T}\exp(u_t - \overline{u}_T)]$. Assumption **A4** and the smoothness of the logarithm function imply that $\widehat{\tau}_T$ and $\widetilde{\tau}_T$ have the same behaviour up to $o_P(T^{-1/2})$. Denoting $\tau = E\ln(z_1^2) = -\ln Ee^{u_t}$, this means $\sqrt{T}(\widehat{\tau}_T - \tau) = \sqrt{T}(\widetilde{\tau}_T - \tau) + o_P(1)$. By Slutsky we only need show that $\widetilde{\Delta}_T = \sqrt{T}(\widetilde{\tau}_T - \tau)$ is asymptotically normal. We have that $\widetilde{\tau}_T = -\ln T - 1\sum_{t=1}^{T} e^{u_t - \bar{u}_T} = \bar{u}_T - \ln T - 1\sum_{t=1}^{T} e^{u_t}$, so $\widetilde{\Delta}_T = \sqrt{T}\bar{u}_T + \sqrt{T}\left[f\left(\frac{1}{T}\sum_{t=1}^{T} e^{u_t}\right) - f(Ee^{u_1})\right]$, where $f(x) = -\ln x$, with $f'(x) = -1/|x|$. By the smoothness of f, the delta method implies that $\widetilde{\Delta}_T = \sqrt{T}\bar{u}_T + f'(Ee^{u_1})\sqrt{T}\left[\frac{1}{T}\sum_{t=1}^{T} e^{u_t} - Ee^{u_1}\right] + o_P(1) = (f'(Ee^{u_1}), 1)\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(e^{u_t} - Ee^{u_1}u_t\right) + o_P(1)$. By the multivariate CLT, we have that $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(e^{u_t} - Ee^{u_1}, u_t)' \xrightarrow{d} (X,Y)' \sim N((0,0)', \mathrm{Cov}\,(e^{u_1}, u_1 e^{u_1}))$. Hence, $\widetilde{\Delta}_T \xrightarrow{d} f'(Ee^{u_1})X + Y$, which is mean zero

normal with variance equal to

$$\zeta^2 = \left(f'(Ee^{u_1})\right)^2 \operatorname{Var} X + \operatorname{Var} Y + 2f'(Ee^{u_1}) \operatorname{Var}(X,Y) = \frac{\operatorname{Var}[\exp(u_1)]}{[E\exp(u_1)]^2} + \operatorname{Var}(u_1) - 2\frac{E[u_1\exp(u_1)]}{E\exp(u_1)}.$$

Using that $\operatorname{Var}[\exp(u_1)] = (Ez_1^4 - 1)/\left\{\exp[E\ln(z_1^2)]\right\}^2$, $E\exp(u_1) = 1/\exp[E\ln(z_1^2)]$, $\operatorname{Var}(u_1) = E[(\ln z_1^2)^2] - [E\ln(z_1^2)]^2$ and $E[u_1\exp(u_1)] = \left\{E[(\ln z_1^2)z_1^2] - E\ln(z_1^2)\right\}/\exp[E\ln(z_1^2)]$, we obtain

$$\zeta^2 = E[(\ln z_1^2)^2] - [E(\ln z_1^2)]^2 + (E(z_1^4) - 1) - 2E[(\ln z_1^2)z_1^2] + 2E(\ln z_1^2).$$

From **A4** we have that $E(z_1^4) < \infty$ and $E[(\ln z_1^2)^2] < \infty$, and the Cauchy-Schwarz inequality implies that $|E[(\ln z_1^2)z_1^2]|^2 \leq (E[(\ln z_1^2)^2])(Ez_1^4)$. So ζ^2 is finite. Finally, the expression simplifies to $\zeta^2 = \operatorname{Var}(z_1^2 - \ln z_1^2)$.