



Munich Personal RePEc Archive

Equation-by-Equation Estimation of a Multivariate Log-GARCH-X Model of Financial Returns

Francq, Christian and Sucarrat, Genaro

Universite de Lille, BI Norwegian Business School

8 October 2015

Online at <https://mpra.ub.uni-muenchen.de/67140/>

MPRA Paper No. 67140, posted 09 Oct 2015 18:12 UTC

An Equation-by-Equation Estimator of a Multivariate Log-GARCH-X Model of Financial Returns

Christian Francq¹ and Genaro Sucarrat²

8th October 2015

Abstract

Estimation of large financial volatility models is plagued by the curse of dimensionality: As the dimension grows, joint estimation of the parameters becomes infeasible in practice. This problem is compounded if covariates or conditioning variables ("X") are added to each volatility equation. In particular, the problem is especially acute for non-exponential volatility models (e.g. GARCH models), since there the variables and parameters are restricted to be positive. Here, we propose an estimator for a multivariate log-GARCH-X model that avoids these problems. The model allows for feedback among the equations, admits several stationary regressors as conditioning variables in the X-part (including leverage terms), and allows for time-varying covariances of unknown form. Strong consistency and asymptotic normality of an equation-by-equation least squares estimator is proved, and the results can be used to undertake inference both within and across equations. The flexibility and usefulness of the estimator is illustrated in two empirical applications

JEL Classification: C13, C22, C32, C51, C58

Keywords: Exponential GARCH, multivariate log-GARCH-X, VARMA-X, Equation-by-Equation Estimation (EBEE), Least Squares

1	Introduction	2
2	Model and notation	3
3	Equation-by-equation estimation	4
3.1	Estimator and strong consistency	6
3.2	Asymptotic normality	11
3.3	Constrained models	15
4	Empirical applications	17
4.1	Volatility spillover	17
4.2	A 50-dimensional log-GARCH-X model with time-varying correlations	20
5	Conclusions	22
	References	25

¹CREST and Université de Lille. Email christian.francq at univ-lille3.fr

²Corresponding author. Department of Economics, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway. Email genaro.sucarrat at bi.no, phone +47+46410779, fax +47+23264788.

1 Introduction

Covariates often help explaining and forecasting financial variability. Examples of such covariates include, amongst other, measures of information arrival (e.g. [Clark \(1973\)](#), [Tauchen and Pitts \(1983\)](#), [Lamoureux and Lastrapes \(1990\)](#)), interest rates (e.g. [Brenner et al. \(1996\)](#), [Hagiwara and Herce \(1999\)](#)), central bank interventions ([Dominguez \(1998\)](#)), bid-ask spreads ([Bollerslev and Melvin \(1994\)](#)), macroeconomic fundamentals ([Apergis and Rezitis \(2011\)](#)), cross-sectional volatility ([Hwang and Satchell \(2005\)](#)) and volatility proxies made up of high-frequency data ([Engle and Gallo \(2006\)](#), [Hansen et al. \(2012\)](#), [Shephard and Sheppard \(2010\)](#)). An example of a study that combines several of these covariates in a single analysis is [Bauwens et al. \(2006\)](#). Despite the widespread use – and undoubtful usefulness – of such additional information in explaining and forecasting volatility, however, there are relatively few results on Autoregressive Conditional Heteroscedasticity (ARCH) models with covariates, i.e. ARCH-X models, where the assumptions on the X-part are non-restrictive and of general practical interest.

In the univariate case, [Han and Kristensen \(2014\)](#) prove the Consistency and Asymptotic Normality (CAN) of the Gaussian QMLE for specifications contained in the GARCH(1,1)-X model, where the X-part consists of a single variable only. [Francq and Thieu \(2015\)](#) also prove CAN for the QMLE, but for a much broader model-class: The Asymmetric Power ARCH(p, q)-X model, where the X-part can contain more than one variable. As is common in GARCH-specifications that are *not* exponential, though, all terms (parameters, the variables in the X-part, etc.) are restricted to be non-negative in both works. [Chen and Song \(2015\)](#) prove CAN of a QMLE for a log-GARCH(1,1) model with no X-part in the classical sense, but where the ARCH-parameter varies over time and is driven by the past values of two covariates that are independent of the standardised error. This independence assumption is somewhat restrictive, however, since it excludes feedback effects between the covariates and the log-volatility process (this is usually not fulfilled in empirical practice). Also, it is not clear what the economic motivation is in making the ARCH-parameter – the part of a log-GARCH(1,1) that usually accounts for the smallest portion of the variation in volatility – time-varying and dependent on past covariates whose properties are usually not fulfilled in empirical practice. Finally, to the best of our knowledge there is no proof of CAN for multivariate GARCH-X models.

[Sucarrat et al. \(2015, \[2010\]\)](#) propose a general framework for the estimation of and inference in univariate and multivariate log-GARCH-X models – with Dynamic Conditional Correlations (DCCs) of unknown form – via the (V)ARMA-X representation. However, they do not prove CAN, neither in the univariate nor in the multivariate case. Here, we adopt their framework, but provide a proof of strong consistency and asymptotic normality under mild assumptions. Specifically, we do so for a least squares Equation-by-Equation (EBE) estimator of a multivariate log-GARCH(1,1)-X model that admits time-varying DCCs of unknown form. Moreover, the assumptions on the X-part are very general: It is not restricted to a single variable, the X-variables are allowed to be subject to feedback effects from volatility (i.e. the X-variables need not be exogenous), and the X-variables need not be independent of the standardised error. The latter means asymmetry or leverage can be accommodated via the X-part. There are several advantages with the VARMA approach. First, it enables theoretical results of unprecedented economic generality and

flexibility. In one of our applications, for example, we illustrate this in an empirical study of volatility spillover among stock markets. In ordinary multivariate GARCH models such tests require complicated restrictions on the parameters, and restrictive assumptions on the correlations (i.e. constant DCCs), see [Conrad and Weber \(2013\)](#), and [Pedersen \(2015\)](#). By contrast, in our model complicated restrictions on the parameters are not needed, and tests are valid under time-varying DCCs of unknown form. Moreover, we can also test whether covariates (e.g. volatility proxies) provide additional – or alternative – channels of volatility spillover. Second, the EBE nature of our estimator, together with the VARMA-X representation, means large systems can readily be estimated with software that is already widely available. We illustrate this in a second empirical application by estimating, in less than a minute, a model of dimension 50 that admits time-varying correlations, and where the X-part contains 5 conditioning variables in each equation. Next, a DCC model of the 1225 correlation paths is fitted. Third, the statistical theory we rely upon is much more tractable than for the EGARCH of [Nelson \(1991\)](#). Indeed, currently the only proof of CAN for a QMLE is for the univariate EGARCH(1,1) without covariates, see [Straumann and Mikosch \(2006\)](#), and [Wintenberger \(2013\)](#). Fourth, estimation via the VARMA-representation is likely to be more efficient when the standardised error is fat-tailed, since the application of the logarithm makes large (in absolute value) observations less influential. Finally, solutions to the log-of-zero (or inlier) problem is available when log-GARCH models are estimated via the (V)ARMA representation, see [Sucarrat and Escribano \(2014\)](#), and [Sucarrat et al. \(2015\)](#).

The rest of the paper is organised into four parts. First, in Section 2, we present the model and its associated notation. Section 3 contains our main theoretical results, i.e. strong consistency and asymptotic normality. Section 4 contains two empirical applications. Finally, Section 5 concludes. A Table and a Figure are located at the end.

2 Model and notation

Let $\boldsymbol{\epsilon}_t = (\epsilon_{1t}, \dots, \epsilon_{Mt})'$ denote a $M \times 1$ vector of random variables, and let $\mathcal{F}_{t-1} = \sigma\{\epsilon_{ju}^2, X_{\ell,t-1} : u < t, j = 1, \dots, M, \ell = 1, \dots, K\}$ be the σ -field generated by the past values of $\boldsymbol{\epsilon}_t^2 = (\epsilon_{1t}^2, \dots, \epsilon_{Mt}^2)'$ and of some K -dimensional vector $\mathbf{X}_t = (X_{1t}, \dots, X_{Kt})'$ of covariates. It should be noted that these need *not* be exogenous (the exact assumptions on \mathbf{X}_t are given below). Assume the existence of the $M \times M$ matrix \mathbf{H}_t such that

$$E(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t' | \mathcal{F}_{t-1}) = \mathbf{H}_t. \quad (2.1)$$

Assuming the nonsingularity of $\mathbf{D}_t^2 = \text{diag}(\mathbf{H}_t)$, let the $M \times 1$ vector

$$\boldsymbol{\eta}_t = \mathbf{D}_t^{-1} \boldsymbol{\epsilon}_t. \quad (2.2)$$

Note that this implies $E(\boldsymbol{\eta}_t^2) = (1, \dots, 1)'$, where $\boldsymbol{\eta}_t^2 = (\eta_{1t}^2, \dots, \eta_{Mt}^2)'$. Let $\boldsymbol{\sigma}_t^2$ be the M -dimensional vector equal to the diagonal of \mathbf{H}_t . If, for some vector $\boldsymbol{\sigma}$, $\ln \boldsymbol{\sigma}$ denotes the vector resulting from applying the natural log on $\boldsymbol{\sigma}$ element-wise, then the M -dimensional log-GARCH(1,1)-X specification with diagonal GARCH matrix and covariate-vector \mathbf{X}_t

is given by

$$\ln \boldsymbol{\sigma}_t^2 = \boldsymbol{\omega}_0 + \boldsymbol{\alpha}_0 \ln \boldsymbol{\epsilon}_{t-1}^2 + \boldsymbol{\beta}_0 \ln \boldsymbol{\sigma}_{t-1}^2 + \boldsymbol{\lambda}_0 \mathbf{X}_{t-1}, \quad (2.3)$$

where $\ln \boldsymbol{\sigma}_t^2 = (\ln \sigma_{1,t}^2, \dots, \ln \sigma_{M,t}^2)'$, $\boldsymbol{\omega}_0 = (\omega_{01}, \dots, \omega_{0M})'$,

$$\boldsymbol{\alpha}_0 = \begin{pmatrix} \alpha_{011} & \cdots & \alpha_{01M} \\ \vdots & \ddots & \vdots \\ \alpha_{0M1} & \cdots & \alpha_{0MM} \end{pmatrix}, \quad \boldsymbol{\lambda}_0 = \begin{pmatrix} \lambda_{011} & \cdots & \lambda_{01K} \\ \vdots & \ddots & \vdots \\ \lambda_{0M1} & \cdots & \lambda_{0MK} \end{pmatrix}$$

and $\boldsymbol{\beta}_0 = \text{diag}(\beta_{011}, \dots, \beta_{0MM})$. In principle, it should be possible to extend the study to log-GARCH-X of higher orders, but at the price of more complicated notations and less explicit assumptions.

The VARMA-X representation of this model is given by

$$\ln \boldsymbol{\epsilon}_t^2 = \mathbf{c}_0 + \boldsymbol{\phi}_0 \ln \boldsymbol{\epsilon}_{t-1}^2 - \boldsymbol{\beta}_0 \mathbf{u}_{t-1} + \boldsymbol{\lambda}_0 \mathbf{X}_{t-1} + \mathbf{u}_t, \quad (2.4)$$

where $\boldsymbol{\phi}_0 = \boldsymbol{\alpha}_0 + \boldsymbol{\beta}_0$,

$$\mathbf{c}_0 = \boldsymbol{\omega}_0 + (\mathbf{I}_M - \boldsymbol{\beta}_0)E(\ln \boldsymbol{\eta}_t^2), \quad \mathbf{u}_t = \ln \boldsymbol{\eta}_t^2 - E(\ln \boldsymbol{\eta}_t^2),$$

and where \mathbf{I}_M is the identity matrix of dimension M . Accordingly, equation j in the VARMA(1,1)-X system can be written as

$$\begin{aligned} \ln \epsilon_{jt}^2 &= c_{0j} + \sum_{\ell=1}^M \phi_{0j\ell} \ln \epsilon_{\ell,t-1}^2 + \sum_{\ell=1}^K \lambda_{0j\ell} X_{\ell,t-1} - \beta_{0jj} u_{j,t-1} + u_{jt}, \\ c_{0j} &= \omega_j + (1 - \beta_{0jj})E(\ln \eta_{jt}^2). \end{aligned} \quad (2.5)$$

The norm of a matrix $M_1 \times M_2$ of generic element $M(i, j)$ is defined by $\|\mathbf{M}\| = \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} |M(i, j)|$. The L_2 -norm of a random variable X is defined by $\|X\|_2 = \sqrt{EX^2}$.

3 Equation-by-equation estimation

Assuming that

$$(\boldsymbol{\eta}'_t, \mathbf{X}'_t)' \text{ is stationary and ergodic} \quad (3.1)$$

and

$$\text{the spectral radius of } \boldsymbol{\phi}_0 \text{ is strictly less than 1, and } |\beta_{0jj}| < 1 \text{ for all } j, \quad (3.2)$$

the equation (2.4) admits a stationary and ergodic solution, explicitly defined by

$$\ln \boldsymbol{\epsilon}_t^2 = \sum_{k=0}^{\infty} \boldsymbol{\phi}_0^k \{ \mathbf{c}_0 + \boldsymbol{\lambda}_0 \mathbf{X}_{t-k-1} - \boldsymbol{\beta}_0 \mathbf{u}_{t-k-1} + \mathbf{u}_{t-k} \}.$$

Moreover, the model is invertible, in the sense that

$$\mathbf{u}_t = \sum_{k=0}^{\infty} \beta_0^k \{ \ln \epsilon_{t-k}^2 - \phi_0 \ln \epsilon_{t-k-1}^2 - \mathbf{c}_0 - \boldsymbol{\lambda}_0 \mathbf{X}_{t-k-1} \}.$$

Remark 3.1 *Note that, contrary to the log-GARCH model, the stationarity and invertibility conditions of the standard multivariate GARCH models are quite complicated (see Boussama et al. (2011) for the BEKK model), or remain unknown (as for the DCC-MGARCH model proposed by Engle, 2002) or are known only in particular cases (see Wintenberger (2013) for the invertibility of the univariate EGARCH(1,1)). An advantage of the log-GARCH is that its ARMA representation has a noise which depends on $\boldsymbol{\eta}_t$ only. In the ARMA representations of the other GARCH formulations, the innovations generally depend on the past observations themselves (they are typically of the form $\mathbf{u}_t = \epsilon_t^2 - E(\epsilon_t^2 | \mathcal{F}_{t-1})$), and thus these ARMA representations are of no use for finding stationarity conditions and are hardly usable for estimation purposes.*

Under the moment conditions

$$E \|\ln \boldsymbol{\eta}_t^2\|^2 < \infty \quad \text{and} \quad E \|\mathbf{X}_t\|^2 < \infty \quad (3.3)$$

we have $E \|\ln \epsilon_t^2\|^2 < \infty$ and $E \|\mathbf{u}_t\|^2 < \infty$, and we can define the Hilbert space \mathcal{H}_{t-1} that is generated by the linear combinations of the $\ln \epsilon_{j_s}^2$'s and the $X_{\ell,s}$'s for $s < t$, $j = 1, \dots, M$, and $\ell = 1, \dots, K$, and by their limits in L^2 . Note that \mathcal{H}_{t-1} is also equal to the Hilbert space generated by $\{X_{\ell s}, u_{j_s} : s < t\}$, and equivalently to the Hilbert space generated by $\{X_{\ell s}, \ln \eta_{j_s}^2 : s < t\}$. One can thus interpret (\mathbf{u}_t) as a linear innovation process:

$$\mathbf{u}_t = \ln \epsilon_t^2 - E(\ln \epsilon_t^2 | \mathcal{H}_{t-1}),$$

where $E(\ln \epsilon_t^2 | \mathcal{H}_{t-1})$ denotes the orthogonal projection of $\ln \epsilon_t^2$ onto \mathcal{H}_{t-1} . Similarly we define

$$\mathbf{v}_t = \mathbf{X}_t - E(\mathbf{X}_t | \mathcal{H}_{t-1}).$$

It will be assumed that

$$\boldsymbol{\Sigma} := E \mathbf{w}_1 \mathbf{w}_1', \quad \text{where } \mathbf{w}_1 = (\mathbf{u}_1', \mathbf{v}_1')', \text{ is a positive definite matrix.} \quad (3.4)$$

Note that this assumption rules out the possibility of exact linear relations between the explanatory variables involved in the log-volatility $\ln \sigma_t^2$, which is obviously a necessary identifiability condition. The following lemma will be used to show the identifiability of the parameters in model (2.5).

Lemma 3.1 *Assume (3.2), (3.3) and (3.4). If for some non random vectors $\boldsymbol{\nu}_1$ of \mathbb{R}^M and $\boldsymbol{\nu}_2$ of \mathbb{R}^K , and for some random variable $\nu_{3,t-1} \in \mathcal{H}_{t-1}$ we have*

$$\boldsymbol{\nu}_1' \ln \epsilon_t^2 + \boldsymbol{\nu}_2' \ln \mathbf{X}_t + \nu_{3,t-1} = 0 \quad \text{a.s.} \quad (3.5)$$

then $\boldsymbol{\nu}_1 = \mathbf{0}_M$, $\boldsymbol{\nu}_2 = \mathbf{0}_K$, and $\nu_{3,t-1} = 0$ almost surely.

Proof of Lemma 3.1. Subtracting the mean, conditionally to \mathcal{H}_{t-1} , on both sides of the equality (3.5), we obtain $\boldsymbol{\nu}'_1 \mathbf{u}_t + \boldsymbol{\nu}'_2 \mathbf{v}_t = 0$ a.s., which entails $\boldsymbol{\nu}_1 = \mathbf{0}_M$ and $\boldsymbol{\nu}_2 = \mathbf{0}_K$ by (3.4), and then that $\nu_{3,t-1} = 0$ a.s. \square

In order to accommodate certain types of dynamic conditional correlation (DCC) models, we assume (instead of the usual assumption that $(\boldsymbol{\eta}_t)$ is iid) that

$$\text{for any } j \in \{1, \dots, M\}, \eta_{jt} \text{ is independent of } \mathcal{F}_{t-1}. \quad (3.6)$$

This is a weak assumption. If, for example, $\boldsymbol{\eta}_t = \mathbf{R}^{1/2}(\Delta_t)\boldsymbol{\xi}_t$, where (Δ_t) and $(\boldsymbol{\xi}_t)$ are two independent processes, $\boldsymbol{\xi}_t$ is independent of \mathcal{F}_{t-1} , $\mathbf{R}(\Delta_t)$ is a correlation matrix for any value of Δ_t , and $(\boldsymbol{\xi}_t)$ is an independent sequence of, say, $\mathcal{N}(0, \mathbf{I}_M)$ -distributed vectors, then (3.6) is satisfied. Indeed, conditionally on Δ_t and \mathcal{F}_{t-1} the variable η_{jt} is $\mathcal{N}(0, 1)$ -distributed. Since this distribution does not depend on \mathcal{F}_{t-1} nor on Δ_t , (3.6) holds true. The $\mathcal{N}(0, \mathbf{I}_M)$ -distribution can be replaced by any other spherical distribution, as shown in Proposition 3.1 of [Francq and Zakoian \(2015\)](#). A special case of (3.6) is

$$(\boldsymbol{\eta}_t) \text{ is an iid sequence (with second order moments)}. \quad (3.7)$$

In this case we have a constant conditional correlation (CCC) model, such that

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R} \mathbf{D}_t, \quad \mathbf{R} := E \boldsymbol{\eta}_t \boldsymbol{\eta}'_t.$$

Note that (3.7) is not satisfied in the previous example where $\boldsymbol{\eta}_t = \mathbf{R}^{1/2}(\Delta_t)\boldsymbol{\xi}_t$, since Δ_t is not independent of \mathcal{F}_{t-1} . Moreover, empirical evidences of non-constant conditional correlations are often found, meaning that (3.7) is generally not satisfied.

3.1 Estimator and strong consistency

Denote by

$$\boldsymbol{\vartheta}_0^{(j)} = (c_{0j}, \phi_{0j1}, \dots, \phi_{0jM}, \beta_{0jj}, \lambda_{0j1}, \dots, \lambda_{0jK})'$$

the vector of the unknown parameters involved in the j -th equation (2.5) of the VARMA-X model. This parameter of dimension $d = M + K + 2$ is assumed to belong to some compact parameter space $\Theta \subset \mathbb{R}^{M+1} \times (-1, 1) \times \mathbb{R}^K$. Let B be the backshift operator. For any $\boldsymbol{\vartheta}^{(j)} = (c, \phi_1, \dots, \phi_M, \beta, \lambda_1, \dots, \lambda_K)' \in \Theta$, let

$$u_{jt}(\boldsymbol{\vartheta}^{(j)}) = \frac{1 - \phi_j B}{1 - \beta B} \ln \epsilon_{jt}^2 - \frac{c}{1 - \beta} - \sum_{\ell \neq j} \frac{\phi_\ell B}{1 - \beta B} \ln \epsilon_{\ell t}^2 - \frac{B}{1 - \beta B} \sum_{\ell=1}^K \lambda_\ell X_{\ell t}.$$

Given observations $\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_n$ and $\mathbf{X}_1, \dots, \mathbf{X}_n$, the $u_{jt}(\boldsymbol{\vartheta}^{(j)})$'s are approximated by the recursions

$$\tilde{u}_{jt}(\boldsymbol{\vartheta}^{(j)}) = \ln \epsilon_{jt}^2 - c - \sum_{\ell=1}^M \phi_\ell \ln \epsilon_{\ell, t-1}^2 - \sum_{\ell=1}^K \lambda_\ell X_{\ell, t-1} + \beta \tilde{u}_{j, t-1}(\boldsymbol{\vartheta}^{(j)}),$$

for $t = 2, \dots, n$, with the initial value $\tilde{u}_{j1}(\boldsymbol{\vartheta}^{(j)}) = 0$. An equation-by-equation least squares estimator of the VARMA-(1,1)-X model is then defined by

$$\hat{\boldsymbol{\vartheta}}_n^{(j)} = \arg \min_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \tilde{Q}_n^{(j)}(\boldsymbol{\vartheta}^{(j)}), \quad \tilde{Q}_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) = \frac{1}{n} \sum_{t=1}^n \tilde{u}_{jt}^2(\boldsymbol{\vartheta}^{(j)}). \quad (3.8)$$

Denote by $\boldsymbol{\vartheta}_0 = \left(\boldsymbol{\vartheta}_0^{(1)'}, \dots, \boldsymbol{\vartheta}_0^{(M)'} \right)'$ the vector of all the parameters of the VARMA-X equation (2.4). This parameter vector belongs to the parameter space Θ^M , whose generic element is denoted by $\boldsymbol{\vartheta} = \left(\boldsymbol{\vartheta}^{(1)'}, \dots, \boldsymbol{\vartheta}^{(M)'} \right)'$. The least squares estimator of the whole parameter $\boldsymbol{\vartheta}_0$ is defined by

$$\hat{\boldsymbol{\vartheta}}_n = \arg \min_{\boldsymbol{\vartheta} \in \Theta^M} \sum_{t=1}^n \tilde{\boldsymbol{u}}_t'(\boldsymbol{\vartheta}) \tilde{\boldsymbol{u}}_t(\boldsymbol{\vartheta}), \quad \tilde{\boldsymbol{u}}_t(\boldsymbol{\vartheta}) = \left(\tilde{u}_{1t}^2(\boldsymbol{\vartheta}^{(1)}), \dots, \tilde{u}_{Mt}^2(\boldsymbol{\vartheta}^{(M)}) \right)'. \quad (3.9)$$

Since $\sum_{t=1}^n \tilde{\boldsymbol{u}}_t'(\boldsymbol{\vartheta}) \tilde{\boldsymbol{u}}_t(\boldsymbol{\vartheta}) = \sum_{j=1}^M \sum_{t=1}^n \tilde{u}_{jt}^2(\boldsymbol{\vartheta}^{(j)})$, one can see that the collection of the equation-by-equation estimators is nothing else than the global least squares estimator:

$$\hat{\boldsymbol{\vartheta}}_n = \left(\hat{\boldsymbol{\vartheta}}_n^{(1)'}, \dots, \hat{\boldsymbol{\vartheta}}_n^{(M)'} \right)'.$$

It is however clearly easier to compute $\hat{\boldsymbol{\vartheta}}_n$ by solving the d -dimensional optimizations (3.8), for $j = 1, \dots, M$, than the Md -dimensional optimization (3.9).

Theorem 3.1 *Assume the log-GARCH(1,1)-X model (2.1)-(2.3) with (3.1), (3.2), (3.3), (3.4) and (3.6). If $\boldsymbol{\vartheta}_0^{(j)}$ belongs to the compact set Θ and $\phi_{0jj} \neq \beta_{0jj}$, then $\hat{\boldsymbol{\vartheta}}_n^{(j)} \rightarrow \boldsymbol{\vartheta}_0^{(j)}$ almost surely as $n \rightarrow \infty$.*

Remark 3.2 *Note that $\phi_{0jj} \neq \beta_{0jj}$ (i.e. $\alpha_{0jj} \neq 0$) is an identifiability condition. It appears naturally when considering (2.5) as an ARMA(1,1)-X model for $\ln \epsilon_{jt}^2$, with covariates $\ln \epsilon_{\ell,t-1}^2$ ($\ell \neq j$) and $X_{\ell,t-1}$ ($\ell \in \{1, \dots, K\}$).*

Proof of Theorem 3.1. We have

$$\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left| u_{jt}(\boldsymbol{\vartheta}^{(j)}) - \tilde{u}_{jt}(\boldsymbol{\vartheta}^{(j)}) \right| = \sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \beta^{t-1} \left| u_{j1}(\boldsymbol{\vartheta}^{(j)}) - \tilde{u}_{j1}(\boldsymbol{\vartheta}^{(j)}) \right| \leq K \rho^t, \quad (3.10)$$

where, here and in the sequel of the paper, K denotes a generic positive random variable which is \mathcal{F}_0 -measurable, and ρ denotes a generic constant belonging to $[0, 1)$. Letting $Q_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) = n^{-1} \sum_{t=1}^n u_{jt}^2(\boldsymbol{\vartheta}^{(j)})$, we then have

$$\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left| Q_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) - \tilde{Q}_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) \right| \leq \frac{K}{n} \sum_{t=1}^{\infty} \rho^t \left(2 \sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left| u_{jt}(\boldsymbol{\vartheta}^{(j)}) \right| + K \rho^t \right) = O\left(\frac{1}{n}\right) \text{ a.s.} \quad (3.11)$$

For the last equality, we use the fact that, under the moments conditions and the compactness assumption, $\left\| \sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} |u_{jt}(\boldsymbol{\vartheta}^{(j)})| \right\|_2 < \infty$, and thus the L^2 -norm of the sum is finite, which entails that the sum is finite almost surely.

By the ergodic theorem we have almost surely

$$\lim_{n \rightarrow \infty} Q_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) = E u_{jt}^2(\boldsymbol{\vartheta}^{(j)}) = E u_{jt}^2 + E \left\{ u_{jt}(\boldsymbol{\vartheta}^{(j)}) - u_{jt} \right\}^2,$$

because u_{jt} is uncorrelated with $u_{jt}(\boldsymbol{\vartheta}^{(j)}) - u_{jt} \in \mathcal{F}_{t-1}$, under assumption (3.6). In view of (3.11), we thus have

$$\lim_{n \rightarrow \infty} \tilde{Q}_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) = \lim_{n \rightarrow \infty} Q_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) \geq \lim_{n \rightarrow \infty} Q_n^{(j)}(\boldsymbol{\vartheta}_0^{(j)}) = \lim_{n \rightarrow \infty} \tilde{Q}_n^{(j)}(\boldsymbol{\vartheta}_0^{(j)}) \quad \text{a.s.}$$

where the inequality is an equality if and only if $P \left\{ u_{jt}(\boldsymbol{\vartheta}^{(j)}) = u_{jt} \right\} = 1$. The last equality is equivalent to

$$\begin{aligned} & \left(\frac{1 - \phi_{0jj}B}{1 - \beta_{0jj}B} - \frac{1 - \phi_j B}{1 - \beta B} \right) \ln \epsilon_{jt}^2 - \left(\frac{c_{0j}}{1 - \beta_{0jj}B} - \frac{c}{1 - \beta B} \right) \\ & - \sum_{\ell \neq j} \left(\frac{\phi_{0j\ell}B}{1 - \beta_{0jj}B} - \frac{\phi_\ell B}{1 - \beta B} \right) \ln \epsilon_{\ell t}^2 - \sum_{\ell=1}^K \left(\frac{\lambda_{0j\ell}}{1 - \beta_{0jj}B} - \frac{\lambda_\ell}{1 - \beta B} \right) X_{\ell,t-1} = 0 \quad \text{a.s.} \end{aligned}$$

By Lemma 3.1, this entails that the four terms displayed in brackets are equal to zero. Under the condition $\phi_{0jl} \neq \beta_{0jj}$, this implies that $\boldsymbol{\vartheta}^{(j)} = \boldsymbol{\vartheta}_0^{(j)}$. We thus have shown that

$$\lim_{n \rightarrow \infty} \tilde{Q}_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) > \lim_{n \rightarrow \infty} \tilde{Q}_n^{(j)}(\boldsymbol{\vartheta}_0^{(j)}),$$

when $\boldsymbol{\vartheta}^{(j)} \neq \boldsymbol{\vartheta}_0^{(j)}$. Using standard arguments (used, for instance, to show (d) on Page 157 in Francq and Zakoian (2010)) the result can be extended to show that for any $\boldsymbol{\vartheta}^{(j)} \neq \boldsymbol{\vartheta}_0^{(j)}$ there exists a neighborhood $V(\boldsymbol{\vartheta}^{(j)})$ of $\boldsymbol{\vartheta}^{(j)}$ such that

$$\liminf_{n \rightarrow \infty} \inf_{\boldsymbol{\vartheta}^* \in V(\boldsymbol{\vartheta}^{(j)}) \cap \Theta} Q_n^{(j)}(\boldsymbol{\vartheta}^*) > \lim_{n \rightarrow \infty} Q_n^{(j)}(\boldsymbol{\vartheta}_0^{(j)}).$$

The conclusion then follows from a compactness argument. \square

To obtain a consistent estimator of the log-GARCH-X parameters, it remains to find a consistent estimator of $E \ln \boldsymbol{\eta}_1^2 := \boldsymbol{\tau} = (\tau_1, \dots, \tau_M)'$. Denote by $\hat{u}_{jt} = \tilde{u}_{jt}(\hat{\boldsymbol{\vartheta}}_n^{(j)})$ the residuals of the j -th ARMA-X equation.

Lemma 3.2 *Let the assumptions of Theorem 3.1, $E \eta_{jt}^4 < \infty$ and $E |\eta_{jt}|^{-s} < \infty$ for some $s > 0$. Almost surely, as $n \rightarrow \infty$ we have*

$$\hat{\tau}_{jn} := -\ln \frac{1}{n} \sum_{t=1}^n e^{\hat{u}_{jt}} \rightarrow \tau_j.$$

Proof of Lemma 3.2. We first study the effect of the initial values. A Taylor expansion and (3.10) show that

$$\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left| e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} - e^{\tilde{u}_{jt}(\boldsymbol{\vartheta}^{(j)})} \right| \leq K\rho^t \sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})}.$$

Since $E \sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} |u_{jt}(\boldsymbol{\vartheta}^{(j)})| < \infty$, we have $\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} |u_{jt}(\boldsymbol{\vartheta}^{(j)})|/t \rightarrow 0$ a.s. as $t \rightarrow \infty$. Therefore

$$\limsup_{t \rightarrow \infty} \frac{1}{t} \ln \sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left| e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} - e^{\tilde{u}_{jt}(\boldsymbol{\vartheta}^{(j)})} \right| \leq \ln \rho < 0,$$

and thus

$$\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left| e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} - e^{\tilde{u}_{jt}(\boldsymbol{\vartheta}^{(j)})} \right| < K\rho^t,$$

from which we deduce

$$\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left| \frac{1}{n} \sum_{t=1}^n e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} - \frac{1}{n} \sum_{t=1}^n e^{\tilde{u}_{jt}(\boldsymbol{\vartheta}^{(j)})} \right| = O\left(\frac{1}{n}\right) \text{ a.s.} \quad (3.12)$$

Now note that

$$\frac{\partial e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})}}{\partial \boldsymbol{\vartheta}^{(j)}} = e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} \frac{\partial u_{jt}(\boldsymbol{\vartheta}^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}}, \quad \frac{\partial u_{jt}(\boldsymbol{\vartheta}^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}} = \mathbf{d}_{t-1}(\boldsymbol{\vartheta}^{(j)}) + \beta \frac{\partial u_{j,t-1}(\boldsymbol{\vartheta}^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}}, \quad (3.13)$$

with

$$\mathbf{d}_{t-1}(\boldsymbol{\vartheta}^{(j)}) = (-1, -\ln \epsilon_{1,t-1}^2, \dots, -\ln \epsilon_{M,t-1}^2, u_{j,t-1}(\boldsymbol{\vartheta}^{(j)}), -X_{1,t-1}, \dots, -X_{K,t-1})'. \quad (3.14)$$

Moreover we have

$$\begin{aligned} \left\| \frac{\partial e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})}}{\partial \boldsymbol{\vartheta}^{(j)}} - \frac{\partial e^{u_{jt}(\boldsymbol{\vartheta}_0^{(j)})}}{\partial \boldsymbol{\vartheta}^{(j)}} \right\| &\leq \left| e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} - e^{u_{jt}(\boldsymbol{\vartheta}_0^{(j)})} \right| \left\| \frac{\partial u_{jt}(\boldsymbol{\vartheta}^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}} \right\| \\ &\quad + e^{u_{jt}(\boldsymbol{\vartheta}_0^{(j)})} \left\| \frac{\partial u_{jt}(\boldsymbol{\vartheta}^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}} - \frac{\partial u_{jt}(\boldsymbol{\vartheta}_0^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}} \right\|. \end{aligned}$$

The compactness of Θ , the fact that $\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} |\beta| < 1$ and (3.3) entail that

$$E \sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left\| \frac{\partial u_{jt}(\boldsymbol{\vartheta}^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}} \right\|^2 < \infty \quad (3.15)$$

By Lemma 2.1 and Remark 2.1 in [Francq and Sucarrat \(2013\)](#), it can be shown that there exists a neighbourhood $V(\boldsymbol{\vartheta}_0^{(j)})$ of $\boldsymbol{\vartheta}_0^{(j)}$ such that

$$E \sup_{\boldsymbol{\vartheta}^{(j)} \in V(\boldsymbol{\vartheta}_0^{(j)})} \left| e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} \right|^2 < \infty. \quad (3.16)$$

Let $V_k(\boldsymbol{\vartheta}_0^{(j)})$ be the ball of center $\boldsymbol{\vartheta}_0^{(j)}$ and radius $1/k$. The dominated convergence theorem and (3.16) imply

$$\lim_{k \rightarrow \infty} E \sup_{\boldsymbol{\vartheta}^{(j)} \in V_k(\boldsymbol{\vartheta}_0^{(j)}) \cap \Theta} \left| e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} - e^{u_{jt}(\boldsymbol{\vartheta}_0^{(j)})} \right|^2 = 0. \quad (3.17)$$

The Lebesgue dominated convergence theorem also shows that

$$\lim_{k \rightarrow \infty} E \sup_{\boldsymbol{\vartheta}^{(j)} \in V_k(\boldsymbol{\vartheta}_0^{(j)}) \cap \Theta} \left\| \frac{\partial u_{jt}(\boldsymbol{\vartheta}^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}} - \frac{\partial u_{jt}(\boldsymbol{\vartheta}_0^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}} \right\| = 0. \quad (3.18)$$

Note that (2.1)-(2.2) entail that $E(\eta_{jt}^2) = 1$, and thus $Ee^{u_{jt}} = e^{-\tau_j}$. Using (3.15) and (3.17) with the Cauchy-Schwarz inequality, and (3.18) with (3.6) and $Ee^{u_{jt}} = e^{-\tau_j}$, we finally obtain

$$\lim_{k \rightarrow \infty} E \sup_{\boldsymbol{\vartheta}^{(j)} \in V_k(\boldsymbol{\vartheta}_0^{(j)}) \cap \Theta} \left\| \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} e^{u_{jt}(\boldsymbol{\vartheta}^{(j)})} - \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} e^{u_{jt}(\boldsymbol{\vartheta}_0^{(j)})} \right\| = 0. \quad (3.19)$$

Note also that, by (3.6), $u_{jt} = u_{jt}(\boldsymbol{\vartheta}_0^{(j)})$ is independent of \mathcal{F}_{t-1} . In view of the strong consistency established in Theorem 3.1, it follows from (3.19) that for any sequence $\boldsymbol{\vartheta}_n$ between $\widehat{\boldsymbol{\vartheta}}_n^{(j)}$ and $\boldsymbol{\vartheta}_0^{(j)}$, almost surely

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} e^{u_{jt}(\boldsymbol{\vartheta}_n)} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} e^{u_{jt}(\boldsymbol{\vartheta}_0^{(j)})} = e^{-\tau_j} E \frac{\partial u_{jt}(\boldsymbol{\vartheta}_0^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)}}. \quad (3.20)$$

Now it suffices to point out that (3.12) and a Taylor expansion entail

$$\frac{1}{n} \sum_{t=1}^n e^{\widehat{u}_{jt}} = \frac{1}{n} \sum_{t=1}^n e^{u_{jt}} + \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} - \boldsymbol{\vartheta}_0^{(j)} \right)' \frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} e^{u_{jt}(\boldsymbol{\vartheta}_n)} + O\left(\frac{1}{n}\right) = e^{-\tau_j} + o(1).$$

□

Let $\boldsymbol{\zeta}_0^{(j)} = (\omega_{0j}, \alpha_{0j1}, \dots, \alpha_{0jM}, \beta_{0jj}, \lambda_{0j1}, \dots, \lambda_{0jK})'$ be the parameter involved in the j -th equation of the log-GARCH model (2.3). The whole log-GARCH parameter $\boldsymbol{\zeta}_0 = (\boldsymbol{\zeta}_0^{(1)'}, \dots, \boldsymbol{\zeta}_0^{(M)'})'$ is a function of the VARMA-X parameter:

$$\boldsymbol{\zeta}_0 = \Psi(\boldsymbol{\varphi}_0) \quad \text{where} \quad \boldsymbol{\varphi}_0 = (\boldsymbol{\vartheta}'_0, \boldsymbol{\tau}')'.$$

The following result is an immediate consequence of Theorem 3.1 and Lemma 3.2.

Corollary 3.1 *Let $\widehat{\boldsymbol{\zeta}}_n = \Psi(\widehat{\boldsymbol{\varphi}}_n)$, with $\widehat{\boldsymbol{\varphi}}_n = \left(\widehat{\boldsymbol{\vartheta}}_n', \widehat{\boldsymbol{\tau}}_n' \right)'$, and $\widehat{\boldsymbol{\tau}}_n = (\widehat{\tau}_{1n}, \dots, \widehat{\tau}_{Mn})'$. Under the assumptions of Lemma 3.2, $\widehat{\boldsymbol{\zeta}}_n$ is a strongly consistent estimator of $\boldsymbol{\zeta}_0$.*

3.2 Asymptotic normality

We now show the asymptotic normality of the equation-by-equation estimator $\widehat{\boldsymbol{\vartheta}}_n$.

Theorem 3.2 *Suppose that the assumptions of Lemma 3.2 hold and that $\boldsymbol{\vartheta}_0^{(j)}$ belongs to the interior of Θ for $j = 1, \dots, M$. As $n \rightarrow \infty$, we have*

$$\sqrt{n} \left(\widehat{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}_0 \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \mathbf{J}^{-1} \mathbf{I} \mathbf{J}^{-1} \right),$$

where \mathbf{J} is a block diagonal matrix with j -th $d \times d$ block $\mathbf{J}^{(j)} = E \frac{\partial u_{jt}}{\partial \boldsymbol{\vartheta}^{(j)}} \frac{\partial u_{jt}}{\partial \boldsymbol{\vartheta}^{(j)'}} (\boldsymbol{\vartheta}_0^{(j)})$ and \mathbf{I} is a matrix whose block (i, j) of size $d \times d$ is $\mathbf{I}(i, j) = E u_{it} u_{jt} E \frac{\partial u_{it}}{\partial \boldsymbol{\vartheta}^{(j)}} (\boldsymbol{\vartheta}_0^{(i)}) \frac{\partial u_{jt}}{\partial \boldsymbol{\vartheta}^{(j)'}} (\boldsymbol{\vartheta}_0^{(j)})$.

Note that the theorem implies

$$\sqrt{n} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} - \boldsymbol{\vartheta}_0^{(j)} \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left\{ 0, E u_{jt}^2 \left(\mathbf{J}^{(j)} \right)^{-1} \right\} \quad (3.21)$$

as $n \rightarrow \infty$, for $j = 1, \dots, M$.

Proof of Theorem 3.2. Similarly to (3.11), it can be seen that

$$\sup_{\boldsymbol{\vartheta}^{(j)} \in \Theta} \left\| \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} Q_n^{(j)} (\boldsymbol{\vartheta}^{(j)}) - \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} \widetilde{Q}_n^{(j)} (\boldsymbol{\vartheta}^{(j)}) \right\| = O \left(\frac{1}{n} \right) \text{ a.s.} \quad (3.22)$$

By arguments similar to those used to prove (e) on Page 174 in [Francq and Zakoïan \(2010\)](#), one can show that for any sequence $\boldsymbol{\vartheta}_n$ between $\widehat{\boldsymbol{\vartheta}}_n^{(j)}$ and $\boldsymbol{\vartheta}_0^{(j)}$ we have

$$\lim_{n \rightarrow \infty} \frac{\partial^2}{\partial \boldsymbol{\vartheta}^{(j)} \partial \boldsymbol{\vartheta}^{(j)'}} Q_n^{(j)} (\boldsymbol{\vartheta}_n) = 2\mathbf{J}^{(j)} \quad \text{a.s.}$$

The existence of $\mathbf{J}^{(j)}$ comes from (3.13)-(3.14) and (3.2)-(3.3). Therefore, by using a Taylor expansion, we obtain

$$o(1) = \sqrt{n} \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} Q_n^{(j)} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} \right) = \sqrt{n} \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} Q_n^{(j)} \left(\boldsymbol{\vartheta}_0^{(j)} \right) + \sqrt{n} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} - \boldsymbol{\vartheta}_0^{(j)} \right) \left\{ 2\mathbf{J}^{(j)} + o(1) \right\}.$$

The central limit theorem of [Billingsley \(1961\)](#) for stationary and square integrable martingale differences entails that

$$\sqrt{n} \begin{pmatrix} \frac{\partial}{\partial \boldsymbol{\vartheta}^{(1)}} Q_n^{(1)} \left(\boldsymbol{\vartheta}_0^{(1)} \right) \\ \vdots \\ \frac{\partial}{\partial \boldsymbol{\vartheta}^{(M)}} Q_n^{(M)} \left(\boldsymbol{\vartheta}_0^{(M)} \right) \end{pmatrix} = \frac{2}{\sqrt{n}} \sum_{t=1}^n \begin{pmatrix} u_{1t} \frac{\partial u_{1t}}{\partial \boldsymbol{\vartheta}^{(1)}} \\ \vdots \\ u_{Mt} \frac{\partial u_{Mt}}{\partial \boldsymbol{\vartheta}^{(M)}} \end{pmatrix}_{\boldsymbol{\vartheta}_0} \xrightarrow{\mathcal{L}} \mathcal{N} (0, 4\mathbf{I})$$

as $n \rightarrow \infty$. It remains to show that the matrices $\mathbf{J}^{(j)}$ are nonsingular. If $\mathbf{J}^{(j)}$ is singular, then there exists a non zero vector $\boldsymbol{\nu}$ of \mathbb{R}^d such that $\boldsymbol{\nu}' \frac{\partial u_{jt}}{\partial \boldsymbol{\vartheta}^{(j)}} = 0$ a.s. By (3.13), this entails $\boldsymbol{\nu}' \mathbf{d}_t(\boldsymbol{\vartheta}_0^{(j)}) = 0$ a.s. In view of Lemma 3.1, this is contradiction with (3.4). The

conclusion follows. \square

The matrices $\mathbf{J}^{(j)}$ and $\mathbf{I}(i, j)$ can be estimated by

$$\widehat{\mathbf{J}}^{(j)} = \frac{1}{n} \sum_{t=1}^n \widehat{\mathbf{Y}}_{jt} \widehat{\mathbf{Y}}_{jt}', \quad \widehat{\mathbf{I}}(i, j) = \frac{1}{n} \sum_{t=1}^n \widehat{u}_{it} \widehat{u}_{jt} \widehat{\mathbf{Y}}_{it} \widehat{\mathbf{Y}}_{jt}', \quad \text{with } \widehat{\mathbf{Y}}_{jt} = \frac{\partial \widetilde{u}_{jt}}{\partial \boldsymbol{\vartheta}^{(j)}}(\widehat{\boldsymbol{\vartheta}}_n^{(j)}),$$

the $\widehat{\mathbf{Y}}_{jt}$ being recursively computed, for $t = 2, \dots, n$, by

$$\widehat{\mathbf{Y}}_{jt} = \widetilde{\mathbf{d}}_{j,t-1} + \widehat{\beta}_n^{(j)} \widehat{\mathbf{Y}}_{j,t-1}, \quad (3.23)$$

with $\widehat{\beta}_n^{(j)}$ being the estimate of β_{0jj} and

$$\widetilde{\mathbf{d}}_{j,t-1} = (-1, -\ln \epsilon_{1,t-1}^2, \dots, -\ln \epsilon_{M,t-1}^2, \widehat{u}_{j,t-1}, -X_{1,t-1}, \dots, -X_{K,t-1})'$$

and the initial value $\widehat{\mathbf{Y}}_{j1} = \mathbf{0}_d$.

Proposition 3.1 *Under the assumptions of Theorem 3.2, for all $i, j = 1, \dots, M$ we have*

$$\widehat{\mathbf{J}}^{(j)} \rightarrow \mathbf{J} \quad \text{and} \quad \widehat{\mathbf{I}}(i, j) \rightarrow \mathbf{I}(i, j) \quad \text{a.s. as } n \rightarrow \infty.$$

Proof of Proposition 3.1. Using already used arguments, it can be shown that the initial values are unimportant. More precisely, we have $\widehat{\mathbf{I}}(i, j) = \mathbf{I}_n(\widehat{\boldsymbol{\vartheta}}_n) + o(1)$ a.s., with

$$\mathbf{I}_n(\boldsymbol{\vartheta}) = \mathbf{I}_{n,i,j}(\boldsymbol{\vartheta}) = \frac{1}{n} \sum_{t=1}^n u_{it}(\boldsymbol{\vartheta}^{(i)}) u_{jt}(\boldsymbol{\vartheta}^{(j)}) \boldsymbol{\Upsilon}_{it}(\boldsymbol{\vartheta}^{(i)}) \boldsymbol{\Upsilon}_{jt}'(\boldsymbol{\vartheta}^{(j)}), \quad \boldsymbol{\Upsilon}_{jt}(\boldsymbol{\vartheta}^{(j)}) = \frac{\partial u_{jt}}{\partial \boldsymbol{\vartheta}^{(j)}}(\boldsymbol{\vartheta}^{(j)}).$$

By the ergodic theorem $\mathbf{I}_n(\boldsymbol{\vartheta}_0) \rightarrow \mathbf{I}(i, j)$ a.s. as $n \rightarrow \infty$. Since $\widehat{\boldsymbol{\vartheta}}_n \rightarrow \boldsymbol{\vartheta}_0$ a.s., it remains to show that for all $\varepsilon > 0$, there exists a neighbourhood $V(\boldsymbol{\vartheta}_0)$ of $\boldsymbol{\vartheta}_0$ such that

$$\lim_{n \rightarrow \infty} \sup_{\boldsymbol{\vartheta} \in V(\boldsymbol{\vartheta}_0)} \|\mathbf{I}_n(\boldsymbol{\vartheta}) - \mathbf{I}_n^{i,j}(\boldsymbol{\vartheta}_0)\| < \varepsilon. \quad (3.24)$$

Let $V_k(\boldsymbol{\vartheta}_0)$ be the ball of centre $\boldsymbol{\vartheta}_0$ and radius $1/k$. Note that

$$\sup_{\boldsymbol{\vartheta} \in V_k(\boldsymbol{\vartheta}_0)} \|\mathbf{I}_n(\boldsymbol{\vartheta}) - \mathbf{I}_n(\boldsymbol{\vartheta}_0)\| \leq \frac{1}{n} \sum_{t=1}^n x_t(k)$$

where

$$x_t(k) = \sup_{\boldsymbol{\vartheta} \in V_k(\boldsymbol{\vartheta}_0)} \|\mathbf{Y}_t(\boldsymbol{\vartheta}) - \mathbf{Y}_t(\boldsymbol{\vartheta}_0)\|, \quad \mathbf{Y}_t(\boldsymbol{\vartheta}) = u_{it}(\boldsymbol{\vartheta}^{(i)}) u_{jt}(\boldsymbol{\vartheta}^{(j)}) \boldsymbol{\Upsilon}_{it}(\boldsymbol{\vartheta}^{(i)}) \boldsymbol{\Upsilon}_{jt}'(\boldsymbol{\vartheta}^{(j)}).$$

Since $u_{it}(\boldsymbol{\vartheta}^{(i)})$ and the components of $\boldsymbol{\Upsilon}_{it}(\boldsymbol{\vartheta}^{(i)})$ admit moments of order 2, uniformly in Θ (see (3.15)), we have

$$E \sup_{\boldsymbol{\vartheta} \in \Theta^M} \|\mathbf{Y}_t(\boldsymbol{\vartheta})\| < \infty. \quad (3.25)$$

The process $\{x_t(k)\}_t$ being stationary and ergodic, the left-hand side of (3.24) is a.s. bounded by $E x_t(k)$. Noting that $x_t(k) \rightarrow 0$ a.s. as $k \rightarrow \infty$, we obtain (3.24) by the dominated convergence theorem and (3.25). The consistency of $\widehat{\mathbf{I}}(i, j)$ is shown. That of $\widehat{\mathbf{J}}^{(j)}$ is obtained similarly. \square

We now give the asymptotic distribution of all the VARMA-X parameters $\widehat{\boldsymbol{\varphi}}_n = (\widehat{\boldsymbol{\vartheta}}_0, \widehat{\boldsymbol{\tau}})'$.

Theorem 3.3 *Under the assumptions of Theorem 3.2 and Lemma 3.2*

$$\sqrt{n}(\widehat{\boldsymbol{\varphi}}_n - \boldsymbol{\varphi}_0) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \boldsymbol{\Sigma}_\varphi),$$

where $\boldsymbol{\Sigma}_\varphi$ can be estimated by $\widehat{\boldsymbol{\Sigma}}_\varphi = \widehat{\mathbf{M}}\widehat{\boldsymbol{\Sigma}}_\mathbf{r}\widehat{\mathbf{M}}'$ with

$$\widehat{\mathbf{M}} = \begin{pmatrix} -\widehat{\mathbf{J}}^{-1} & \mathbf{0}_{M_d \times M} \\ \widehat{\mathbf{D}}\widehat{\mathbf{J}}^{-1} & \widehat{\mathbf{E}} \end{pmatrix}, \quad \widehat{\mathbf{D}} = \begin{pmatrix} \widehat{\mathbf{D}}'_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \widehat{\mathbf{D}}'_M \end{pmatrix} \quad \text{with } \widehat{\mathbf{D}}_j = \frac{1}{n} \sum_{t=1}^n \widehat{\boldsymbol{\Upsilon}}_{jt},$$

$\widehat{\mathbf{E}} = -\text{diag}(e^{\widehat{\tau}_1}, \dots, e^{\widehat{\tau}_M})$ and $\widehat{\boldsymbol{\Sigma}}_\mathbf{r} = \frac{1}{n} \sum_{t=1}^n \widehat{\boldsymbol{\Upsilon}}_t \widehat{\boldsymbol{\Upsilon}}_t'$ with

$$\widehat{\boldsymbol{\Upsilon}}_t = \left(\widehat{u}_{1t} \widehat{\boldsymbol{\Upsilon}}'_{1t}, \dots, \widehat{u}_{Mt} \widehat{\boldsymbol{\Upsilon}}'_{Mt}, e^{\widehat{u}_{1t}} - e^{-\widehat{\tau}_1}, \dots, e^{\widehat{u}_{Mt}} - e^{-\widehat{\tau}_M} \right)'$$

Proof of Theorem 3.3. We first show that the initial values are asymptotically negligible. In view of (3.12) and the almost sure convergence of $\frac{1}{n} \sum_{t=1}^n e^{\widehat{u}_{jt}}$ to $e^{-\tau_1} > 0$, a Taylor expansion shows that

$$\left| \widehat{\tau}_{jn} + \ln \frac{1}{n} \sum_{t=1}^n e^{u_{jt}(\widehat{\boldsymbol{\vartheta}}_n^{(j)})} \right| \leq K \left| \frac{1}{n} \sum_{t=1}^n e^{u_{jt}(\widehat{\boldsymbol{\vartheta}}_n^{(j)})} - \frac{1}{n} \sum_{t=1}^n e^{\widetilde{u}_{jt}(\widehat{\boldsymbol{\vartheta}}_n^{(j)})} \right| = o(n^{-1/2}) \quad \text{a.s.}$$

Doing again a Taylor expansion, we obtain

$$\ln \frac{1}{n} \sum_{t=1}^n e^{u_{jt}(\widehat{\boldsymbol{\vartheta}}_n^{(j)})} = \ln \frac{1}{n} \sum_{t=1}^n e^{u_{jt}(\boldsymbol{\vartheta}_0^{(j)})} + \frac{1}{\frac{1}{n} \sum_{t=1}^n e^{u_{jt}(\boldsymbol{\vartheta}_n)} \frac{1}{n} \sum_{t=1}^n \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)'}} e^{u_{jt}(\boldsymbol{\vartheta}_n)} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} - \boldsymbol{\vartheta}_0^{(j)} \right)$$

for some sequence $\boldsymbol{\vartheta}_n$ between $\widehat{\boldsymbol{\vartheta}}_n^{(j)}$ and $\boldsymbol{\vartheta}_0^{(j)}$. In view of (3.20) and Theorem 3.2, it follows that

$$\sqrt{n}(\widehat{\tau}_{jn} - \tau_j) = \sqrt{n} \left(-\ln \frac{1}{n} \sum_{t=1}^n e^{u_{jt}} - \tau_j \right) - \mathbf{D}'_j \sqrt{n} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} - \boldsymbol{\vartheta}_0^{(j)} \right) + o_P(1),$$

where $\mathbf{D}_j = E \frac{\partial u_{j1}(\boldsymbol{\vartheta}_0^{(j)})}{\partial \boldsymbol{\vartheta}^{(j)'}}$. Now, using Lemma 3.2 we have

$$\ln \frac{1}{n} \sum_{t=1}^n e^{u_{jt}} = -\tau_j + \frac{1}{e^{-\tau_j} + o(1)} \frac{1}{n} \sum_{t=1}^n (e^{u_{jt}} - e^{-\tau_j}).$$

Putting the results together, we obtain

$$\sqrt{n}(\widehat{\tau}_{jn} - \tau_j) = -e^{\tau_j} \frac{1}{\sqrt{n}} \sum_{t=1}^n (e^{u_{jt}} - e^{-\tau_j}) - \mathbf{D}'_j \sqrt{n} (\widehat{\boldsymbol{\vartheta}}_n^{(j)} - \boldsymbol{\vartheta}_0^{(j)}) + o_P(1).$$

From the proof of Theorem 3.2, it follows that

$$\sqrt{n}(\widehat{\boldsymbol{\varphi}}_n - \boldsymbol{\varphi}_0) = \mathbf{M} \frac{1}{\sqrt{n}} \sum_{t=1}^n \boldsymbol{\Upsilon}_t + o_P(1)$$

where

$$\boldsymbol{\Upsilon}_t = \begin{pmatrix} u_{1t} \frac{\partial u_{1t}}{\partial \boldsymbol{\vartheta}^{(1)}} \\ \vdots \\ u_{Mt} \frac{\partial u_{Mt}}{\partial \boldsymbol{\vartheta}^{(M)}} \\ e^{u_{1t}} - e^{-\tau_1} \\ \vdots \\ e^{u_{Mt}} - e^{-\tau_M} \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} -\mathbf{J}^{-1} & \mathbf{0}_{Md \times M} \\ \mathbf{D}\mathbf{J}^{-1} & \mathbf{E} \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} \mathbf{D}'_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \mathbf{D}'_M \end{pmatrix}$$

and $\mathbf{E} = -\text{diag}(e^{\tau_1}, \dots, e^{\tau_M})$. By already given arguments, $\frac{1}{\sqrt{n}} \sum_{t=1}^n \boldsymbol{\Upsilon}_t \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\Upsilon}})$, where

$$\boldsymbol{\Sigma}_{\boldsymbol{\Upsilon}} = \begin{pmatrix} \mathbf{I} & \boldsymbol{\Sigma}_{\boldsymbol{\vartheta}, \tau} \\ \boldsymbol{\Sigma}'_{\boldsymbol{\vartheta}, \tau} & \boldsymbol{\Sigma}_{\tau} \end{pmatrix}, \quad \boldsymbol{\Sigma}_{\boldsymbol{\vartheta}, \tau} = \begin{pmatrix} \text{cov}(u_{1t}, e^{u_{1t}})D_1 & \cdots & \text{cov}(u_{1t}, e^{u_{Mt}})D_1 \\ \vdots & & \\ \text{cov}(u_{Mt}, e^{u_{1t}})D_M & \cdots & \text{cov}(u_{Mt}, e^{u_{Mt}})D_M \end{pmatrix}$$

and

$$\boldsymbol{\Sigma}_{\tau} = \begin{pmatrix} \text{cov}(e^{u_{it}}, e^{u_{jt}}) \end{pmatrix}.$$

□

To deduce the asymptotic distribution of $\widehat{\boldsymbol{\zeta}}_n$ from that of $\widehat{\boldsymbol{\varphi}}_n$ we need to compute the $Md \times M(d+1)$ matrix $\frac{\partial \Psi(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'}$. For instance, in the case $K=1$ and $M=2$, we have

$$\frac{\partial \Psi(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'} = \begin{pmatrix} 1 & 0 & 0 & \tau_1 & 0 & 0 & 0 & 0 & 0 & 0 & \beta_{011} - 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \tau_2 & 0 & 0 & \beta_{022} - 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

More generally, for $j = 1, \dots, M$, the line $d(j-1) + 1$ of that matrix is given by

$$\frac{\partial \Psi_{d(j-1)+1}(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'} = \left(0'_{d(j-1)} \quad 1 \quad 0_M \quad \tau_j \quad 0_K \quad 0_{d(M-j)} \quad 0_{j-1} \quad \beta_{0jj} - 1 \quad 0_{M-j} \right),$$

the line $d(j-1) + 1 + j$ is given by

$$\frac{\partial \Psi_{d(j-1)+1+j}(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'} = \left(0'_{d(j-1)} \quad 0_j \quad 1 \quad 0_{M-j} \quad -1 \quad 0_K \quad 0_{d(M-j)} \quad 0_M \right),$$

the line $d(j-1) + 1 + k$ for $k = 1, \dots, j-1$ or $k = j+1, \dots, M$ is given by

$$\frac{\partial \Psi_{d(j-1)+1+k}(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'} = \left(0'_{d(j-1)} \quad 0_k \quad 1 \quad 0_{M-k} \quad 0 \quad 0_K \quad 0_{d(M-j)} \quad 0_M \right),$$

the line $d(j-1) + M + 2$ is given by

$$\frac{\partial \Psi_{d(j-1)+M+2}(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'} = \left(0'_{d(j-1)} \quad 0_{M+1} \quad 1 \quad 0_K \quad 0_{d(M-j)} \quad 0_M \right),$$

and the line $d(j-1) + M + 2 + k$ for $k = 1, \dots, K$ is given by

$$\frac{\partial \Psi_{d(j-1)+M+2+k}(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'} = \left(0'_{d(j-1)} \quad 0_{M+k+1} \quad 1 \quad 0_{K-k} \quad 0_{d(M-j)} \quad 0_M \right).$$

A direct application of the delta method then gives the following result.

Corollary 3.2 *Under the assumptions of Theorem 3.3*

$$\sqrt{n} \left(\widehat{\boldsymbol{\zeta}}_n - \boldsymbol{\zeta}_0 \right) \xrightarrow{L} \mathcal{N} \left(0, \boldsymbol{\Sigma}_{\boldsymbol{\zeta}} := \frac{\partial \Psi(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}'} \boldsymbol{\Sigma}_{\boldsymbol{\varphi}} \frac{\partial \Psi'(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}} \right),$$

as $n \rightarrow \infty$.

3.3 Constrained models

The asymptotic results of the previous section can readily be used to test the significance of the log-GARCH parameters (see Section 4.1 for an illustration). Such tests may lead to the estimation of new models in which some coefficients are constrained to be zero. Similarly, even if estimation of a large system of dimension M is performed equation-by-equation, a general model of the form (2.3) remains intractable when M is large (in each equation, the number of parameters is $M + K + 2$). For these reasons, one could want to impose constraints, such as a diagonal or block-diagonal form for the matrix $\boldsymbol{\alpha}_0$ (see Section 4.2 for an illustration).

A way to introduce very general constraints is to assume that the VARMA-X model (2.4) is parametrised by a vector $\boldsymbol{\vartheta}_0$, that does not necessarily correspond to the parameter defined in Section 3.1, and may be of lower dimension. We thus assume that $\boldsymbol{c}_0 = \boldsymbol{c}(\boldsymbol{\vartheta}_0)$, $\boldsymbol{\phi}_0 = \boldsymbol{\phi}(\boldsymbol{\vartheta}_0)$, $\boldsymbol{\beta}_0 = \boldsymbol{\beta}(\boldsymbol{\vartheta}_0)$ and $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}(\boldsymbol{\vartheta}_0)$, and that $\boldsymbol{\vartheta}_0 = \left(\boldsymbol{\vartheta}_0^{(1)'} , \dots , \boldsymbol{\vartheta}_0^{(M)'} \right)'$ where $\boldsymbol{\vartheta}_0^{(j)}$

is the vector of the unknown parameters involved in the j -th VARMA-X equation. The parameter $\boldsymbol{\vartheta}_0^{(j)}$ belongs to some compact parameter space $\Theta^{(j)}$, whose generic element $\boldsymbol{\vartheta}^{(j)}$ has typically less than $M + K + 2$ components. If, for instance, the matrix $\boldsymbol{\alpha}_0$ is assumed to be diagonal, then one can set

$$\boldsymbol{\vartheta}^{(j)} = (c_j, \alpha_{jj} + \beta_{jj}, \beta_{jj}, \lambda_{j1}, \dots, \lambda_{jK})', \quad \Theta^{(j)} = \Theta \subset \mathbb{R}^{K+3}.$$

In the general case, it is assumed that the parametrisation satisfies

$$\text{the application } \boldsymbol{\vartheta} \mapsto \{\mathbf{c}(\boldsymbol{\vartheta}), \boldsymbol{\phi}(\boldsymbol{\vartheta}), \boldsymbol{\beta}(\boldsymbol{\vartheta}), \boldsymbol{\lambda}(\boldsymbol{\vartheta})\} \text{ is injective} \quad (3.26)$$

and

$$\text{admits continuous third order derivatives in } \Theta^{(1)} \times \dots \times \Theta^{(M)}. \quad (3.27)$$

Assume also that, denoting by $\beta_{jj}(\boldsymbol{\vartheta})$ the j -th diagonal term of $\boldsymbol{\beta}(\boldsymbol{\vartheta})$,

$$|\beta_{jj}(\boldsymbol{\vartheta})| < 1, \quad \forall \boldsymbol{\vartheta} \in \Theta^{(1)} \times \dots \times \Theta^{(M)}. \quad (3.28)$$

With this change of notation, the estimator of $\boldsymbol{\vartheta}_0^{(j)}$ can still be defined by (3.8), replacing Θ by $\Theta^{(j)}$ if necessary. The asymptotic behaviour of $\widehat{\boldsymbol{\vartheta}}_n^{(j)}$ is unchanged.

More precisely, the strong consistency of $\widehat{\boldsymbol{\vartheta}}_n^{(j)}$ to $\boldsymbol{\vartheta}_0^{(j)}$ holds true under the previous assumptions (3.1), (3.2), (3.3), (3.4), (3.6), the invertibility condition (3.28) and the identifiability conditions (3.26) and $\alpha_{jj}(\boldsymbol{\vartheta}_0) \neq 0$. The asymptotic normality of Theorem 3.2 continues to hold under the additional assumption that $\boldsymbol{\vartheta}_0^{(j)}$ belongs to the interior of $\Theta^{(j)}$ for all j , and under the smoothness condition (3.27). The output of Proposition 3.1 remains also valid if, in (3.23), the definition of $\widetilde{\boldsymbol{d}}_{j,t-1}$ is changed into

$$\begin{aligned} \widetilde{\boldsymbol{d}}_{j,t-1} = & -\frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} c_j \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} \right) - \sum_{\ell=1}^M \ln \epsilon_{\ell,t-1}^2 \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} \phi_{j\ell} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} \right) \\ & - \sum_{\ell=1}^K X_{\ell,t-1} \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} \lambda_{j\ell} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} \right) + \widetilde{u}_{j,t-1} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} \right) \frac{\partial}{\partial \boldsymbol{\vartheta}^{(j)}} \beta_{jj} \left(\widehat{\boldsymbol{\vartheta}}_n^{(j)} \right). \end{aligned}$$

With this modification, Theorem 3.3 directly applies.

As an illustration, consider for instance the case where $\boldsymbol{\alpha}_0$ is assumed to be diagonal. Moreover, assume that one wants to estimate the model with $K = 2$ covariates under the constraint $\lambda_{j1} = \lambda_{j2}$ for all j . The condition (3.26)-(3.27) is satisfied with $\boldsymbol{\vartheta}^{(j)} = (c_j, \phi_{jj}, \beta_{jj}, \lambda_{j1}) = (c, \phi_j, \beta, \lambda_1)$. If $\Theta^{(j)}$ is assumed to be a compact subset of $\mathbb{R}^2 \times (-1, 1) \times \mathbb{R}$, the condition (3.28) also holds true. In particular, we have (3.21) where $\boldsymbol{J}^{(j)}$ is a 4×4 matrix which can be consistently estimated as in Proposition 3.1 with

$$\widetilde{\boldsymbol{d}}_{j,t-1} = (-1, -\ln \epsilon_{j,t-1}^2, \widehat{u}_{j,t-1}, -X_{1,t-1} - X_{2,t-1})'.$$

4 Empirical applications

4.1 Volatility spillover

How and to what extent volatility in one financial markets may spill over to others is of importance for both policymakers and investors. In ordinary multivariate GARCH models the study of volatility spillover rests upon complicated conditions on the parameters, and on the restrictive assumption that the DCCs are constant, see [Conrad and Weber \(2013\)](#), and [Pedersen \(2015\)](#). By contrast, in our model complicated assumptions on the parameters are not required (due to the exponential volatility specification), and our tests are valid under time-varying DCCs of unknown form. Indeed, a variety of tests can be undertaken either equation-by-equation, or jointly for the whole system. Furthermore, we can also test whether covariates (e.g. volatility proxies) provide additional – or alternative – channels of volatility spillover.

We illustrate the flexibility and generality of the results from the previous section in a study of how European and US stock market volatility affect each other. For illustration purposes we restrict our attention to only two indices, the FTSE100 and the SP100, and initially we only include two variables in the X-part: The log of a range-based volatility proxy for FTSE100 and SP100, respectively. The range, i.e. the difference between the maximum and minimum prices, is often available, and constitutes a volatility proxy, see e.g. [Parkinson \(1980\)](#) and [Garman and Klass \(1980\)](#). The log of our range-based volatility-proxy is computed as $\ln[(high_t - low_t) \cdot 100]^2$, where $high_t$ is the natural log of the maximum value during day t , and where low_t is the natural log of the minimum value during day t . The source of our data is Bloomberg and goes from 2 January 1998 to 1 June 2015, a total 4297 observations before differencing and lagging. Initially, before we add more variables to the X-part, all of our estimated models will be contained in the two-dimensional log-GARCH(1,1)-X model

$$\begin{aligned} \ln \sigma_{1,t}^2 &= \omega_{01} + \alpha_{011} \ln \epsilon_{1,t-1}^2 + \alpha_{012} \ln \epsilon_{2,t-1}^2 + \beta_{01} \ln \sigma_{1,t-1}^2 \\ &\quad + \lambda_{011} X_{1,t-1} + \lambda_{012} X_{2,t-1}, \end{aligned} \tag{4.1}$$

$$\begin{aligned} \ln \sigma_{2,t}^2 &= \omega_{02} + \alpha_{021} \ln \epsilon_{1,t-1}^2 + \alpha_{022} \ln \epsilon_{2,t-1}^2 + \beta_{02} \ln \sigma_{2,t-1}^2 \\ &\quad + \lambda_{021} X_{1,t-1} + \lambda_{022} X_{2,t-1}, \end{aligned} \tag{4.2}$$

where $\epsilon_{1,t}$ and $\epsilon_{2,t}$ denote daily European and US return (in percent) at day t , and $X_{1,t}$ and $X_{2,t}$ are the logs of the volatility proxies at day t . The VARMA-X representation of this model is

$$\begin{aligned} \ln \epsilon_{1,t}^2 &= c_{01} + \phi_{011} \ln \epsilon_{1,t-1}^2 + \alpha_{012} \ln \epsilon_{2,t-1}^2 - \beta_{01} u_{1,t-1} \\ &\quad + \lambda_{011} X_{1,t-1} + \lambda_{012} X_{2,t-1} + u_{1,t}, \end{aligned} \tag{4.3}$$

$$\begin{aligned} \ln \epsilon_{2,t}^2 &= c_{02} + \alpha_{021} \ln \epsilon_{1,t-1}^2 + \phi_{022} \ln \epsilon_{2,t-1}^2 - \beta_{02} u_{2,t-1} \\ &\quad + \lambda_{021} X_{1,t-1} + \lambda_{022} X_{2,t-1} + u_{2,t}, \end{aligned} \tag{4.4}$$

where, for equation $j = 1, 2$,

$$c_{0j} = \omega_{0j} + (1 - \beta_{0j})E(\ln \eta_{j,t}^2), \quad \phi_{0jj} = \alpha_{0jj} + \beta_{0j} \quad \text{and} \quad u_{j,t} = \ln \eta_{j,t}^2 - E(\ln \eta_{j,t}^2).$$

We start by estimating two univariate log-GARCH(1,1) models for comparison purposes:

$$\ln \hat{\sigma}_{1,t}^2 = 0.066 + \frac{0.047}{(0.006)} \ln \epsilon_{1,t-1}^2 + \frac{0.943}{(0.008)} \ln \hat{\sigma}_{1,t-1}^2, \quad (4.5)$$

$$\ln \hat{\sigma}_{2,t}^2 = 0.070 + \frac{0.046}{(0.006)} \ln \epsilon_{2,t-1}^2 + \frac{0.947}{(0.007)} \ln \hat{\sigma}_{2,t-1}^2. \quad (4.6)$$

The numbers in parentheses are standard errors of the estimates (we explain how they are computed below). The ARCH and GARCH estimates are in the usual range: The ARCH parameters are close to 0.05, the GARCH parameters are close to 0.95 and their sum in each equation is close to 1. Next we estimate the two-dimensional log-GARCH(1,1)-X model using our EBEE, which gives

$$\begin{aligned} \ln \hat{\sigma}_{1,t}^2 &= -0.191 - \frac{0.020}{(0.014)} \ln \epsilon_{1,t-1}^2 - \frac{0.012}{(0.014)} \ln \epsilon_{2,t-1}^2 + \frac{0.674}{(0.048)} \ln \hat{\sigma}_{1,t-1}^2 \\ &\quad + \frac{0.158}{(0.032)} X_{1,t-1} + \frac{0.172}{(0.042)} X_{2,t-1}, \end{aligned} \quad (4.7)$$

$$\begin{aligned} \ln \hat{\sigma}_{2,t}^2 &= -0.232 - \frac{0.010}{(0.013)} \ln \epsilon_{1,t-1}^2 - \frac{0.056}{(0.014)} \ln \epsilon_{2,t-1}^2 + \frac{0.751}{(0.032)} \ln \hat{\sigma}_{2,t-1}^2 \\ &\quad + \frac{0.081}{(0.030)} X_{1,t-1} + \frac{0.258}{(0.035)} X_{2,t-1}. \end{aligned} \quad (4.8)$$

It is noteworthy that all ARCH effects become negative – currently there are no QMLE results for ordinary (i.e. non-exponential) GARCH models in the presence of negative ARCH effects (Pedersen (2015)), and that only one of the ARCH effects – that of SP100 on its own log-volatility – is significant according to usual significance levels. It is also noteworthy that the GARCH effects fall to 0.674 and 0.751, respectively. This is in line with the findings of Lamoureux and Lastrapes (1990). Both volatility proxies are significant in both equations according to t -tests at usual significance levels, so all-in-all the t -tests suggest the volatility spill-over is via the volatility proxies, not the ARCH effects.

The standard errors can be computed in at least two way. The first, which is usually the simplest in practice, exploits that numerical software often provides utility functions for the numerical computation of the Hessian of the criterion function. In numerical software the least squares estimator of an ARMA-X specification is typically implemented by minimising $\sum_{t=1}^n \tilde{u}_{jt}^2(\boldsymbol{\vartheta}^{(j)})$ (rather than the average $Q_n^{(j)}(\boldsymbol{\vartheta}^{(j)}) = n^{-1} \sum_{t=1}^n \tilde{u}_{jt}^2(\boldsymbol{\vartheta}^{(j)})$). Let $\widehat{\mathbf{H}}_n^{(j)}(\widehat{\boldsymbol{\vartheta}}_n^{(j)})$ denote a numerical estimate of the Hessian of the criterion function $\sum_{t=1}^n \tilde{u}_{jt}^2(\widehat{\boldsymbol{\vartheta}}_n^{(j)})$ about the least squares estimate $\widehat{\boldsymbol{\vartheta}}_n^{(j)}$. This means $n^{-1} \widehat{\mathbf{H}}_n^{(j)}(\widehat{\boldsymbol{\vartheta}}_n^{(j)})$ provides an estimate of $2J^{(j)}$, since (see the proof of Theorem 3.2)

$$\lim_{n \rightarrow \infty} \frac{\partial^2}{\partial \boldsymbol{\vartheta}^{(j)} \partial \boldsymbol{\vartheta}^{(j)'}} Q_n^{(j)}(\boldsymbol{\vartheta}_n) = 2J^{(j)} \quad \text{a.s.}$$

The expression

$$\left(\frac{1}{n} \sum_{t=1}^n \tilde{u}_{jt}^2 \right) \cdot 2n \cdot \left(\widehat{\mathbf{H}}_n^{(j)}(\widehat{\boldsymbol{\vartheta}}_n^{(j)}) \right)^{-1}$$

thus provides an estimate of the asymptotic variance-covariance matrix $Eu_{jt}^2 (J^{(j)})^{-1}$ for equation j . Finally, to obtain the empirical variance-covariance matrix of the log-GARCH-X parameters in equation j , the relationships between the ARMA-X and log-GARCH-X parameters are used. The standard errors in all of the equations above have been computed in this way.³ The second way the standard errors can be computed is by using the formulas following Theorem 3.2, and whose strong consistency is ensured by Proposition 3.1. This is necessary if the joint variance-covariance matrix of all the parameters in the M -dimensional VARMA-X system of equations is needed. When studying volatility spillover effects we are indeed interested in the joint variance-covariance of all the parameters, so now we use these formulas instead. The null-hypothesis of no spillover-effects between European and US markets corresponds to

$$H_0 : \alpha_{012} = \lambda_{012} = \alpha_{021} = \lambda_{021} = 0. \quad (4.9)$$

These are linear restrictions on a subset of the parameters of the VARMA-X representation. The associated Wald-statistic is distributed as a Chi-squared with 4 degrees of freedom, and turns out to be huge: 7601.2. So the null of no spill-over is resoundingly rejected at common significance levels.

In empirical work it is usually not of interest to test the log-volatility intercepts ω_{01} and ω_{02} . There is, however, one exception: When one would like to inquire whether a volatility proxy is a "perfect" predictor of volatility. That is, whether we have that $\ln \sigma_{j,t}^2 = X_{j,t-1}$. This would imply that there is absolutely no spillover. In the first equation of the two-dimensional log-GARCH(1,1)-X specification this amounts to a test of whether

$$H_0 : (\omega_{01}, \alpha_{011}, \alpha_{012}, \beta_{01}, \lambda_{011}, \lambda_{012})' = (0, 0, 0, 0, 1, 0)'$$

Since the log-volatility intercept ω_{01} is one of the restricted parameters, we need the results of Theorem 3.3 and computations similar to those in Corollary 3.2. Theorem 3.3 implies, for the ARMA-X parameter $\varphi_0^{(j)} = (\vartheta_0^{(j)}, \tau_j)'$ of equation j , that

$$\sqrt{n} \left(\widehat{\varphi}_n^{(j)} - \varphi_0^{(j)} \right) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \Sigma_{\varphi_j} \right).$$

In terms of the ARMA-X parameters of equation 1 the claims in H_0 correspond to a set of restrictions $a_1(\varphi_0^{(1)}) = \mathbf{0}$, with

$$a_1(\varphi_0^{(1)}) = (c_1 - (1 - \beta_{11})\tau_1, \phi_{11} - \beta_{11}, \phi_{12}, \beta_{11}, \lambda_{11} - 1, \lambda_{12})'$$

The first of these restrictions is non-linear, so the matrix of first derivatives with respect

³The R package `lgarch` version 0.6, see [Sucarrat \(2015\)](#), was used.

to the ARMA-X parameter $\varphi_0^{(1)}$ is

$$\frac{\partial a_1(\varphi^{(1)})}{\partial \varphi^{(1)}} = \begin{pmatrix} 1 & 0 & 0 & \tau_1 & 0 & 0 & (\beta_1 - 1) \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

The Wald statistic, which is distributed as a Chi-squared with 6 degrees of freedom, is thus

$$n a_1(\hat{\varphi}_0^{(1)})' \left(\frac{\partial a_1(\hat{\varphi}^{(1)})}{\partial \varphi^{(1)}} \hat{\Sigma}_{\varphi^{(j)}} \frac{\partial a_1(\hat{\varphi}^{(1)})}{\partial \varphi^{(1)'}} \right)^{-1} a_1(\hat{\varphi}_0^{(1)}) = 30796.09.$$

The value is huge, so the hypothesis that the range-based volatility proxy is a “perfect” predictor of volatility in equation 1 is resoundingly rejected.

To further illustrate the computational attraction of our estimator we add more variables to the X-part. It is often the case that additional explanatory information is readily available, for example volume and leverage. For ordinary or non-exponential GARCH models strong non-negative restrictions on parameters are needed if one were to include this additional information. In the multivariate log-GARCH(1,1)-X model, by contrast, we readily obtain the following estimates by means of our EBEE:

$$\begin{aligned} \ln \hat{\sigma}_{1,t}^2 &= -0.407 - 0.011 \ln \epsilon_{1,t-1}^2 - 0.005 \ln \epsilon_{2,t-1}^2 + 0.808 \ln \hat{\sigma}_{1,t-1}^2 + 0.118 X_{1,t-1} \\ &\quad + 0.066 X_{2,t-1} + 0.143 X_{3,t-1} + 0.210 X_{4,t-1} + 0.004 X_{5,t-1} + 0.002 X_{6,t-1} \\ &\quad - 0.058 X_{7,t-1} + 0.269 X_{8,t-1} \end{aligned} \quad (4.10)$$

$$\begin{aligned} \ln \hat{\sigma}_{1,t}^2 &= -0.202 - 0.002 \ln \epsilon_{1,t-1}^2 - 0.050 \ln \epsilon_{2,t-1}^2 + 0.772 \ln \hat{\sigma}_{1,t-1}^2 + 0.070 X_{1,t-1} \\ &\quad + 0.204 X_{2,t-1} + 0.252 X_{3,t-1} + 0.261 X_{4,t-1} - 0.082 X_{5,t-1} + 0.074 X_{6,t-1} \\ &\quad - 0.018 X_{7,t-1} - 0.215 X_{8,t-1}. \end{aligned} \quad (4.11)$$

The $X_{3,t}$ and $X_{4,t}$ are leverage-terms defined as $I_{\eta_{1,t} < 0}$ and $I_{\eta_{2,t} < 0}$, $X_{5,t}$ and $X_{6,t}$ are the log of volume, and $X_{7,t}$ and $X_{8,t}$ are the first-difference of log-volume. The leverage-terms are highly significant at usual significant levels in both equations according to t -tests, but most of the other additional X-variables are not ($X_{5,t}$ in equation 2 is the only significant term at common significance levels).

4.2 A 50-dimensional log-GARCH-X model with time-varying correlations

Estimation of large multivariate GARCH models is plagued by the curse of dimensionality: As the dimension grows, it becomes computationally infeasible to reliably estimate the parameters jointly. Our estimator sidesteps this problem by estimating the system

equation-by-equation. To illustrate we estimate a constrained version (see Section 3.3) of an M -dimensional log-GARCH-X model of the returns of the stocks that make up the Eurostoxx50 index. The source of our individual stock price data is Bloomberg and goes from 2 January 1998 to 11 June 2015, a total of up to 4474 observations before differencing, lagging and data-cleaning (e.g. removal of missing X-values, etc.). Estimation of the 50 equations is fast, since it takes less than a minute on a seven year old laptop.⁴ The j th. equation is given by

$$\begin{aligned} \ln \sigma_{jt}^2 = & \omega_j + \alpha_{j1} \ln \epsilon_{j,t-1}^2 + \beta_j \ln \sigma_{j,t-1}^2 + \lambda_{j1} X_{j1,t-1} \\ & + \lambda_{j2} X_{2,t-1} + \lambda_{j3} X_{3,t-1} + \lambda_{j4} X_{4,t-1} + \lambda_{j5} X_{5,t-1}, \end{aligned} \quad (4.12)$$

where $X_{j1,t}$ is a leverage-term for return $\epsilon_{j,t}$ computed as $X_{j1,t} = I_{\eta_{1,t} < 0}$, $X_{2,t}$ is the Eurostoxx50 return, $X_{3,t}$ is a leverage-term for the Eurostoxx50 return, $X_{4,t}$ is the log-volume of Eurostoxx50 and $X_{5,t}$ is a range-based volatility proxy of Eurostoxx50 return (computed in the same way as earlier). The constraint we impose on the multivariate system is thus that α_0 is diagonal.⁵ Table 1 contains the results, and an asterisk (*) indicates significance at the 10% level. The ARCH (α_{j1}), GARCH (β_j) and leverage (λ_{j1}) parameters behave as expected. All the ARCH parameters are positive, in the 0.01 to 0.06 range and significant (except that of $j = 22$) at 10%, all the GARCH parameters are positive, in the 0.79 to 0.97 range and significant at 10%, and all the leverage parameters are positive and significant except for $j = 19, 22, 30, 37$ and 50. The first unexpected result is the impact of $X_{2,t}$, which is defined as $\ln \epsilon_t^2$ where ϵ_t is the daily log-return (in percent) of the Eurostoxx50 index. The nineteen terms that are significant all have a negative effect, while one would maybe have expected a positive one. The effect is small, however, since the largest – in absolute value – is 0.023 (for $j = 8$). An economic explanation for the negative impact is, possibly, that higher Eurostoxx50 volatility reduces volatility for certain types of stocks. Further investigation is needed in order to shed light on this hypothesis. The leverage term of the Eurostoxx50 index, $X_{3,t}$, is positive and significant for all but 1 (when $j = 27$) out of the 50 stocks, and economically quite substantial when compared with the ARCH effects (i.e. α_{j1} and λ_{0j2}). Twenty-eight of the stocks are significantly affected by Eurostoxx50 volume, $X_{4,t}$, and in all cases but one (when $j = 28$) is the effect negative. In other words, higher volume – i.e. greater liquidity – seems on average to reduce volatility. Finally, the Eurostoxx50 volatility proxy, $X_{5,t}$, has a positive and significant impact for all stocks, except $j = 4$ and $j = 46$.

To obtain estimates of the conditional correlations we fit the corrected DCC (cDCC) model of Aielli (2013), which is a modified version of Engle’s (2002) DCC. The cDCC model is given by

$$\mathbf{R}_t = \mathbf{Q}_t^{*-1/2} \mathbf{Q}_t \mathbf{Q}_t^{*-1/2}, \quad \mathbf{Q}_t = (1 - \gamma_0 - \delta_0) \mathbf{S}_0 + \gamma_0 \mathbf{Q}_{t-1}^{*1/2} \boldsymbol{\eta}_{t-1} \boldsymbol{\eta}'_{t-1} \mathbf{Q}_{t-1}^{*1/2} + \delta_0 \mathbf{Q}_{t-1} \quad (4.13)$$

where $\gamma_0, \delta_0 \geq 0$, $\gamma_0 + \delta_0 < 1$, \mathbf{S}_0 is a correlation matrix, \mathbf{Q}_t^* is a diagonal matrix with the elements from the diagonal of \mathbf{Q}_t and $\boldsymbol{\eta}_t = \mathbf{D}_t^{-1} \boldsymbol{\epsilon}_t$. Estimation of γ_0 and δ_0 by Gaussian

⁴The model is estimated with the R package `lgarch` version 0.6 under R version 3.1.0 running on Windows 7 (64-bit) on a Lenovo X61s.

⁵Otherwise each of the j equations would have $M + K + 2 = 57$ parameters.

QML yields the estimator

$$(\hat{\gamma}, \hat{\delta}) = \arg \min_{(\gamma, \delta)} \sum_{t=1}^n \left(\ln |\hat{\mathbf{R}}_t| + \hat{\boldsymbol{\eta}}_t' \hat{\mathbf{R}}_t^{-1} \hat{\boldsymbol{\eta}}_t \right), \quad (4.14)$$

where $\hat{\boldsymbol{\eta}}_t$ are the standardised residuals from the 50 log-GARCH models,

$$\begin{aligned} \hat{\mathbf{R}}_t &= \hat{\mathbf{Q}}_t^{*-1/2} \hat{\mathbf{Q}}_t \hat{\mathbf{Q}}_t^{*-1/2}, \quad \hat{\mathbf{Q}}_t = (1 - \hat{\gamma} - \hat{\delta}) \mathbf{S}_n + \hat{\gamma} \hat{\mathbf{Q}}_{t-1}^{*1/2} \hat{\boldsymbol{\eta}}_{t-1} \hat{\boldsymbol{\eta}}_{t-1}' \hat{\mathbf{Q}}_{t-1}^{*1/2} + \hat{\delta} \hat{\mathbf{Q}}_{t-1} \\ \mathbf{S}_n &= \frac{1}{n} \sum_{t=1}^n \hat{\mathbf{Q}}_t^{*1/2} \hat{\boldsymbol{\eta}}_t \hat{\boldsymbol{\eta}}_t' \hat{\mathbf{Q}}_t^{*1/2}, \quad \hat{\mathbf{Q}}_t^* = \text{diag}(\hat{q}_{11t}, \dots, \hat{q}_{mmt}) \\ \hat{q}_{iit} &= (1 - \hat{\gamma} - \hat{\delta}) + \hat{\gamma} \hat{\eta}_{i,t-1}^2 + \hat{\delta} \hat{q}_{i,t-1} \quad \text{for } i = 1, \dots, m. \end{aligned}$$

The estimates of γ_0 and δ_0 are 0.003 and 0.959, respectively, which suggests the correlations are very persistent. Figure 1 contains histograms of selected descriptive statistics of the $50 \cdot 49/2 = 1225$ fitted correlations paths. Graphically the empirical distribution of the unconditional correlations is bell-shaped and symmetric around 0.48 (the mean and median are virtually identical, 0.485 and 0.484, respectively), and the maximum and minimum unconditional correlations are 0.89 and 0.16. The distributions of the minima and maxima of the conditional correlations are also unimodal, but somewhat skewed to the left and right, respectively. Moreover, the distribution of the minima reveals that some paths in fact cross the zero line.

5 Conclusions

We derive an equation-by-equation estimator (EBEE) of a multivariate log-GARCH-X model that admits feedback effects between the equations, and Dynamic Conditional Correlations (DCCs). Our least squares EBEE does not rely on financial returns being distributed according to a specific conditional distribution (e.g. the multivariate normal). Equation-by-equation estimation is particularly attractive when the dimensionality of the system is large, or when the number of covariates is large, or both, since then estimation often becomes numerically infeasible due to the ‘‘curse of dimensionality’’. The vector of covariates (‘‘X’’) is assumed to be stationary and ergodic, which means many of the variables that are believed to have an impact on volatility, and which are readily available (e.g. leverage, volume and volatility proxies), can be included as conditioning variables. Both strong consistency and asymptotic normality of our EBEE is proved under mild assumptions, and consistency of the estimators of the terms involved in the asymptotic variance-covariance matrices is also proved.

Two empirical applications illustrate the usefulness of the results. In the first we show how the volatility-spillover hypothesis can be tested when it involves restrictions in several equations – the null of no spillover is resoundingly rejected, and we also illustrate how the claim that a volatility proxy is a perfect predictor of future volatility can be tested in a model that conditions on a variety of competing effects (also here is the null rejected). In the second application we illustrate how a high-dimensional multivariate log-GARCH-X

model can readily be estimated in minutes. The model is of the constituent returns of the Eurostoxx50 index, i.e. the dimension is 50, and the X-part contains a leverage term, past values of Eurostoxx50 variability together with its leverage, volume and a range-based volatility proxy. One or more of these are found to be significant in most of the equations. Next we fit the corrected DCC of Aielli (2013) to the time-varying correlations.

References

- Aielli, G. P. (2013). Dynamic Conditional Correlations: On Properties and Estimation. *Journal of Business and Economic Statistics* 31, 282–299. <http://dx.doi.org/10.1080/07350015.2013.771027>.
- Apergis, N. and A. Reztis (2011). Food Price Volatility and Macroeconomic Factors: Evidence from GARCH and GARCH-X Estimates. *Journal of Agricultural and Applied Economics* 43, 95–110.
- Bauwens, L., D. Rime, and G. Sucarrat (2006). Exchange Rate Volatility and the Mixture of Distribution Hypothesis. *Empirical Economics* 30, 889–911.
- Billingsley, P. (1961). Statistical methods in Markov chains. *The Annals of Mathematical Statistics* 32, 12–40.
- Bollerslev, T. and M. Melvin (1994). Bid-Ask spreads and the volatility in the foreign exchange market: An empirical analysis. *Journal of International Economics* 36, 355–372.
- Boussama, F., F. Fuchs, and R. Stelzer (2011). Stationarity and Geometric Ergodicity of BEKK Multivariate GARCH Models. *Stochastic Processes and Their Applications* 121, 2331–2360.
- Brenner, R., R. Harjes, and K. Kroner (1996). Another look at models of the short-term interest rate. *Journal of Financial and Quantitative Analysis* 31, 85–107.
- Chen, M. and Q. Song (2015). Semi-parametric estimation and forecasting for exogenous log-GARCH models. *TEST*. DOI: <http://dx.doi.org/10.1007/s11749-015-0442-6>.
- Clark, P. (1973). A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* 41, 135–155.
- Conrad, C. and E. Weber (2013). Measuring persistence in volatility spillovers. Discussion Paper 543, Department of Economics, University of Heidelberg. <http://archiv.ub.uni-heidelberg.de/volltextserver/14865/>.
- Dominguez, K. M. (1998). Central bank interventions and exchange rate volatility. *Journal of International Money and Finance* 17, 161–190.

- Engle, R. (2002). Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business and Economic Statistics* 20, 339–350.
- Engle, R. and G. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics* 131, 3–27.
- Francq, C. and G. Sucarrat (2013). An Exponential Chi-Squared QMLE for Log-GARCH Models Via the ARMA Representation. <http://mpr.aub.uni-muenchen.de/51783/>.
- Francq, C. and L. Q. Thieu (2015). Qml inference for volatility models with covariates. <http://mpr.aub.uni-muenchen.de/63198/>.
- Francq, C. and J.-M. Zakoïan (2010). *GARCH Models*. New York: Marcel Dekker.
- Francq, C. and J.-M. Zakoïan (2015). Estimating multivariate GARCH and stochastic correlation models equation by equation. Forthcoming in *The Journal of the Royal Statistical Society. Series B*. Working paper version: MPRA Paper No. 54250. Online at <http://mpr.aub.uni-muenchen.de/54250/>.
- Garman, M. B. and M. J. Klass (1980). On the Estimation of Securities Price Volatilities from Historical Data. *Journal of Business* 53, 67–78.
- Hagiwara, M. and M. Herce (1999). Endogenous exchange rate volatility, trading volume and interest rate differentials in a model of portfolio selection. *Review of International Economics* 7, 202–218.
- Han, H. and D. Kristensen (2014). Asymptotic theory for the QMLE in GARCH-X models with stationary and non-stationary covariates. *Journal of Business and Economic Statistics*, ??–??
- Hansen, P. R., Z. Huan, and H. H. Shek (2012). Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility. *Journal of Applied Econometrics* 27, 877–906.
- Hwang, S. and S. Satchell (2005). GARCH model with cross-sectional volatility: GARCHX models. *Applied Financial Economics* 15, 203–216.
- Lamoureux, C. G. and W. D. Lastrapes (1990). Heteroscedasticity in Stock Return Data: Volume versus GARCH Effects. *The Journal of Finance*, 221–229.
- Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59, 347–370.
- Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *Journal of Business* 53, 61–65.
- Pedersen, R. S. (2015). Inference and testing on the boundary in extended constant correlation GARCH models. Working Paper.

- Shephard, N. and K. Sheppard (2010). Realising the future: forecasting with high frequency based volatility (HEAVY). *Journal of Applied Econometrics* 25, 197–231.
- Straumann, D. and T. Mikosch (2006). Quasi-Maximum-Likelihood Estimation in Conditionally Heteroscedastic Time Series: A Stochastic Recurrence Equations Approach. *The Annals of Statistics* 34, 2449–2495.
- Sucarrat, G. (2015). *lgarch: Simulation and estimation of log-GARCH models*. R package version 0.6. <http://cran.r-project.org/web/packages/lgarch/>.
- Sucarrat, G., G. Bjørnes, and S. Grønneberg (2015). Models of Financial Return With Time-Varying Zero Probability. Work in progress.
- Sucarrat, G. and Á. Escibano (2014). Unbiased Estimation of Log-GARCH Models in the Presence of Zero Returns. MPRA Paper No. 59040. <http://mpa.ub.uni-muenchen.de/59040/>.
- Sucarrat, G., S. Grønneberg, and Á. Escibano (2015). Estimation and Inference in Univariate and Multivariate Log-GARCH-X Models When the Conditional Density is Unknown. MPRA Paper No. 62352. Online at <http://mpa.ub.uni-muenchen.de/62352/>. An earlier version circulated as “The Power Log-GARCH Model”, *Universidad Carlos III de Madrid Working Paper 10-13 in the Economic Series*, June 2010.
- Tauchen, G. and M. Pitts (1983). The Price Variability-Volume Relationship on Speculative Markets. *Econometrica* 51, 485–505.
- Wintenberger, O. (2013). Continuous Invertibility and Stable QML Estimation of the EGARCH(1,1) model. *Scandinavian Journal of Statistics* 40, 846–867.

Table 1: A multivariate Log-GARCH(1,1)-X model of the Eurostoxx50 constituent returns (see Section 4.2)

j	$\hat{\omega}_{0j}$ (s.e.)	$\hat{\alpha}_{0jj}$ (s.e.)	$\hat{\beta}_{0jj}$ (s.e.)	$\hat{\lambda}_{0j1}$ (s.e.)	$\hat{\lambda}_{0j2}$ (s.e.)	$\hat{\lambda}_{0j3}$ (s.e.)	$\hat{\lambda}_{0j4}$ (s.e.)	$\hat{\lambda}_{0j5}$ (s.e.)
1	1.397	0.037 (0.009)*	0.821 (0.039)*	0.135 (0.045)*	-0.006 (0.010)	0.173 (0.043)*	-0.072 (0.021)*	0.090 (0.029)*
2	0.209	0.047 (0.009)*	0.904 (0.025)*	0.051 (0.033)*	-0.006 (0.007)	0.105 (0.038)*	-0.009 (0.009)	0.033 (0.015)*
3	0.167	0.044 (0.006)*	0.905 (0.014)*	0.116 (0.034)*	0.004 (0.007)	0.092 (0.036)*	-0.008 (0.006)*	0.038 (0.013)*
4	-0.132	0.033 (0.006)*	0.944 (0.013)*	0.053 (0.031)*	-0.001 (0.006)	0.104 (0.030)*	0.006 (0.006)	0.011 (0.009)
5	0.239	0.027 (0.005)*	0.961 (0.007)*	0.079 (0.020)*	-0.004 (0.004)	0.048 (0.023)*	-0.013 (0.004)*	0.011 (0.006)*
6	-0.018	0.044 (0.006)*	0.935 (0.011)*	0.078 (0.025)*	-0.007 (0.006)	0.067 (0.028)*	0.000 (0.004)	0.018 (0.009)*
7	-0.078	0.028 (0.007)*	0.915 (0.021)*	0.109 (0.031)*	-0.005 (0.007)	0.150 (0.035)*	0.001 (0.005)	0.051 (0.013)*
8	0.103	0.034 (0.007)*	0.859 (0.029)*	0.177 (0.040)*	-0.023 (0.009)*	0.102 (0.040)*	-0.009 (0.008)	0.092 (0.020)*
9	0.185	0.028 (0.007)*	0.863 (0.029)*	0.075 (0.039)*	0.006 (0.008)	0.093 (0.039)*	-0.007 (0.008)	0.061 (0.022)*
10	0.055	0.039 (0.006)*	0.896 (0.017)*	0.135 (0.033)*	-0.017 (0.007)*	0.088 (0.036)*	-0.005 (0.006)	0.071 (0.018)*
11	0.108	0.034 (0.008)*	0.886 (0.019)*	0.195 (0.038)*	-0.019 (0.008)*	0.080 (0.039)*	-0.009 (0.006)*	0.083 (0.019)*
12	0.263	0.031 (0.006)*	0.933 (0.016)*	0.037 (0.027)*	-0.005 (0.006)	0.099 (0.030)*	-0.013 (0.006)*	0.030 (0.012)*
13	-0.042	0.048 (0.008)*	0.908 (0.017)*	0.119 (0.035)*	-0.013 (0.007)*	0.075 (0.035)*	0.001 (0.005)	0.042 (0.014)*
14	0.222	0.039 (0.006)*	0.931 (0.013)*	0.048 (0.028)*	0.001 (0.006)	0.110 (0.030)*	-0.011 (0.005)*	0.017 (0.010)*
15	-0.105	0.035 (0.007)*	0.896 (0.018)*	0.066 (0.032)*	-0.012 (0.008)*	0.107 (0.037)*	0.005 (0.007)	0.062 (0.017)*
16	0.308	0.048 (0.008)*	0.920 (0.017)*	0.066 (0.030)*	-0.005 (0.007)	0.078 (0.030)*	-0.015 (0.006)*	0.016 (0.011)*
17	0.057	0.032 (0.007)*	0.909 (0.020)*	0.084 (0.031)*	-0.003 (0.007)	0.115 (0.037)*	-0.003 (0.005)	0.053 (0.017)*
18	0.006	0.034 (0.007)*	0.904 (0.024)*	0.069 (0.034)*	-0.004 (0.007)	0.097 (0.039)*	-0.001 (0.009)	0.044 (0.018)*
19	0.538	0.046 (0.006)*	0.928 (0.011)*	0.029 (0.029)	-0.002 (0.006)	0.129 (0.030)*	-0.026 (0.007)*	0.019 (0.009)*
20	0.371	0.059 (0.008)*	0.875 (0.022)*	0.137 (0.037)*	-0.021 (0.009)*	0.102 (0.037)*	-0.020 (0.008)*	0.059 (0.018)*
21	-0.162	0.056 (0.007)*	0.922 (0.013)*	0.104 (0.031)*	-0.017 (0.006)*	0.089 (0.033)*	0.006 (0.006)	0.024 (0.009)*
22	-0.129	0.012 (0.012)	0.821 (0.055)*	0.066 (0.056)	-0.007 (0.012)	0.116 (0.060)*	0.007 (0.029)	0.117 (0.044)*
23	0.160	0.037 (0.006)*	0.919 (0.014)*	0.118 (0.030)*	0.004 (0.006)	0.056 (0.031)*	-0.008 (0.005)*	0.019 (0.010)*
24	1.139	0.049 (0.008)*	0.881 (0.028)*	0.069 (0.038)*	0.000 (0.008)	0.152 (0.038)*	-0.057 (0.018)*	0.030 (0.015)*
25	-0.040	0.045 (0.007)*	0.924 (0.014)*	0.067 (0.028)*	-0.009 (0.006)*	0.099 (0.032)*	0.002 (0.005)	0.031 (0.012)*
26	-0.046	0.053 (0.007)*	0.929 (0.010)*	0.110 (0.026)*	-0.012 (0.006)*	0.060 (0.028)*	0.001 (0.004)	0.019 (0.008)*
27	-0.053	0.036 (0.007)*	0.919 (0.019)*	0.091 (0.034)*	-0.008 (0.007)	0.037 (0.035)	0.003 (0.009)	0.035 (0.012)*
28	-0.200	0.052 (0.006)*	0.917 (0.012)*	0.102 (0.033)*	-0.010 (0.006)*	0.150 (0.034)*	0.008 (0.006)*	0.035 (0.012)*
29	-0.039	0.046 (0.006)*	0.936 (0.010)*	0.080 (0.025)*	-0.011 (0.006)*	0.071 (0.028)*	0.001 (0.004)	0.021 (0.008)*
30	2.356	0.033 (0.009)*	0.795 (0.054)*	0.008 (0.046)	-0.001 (0.010)	0.173 (0.045)*	-0.115 (0.034)*	0.098 (0.036)*
31	0.332	0.035 (0.006)*	0.907 (0.020)*	0.072 (0.033)*	-0.003 (0.007)	0.111 (0.034)*	-0.017 (0.007)*	0.042 (0.016)*
32	0.487	0.037 (0.006)*	0.924 (0.014)*	0.054 (0.028)*	-0.001 (0.006)	0.088 (0.031)*	-0.024 (0.008)*	0.033 (0.012)*
33	0.279	0.033 (0.006)*	0.941 (0.013)*	0.043 (0.026)*	-0.007 (0.006)	0.075 (0.029)*	-0.013 (0.005)*	0.022 (0.009)*
34	0.514	0.047 (0.007)*	0.931 (0.012)*	0.046 (0.027)*	-0.008 (0.006)*	0.129 (0.031)*	-0.026 (0.008)*	0.018 (0.009)*
35	0.547	0.033 (0.006)*	0.925 (0.014)*	0.087 (0.029)*	-0.012 (0.007)*	0.094 (0.031)*	-0.029 (0.008)*	0.046 (0.014)*
36	0.006	0.048 (0.007)*	0.921 (0.014)*	0.081 (0.028)*	-0.014 (0.006)*	0.106 (0.031)*	-0.002 (0.004)	0.026 (0.010)*
37	0.344	0.052 (0.007)*	0.903 (0.018)*	0.036 (0.031)	-0.010 (0.007)*	0.067 (0.033)*	-0.015 (0.007)*	0.035 (0.013)*
38	0.004	0.031 (0.006)*	0.939 (0.013)*	0.150 (0.030)*	-0.011 (0.006)*	0.065 (0.028)*	-0.003 (0.004)	0.030 (0.011)*
39	1.099	0.042 (0.008)*	0.878 (0.030)*	0.091 (0.038)*	0.000 (0.007)	0.136 (0.035)*	-0.055 (0.018)*	0.044 (0.019)*
40	0.951	0.033 (0.007)*	0.928 (0.013)*	0.140 (0.030)*	0.001 (0.006)	0.114 (0.031)*	-0.049 (0.011)*	0.026 (0.010)*
41	0.190	0.038 (0.006)*	0.921 (0.017)*	0.067 (0.031)*	0.001 (0.006)	0.123 (0.032)*	-0.010 (0.005)*	0.024 (0.013)*
42	0.272	0.039 (0.006)*	0.932 (0.011)*	0.083 (0.032)*	-0.009 (0.006)*	0.105 (0.033)*	-0.015 (0.005)*	0.029 (0.011)*
43	0.628	0.051 (0.009)*	0.898 (0.050)*	0.088 (0.034)*	-0.015 (0.007)*	0.112 (0.035)*	-0.033 (0.010)*	0.046 (0.015)*
44	1.374	0.036 (0.009)*	0.850 (0.022)*	0.151 (0.043)*	-0.013 (0.011)	0.105 (0.042)*	-0.071 (0.028)*	0.083 (0.038)*
45	-0.217	0.041 (0.009)*	0.914 (0.022)*	0.193 (0.051)*	-0.014 (0.009)*	0.136 (0.049)*	0.005 (0.019)	0.042 (0.019)*
46	-0.009	0.053 (0.006)*	0.940 (0.007)*	0.081 (0.022)*	-0.002 (0.005)	0.082 (0.027)*	0.000 (0.003)	0.001 (0.006)
47	0.718	0.044 (0.011)*	0.887 (0.030)*	0.086 (0.038)*	0.005 (0.007)	0.054 (0.033)*	-0.035 (0.011)*	0.037 (0.017)*
48	0.123	0.056 (0.011)*	0.902 (0.030)*	0.131 (0.038)*	-0.008 (0.008)	0.087 (0.036)*	-0.006 (0.006)	0.022 (0.016)*
49	0.263	0.045 (0.006)*	0.924 (0.012)*	0.078 (0.030)*	-0.001 (0.006)	0.110 (0.032)*	-0.013 (0.005)*	0.019 (0.010)*
50	0.300	0.045 (0.008)*	0.898 (0.021)*	0.005 (0.029)	0.002 (0.007)	0.087 (0.034)*	-0.011 (0.007)*	0.034 (0.015)*

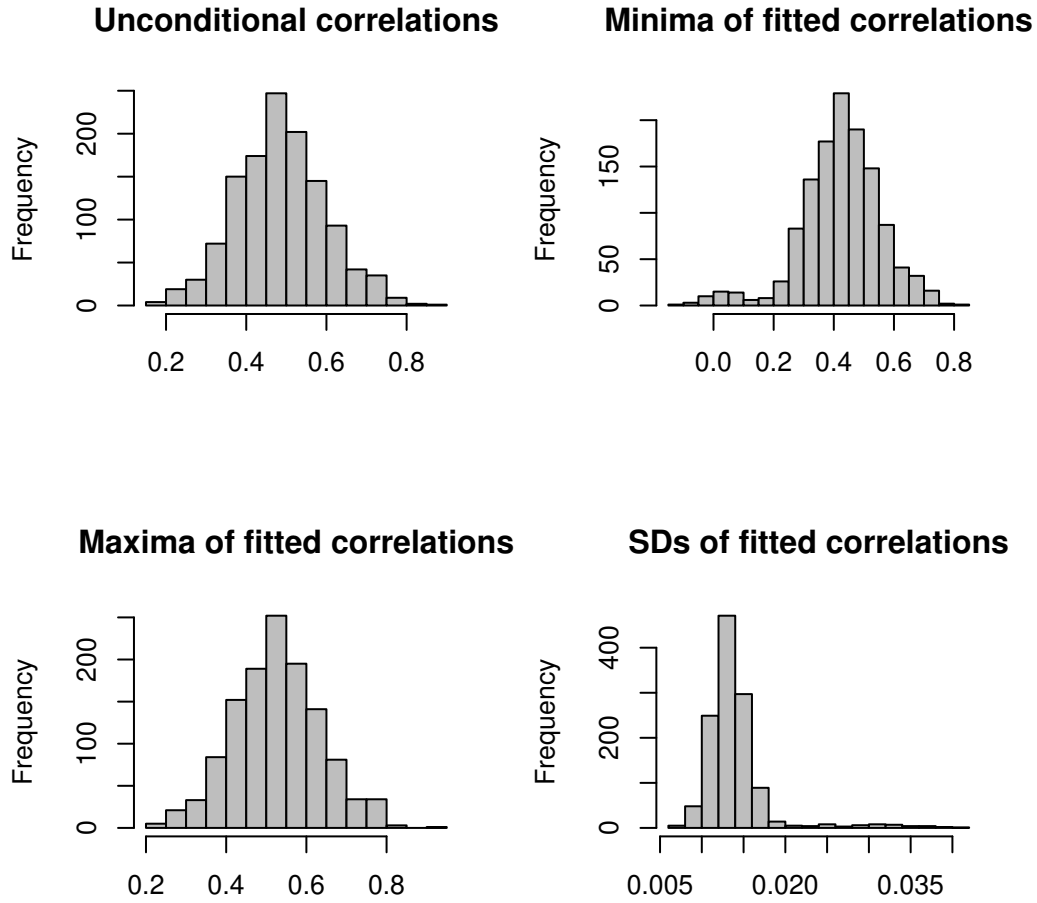


Figure 1: Histograms of the estimated unconditional correlations, the minima and maxima of the fitted conditional correlations, and the sample standard deviations of the 1225 fitted conditional correlations of the 50-dimensional model of the Eurostoxx50 index (see Section 4.2)