Inconsistency of the QMLE and asymptotic normality of the weighted LSE for a class of conditionally heteroscedastic models.

Christian Francq and Jean-Michel Zakoian

Université Lille III, EQUIPPE-GREMARS, CREST and EQUIPPE-GREMARS

2009

Online at http://mpra.ub.uni-muenchen.de/15147/
MPRA Paper No. 15147, posted 11. May 2009 01:45 UTC
Inconsistency of the QMLE and asymptotic normality of the weighted LSE for a class of conditionally heteroscedastic models.

Christian Francq, Jean-Michel Zakoïan

Abstract. This paper considers a class of finite-order autoregressive linear ARCH models. The model captures the leverage effect, allows the volatility to be zero and to reach its minimum for non-zero innovations, and is appropriate for long-memory modeling when infinite orders are allowed. It is shown that the quasi-maximum likelihood estimator is, in general, inconsistent. To solve this problem, we propose a self-weighted least-squares estimator and show that this estimator is asymptotically normal. Furthermore, a score test for conditional homoscedasticity and diagnostic portmanteau tests are developed. The latter have an asymptotic distribution which is far from the standard chi-square. Simulation experiments are carried out to assess the performance of the proposed estimator.

Keywords: Conditional homoscedasticity testing; Inconsistent estimator; Leverage effect; Linear ARCH; Quasi-maximum likelihood; Weighted least-squares.

*Université Lille III, EQUIPPE-GREMARS, BP 60 149, 59653 Villeneuve d’Ascq cedex, France. E-mail: christian.francq@univ-lille3.fr
†CREST and EQUIPPE-GREMARS, 15 Bd G. Péri, 92245 Malakoff Cedex, France. E-mail: zakoian@ensae.fr
1 Introduction

In recent years, the econometric literature on GARCH inference has been marked by formidable improvements, in different directions. For the standard class of Engle (1982) and Bollerslev (1986), optimal conditions for the consistency and asymptotic normality of the quasi-maximum likelihood estimation (QMLE) seem to have been obtained (see Berkes, Horváth and Kokoszka (2003), Francq and Zakoïan (2004)). The main finding is that the strict stationarity is essentially sufficient, and no moment on the observed process is required, for the asymptotic normality of the QMLE of GARCH models. For the existence of the information matrix, fourth-moment conditions have to be imposed on the underlying iid process, however. For an ARMA-GARCH model, fourth-moment conditions have to be imposed on the observed process. Alternative estimation methods have been considered when such moments do not exist (see e.g. Hall and Yao (2003), Horváth and Liese (2004), Ling (2007)).

Despite those theoretical improvements, the statistical inference in standard GARCH models remains problematic. The main complication, in the inference on GARCH models, results from the positivity constraints on the coefficients. The QMLE is a constrained estimator and, as a consequence, its asymptotic distribution when the parameter lies on the boundary is non standard. For the same reason, standard tests of nullity on GARCH coefficients (such as the Wald and quasi-likelihood ratio tests) have to be corrected (see Francq and Zakoïan (2008)).

Robinson (1991), Giraitis, Robinson and Surgailis (2000), Giraitis and Surgailis (2002), Berkes and Horváth (2003) and Giraitis, Leipus, Robinson and Surgailis (2004) proposed and analyzed a long memory alternative to the standard GARCH, called "linear ARCH" (LARCH), defined by

\[ u_t = \sigma_t \epsilon_t, \quad \sigma_t = b_0 + \sum_{i=0}^{\infty} b_i u_{t-i}, \quad \epsilon_t \text{ iid } (0, 1). \]  

(1.1)
Under appropriate conditions, this model is consistent with long memory in $u_t^2$, whereas an infinite order ARCH model fails to capture this property. From another point of view, this model has the advantage over standard ARCH formulations to be free of any positivity constraint on the volatility coefficients. Moreover, it is amenable to multivariate extensions (see Doukhan, Teyssière and Winant, 2006). Finite-order LARCH models were considered in Francq, Makarova and Zakoïan (2007) (hereinafter FMZ) in the purpose of analyzing the properties if unit root tests in presence of conditional heteroscedasticity. M-estimators of the location parameter when the error process is LARCH has been considered by Beran (2006). To our knowledge, only two working documents deal with the estimation of the full parameter in LARCH models. Beran and Schützner (2008) consider in particular the estimation of the parameters $C$ and $d$ when the LARCH($\infty$) have the form $b_i = C_i^d$, both in the short and long memory cases. One of the estimators considered by these authors is a modified conditional maximum likelihood estimator, that will be commented later. Truquet (2008) employs the same estimator, but focuses on the short memory case and considers the estimation of general LARCH($q$) models with finite order $q$.

The present paper attempts to contribute further to the statistical inference of finite-order LARCH models. As counterpart of the model flexibility, QMLE encounters serious difficulties which can only be avoided by strict conditions on the parameter space. It is also an aim of this paper to show that the behavior of the QMLE can be very pathological in certain situations and that phrases such that "QMLE is consistent under usual regularity conditions" should be taken with caution in general. It will be seen that, for the LARCH models, an approach which is more fruitful than the QMLE is to consider weighted least-squares estimation (WLSE), as was done by Horváth and Liese (2004) and Ling (2007) in the context of ARCH and ARMA-GARCH models.
The paper is organized as follows. In Section 2, we give the basic assumptions on the model and we establish the consistency and asymptotic normality of the QMLE. Section 3 illustrates the possible inconsistency of the QMLE when the stringent conditions used for the first theorem are in failure. Section 4 is devoted to the weighted least-squares estimation. Section 5 considers specification testing. Diagnostic checks are studied in Section 6. Section 7 reports simulation results. Concluding remarks are given in Section 8 and all proofs are relegated to Appendix A. Throughout the paper, $L\rightarrow$ denotes convergence in distribution. The spectral radius of a square matrix $A$ is denoted by $\rho(A)$ and $\otimes$ denotes the Kronecker product of matrices.

2 Model specification and QML estimation

The AR($p$)-LARCH($q$) model considered in this paper assumes that

$$
\begin{align*}
    x_t &= \psi_{01}x_{t-1} + \cdots + \psi_{0p}x_{t-p} + u_t, \\
    u_t &= (1 + b_{01}u_{t-1} + \cdots + b_{0q}u_{t-q})\epsilon_t, \\
    \epsilon_t \iid (0, \sigma^2_0\epsilon_t), &\quad \sigma_{0\epsilon} > 0
\end{align*}
$$

(2.1)

where $\psi_{01}, \ldots, \psi_{0p}, b_{01}, \ldots, b_{0q}$ are unknown real numbers.

The model for $(u_t)$ is a particular case of quadratic ARCH, as introduced by Sentana (1995). Apart from the absence of positivity constraints on the coefficients, this formulation has several distinctive feature compared to the standard ARCH. The volatility is not bounded below by a positive constant, it is able to capture the so-called leverage effect and it is not minimum at zero (see FMZ). This is illustrated in Figure 1 for the LARCH(1) model.

Let

$$
    A_{0t} = \begin{pmatrix}
        b_{1:q-1}\epsilon_t & b_{0q}\epsilon_t \\
        I_{q-1} & 0_{q-1}
    \end{pmatrix},
$$

4
Figure 1: News impact curve of $u_t$ in Model (2.1) with $q = 1$, $b_{01} = -0.2$ and $\sigma_\epsilon = 1$ (full line) compared with the news impact curve of the ARCH(1) model $u_t = \sqrt{1 + b_{01}^2 u_{t-1}^2} \epsilon_t$ (dotted line). Source FMZ.

where $b_{1:q-1} = (b_{01}, \ldots, b_{0q-1})$ and $I_k$ is the $k \times k$ identity matrix. By convention $A_0 t = b_{01} \epsilon_t$ when $q = 1$. Let $\gamma(A_0)$ be the top-Lyapunov exponent of the sequence $A_0 = (A_0 t)$, that is, for any norm $\| \cdot \|$ on the space of the $q \times q$ matrices, $\gamma(A_0) = \lim_{t \to \infty} \frac{1}{t} \log \| A_0 t A_0 t-1 \ldots A_01 \|$ a.s. In the above-mentioned paper, it was shown, following the approach of Bougerol and Picard (1992a, 1992b) that the second equation of (2.1) admits a strictly stationary solution $(u_t)$ if and only if

$$A_1: \quad \gamma(A_0) < 0.$$ 

In the case $q = 1$, this condition reduces to $|b_{01}| < \exp\{-E \log |\epsilon_1|\}$. Under $A_1$, the strictly stationary solution is unique, nonanticipative and ergodic. This solution admits a second order moment if and only if $\sum_{i=1}^q b_{0i}^2 \sigma_{0i}^2 < 1$. In this case, the solution is a conditionally heteroskedastic white noise. We also make the following standard assumption on the AR part.

$$A_2: \quad \text{the zeroes of the polynomial } \psi_0(z) := 1 - \sum_{i=1}^p \psi_{0i} z^i \text{ are outside the unit disk.}$$
We now turn to the QMLE of
\[ \theta_0 = (\psi_{01}, \ldots, \psi_{0p}, b_{01}, \ldots, b_{0q}, \sigma^2_0). \]

Assume we observe \( x_{-q-p+1}, x_{-q-p+2}, \ldots, x_n \) generated by Model (2.1), where the first \( p+q \) variables are considered as initial values. We consider a parameter space \( \Theta \subset \mathbb{R}^{p+q} \times (0, \infty) \) and we denote by \( \theta = (\psi_1, \ldots, \psi_p, b_1, \ldots, b_q, \sigma^2_0) \) a generic element of \( \Theta \). We assume

**A3:** \( \theta_0 \in \Theta \) and \( \Theta \) is a compact set,

and the identifiability condition

**A4:** the support of the law of \( \epsilon_t \) does not reduce to a set of 2 points.

Let \( u_t(\theta) = x_t - \sum_{i=1}^p \psi_i x_{t-i} \) and

\[ \sigma_t^2(\theta) = \sigma^2_0 \left( 1 + b_1 u_{t-1}(\theta) + \cdots + b_q u_{t-q}(\theta) \right)^2. \]

Denoting by \( L_n(\theta) \) the quasi-likelihood, a QMLE of \( \theta \) is a measurable solution of

\[ \hat{\theta}_n = \arg \max_{\theta \in \Theta} L_n(\theta) = \arg \min_{\theta \in \Theta} I_n(\theta), \tag{2.2} \]

where

\[ I_n(\theta) = n^{-1} \sum_{t=1}^n \ell_t(\theta), \quad \text{and} \quad \ell_t(\theta) = \frac{u_t^2(\theta)}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta) \in [-\infty, \infty], \tag{2.3} \]

with the conventions \( 1/0 + \log 0 = +\infty, 0/0 + \log 0 = -\infty \) and \( +\infty - \infty = +\infty \).

These conventions are required because \( u_t(\theta) \) and \( \sigma_t^2(\theta) \) may be equal to zero. When \( \sigma_t^2(\theta) = 0 \) and \( u_t(\theta) \neq 0 \), the value \( \theta \) can be precluded for the parameter. This justifies the conventions, which lead to \( I_n(\theta) = \infty \) for such values of \( \theta \). The following "high-level" assumption, to be discussed below, can be made to avoid such problems.
A5: The variable $\inf_{\theta \in \Theta} \sigma_t^2(\theta)$ is almost surely (a.s.) bounded away from 0.

Consider the case where $p = 0$, $q = 1$ and $\epsilon_t$ has a compact support $[-c,c]$. This case is quite artificial, and is just given for illustrating A5. When $|b_{01}c| < 1$, the white noise $u_t = \epsilon_t + \sum_{i=1}^{\infty} b_{01}^i \epsilon_{t-1} \cdots \epsilon_{t-i}$ belongs to $[-c/(1-b_{01}c), c/(1-b_{01}c)]$ with probability one. Thus, it is easy to see that A5 holds when $\{\sup_{\theta \in \Theta} |b_{1}|c < 1/2$. We will consider later the case where A5 does not hold. To establish the asymptotic normality, we need the following additional assumptions.

A6: $\theta_0$ belongs to the interior of $\Theta$,

A7: $E \epsilon_1^4 < \infty$ and $\rho \{ E(A_{01} \otimes A_{01} \otimes A_{01} \otimes A_{01}) \} < 1$.

It can be shown that Assumption A7 entails the existence of $Eu_1^4$ and, under A2, that of $Ex_1^4$. When $q = 1$, the condition is simply $b_{01}^4 E \epsilon_1^4 < 1$. Writing $A_{0t} = B\epsilon_t + J$, where $B$ and $J$ are non-random matrices, the second part of A7 takes the more explicit form:

$$
\rho \left\{ \sum_{j=1}^{4} \sum_{i \in \{0,1\}} E(\epsilon_1^{i_1+\cdots+i_4})(B^{i_1} + J^{1-i_1}) \otimes \cdots \otimes (B^{i_4} + J^{1-i_4}) \right\} < 1.
$$

Theorem 2.1 Under A1–A5 we have $\hat{\theta}_n \rightarrow \theta_0$ a.s. as $n \rightarrow \infty$. Under the additional Assumptions A6-A7, $\sqrt{n}(\hat{\theta}_n - \theta_0)$ is asymptotically distributed as $N(0, \Sigma)$, where $\Sigma = J^{-1} I J^{-1}$,

$$
I = E \left( \frac{\partial \ell_1(\theta_0)}{\partial \theta} \frac{\partial \ell_1(\theta_0)}{\partial \theta'} \right), \quad J = E \left( \frac{\partial^2 \ell_1(\theta_0)}{\partial \theta \partial \theta'} \right).
$$

3 Inconsistency of the QML estimator

Assumption A5 is essential for the consistency of the QMLE. For illustration purposes, consider the simplest version of Model (2.1), i.e. the AR(0)-LARCH(1)
When $\epsilon_t$ follows a uniform distribution on $[-1/2, 1/2]$ say, Assumption A5 is satisfied for sufficiently small $\Theta \subset (-2, 2) \times (0, \infty)$ because $\sigma_t(\theta)/\sigma_\epsilon \in (0, 2)$. The likelihood is then well-behaved (see the left panel in Figure 2). On the other hand, when $\epsilon_t$ has a continuous distribution with a non compact support, Assumption A5 is not satisfied because $\sigma^2_t(\theta) = \sigma^2_\epsilon(1 + bu_{t-1})^2$ cancels for $\theta = (-1/x_{t-1}, \sigma^2_\epsilon)$. Moreover, when $x_t \neq 0$ the true value $b_0$ cannot be equal to $-1/x_{t-1}$, which explains that the likelihood is null at these points (see the right panel of Figure 2). It should be noted that the non-smoothness of the likelihood is not due to the small sample size $n = 10$. On the contrary, the number of points where the likelihood vanishes increases with $n$, which would entail enormous computational burden for any reasonable sample size.

For more general models, we can even show the inconsistency of the QMLE when B4 is violated.

**Proposition 3.1** Consider the general AR(p)-LARCH(q) model (2.1) with $pq \neq 0$. If the distribution of $\epsilon_t$ is absolutely continuous with respect to the Lebesgue measure, for $\Theta$ sufficiently large, there exists an infinite number of sequences $(\hat{\theta}_n)$ of QMLE, and these sequences are generally inconsistent.

**Remark 3.1** This inconsistency result is very general for the model considered in this paper. It applies in particular when $\epsilon_t$ is gaussian. This shows that Assumption B4, though restrictive, is essential for the consistency result of Theorem 2.1.

**Remark 3.2** The inconsistency of the QMLE may seem surprising. In the iid case, frameworks where the QMLE is inconsistent include that of a mixture of two gaussian distributions (Kiefer and Wolfowitz (1956), Redner and Walker (1984)),
a one-parameter mixture (Ferguson (1982)), life distributions (Boyles, Marshall and Proschan (1985)), distributions with nuisance parameters (Neyman and Scott (1948)), the Rasch model (Ghosh, (1985)). In dynamic models however, examples of inconsistency seem much less frequent.

Figure 2: Likelihood (as a function of $b$ with $\sigma^2_t$ fixed) of a simulation of length $n = 10$ of Model (3.1) with $b_0 = 0.5$ and, in the left panel $\epsilon_t \sim U[-1/2, 1/2]$, and in the right panel $\epsilon_t \sim N(0, 1)$.

4 Weighted least squares estimators

We have seen that the QMLE is in failure without restrictive assumptions on the distribution of $\epsilon_t$. Another popular estimation method in time series is the least squares procedure. To avoid unnecessary moment conditions and to gain in efficiency we will consider Weighted Least Squares Estimators (WLSE). The asymptotic properties of weighted M-estimators have been studied by Horváth and Liese (2004), in the context of ARCH models. The asymptotic properties of weighted LSE and QMLE have been studied, in the context of ARMA-GARCH models, by Ling (2005).
4.1 WSLE of the AR parameter

The WLSE of the AR parameter \( \psi = (\psi_1, \ldots, \psi_p)' \) are defined by

\[
\hat{\psi}_{WLS} = \arg \min_{\psi \in \Theta_\psi} \frac{1}{n} \sum_{t=1}^{n} \omega_t u_t^2(\psi), \quad u_t(\psi) = x_t - \sum_{i=1}^{p} \psi_i x_{t-i}, \tag{4.1}
\]

where \( \Theta_\psi \) is the compact parameter space of the AR coefficients and the \( \omega_t \)'s are weights, which are allowed to depend on the past values \( \{x_s, s < t\} \) but not on \( \psi \).

For simplicity, we assume that \( \omega_t \) only depends on \( r \) past values:

\[ A8: \quad \omega_t = f(x_{t-1}, \ldots, x_{t-r}) \text{ for some function } f : \mathbb{R}^r \rightarrow (0, +\infty) \text{ and some integer } r \geq 1. \]

The initial values \( x_{1-r}, \ldots, x_0 \) required to compute \( \omega_1 \) are supposed to be available.

An attractive feature of the WLSE is that the minimization problem (4.1) does not require optimization routine. Under \( A6 \), the solution is explicitly given by

\[
\hat{\psi}_{WLS} = (X'\Omega X)^{-1} X'\Omega Y, \tag{4.2}
\]

where \( \Omega = \text{Diag}(\omega_1, \ldots, \omega_n) \), \( X \) is a \( n \times p \) matrix with generic term \( x_{i-j} \) and \( Y' = (x_1, \ldots, x_n) \). We introduce the following conditions.

\[ A9: \quad E\omega_1 \sum_{i=1}^{p} x_{1-i}^2 < \infty \text{ and } E\omega_1|\sigma_1(\theta_0)| \sum_{i=1}^{p} |x_{1-i}| < \infty. \]

\[ A10: \quad E\omega_1^2 \sigma_1^2(\theta_0) \sum_{i=1}^{p} x_{1-i}^2 < \infty. \]

We also introduce the notation \( X_t' = (x_{t-1}, \cdots, x_{t-p}) \).

**Theorem 4.1** Under \( A1, A2, A8, A9 \), \( \hat{\psi}_{WLS} \rightarrow \psi_0 \) a.s. as \( n \rightarrow \infty \). If, in addition, \( A10 \) holds, then

\[
\sqrt{n}(\hat{\psi}_{WLS} - \psi_0) \xrightarrow{d} N(0, \Sigma_{WLS}),
\]

where \( \Sigma_{WLS} = A_\psi^{-1} B_\psi A_\psi^{-1}, \quad A_\psi = E(\omega_1 X_1 X_1'), \quad B_\psi = E(\omega_1^2 \sigma_1^2(\theta_0) X_1 X_1'). \)
Remark 4.1 When applied with $\omega_t \equiv 1$, the Weighted Least Squares (WLS) procedure yields the usual least squares estimator (LSE) and, for the asymptotic normality, the fourth-order moments are required. Such moment conditions can be avoided by choosing, for instance, $\omega_t^{-1} = c_0 + \sum_{i=1}^{q+p} c_i x_{t-i}$, where the $c_i$ are strictly positive constants. In this case, no moment is needed since A9 and A10 are always satisfied.

Remark 4.2 Under A5, it is well-known that the optimal choice of the weighting matrix is

$$\Omega^* = \text{Diag}(1/\sigma_1^2(\theta_0), \ldots, 1/\sigma_n^2(\theta_0)).$$

Of course the resulting estimator is infeasible because $\sigma_t^2(\theta_0)$ depends on the unknown $b_{0i}$ coefficients.

4.2 WSLE of the LARCH parameter

We now consider the estimation of the LARCH coefficients. Let $\hat{u}_t = u_t(\hat{\psi})$, $t = 1 - q, \ldots, n$, where $\hat{\psi}$ denotes any consistent estimator of $\psi$. The WLS estimators of the volatility parameter $\beta = (b_1, \ldots, b_q, \sigma^2_\epsilon) \in \Theta_\beta$ are defined by

$$\hat{\beta}_{WLS} = \arg \min_{\beta \in \Theta_\beta} \frac{1}{n} \sum_{t=1}^{n} \tau_t \nu_t^2(\hat{\psi}, \beta), \quad \nu_t(\psi, \beta) = u_t^2(\psi) - \sigma_t^2(\psi, \beta)$$

where the positive weights $\tau_t \in \mathcal{F}_{t-1}$, the $\sigma$-field generated by $\epsilon_{t-i}, i > 0$. We introduce the following conditions.

A11: $E\epsilon_1^4 < \infty$ and $E\tau_1 \sigma_1^4(\theta_0) < \infty$.

A12: $E\sup_{\theta \in \mathcal{V}(\theta_0)} \left\| \tau_1 \frac{\partial u_t^2(\theta)}{\partial \theta} \frac{\partial u_t^2(\theta)}{\partial \theta'} \right\| < \infty$ for some neighborhood $\mathcal{V}(\theta_0)$ of $\theta_0$, $E\tau_1 |x_{1-i}|^\ell < \infty$, $E\tau_1 \sigma_1^4(\theta_0) |x_{1-i}|^{\ell} < \infty$, and $E\tau_1 \omega_1 |\sigma_1(\theta_0)|^3 |x_{1-i}|^{\ell'} < \infty$ for all $1 \leq i \leq p + q$, all $0 \leq \ell \leq 4$ and all $0 \leq \ell' \leq 3$. 

11
Theorem 4.2 Under $A1 - A3$, $A5$, $A8$ with $\omega_t$ replaced by $\tau_t$, and $A11$, $\hat{\beta}_{WLS} \to \beta_0$ a.s. as $n \to \infty$.

If, in addition, $A9$, $A10$, $A12$ hold and $\hat{\psi} = \hat{\psi}_{WLS}$,

$$\sqrt{n} \left( \frac{\hat{\psi}_{WLS} - \psi_0}{\hat{\beta}_{WLS} - \beta_0} \right) \overset{\mathcal{D}}{\to} \mathcal{N} \left( 0, \Sigma_{WLS} := \begin{pmatrix} \Sigma_{WLS}^{\psi} & \Sigma_{WLS}^{\psi \beta} \\ \Sigma_{WLS}^{\beta \psi} & \Sigma_{WLS}^{\beta} \end{pmatrix} \right),$$

where

$$\Sigma_{WLS}^{\beta} = A_{\beta}^{-1} \left\{ B_{\beta} + A_{\beta \psi} A_{\psi}^{-1} B_{\psi} A_{\beta \psi} + B_{\beta \psi} A_{\psi}^{-1} A_{\beta \psi}^\prime + A_{\beta \psi} A_{\psi}^{-1} B_{\psi} A_{\psi}^{-1} A_{\beta \psi}^\prime \right\} A_{\beta}^{-1},$$

$$\Sigma_{WLS}^{\psi} = A_{\psi}^{-1} \left\{ B_{\psi}^\prime + B_{\psi} A_{\psi}^{-1} A_{\psi}^\prime \right\} A_{\beta}^{-1} = \left( \Sigma_{WLS}^{\beta \psi} \right)^\prime,$$

with $\mu_4 = E\epsilon_1^4 / \sigma_\epsilon^4$, $Y_t = \frac{\partial \sigma_t^2}{\partial \beta_0}$, $Z_t = \frac{\partial \nu_t}{\partial \psi_0}$ and

$$A_{\beta} = E \left( \tau_1 Y_1 Y_1^\prime \right), \quad A_{\beta \psi} = E \left( \tau_1 Y_1 Z_1^\prime \right),$$

$$B_{\beta} = (\mu_4 - 1) E \left( \tau_1^2 \sigma_1^4 (\theta_0) Y_1 Y_1^\prime \right), \quad B_{\beta \psi} = \frac{E\epsilon_1^3}{\sigma_\epsilon^3} E \left( \tau_1 \omega_1 \sigma_1^3 (\theta_0) Y_1 X_1^\prime \right).$$

Remark 4.3 A remark similar to 4.1 holds. When $\omega_t$ and $\tau_t$ are (strictly positive) constants, eighth-order moments are required for the asymptotic normality. Choosing, for instance, $\omega_t^{-1} = c_0 + \sum_{i=1}^{q+p} c_i x_{t-i}^2$ and $\tau_t^{-1} = c_0^* + \sum_{i=1}^{q+p} c_i^* x_{t-i}^4$ where the $c_i$ and $c_i^*$ are strictly positive constants, no moment is needed on the observed process.

Remark 4.4 When the distribution of $\epsilon_t$ is symmetric, it can be seen that $\Sigma_{WLS}^{\psi \beta} = 0$ and $\Sigma_{WLS}^{\beta} = A_{\beta}^{-1} B_{\beta} A_{\beta}^{-1}$. In this case, under $A5$, the optimal weights are $\tau_t = 1 / \sigma_1^4 (\theta_0)$ (see Remark 4.2).

4.3 Choice of the weights

As argued by Horváth and Liese (2004), a natural choice of the weight functions is

$$\omega_t = \frac{1}{1 + \|X_t^*\|^2}, \quad \tau_t = \frac{1}{1 + \|X_t^*\|^4}, \quad (4.4)$$
where \( X_t^* = (x_{t-1}, \ldots, x_{t-p-q})' \). Many other sequences of weights satisfy A8 – A12. In the spirit of Ling (2007), and in connection to Huber's robust estimator for the regression model, one can consider sequences of weights of the form

\[
\omega_t = \frac{1}{\max \left\{ 1, C^{-1} \left( \sum_{i=1}^{p+q+1} |x_{t-i}| 1\{|x_{t-i}|>C\} \right) \right\}^2}, \quad \tau_t = \omega_t^2,
\]

(4.5)

where \( C \) is a positive constant. For the numerical illustrations we follow the suggestion of Ling (2007), taking \( C \) as the 90\% quantile of the absolute values of the observations \( |x_1|, \ldots, |x_n| \). In view of the remarks 4.2 and 4.4, one can also propose weights of the form

\[
\omega_t = \frac{1}{\hat{h}_t}, \quad \tau_t = \omega_t^2,
\]

(4.6)

where \( \hat{h}_t \) is a strictly positive proxy of the volatility. In the sequel we choose \( \hat{h}_t \) as being the implied volatility based on a standard ARCH\((p+q)\) model.

5 Specification Testing

As we have seen, the QML estimator has a pathological behavior in our framework, so we cannot consider the standard tests (Wald, score, likelihood ratio). Instead, we will base our tests on the WLS criterion. For notational convenience we will omit the subscript "WLS" in the estimators.

5.1 Wald tests

To test an assumption of the form \( R\theta_0 = r \), where \( r \in \mathbb{R}^d \) and \( R \) is a full row-rank \( d \times (p+q+1) \) matrix, the asymptotic normality results of Theorem 4.2 can be used. Under \( H_0 \) and the assumptions of this theorem, the Wald-type statistics

\[
W_n = n(R\hat{\theta} - r)'(R\hat{\Sigma}R')^{-1}(R\hat{\theta} - r) \xrightarrow{\mathcal{L}} \chi_d^2,
\]
where \( \hat{\theta} = (\hat{\psi}', \hat{\beta}')' \), and \( \hat{\Sigma} \) denotes any consistent estimator of \( \Sigma \). Empirical estimates of \( A_\beta, A_{\beta\psi}, B_\beta, B_{\beta\psi} \) can be considered to construct such an estimator.

To test the nullity of all the coefficients \( b_i \) it seems much more appropriate to consider a score-type test, which does not require estimating the general model. This is considered in the next section.

### 5.2 Testing for conditional homoscedasticity

The aim is to test for

\[
H_0 : b_0 = 0
\]

where \( b_0 = (b_{01}, \ldots, b_{0q})' \). Under \( H_0 \) the model reduces to a simple AR\((p)\) model with independent errors. Let \( \hat{\theta}^c = (\hat{\psi}', 0_p, \hat{\sigma}_c^2)' \) denote the estimator constrained by \( H_0 \), where \( \hat{\psi} \) is defined in (4.2) and \( \hat{\sigma}_c^2 \) is the constrained WLS estimator of \( \sigma^2 \) defined by

\[
\hat{\sigma}_c^2 = \frac{1}{\sum_{t=1}^n \tau_t} \sum_{t=1}^n \tau_t \hat{u}_t^2. \tag{5.1}
\]

A Rao score-type (or Lagrange multiplier) statistic is based on the derivative of the second-step criterion at \( \hat{\theta}^c \). To derive the statistic, we start by evaluating the asymptotic distribution of this derivative under \( H_0 \). Let

\[
A_\beta = \begin{pmatrix} A_b & A_{b\sigma} \\ A_{\sigma b} & A_\sigma \end{pmatrix}, \quad A_* = -A_b + \frac{1}{A_\sigma} A_{b\sigma} A_{\sigma b}.
\]

Under the assumptions of Theorem 4.2, we have

\[
\Delta_n := \frac{1}{\sqrt{n}} \sum_{t=1}^n \tau_t \frac{\partial \nu^2(\hat{\psi}, 0_q, \hat{\sigma}_c^2)}{\partial b} \overset{d}{\to} \mathcal{N}(0, \Sigma_\Delta := A_* \Sigma_b A_*'), \tag{5.2}
\]

where \( \Sigma_b \) is the top-left \( q \times q \) block of the matrix \( \Sigma^\beta \). A Rao score-type statistic is then given by

\[
R_n = (\Delta_n^c)' \hat{\Sigma}_\Delta^{-1} \Delta_n^c
\]
where \( \hat{\Sigma}_\Delta \) denotes any \( H_0 \)-consistent estimator of \( \Sigma_\Delta \). This statistic follows asymptotically a \( \chi^2_q \) distribution under the null and the critical region at the asymptotic level \( \alpha \) is given by

\[
\{ R_n > \chi^2_q(1 - \alpha) \}
\]

where \( \chi^2_q(1 - \alpha) \) denotes the \( 1 - \alpha \) quantile of the \( \chi^2_q \) distribution.

We will now derive an explicit form for this statistic. It is known that, under quite general assumptions, a version of the score test statistic based on the LSE can be interpreted as the uncentred coefficient of determination of the regression of the constant 1 on the components of the score vector (see for instance Godfrey, 1988, p.15). We will show that a similar interpretation holds for the statistic \( R_n \) based on the WLSE. First notice that

\[
\Delta^c_n = -\frac{4\hat{\sigma}_\epsilon^{2c}}{\sqrt{n}} \sum_{t=1}^{n} \tau_t(\hat{u}_t^2 - \hat{\sigma}_\epsilon^{2c})\hat{u}_{t-1}
\]

where \( \hat{u}_{t-1} = (\hat{u}_{t-1}, \ldots, \hat{u}_{t-q})' \). Note also that, under the null,

\[
\Sigma_\Delta = 16\sigma_\epsilon^4 \text{Var} \epsilon^2_1 E (\tau_1^2 \hat{u}_0^2, \ldots, \tau_n^2 \hat{u}_n^2) ,
\]

where \( \hat{u}_t = (u_t - \bar{u})' \). Writing \( \Delta^c_n = -4\hat{\sigma}_\epsilon^{2c} n^{-1/2} U'V \) with

\[
U' = (\tau_1 \hat{u}_0, \ldots, \tau_n \hat{u}_n), \quad V = (\hat{u}_1^2 - \hat{\sigma}_\epsilon^{2c}, \ldots, \hat{u}_n^2 - \hat{\sigma}_\epsilon^{2c})'
\]

and using the estimator of \( \Sigma_\Delta \) defined by

\[
\hat{\Sigma}_\Delta = 16(\hat{\sigma}_\epsilon^{2c})^2 n^{-1} V'V n^{-1} U'U,
\]

we obtain the test statistic

\[
R_n = n \frac{V'U (U'U)^{-1} U'V}{V'V} ,
\]

which is \( n \) times the uncentred coefficient of determination of the regression of

\( \hat{u}_t^2 - \hat{\sigma}_\epsilon^{2c} \) on \( \tau_t \hat{u}_{t-1}, \ldots, \tau_t \hat{u}_{t-q} \).
This test has of course similarities with the standard test for conditional heteroskedasticity of \((u_t)\) in the ARCH\((q)\) (or GARCH\((p,q)\)) framework. In this case, a Rao-score test statistic is \(n\) times the \(R^2\) of the regression of \(u_t^2\) over a constant and \(u_{t-1}^2, \ldots, u_{t-q}^2\).

6 Diagnostic checks

In this section we develop some diagnostic tools for the AR\((p)\)-LARCH\((q)\) model (2.1). We first consider adequacy of the AR equation.

6.1 Diagnostic checking for the AR part

Conventional ways of testing adequacy of linear models involve checks that the residuals are approximately uncorrelated. To this aim the portmanteau tests of Box-Pierce (1970) and Ljung-Box (1978) are the most popular tools. We only consider the Ljung-Box statistic (hereafter LB) which has the same asymptotic behavior as the Box-Pierce statistic, but is the most widely used by practitioners. The LB statistic is defined by

\[
Q_{\hat{u}} = n(n + 2) \sum_{h=1}^{m} \frac{\hat{\rho}_u^2(h)}{n - h} \tag{6.1}
\]

where \(\hat{\rho}_u(h)\) is the residual autocorrelation at lag \(h\) and \(m\) is a fixed integer.

The standard test procedure consists, for \(m > p\), in rejecting the AR\((p)\) model if \(Q_{\hat{u}} > \chi^2_{m-p}(1 - \alpha)\). The procedure is (approximately) valid when (i) the residuals are obtained by least-squares, and (ii) the error terms of the AR equation are iid. Because none of these conditions is satisfied in our framework, the standard portmanteau tests require an adaptation. In the more general setting of weak ARMA models, Francq, Roy and Zakoïan (2005) relaxed condition (ii), but we cannot directly use their results because we consider here WLS estimators.
For $p > 0$, let $\hat{u}_t = u_t(\hat{\psi}_{WLS}) = u_t(\hat{\psi})$, $t = 1 - q, \ldots, n$, be the AR($p$) residuals, where $\hat{\psi}_{WLS} = \hat{\psi}$ is the WLS estimator defined in (4.2). For $p = 0$, one can set $\hat{u}_t = u_t = x_t$. The residuals autocovariances and autocorrelations are defined by

$$\hat{\gamma}_u(\ell) = \frac{1}{n} \sum_{t=1}^{n-\ell} \hat{u}_t \hat{u}_{t+\ell} \quad \text{and} \quad \hat{\rho}_u(\ell) = \frac{\hat{\gamma}_u(\ell)}{\hat{\gamma}_u(0)}.$$  \hspace{1cm} (6.2)

Let $\hat{\rho}_m = (\hat{\rho}_1, \ldots, \hat{\rho}_m)'$ and $U_t = (u_{t-1}, \ldots, u_{t-m})'$. We denote by $\phi_i^s$ the coefficients defined by

$$\psi^{-1}(z) = \sum_{i=0}^{\infty} \phi_i^s z^i, \quad |z| \leq 1.$$

Take $\phi_i^s = 0$ when $i < 0$. Let $\lambda_i = (\phi_i^s_{i-1}, \ldots, \phi_i^s_{i-p})' \in \mathbb{R}^p$ and let the $p \times m$ matrix

$$\Lambda = (\lambda_1 \lambda_2 \cdots \lambda_m).$$  \hspace{1cm} (6.3)

The following lemma gives the asymptotic distribution of a vector of residual autocorrelations of an AR($p$) model, when the Data Generating Process (DGP) actually follows an AR($p$)-LARCH($q$) model.

**Lemma 6.1** Under the assumptions of Theorem 4.1, \(\sqrt{n}\hat{\rho}_m \xrightarrow{d} \mathcal{N}(0, \Sigma_{\hat{\rho}_m})\), where

$$\Sigma_{\hat{\rho}_m} = \frac{1}{\sigma_u^4} E(u_1^2 U_1 U_1') \quad \text{when} \quad p = 0,$$

and when $p > 0$,

$$\Sigma_{\hat{\rho}_m} = \Lambda A_{\psi}^{-1} B_{\psi} A_{\psi}^{-1} \Lambda + \frac{1}{\sigma_u^4} E(u_1^2 U_1 U_1')$$

$$- \frac{1}{\sigma_u^4} \left\{ \Lambda A_{\psi}^{-1} E(\omega_1 u_1^2 X_1 U_1') + E(\omega_1 u_1^2 U_1 X_1') A_{\psi}^{-1} \Lambda \right\},$$  \hspace{1cm} (6.4)

where $\sigma_u^2 = Eu_1^2$.

The following theorem is an obvious consequence of Lemma 6.1.
Theorem 6.1 Suppose that the assumptions of Theorem 4.1 hold, in particular that the AR order is correctly specified. Then the portmanteau statistic \( Q_m^\hat{\Sigma} \overset{D}{\to} \sum_{i=1}^m \hat{\xi}_{i,m} Z_i^2 \), where \( \hat{\xi}_{i,m} = (\hat{\xi}_{1,m}, \ldots, \hat{\xi}_{m,m})' \) is the eigenvalues vector of the matrix \( \hat{\Sigma}_m^\hat{\rho} \) and \( Z_1, \ldots, Z_m \) are independent \( N(0,1) \) variables.

It should be noted that an estimator \( \hat{\Sigma}_m^\hat{\rho} \) of \( \Sigma_m^\rho \) can be straightforwardly obtained from the estimation of the sole AR part in model (2.1). Indeed, by inversion of the estimated AR polynomial, an estimator of \( \hat{\Lambda} \) is obtained. The matrices \( \hat{A}\psi \) and \( \hat{B}\psi \) can be estimated by

\[
\hat{A}_\psi = \frac{1}{n} \sum_{t=r/p+1}^{n+1} \omega_t X_t X_t', \quad \hat{B}_\psi = \frac{1}{n} \sum_{t=r/p+1}^{n+1} \omega_t^2 \hat{u}_t^2 X_t X_t', \quad (6.5)
\]

noting that \( E\omega_t^2 \sigma_t^2(b_0) X_t X_t' = E\omega_t^2 \hat{u}_t^2 X_t X_t' \). Similarly the other matrices involved in the right-hand side of (6.4) have the form of expectations and can therefore be estimated by empirical means (with \( U_t \) replaced by \( \hat{U}_t = (\hat{u}_{t-1}, \ldots, \hat{u}_{t-m})' \)). Finally \( \sigma_u^2 \) is estimated by the empirical mean of the \( \hat{u}_t^2 \). Thus the diagnostic checking of the AR part can be made at the end of the first stage of the WLS procedure, and does not require estimating the LARCH parameter \( \beta \). The distribution of the quadratic form \( \sum_{i=1}^m \hat{\xi}_{i,m} Z_i^2 \), where \( \hat{\xi}_{i,m} \) are the eigenvalues of the matrix \( \hat{\Sigma}_m^\hat{\rho} \), can then be computed using the algorithm by Imhof (1961).

Remark 6.1 When \( q = 0 \) and \( \omega_t = 1 \), i.e. when a standard AR model is estimated by LS, it is well known that the asymptotic distribution of \( Q_m^\hat{\rho} \) can be approximated by a \( \chi^2_{m-p} \). No such simplification seems to hold with the general WLS, even in the case \( q = 0 \). Similarly the law does not reduce to a \( \chi^2 \) when \( \omega_t = 1 \) and \( q > 0 \) (see the remark below), which is in accordance with the results obtained by Francq et al. (2005) in the general framework of weak ARMA models.

Remark 6.2 It can be noticed that when \( p = 0 \) and \( b_0 = (b_{01}, \ldots, b_{0q}) = 0 \), the process \( (X_t) \) is an iid white noise and the asymptotic distribution of the portman-
teau statistic is the usual $\chi^2_m$ distribution, because $\Sigma_{\hat{\rho}_m}$ reduces to the $m \times m$ identity matrix. Still when $p = 0$ but $b \neq 0$, the matrix $\Sigma_{\hat{\rho}_m}$ is not the identity matrix. For instance if $q = 1$ and the distribution of $\epsilon_t$ is symmetric, elementary computations show that the first diagonal term of $\Sigma_{\hat{\rho}_m}$ is

$$
1 - b_0^2 \frac{\sigma_0^2}{\epsilon_1^2} \left\{ 1 + \frac{b_0^2 E \epsilon_1^4}{\sigma_0^2} \right\} \neq 1 \text{ when } b_0 \neq 0,
$$

so that $Q_{\hat{\rho}}^m$ does not asymptotically follow the $\chi^2_m$ distribution.

**Remark 6.3** Note that when $\Sigma_{\hat{\rho}_m}$ is regular, the modified Box-Pierce statistic

$$
\tilde{Q}_{\hat{\rho}}^m := n \hat{\rho}_m^\prime \hat{\Sigma}_{\hat{\rho}_m}^{-1} \hat{\rho}_m
$$

asymptotically follows a $\chi^2_m$ distribution, under the null hypothesis of adequacy of the order $p$ for the AR part. Since the asymptotic distribution of $\tilde{Q}_{\hat{\rho}}^m$ is simpler than that of $Q_{\hat{\rho}}^m$, the former seems more attractive for testing the overall significance of $\hat{\rho}(h), h = 1, \ldots, m$. Note however that the regularity assumption on $\Sigma_{\hat{\rho}_m}$ is not very explicit, because the invertibility of this matrix depends on the unknown coefficients and on the choice of the weights in the estimation procedure.

### 6.2 Diagnostic checking for the LARCH part

As proposed by Higgins and Bera (1992), the adequacy of ARCH-type models can be assessed by means of the Box-Pierce statistic $Q_{\hat{\rho}}^m$ on the first $m$ squared standardized residual autocorrelations. The asymptotic distribution of $Q_{\hat{\rho}}^m$ has been given by Li and Mak (1994), under regularity conditions which do not hold in our framework. Because we use WLS estimators instead of the maximum-likelihood estimator, the asymptotic distribution of $Q_{\hat{\rho}}^m$ will be different than that obtained by Li and Mak (1994).

Recall that the WLS estimator defined in Theorem 4.2 is denoted by $\hat{\theta} = (\hat{\psi}', \hat{\beta}')'$, with $\hat{\psi} = \hat{\psi}_{WLS} = (\hat{\psi}_1, \ldots, \hat{\psi}_p)'$ and $\hat{\beta} = \hat{\beta}_{WLS} = (\hat{b}_1, \ldots, \hat{b}_q, \hat{\sigma}_\epsilon^2)'$. The
autocovariances and autocorrelations of the squared (standardized) residuals are defined by

\[
\hat{\gamma}_{\epsilon^2}(\ell) = \frac{1}{n} \sum_{t=\ell+1}^{n} \left( \hat{\epsilon}_t^2 - \overline{\epsilon}^2 \right) \left( \hat{\epsilon}_{t-\ell}^2 - \overline{\epsilon}^2 \right) \quad \text{and} \quad \hat{\rho}_{\epsilon^2}(\ell) = \frac{\hat{\gamma}_{\epsilon^2}(\ell)}{\hat{\gamma}_{\epsilon^2}(0)},
\]

for \(0 \leq \ell < n\), where for \(q > 0\)

\[
\dot{\epsilon}_t = \epsilon_t(\hat{\theta}), \quad \epsilon_t(\theta) = \frac{u_t(\psi)}{1 + \sum_{i=1}^{q} b_i u_{t-i}(\psi)}, \quad \overline{\epsilon}^2 = \frac{1}{n} \sum_{t=1}^{n} \hat{\epsilon}_t^2.
\]

(6.6)

When \(q = 0\), we set \(\epsilon_t(\theta) = u_t(\psi)\). In order to guarantee that \(\dot{\epsilon}_t\) be almost surely well defined, at least for \(n\) large enough, we make the following assumption

\[
P \left( 1 + \sum_{i=1}^{q} b_0 u_{t-i} = 0 \right) = 0.
\]

(6.8)

Note that (6.8) is satisfied when the distribution of \(\epsilon_t\) has a density with respect to the Lebesgue measure. This assumption entails the (almost sure) existence of \((\partial \epsilon_t / \partial \theta)(\theta_0)\). Let \(\hat{\rho}_{m}^2 = (\hat{\rho}_{\epsilon^2}(1), \ldots, \hat{\rho}_{\epsilon^2}(m))'\) and

\[
V_t = \left( \epsilon_t^2 - \sigma_0^2, \epsilon_{t-1}^2 - \sigma_0^2, \ldots, \epsilon_{t-m}^2 - \sigma_0^2 \right)'.
\]

We also define the matrices

\[
S = \begin{pmatrix}
A^{-1} \epsilon_1 E(\omega_1 u_1 X_1 V_1') \\
A^{-1} A_{\psi} A^{-1} \epsilon_1 E(\omega_1 u_1 X_1 V_1') + A^{-1} \epsilon_1 E(\tau_1 \nu_1 \frac{\partial \sigma_1^2(\psi_0, \theta_0)}{\partial \beta} V_1')
\end{pmatrix}
\]

and

\[
\Lambda_{\epsilon^2} = \left( \lambda_1^2, \ldots, \lambda_m^2 \right)', \quad \text{where} \quad \lambda_\ell^2 = 2 E \epsilon_1 \frac{\partial \epsilon_1}{\partial \theta}(\theta_0)(\epsilon_{1-\ell}^2 - \sigma_0^2).
\]

(6.7)

The existence of these matrices requires moment conditions. Note that \(S = 0\) when \(E \epsilon_t^3 = 0\). We also need to reinforce Assumption (6.8). Thus we make the following assumptions.

**A13:** If \(q > 0\), there exist a neighborhood \(V(\theta_0)\) of \(\theta_0\) and a positive number \(\iota > 0\) such that

\[
P \left( \inf_{\theta \in V(\theta_0)} \left| 1 + \sum_{i=1}^{q} b_i u_{t-i}(\psi) \right| > \iota \right) = 1.
\]
A14: \( \text{Ex}_t^0 < \infty. \)

With these notations and assumptions we have the following result.

**Theorem 6.2** Suppose that the assumptions of Theorem 4.2 hold, in particular that the AR order \( p \) and the LARCH order \( q \) are correctly specified. Assume also that the assumptions A13 and A14 hold true. Then \( \sqrt{n} \hat{\rho}_m^2 \xrightarrow{L} N(0, \Sigma_{\hat{\rho}_m^2}) \), where

\[
\Sigma_{\hat{\rho}_m^2} = \frac{1}{\sigma_t^8(\mu_4 - 1)^2} \left\{ \sigma_t^8 \mu_4^2 I_m + \Lambda^2 \Sigma_{WLS} \Lambda^2 + S' \Lambda^2 S + \Lambda^2 S \right\}
\]

when \( q \neq 0 \), and

\[
\Sigma_{\hat{\rho}_m^2} = I_m \quad (6.9)
\]

when \( q = 0 \).

Moreover the portmanteau statistic

\[
Q_m^{\hat{\epsilon}^2} := n(n + 2) \sum_{h=1}^{m} \frac{\hat{\rho}_m^2(h)}{n - h} \xrightarrow{L} \sum_{i=1}^{m} \xi_i Z_i^2,
\]

where \( \xi_1^2, \ldots, \xi_m^2 \) are the eigenvalues of the matrix \( \Sigma_{\hat{\rho}_m^2} \) and \( Z_1, \ldots, Z_m \) are independent \( N(0, 1) \) variables.

**Remark 6.4** Assumption A13 is restrictive, but seems unavoidable since the portmanteau statistics relies on rescaled residuals in which the inverses of \( \sigma_t(\theta) \) are taken in a neighborhood of \( \theta_0 \). However, simulation experiments show that the portmanteau test behaves well in finite sample when (most of) the \( 1 + \sum_{i=1}^{q} \hat{b}_i \hat{u}_{t-i} \) are far enough from 0.

**Remark 6.5** In Remark 6.1 it was seen that the asymptotic distribution of \( Q_m^{\hat{\epsilon}^2} \) depends, in a complicated way, of the weights and the coefficients, even in the case \( q = 0 \). By contrast, (6.9) shows that the asymptotic distribution of \( Q_m^{\hat{\epsilon}^2} \) is \( \chi_m^2 \) when the DGP is an AR model with iid innovations, whatever the AR order \( p \) and
whatever the weights $\omega_t$. The $\chi^2_m$-asymptotic distribution for $Q^2_m$ was obtained by McLeod and Li (1983) in the case $q = 0$ and $\omega_t = 1$, which corresponds to the standard LSE.

**Remark 6.6** A remark similar to 6.3 holds. When $\Sigma_{\hat{\rho}^2_m}$ is regular and $\hat{\Sigma}_{\hat{\rho}^2_m}$ denotes any consistent estimator of $\Sigma_{\hat{\rho}^2_m}$, the modified statistic

$$\tilde{Q}^2_m := n \hat{\rho}^2_m \hat{\Sigma}_{\hat{\rho}^2_m}^{-1} \hat{\rho}^2_m$$

asymptotically follows a $\chi^2_m$ distribution, under the null hypothesis of adequacy of the orders $p$ and $q$.

## 7 Numerical Illustration

### 7.1 Monte Carlo study

This section examines the performance of the asymptotic estimation results in finite samples through Monte Carlo experiments. Data are generated through the AR(1)-LARCH(1) model

$$x_t = \psi_{t-1} + u_t, \quad u_t = (1 + b_{t-1}) \epsilon_t, \quad \epsilon_t \text{ iid } \mathcal{N}(0, \sigma_0^2), \quad \sigma_0^2 > 0. \quad (7.1)$$

Table 1 compares the distribution of the QML, LS and WLS estimates of the 3 parameters $\psi_{01}, b_{01}$ and $\sigma_0^2$ over $N = 500$ independent simulations of the model, for the sample sizes $n = 100$ and $n = 1,000$. We used the version of the WLSE defined by the weights (4.6) based on an ARCH proxy of the volatility. The failure of the QMLE is perfectly explained by Proposition 3.1, since Assumption A5 is not satisfied by the DGP. With the particular choice of parameters of these simulations experiments, the LSE and WLSE provide very close results.
Table 1: Comparison of the QML, LS and WLS estimators of the AR(1)-LARCH(1) model (7.1). The number of replications is $N = 500$.

<table>
<thead>
<tr>
<th></th>
<th>QMLE</th>
<th>LSE</th>
<th>WLSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 100$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{01} = 0.9$</td>
<td>-136.71, 29.69</td>
<td>0.58, 1.14</td>
<td>0.69, 1</td>
</tr>
<tr>
<td>$b_{01} = -0.5$</td>
<td>-101.51, 61.91</td>
<td>-0.13, 0.18</td>
<td>-0.98, -0.13</td>
</tr>
<tr>
<td>$\sigma^2_{0\epsilon} = 1$</td>
<td>-0.09, 48.21</td>
<td>0.44, 6.15</td>
<td>0.53, 2.14</td>
</tr>
<tr>
<td>$n = 1000$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_{01} = 0.9$</td>
<td>-166.42, 34.11</td>
<td>0.72, 0.86</td>
<td>0.72, 0.86</td>
</tr>
<tr>
<td>$b_{01} = -0.5$</td>
<td>-215.38, 942.05</td>
<td>-0.62, -0.34</td>
<td>-0.62, -0.34</td>
</tr>
<tr>
<td>$\sigma^2_{0\epsilon} = 1$</td>
<td>2.25, 6.53</td>
<td>0.53, 1.43</td>
<td>0.82, 1.27</td>
</tr>
</tbody>
</table>

Table 2 compares the performance of four versions of the WLSE: the LSE in which the weights are constant, the WLSE based on an ARCH proxy of the volatility, the WLSE$^{HL}$ with the weights (4.4) of Horváth and Liese (2004), and the WLSE$L$ defined by the weights (4.5) proposed by Ling (2007) in a similar context. With the value $b_{01} = -0.54$ the simulated process $(x_t)$ admits moments of order eight, with $b_{01} = -0.63$ we have $E x_t^6 < \infty$ but $E x_t^8 = \infty$, with $b_{01} = -0.75$ we have $E x_t^4 < \infty$ and $E x_t^6 = \infty$, with $b_{01} = -0.99$ we have $E x_t^2 < \infty$ and $E x_t^4 = \infty$, and with $b_{01} = -1.1$ the second order moments do not exist. In the table, the best (i.e. minimal) root mean squared error (RMSE) and the best bias of estimation are displayed in bold. As expected the performance of the four versions is equivalent when the DGP admits moments of hight order, and the performance of the LSE decreases dramatically when $|b_{01}|$ increases. Overall the behavior of the WLSE and WLSE$^{HL}$ remains satisfactory whatever the value of $b_{01}$, with a slight advantage for the WLSE in terms of RMSE. We thus used this WLSE version for the application.
Table 2: Comparison of four different versions of the WLS estimator. The DGP is an AR(1)-LARCH(1) process with a gaussian iid noise $\epsilon_t$. The number of replications is $N = 500$ and the length of the simulations is $n = 100$.

<table>
<thead>
<tr>
<th></th>
<th>LSE</th>
<th>WLSE</th>
<th>WLSE$^H$</th>
<th>WLSE$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_{01} = 0.9$</td>
<td>-0.020 0.057</td>
<td>0.016 <strong>0.052</strong></td>
<td><strong>0.006</strong> 0.069</td>
<td>0.010 0.053</td>
</tr>
<tr>
<td>$b_{01} = -0.54$</td>
<td>0.294 1.967</td>
<td>-0.071 <strong>0.205</strong></td>
<td><strong>0.011</strong> 0.340</td>
<td>-0.082 0.223</td>
</tr>
<tr>
<td>$\sigma^2_{\theta_0} = 1$</td>
<td>0.127 0.340</td>
<td>-0.045 0.336</td>
<td><strong>-0.029</strong> 0.387</td>
<td>-0.083 <strong>0.291</strong></td>
</tr>
<tr>
<td>$\psi_{01} = 0.9$</td>
<td>-0.022 0.061</td>
<td>0.016 <strong>0.053</strong></td>
<td><strong>0.007</strong> 0.072</td>
<td>0.010 0.055</td>
</tr>
<tr>
<td>$b_{01} = -0.63$</td>
<td>0.383 2.218</td>
<td>-0.079 <strong>0.226</strong></td>
<td><strong>-0.014</strong> 0.338</td>
<td>-0.096 0.481</td>
</tr>
<tr>
<td>$\sigma^2_{\theta_0} = 1$</td>
<td>0.210 0.497</td>
<td>-0.067 <strong>0.333</strong></td>
<td><strong>-0.059</strong> 0.427</td>
<td>-0.139 0.392</td>
</tr>
<tr>
<td>$\psi_{01} = 0.9$</td>
<td>-0.026 0.068</td>
<td>0.016 <strong>0.054</strong></td>
<td><strong>0.008</strong> 0.077</td>
<td>0.01 0.058</td>
</tr>
<tr>
<td>$b_{01} = -0.75$</td>
<td>0.495 4.315</td>
<td>-0.059 <strong>0.277</strong></td>
<td><strong>-0.038</strong> 0.363</td>
<td><strong>0.021</strong> 2.411</td>
</tr>
<tr>
<td>$\sigma^2_{\theta_0} = 1$</td>
<td>0.403 1.109</td>
<td><strong>-0.066</strong> 0.355</td>
<td>-0.107 0.497</td>
<td>-0.238 0.621</td>
</tr>
<tr>
<td>$\psi_{01} = 0.9$</td>
<td>-0.035 0.094</td>
<td>0.012 <strong>0.054</strong></td>
<td><strong>0.004</strong> 0.094</td>
<td>0.010 0.070</td>
</tr>
<tr>
<td>$b_{01} = -0.99$</td>
<td>2.200 9.022</td>
<td>-0.069 <strong>0.282</strong></td>
<td><strong>-0.009</strong> 0.576</td>
<td>1.864 8.840</td>
</tr>
<tr>
<td>$\sigma^2_{\theta_0} = 1$</td>
<td>2.864 11.589</td>
<td><strong>-0.069</strong> 0.282</td>
<td>-0.241 0.828</td>
<td>-1.400 7.050</td>
</tr>
<tr>
<td>$\psi_{01} = 0.9$</td>
<td>-0.040 0.110</td>
<td>0.012 <strong>0.067</strong></td>
<td><strong>0.004</strong> 0.110</td>
<td>0.010 0.080</td>
</tr>
<tr>
<td>$b_{01} = -1.1$</td>
<td>2.417 9.138</td>
<td><strong>-0.065</strong> 0.304</td>
<td>0.254 2.665</td>
<td>4.372 12.547</td>
</tr>
<tr>
<td>$\sigma^2_{\theta_0} = 1$</td>
<td>13.896 65.483</td>
<td><strong>-0.096</strong> 0.708</td>
<td>-0.286 1.035</td>
<td>-5.591 44.282</td>
</tr>
</tbody>
</table>
Table 3: Test of conditional homoscedasticity against a LARCH(q) model for stock market indices.

<table>
<thead>
<tr>
<th></th>
<th>R_n</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC</td>
<td></td>
<td>1.0</td>
<td>18.9</td>
<td>24.9</td>
<td>31</td>
<td>31</td>
<td>35.8</td>
<td>40</td>
<td>55.6</td>
<td>56.2</td>
<td>58.8</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.025</td>
<td>0.006</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Shanghai</td>
<td>R_n</td>
<td>0.5</td>
<td>0.6</td>
<td>1.7</td>
<td>5.1</td>
<td>8.4</td>
<td>8.8</td>
<td>8.8</td>
<td>12.4</td>
<td>15</td>
<td>15.4</td>
<td>16</td>
<td>17.1</td>
</tr>
<tr>
<td>p-value</td>
<td>0.479</td>
<td>0.728</td>
<td>0.643</td>
<td>0.28</td>
<td>0.136</td>
<td>0.186</td>
<td>0.267</td>
<td>0.132</td>
<td>0.092</td>
<td>0.12</td>
<td>0.142</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td>DAX</td>
<td>R_n</td>
<td>8.3</td>
<td>14.4</td>
<td>17.7</td>
<td>19.3</td>
<td>21</td>
<td>21</td>
<td>22.4</td>
<td>23.1</td>
<td>30.1</td>
<td>30.3</td>
<td>36.9</td>
<td>37.9</td>
</tr>
<tr>
<td>p-value</td>
<td>0.004</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
<td>0.002</td>
<td>0.003</td>
<td>0</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DJA</td>
<td>R_n</td>
<td>5.5</td>
<td>23.9</td>
<td>26</td>
<td>26.2</td>
<td>29.8</td>
<td>30.8</td>
<td>36.7</td>
<td>38.7</td>
<td>41</td>
<td>45.1</td>
<td>45.7</td>
<td>50.2</td>
</tr>
<tr>
<td>p-value</td>
<td>0.019</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DJT</td>
<td>R_n</td>
<td>1.1</td>
<td>8.6</td>
<td>11.1</td>
<td>11.2</td>
<td>11.9</td>
<td>14.2</td>
<td>16.2</td>
<td>16.3</td>
<td>22.2</td>
<td>22.6</td>
<td>22.7</td>
<td>26.9</td>
</tr>
<tr>
<td>p-value</td>
<td>0.303</td>
<td>0.011</td>
<td>0.025</td>
<td>0.025</td>
<td>0.036</td>
<td>0.028</td>
<td>0.023</td>
<td>0.039</td>
<td>0.008</td>
<td>0.012</td>
<td>0.019</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>FTSE</td>
<td>R_n</td>
<td>6.3</td>
<td>12.9</td>
<td>15.8</td>
<td>21</td>
<td>25.5</td>
<td>25.7</td>
<td>33.4</td>
<td>35.5</td>
<td>51.4</td>
<td>52.5</td>
<td>53</td>
<td>54</td>
</tr>
<tr>
<td>p-value</td>
<td>0.012</td>
<td>0.002</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nasdaq</td>
<td>R_n</td>
<td>3.2</td>
<td>8.1</td>
<td>8.2</td>
<td>8.3</td>
<td>11.5</td>
<td>11.5</td>
<td>11.6</td>
<td>11.6</td>
<td>12</td>
<td>12.6</td>
<td>12.6</td>
<td>14.9</td>
</tr>
<tr>
<td>p-value</td>
<td>0.075</td>
<td>0.018</td>
<td>0.043</td>
<td>0.08</td>
<td>0.043</td>
<td>0.074</td>
<td>0.116</td>
<td>0.172</td>
<td>0.216</td>
<td>0.247</td>
<td>0.319</td>
<td>0.248</td>
<td></td>
</tr>
<tr>
<td>Nikkei</td>
<td>R_n</td>
<td>11.6</td>
<td>28.5</td>
<td>32</td>
<td>32.1</td>
<td>44</td>
<td>45.8</td>
<td>50.7</td>
<td>53.1</td>
<td>57.2</td>
<td>58.5</td>
<td>59.3</td>
<td>64.1</td>
</tr>
<tr>
<td>p-value</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SP 500</td>
<td>R_n</td>
<td>6.7</td>
<td>27.8</td>
<td>29.6</td>
<td>29.6</td>
<td>38.1</td>
<td>45.1</td>
<td>47.1</td>
<td>48.4</td>
<td>55.4</td>
<td>61.6</td>
<td>62.1</td>
<td>66.8</td>
</tr>
<tr>
<td>p-value</td>
<td>0.009</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

of the next section.

7.2 Application to financial series

8 Conclusion

LARCH is an attractive class of models for conditional heteroscedasticity, which is able to capture different effects of the volatility, keeping the parsimony of the standard ARCH and avoiding the positivity constraints on the coefficients. However, the QMLE is not recommended for these models. This paper has shown that this method produces inconsistent estimator. The theoretical results were confirmed by
Table 4: LARCH(5) models for stock market indices.

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
<th>$b_4$</th>
<th>$b_5$</th>
<th>$\sigma^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC</td>
<td></td>
<td>-0.086</td>
<td>-0.075</td>
<td>-0.159</td>
<td>-0.136</td>
<td>-0.123</td>
<td>1.424</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.013</td>
<td>0.013</td>
<td>0.014</td>
<td>0.014</td>
<td>0.013</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-6.66</td>
<td>-5.65</td>
<td>-11.48</td>
<td>-10.04</td>
<td>-9.22</td>
<td></td>
</tr>
<tr>
<td>Shanghai</td>
<td>Estimate</td>
<td>-0.084</td>
<td>-0.074</td>
<td>-0.104</td>
<td>-0.096</td>
<td>-0.11</td>
<td>1.878</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.025</td>
<td>0.025</td>
<td>0.025</td>
<td>0.025</td>
<td>0.025</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-3.4</td>
<td>-2.99</td>
<td>-4.15</td>
<td>-3.85</td>
<td>-4.35</td>
<td></td>
</tr>
<tr>
<td>DAX</td>
<td></td>
<td>-0.141</td>
<td>-0.209</td>
<td>-0.164</td>
<td>-0.206</td>
<td>-0.139</td>
<td>1.228</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.017</td>
<td>0.018</td>
<td>0.018</td>
<td>0.019</td>
<td>0.018</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-8.29</td>
<td>-11.43</td>
<td>-9.1</td>
<td>-10.98</td>
<td>-7.66</td>
<td></td>
</tr>
<tr>
<td>DJA</td>
<td></td>
<td>-0.219</td>
<td>-0.5</td>
<td>-0.421</td>
<td>0.218</td>
<td>-0.071</td>
<td>0.453</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.036</td>
<td>0.045</td>
<td>0.043</td>
<td>0.037</td>
<td>0.034</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-6.06</td>
<td>-11.08</td>
<td>-9.91</td>
<td>5.93</td>
<td>-2.09</td>
<td></td>
</tr>
<tr>
<td>DJT</td>
<td></td>
<td>-0.034</td>
<td>-0.132</td>
<td>-0.114</td>
<td>0.044</td>
<td>-0.041</td>
<td>1.577</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.019</td>
<td>0.021</td>
<td>0.02</td>
<td>0.019</td>
<td>0.018</td>
<td>0.062</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-1.78</td>
<td>-6.42</td>
<td>-5.8</td>
<td>2.33</td>
<td>-2.25</td>
<td></td>
</tr>
<tr>
<td>FTSE</td>
<td></td>
<td>-0.186</td>
<td>-0.113</td>
<td>-0.218</td>
<td>-0.211</td>
<td>-0.213</td>
<td>0.871</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.018</td>
<td>0.018</td>
<td>0.018</td>
<td>0.019</td>
<td>0.018</td>
<td>0.022</td>
</tr>
<tr>
<td>Nasdaq</td>
<td></td>
<td>-0.344</td>
<td>-0.673</td>
<td>-0.099</td>
<td>-0.034</td>
<td>-0.051</td>
<td>0.691</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.024</td>
<td>0.03</td>
<td>0.022</td>
<td>0.022</td>
<td>0.023</td>
<td>0.025</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-14.33</td>
<td>-22.25</td>
<td>-4.43</td>
<td>-1.51</td>
<td>-2.26</td>
<td></td>
</tr>
<tr>
<td>Nikkei</td>
<td></td>
<td>-0.042</td>
<td>-0.064</td>
<td>-0.056</td>
<td>-0.035</td>
<td>-0.055</td>
<td>1.762</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.013</td>
<td>0.014</td>
<td>0.014</td>
<td>0.013</td>
<td>0.014</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-3.19</td>
<td>-4.7</td>
<td>-4.11</td>
<td>-2.62</td>
<td>-4.06</td>
<td></td>
</tr>
<tr>
<td>SP 500</td>
<td></td>
<td>-0.323</td>
<td>-0.545</td>
<td>-0.257</td>
<td>0.086</td>
<td>-0.081</td>
<td>0.531</td>
</tr>
<tr>
<td></td>
<td>Standard Error</td>
<td>0.028</td>
<td>0.033</td>
<td>0.027</td>
<td>0.026</td>
<td>0.025</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>t-ratio</td>
<td>-11.69</td>
<td>-16.63</td>
<td>-9.5</td>
<td>3.35</td>
<td>-3.22</td>
<td></td>
</tr>
</tbody>
</table>
Table 5: Portmanteau test of adequacy of the AR(0) model (absence of linear part) for
the linear dynamics of nine stock market returns.

<table>
<thead>
<tr>
<th></th>
<th>$m$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC</td>
<td>$\hat{Q}^u_m$</td>
<td>0.1</td>
<td>0.4</td>
<td>5.1</td>
<td>5.8</td>
<td>10.6</td>
<td>11.3</td>
<td>12.6</td>
<td>12.8</td>
<td>12.8</td>
<td>12.9</td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.816</td>
<td>0.824</td>
<td>0.163</td>
<td>0.212</td>
<td>0.059</td>
<td>0.080</td>
<td>0.12</td>
<td>0.173</td>
<td>0.227</td>
<td>0.267</td>
<td></td>
</tr>
<tr>
<td>Shanghai</td>
<td>$\hat{Q}^u_m$</td>
<td>0.0</td>
<td>1.1</td>
<td>3.3</td>
<td>5.8</td>
<td>6.1</td>
<td>8.3</td>
<td>8.6</td>
<td>8.7</td>
<td>8.7</td>
<td>9.2</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.853</td>
<td>0.577</td>
<td>0.351</td>
<td>0.218</td>
<td>0.292</td>
<td>0.219</td>
<td>0.283</td>
<td>0.371</td>
<td>0.463</td>
<td>0.509</td>
<td>0.427</td>
</tr>
<tr>
<td>DAX</td>
<td>$\hat{Q}^u_m$</td>
<td>0.2</td>
<td>0.2</td>
<td>3.5</td>
<td>6.0</td>
<td>7.3</td>
<td>10.4</td>
<td>10.6</td>
<td>11.3</td>
<td>11.3</td>
<td>11.9</td>
<td>13.6</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.634</td>
<td>0.893</td>
<td>0.316</td>
<td>0.202</td>
<td>0.199</td>
<td>0.107</td>
<td>0.156</td>
<td>0.186</td>
<td>0.256</td>
<td>0.294</td>
<td>0.255</td>
</tr>
<tr>
<td>DIA</td>
<td>$\hat{Q}^u_m$</td>
<td>0.6</td>
<td>1.2</td>
<td>1.2</td>
<td>1.4</td>
<td>1.9</td>
<td>4.2</td>
<td>8.4</td>
<td>8.5</td>
<td>9.0</td>
<td>9.5</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.458</td>
<td>0.547</td>
<td>0.751</td>
<td>0.847</td>
<td>0.859</td>
<td>0.65</td>
<td>0.297</td>
<td>0.384</td>
<td>0.435</td>
<td>0.486</td>
<td>0.51</td>
</tr>
<tr>
<td>DJT</td>
<td>$\hat{Q}^u_m$</td>
<td>8.1</td>
<td>10.3</td>
<td>11.3</td>
<td>12.6</td>
<td>12.8</td>
<td>17.3</td>
<td>20.8</td>
<td>21.4</td>
<td>21.4</td>
<td>21.4</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.004</td>
<td>0.006</td>
<td>0.01</td>
<td>0.013</td>
<td>0.025</td>
<td>0.008</td>
<td>0.004</td>
<td>0.006</td>
<td>0.011</td>
<td>0.018</td>
<td>0.029</td>
</tr>
<tr>
<td>FTSE</td>
<td>$\hat{Q}^u_m$</td>
<td>1.1</td>
<td>1.8</td>
<td>14.4</td>
<td>16.1</td>
<td>17.7</td>
<td>19.9</td>
<td>20.0</td>
<td>20.6</td>
<td>20.8</td>
<td>21.1</td>
<td>23.3</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.303</td>
<td>0.399</td>
<td>0.002</td>
<td>0.003</td>
<td>0.003</td>
<td>0.003</td>
<td>0.005</td>
<td>0.008</td>
<td>0.014</td>
<td>0.02</td>
<td>0.016</td>
</tr>
<tr>
<td>Nasdaq</td>
<td>$\hat{Q}^u_m$</td>
<td>1.4</td>
<td>4.0</td>
<td>4.3</td>
<td>4.3</td>
<td>4.7</td>
<td>4.9</td>
<td>5.5</td>
<td>7.1</td>
<td>7.2</td>
<td>7.3</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.243</td>
<td>0.138</td>
<td>0.265</td>
<td>0.367</td>
<td>0.449</td>
<td>0.555</td>
<td>0.6</td>
<td>0.528</td>
<td>0.614</td>
<td>0.694</td>
<td>0.771</td>
</tr>
<tr>
<td>Nikkei</td>
<td>$\hat{Q}^u_m$</td>
<td>0.4</td>
<td>9.3</td>
<td>9.5</td>
<td>9.5</td>
<td>9.5</td>
<td>10.9</td>
<td>10.9</td>
<td>11.2</td>
<td>11.5</td>
<td>14.5</td>
<td>14.7</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.532</td>
<td>0.01</td>
<td>0.024</td>
<td>0.05</td>
<td>0.091</td>
<td>0.091</td>
<td>0.142</td>
<td>0.192</td>
<td>0.242</td>
<td>0.152</td>
<td>0.195</td>
</tr>
<tr>
<td>SP 500</td>
<td>$\hat{Q}^u_m$</td>
<td>0.6</td>
<td>1.4</td>
<td>2.6</td>
<td>2.6</td>
<td>4.6</td>
<td>6.2</td>
<td>9.6</td>
<td>9.6</td>
<td>9.8</td>
<td>10.2</td>
<td>10.4</td>
</tr>
<tr>
<td></td>
<td>$p$-val</td>
<td>0.431</td>
<td>0.499</td>
<td>0.456</td>
<td>0.623</td>
<td>0.461</td>
<td>0.403</td>
<td>0.215</td>
<td>0.292</td>
<td>0.369</td>
<td>0.42</td>
<td>0.493</td>
</tr>
</tbody>
</table>
finite-sample experiments. It is interesting to note that a major estimation technique, which is very robust under change of the distribution of the iid noise, fails for a class of conditionally heteroscedastic models. To our knowledge, this is the only example of failure of the QMLE, in GARCH-type models, that is not due to the lack of a moment condition.

To overcome this problem, we proposed a self-weighted LSE. For AR-LARCH models, this estimator was shown to be asymptotically normal under moment conditions depending on the choice of weights for the AR and ARCH parts. These results were used to construct Wald and score tests for testing conditional homoscedasticity. Furthermore, diagnostic portmanteau tests were developed. Their asymptotic distribution was shown to be far from the standard chi-square. It is possible to extend the class to GARCH-type models, allowing the volatility to depend on its own past values. This is left for future research.

Appendix: Proofs

A.1 Proof of Theorem 2.1

The scheme of the proof is standard (see e.g. Francq and Zakoïan, 2004, Theorems 2.1 and 3.1), and consists in showing

\begin{enumerate}
  \item[(i)] \( u_t(\theta) = u_t(\theta_0) \) and \( \sigma_t^2(\theta) = \sigma_t^2(\theta_0) \) \( P_{\theta_0} \) a.s. for all \( t \Rightarrow \theta = \theta_0 \),
  \item[(ii)] \( E|\ell_t(\theta_0)| < \infty \), and if \( \theta \neq \theta_0 \), \( E\ell_t(\theta) > E\ell_t(\theta_0) \),
  \item[(iii)] any \( \theta \neq \theta_0 \) has a neighborhood \( V(\theta) \) such that
    \[
    \liminf_{n \to \infty} \inf_{\theta^* \in V(\theta)} I_{n}(\theta^*) > E\ell_1(\theta_0), \text{ a.s.}
    \]
\end{enumerate}

We first prove i). In view of A2 and A5, we have \( \sigma_t^2(\theta_0) = \text{Var}(x_t \mid \mathcal{F}_{t-1}) > 0 \) with probability 1, and it can be shown that \( u_t(\theta) = u_t(\theta_0) \) entails that the first
$p$ components of $\theta$ and $\theta_0$ are the same. Let $\theta$ such that $\sigma_t^2 = \sigma_t^2(\theta) = \sigma_t^2(\theta_0) \neq 0$ and $u_t = u_t(\theta) = u_t(\theta_0)$ a.s. Writing $\sigma_t(\theta) = \sigma_t\{b_1u_{t-1} + v_{t-2}(\theta)\}$ where $v_{t-2}(\theta) = 1 + \sum_{i=2}^q b_iu_{t-i}$, we have

$$
\sigma_t^2(\theta) = \sigma_t^2(\theta) \\
\Leftrightarrow \sigma_0\{b_1u_{t-1} + v_{t-2}(\theta_0)\}^2 = \sigma_0\{b_1u_{t-1} + v_{t-2}(\theta)\}^2 \\
\Leftrightarrow (\sigma_t^2b_1^2 - \sigma_0^2b_0^2)\eta_{t-1}^2 + 2\sigma_{t-1}\{\sigma_t^2b_1v_{t-2}(\theta) - \sigma_0b_0v_{t-2}(\theta)\}\eta_{t-1} \\
+\{\sigma_t^2v_{t-2}(\theta) - \sigma_0^2v_{t-2}(\theta_0)\} = a_{t-2}\eta_{t-1}^2 + b_{t-2}\eta_{t-1} + c_{t-2} = 0.
$$

By taking the expectation of the last equality conditional on $\mathcal{F}_{t-2}$ we get $a_{t-2} + c_{t-2} = 0$. We thus have

$$
a_{t-2}(\eta_{t-1}^2 - 1) = -b_{t-2}\eta_{t-1} \quad \text{a.s.} \quad \text{(A.1)}
$$

Suppose that $\sigma_t^2b_1^2 \neq \sigma_0^2b_0^2$, that is $a_{t-2} \neq 0$ a.s. It follows that $\eta_{t-1} \neq 0$ and $(\eta_{t-1}^2 - 1)/\eta_{t-1} = -b_{t-2}/a_{t-2}$ a.s. Because the two sides of this equality involve independent variables, these variables are constant. Hence there is a constant $c$ such that $\eta_{t-1}^2 - 1 = c\eta_{t-1}$, but this contradicts A5. We thus have proved that $\sigma_t^2b_1^2 = \sigma_0^2b_0^2$. If $b_1 = 0$ we have $b_1 = b_0$. Now suppose $b_0 \neq 0$. Since $a_{t-2} = 0$ a.s. we have, from (A.1),

$$
b_{t-2} = 0 = \{\sigma_t^2b_1v_{t-2}(\theta) - \sigma_0b_0v_{t-2}(\theta_0)\}\sigma_{t-1}\eta_{t-1}.
$$

Multiplying the last equation by $\eta_{t-1}$ and taking the expectation conditional to $\mathcal{F}_{t-2}$ yields

$$
\sigma_t^2b_1\sigma_{t-1}v_{t-2}(\theta) = \sigma_0^2b_0\sigma_{t-1}v_{t-2}(\theta_0)
$$

and thus, since by assumption $\sigma_{t-1} \neq 0$ and since we have $\sigma_t^2b_1^2 = \sigma_0^2b_0^2$,

$$
b_0v_{t-2}(\theta) = b_1v_{t-2}(\theta_0)
$$
which, by taking the expectation, implies \( b_{01} = b_1 \). Proceeding similarly we get, recursively, \( b_{0i} = b_i \) for all \( i \). Finally, \( \sigma_c = \sigma_{0c} \) and \( \theta = \theta_0 \).

Now we turn to ii). Note that, by \textbf{A1} and \textbf{A2}, the process \((x_t)\) is stationary and ergodic (see \textit{e.g.} Billingsley (1995, Theorem 36.4)). Since \( \ell_i(\theta) \) is a measurable function of \( x_t, \ldots, x_{t-p-q} \), the process \( \{\ell_t(\theta)\} \) is also stationary and ergodic. Moreover, in view of \textbf{A5}, \( E\ell_t(\theta) \) exists in \( \mathbb{R} \cup \{+\infty\} \). Thus the objective function \( l_n(\theta) \) converges a.s. to \( E\ell_t(\theta) \) as \( n \to \infty \). In FMZ it was shown that under \textbf{A1}, \( E\sigma_t^{2s}(\theta_0) < \infty \) for some sufficiently small \( s > 0 \). It follows that \( E\ell_t(\theta_0) = 1 + \frac{1}{2} E\log \sigma_t^{2s}(\theta_0) \) exists in \( \mathbb{R} \). The limit criterion is minimum at the true value because

\[
E\ell_t(\theta) - E\ell_t(\theta_0) = E \left\{ \log \frac{\sigma_t^2(\theta)}{\sigma_t^2(\theta_0)} + \frac{\sigma_t^2(\theta_0)}{\sigma_t^2(\theta)} - 1 \right\} + E \frac{\{u_t(\theta) - u_t(\theta_0)\}^2}{\sigma_t^2(\theta)} + E \frac{2\epsilon_t\sigma_t(\theta_0)\{u_t(\theta) - u_t(\theta_0)\}}{\sigma_{0c}\sigma_t^2(\theta)} \geq 0
\]

using the fact that the last expectation is null (\( \epsilon_t \) being orthogonal to the random variable \( \sigma_t(\theta_0)\{u_t(\theta) - u_t(\theta_0)\}\sigma_t^{-2}(\theta) \in \mathcal{F}_{t-1} \)), and using the elementary inequality \( \log x \leq x - 1 \). Moreover the inequality is an equality if and only if \( u_t(\theta) - u_t(\theta_0) = 0 \) and \( \sigma_t^2(\theta_0) = \sigma_t^2(\theta) \) with probability 1, which by ii) implies \( \theta = \theta_0 \).

As in Francq and Zakoïan (2004) we can show that the ergodic theorem and the continuity of \( \theta \mapsto E\ell_1(\theta) \) entail iii). A standard compactness argument allows to complete the proof of the consistency.

Now we turn to the asymptotic normality. It is easy to see that the proof follows from the following properties:

\[
i) \quad \mathbb{E} \left\| \frac{\partial \ell_t(\theta_0)}{\partial \theta} \frac{\partial \ell_t(\theta_0)}{\partial \theta'} \right\| < \infty \quad \text{and} \quad n^{-1/2} \sum_{t=1}^n \frac{\partial \ell_t}{\partial \theta}(\theta_0) \Rightarrow \mathcal{N}(0, I),
\]

\[
ii) \quad \mathbb{E} \left\| \frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'} \right\| < \infty \quad \text{and} \quad n^{-1} \sum_{t=1}^n \frac{\partial^2 \ell_t}{\partial \theta_i \partial \theta_j}(\theta^*) \to \mathcal{J}(i,j) \quad \text{a.s.,}
\]

30
for any \( \theta^* \) between \( \hat{\theta}_n \) and \( \theta_0 \).

\( iii \) \( I \) and \( J \) are not singular.

Differentiating (2.3) we obtain

\[
\frac{\partial \ell_t(\theta)}{\partial \theta} = \left\{ 1 - \frac{u_t^2(\theta)}{\sigma_t^2(\theta)} \right\} \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} + 2 \frac{u_t(\theta)}{\sigma_t^2(\theta)} \frac{\partial u_t(\theta)}{\partial \theta} = \left\{ 1 - \frac{u_t^2(\theta)}{\sigma_t^2(\theta)} \right\} \frac{2}{1 + \sum_{i=1}^q b_i u_{t-i}(\theta)} \begin{pmatrix}
- \sum_{i=1}^q b_i X_{t-i} \\
u_{t-1}(\theta) \\
\vdots \\
u_{t-q}(\theta) \\
\frac{1 + \sum_{i=1}^q b_i u_{t-i}(\theta)}{2 \sigma_t^2(\theta)}
\end{pmatrix}
+ 2 \frac{u_t(\theta)}{\sigma_t^2(\theta)} \begin{pmatrix}
-X_t \\
0_{q+1}
\end{pmatrix}
\]

with \( X_t = (x_{t-1}, \ldots, x_{t-p})' \). Noting that \( \left\{ 1 - \frac{u_t^2(\theta_0)}{\sigma_t^2(\theta_0)} \right\} = 1 - \epsilon_t^2 / \sigma_t^2 \) and \( u_t(\theta_0) / \sigma_t(\theta_0) = \epsilon_t / \sigma_t \) are centered and independent of the other random variables involved in \( \partial \ell_t(\theta_0) / \partial \theta \), it can be shown that, under \( A_2, A_5 \) and \( A_7 \), \( (\partial \ell_t(\theta_0) / \partial \theta, F_t) \) is a square integrable stationary martingale difference. Thus i) comes from the Central Limit Theorem (CLT) of Billingsley (1961).

Differentiating (A.2) we obtain

\[
\frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta'} = \left( 1 - \frac{u_t^2(\theta)}{\sigma_t^2(\theta)} \right) \frac{1}{\sigma_t^2(\theta)} \frac{\partial^2 \sigma_t^2(\theta)}{\partial \theta \partial \theta'} + \left( \frac{2 u_t^2(\theta)}{\sigma_t^2(\theta)} - 1 \right) \frac{1}{\sigma_t^2(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} + \frac{2}{\sigma_t^2(\theta)} \frac{\partial u_t(\theta)}{\partial \theta} \frac{\partial u_t(\theta)}{\partial \theta'} + \frac{2 u_t(\theta)}{\sigma_t^2(\theta)} \frac{\partial^2 u_t(\theta)}{\partial \theta \partial \theta'} - 2 \frac{u_t(\theta)}{\sigma_t^2(\theta)} \left( \frac{\partial u_t(\theta)}{\partial \theta} \frac{\partial^2 \sigma_t^2(\theta)}{\partial \theta \partial \theta'} + \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial u_t(\theta)}{\partial \theta'} \right).
\]

Using the Hölder inequality, the compactness assumption \( A_3 \), the existence of fourth-order moments for \( x_t \) and \( u_t(\theta) \) and Assumption \( A_5 \), it can be shown that

\[
\sup_{\theta \in \Theta} \left\| \frac{\partial \ell_t(\theta)}{\partial \theta} \right\|_{4/3} < \infty.
\]
With the same arguments it can be shown that
\[ \sup_{\theta \in \Theta} \left| \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta'} \right|_1 < \infty. \] (A.3)

The continuity of \( \theta \mapsto \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta'} \), the ergodic theorem and the dominated convergence theorem now entail that for any \( \varepsilon > 0 \) there exists a neighborhood \( V(\theta_0) \) of \( \theta_0 \) such that, a.s.
\[ \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta \in V(\theta_0)} \left| \frac{\partial^2 \ell_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'} \right| \leq \varepsilon. \] (A.4)

A direct application of the ergodic theorem entails
\[ \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \frac{\partial^2 \ell_t(\theta_0)}{\partial \theta \partial \theta'} = J \quad \text{a.s.} \] (A.5)

Thus ii) comes from (A.3), (A.4), (A.5) and the strong consistency of \( \hat{\theta}_n \).

The arguments used by Francq and Zakoïan (2004, p 631) show that if \( I \) is singular then there exists \( \lambda = (\lambda_1', \lambda_2')' \), with \( \lambda_1 \in \mathbb{R}^p \) and \( \lambda_2 \in \mathbb{R}^{q+1} \), such that a.s.
\[ \lambda_1' \frac{\partial u_t(\theta_0)}{\partial \theta} = 0 \quad \text{and} \quad \lambda_2' \frac{\partial^2 \sigma_t^2(\theta_0)}{\partial \theta} = 0. \] (A.6)

Because \( \frac{\partial u_t(\theta_0)}{\partial \theta} = (-X_t', 0_{q+1}')' \) the first equality entails \( \lambda_1 = 0 \), and the second equality reduces to
\[
0 = \lambda_2' \begin{pmatrix}
\frac{\partial^2 \sigma_t^2(\theta_0)}{\partial \theta_1^2} \\
\vdots \\
\frac{\partial^2 \sigma_t^2(\theta_0)}{\partial \theta_q^2} \\
\frac{\partial^2 \sigma_t^2(\theta_0)}{\partial \sigma_{\epsilon}^2}
\end{pmatrix} = \lambda_2' \begin{pmatrix}
2\sigma^2_{0 \epsilon} \left(1 + \sum_{i=1}^{q} b_{0i} u_{t-i} \right) u_{t-1} \\
\vdots \\
2\sigma^2_{0 \epsilon} \left(1 + \sum_{i=1}^{q} b_{0i} u_{t-i} \right) u_{t-q} \\
(1 + \sum_{i=1}^{q} b_{0i} u_{t-i})^2
\end{pmatrix} \quad \text{a.s.}
\]

Using the stationarity, we deduce that, conditional on \( \{ \epsilon_u, u < t \} \) there exists a polynomial of degree 2, \( P_2(x) = a_0 + a_1 x + a_2 x^2 \), such that \( P_2(u_t) = 0 \), which contradicts A5. Moreover
\[
\mathcal{J} = E \left( \frac{1}{\sigma_t^2} \frac{\partial \sigma_t^2}{\partial \theta} \frac{\partial \sigma_t^2}{\partial \theta'}(\theta_0) \right) + 2E \left( \frac{1}{\sigma_t^2} \frac{\partial u_t}{\partial \theta} \frac{\partial u_t}{\partial \theta'}(\theta_0) \right) : = \mathcal{A} + \mathcal{B}
\]
where \( \mathcal{A} \) is strictly positive definite, by the previous arguments, and \( \mathcal{B} \) is positive semi-definite. Thus \( \mathcal{I} \) and \( \mathcal{J} \) are invertible.
A.2 Proof of Proposition 3.1

For any fixed integer $t_0$, with probability one we have $x_{t_0-1} \neq 0$, $x_{t_0}/x_{t_0-1} \neq \psi_1$ and $x_{t_0-1}^2 - x_{t_0} x_{t_0-2} \neq 0$. For $\Theta$ sufficiently large

$$\theta(t_0) := \left( \frac{x_{t_0}}{x_{t_0-1}}, 0'_{p-1}, -\frac{1}{x_{t_0-1} - x_{t_0} x_{t_0-2}}, 0'_{q-1}, 1 \right) \in \Theta.$$ 

Note that $u_t \{ \theta(t_0) \} = \sigma_t^2 \{ \theta(t_0) \} = 0$. It follows that, with the conventions given after (2.3), $L_n \{ \theta(t_0) \} = +\infty$. The measurable sequences $(\hat{\theta}_n)_{n \geq 1}$ such that $\hat{\theta}_n = \theta(t_0)$ for all $n \geq t_0$ are inconsistent sequences of QMLE.

A.3 Proof of Theorem 4.1.

Writing $Y = X\psi_0 + U$ with $U' = (u_1, \ldots, u_n)$, we have

$$\hat{\psi}_{WLS} = (X'\Omega X)^{-1} X'\Omega (X\psi_0 + U) = \psi_0 + (X'\Omega X)^{-1} X'\Omega U = \psi_0 + o(1)$$

a.s., because in view of the ergodic theorem

$$n^{-1}X'\Omega X \rightarrow A_{\psi}, \quad n^{-1}X'\Omega U \rightarrow E\omega_t u_t X_t = E\epsilon_t \sigma_t^{-1} E\sigma_t(\theta_0) \omega_t X_t = 0.$$ 

The consistency is shown. Applying the CLT of Billingsley (1961) to the square integrable stationary martingale difference $(\omega_t u_t X_t, F_t)$, we obtain that $n^{-1/2}X'\Omega U$ converges in law to the $\mathcal{N}(0, B_{\psi})$ distribution. To complete the proof, it remains to show that $A_{\psi}$ is invertible. If $A_{\psi}$ was singular then there would exist $\lambda \neq 0 \in \mathbb{R}^p$ such that $\lambda'\sqrt{\omega_t} X_t = 0$ which would imply $\lambda' X_t = 0$ with probability one. This would entail that $x_t$, $u_t$ and $\epsilon_t$ belong to $F_{t-1}$, and $\epsilon_t$ would be independent of $\epsilon_{t-1}$. This is clearly impossible because $E\epsilon_t = 0$ and $E\epsilon_t^2 \neq 0$. Thus $A_{\psi}$ is invertible, and the proof is complete.
A.4 Proof of Theorem 4.2.

Let

\[ \hat{Q}_n(\beta) = \frac{1}{n} \sum_{t=1}^{n} \tau_t \nu_t^2(\hat{\psi}, \beta), \quad Q_n(\beta) = \frac{1}{n} \sum_{t=1}^{n} \tau_t \nu_t^2(\psi_0, \beta). \]

We first show that

\[ \lim_{n \to \infty} \sup_{\beta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^{n} \tau_t \nu_t^2(\hat{\psi}, \beta) - \frac{1}{n} \sum_{t=1}^{n} \tau_t \nu_t^2(\psi_0, \beta) \right| = 0 \quad \text{a.s. (A.7)} \]

We have, for some constant $K$

\[
\left| \nu_t^2(\hat{\psi}, \beta) - \nu_t^2(\psi_0, \beta) \right|
\leq \left| \nu_t(\hat{\psi}, \beta) - \nu_t(\psi_0, \beta) \right| 2 \sup_{\theta \in \Theta} \left| \nu_t(\psi, \beta) \right|
\leq K \left\{ \left| u_t(\hat{\psi}) - u_t \right| \sup_{\psi \in \Theta} \left| u_t(\psi) \right|
+ \left( \sum_{i=1}^{q} \left| b_i \right| \left| u_{t-i}(\hat{\psi}) - u_{t-i} \right| \right) \sup_{\theta \in \Theta} \left| \sigma_t^2(\psi, \beta) \right| \right\} \sup_{\theta \in \Theta} \left| \nu_t(\psi, \beta) \right|
\]

and

\[ \left| u_t(\hat{\psi}) - u_t \right| \leq \sum_{i=1}^{p} \left| \hat{\psi}_i - \psi_0 \right| \left| x_{t-i} \right|. \]

It follows that

\[ \sup_{\beta \in \Theta} \left| \nu_t^2(\hat{\psi}, \beta) - \nu_t^2(\psi_0, \beta) \right| \leq M_t \left\| \hat{\psi} - \psi_0 \right\|, \]

where $(M_t)$ is a strictly stationary process. For $t$ fixed, the strong consistency of $\hat{\psi}$ implies $M_t \left\| \hat{\psi} - \psi_0 \right\| \to 0$ a.s. Therefore the Cesaro sum $n^{-1} \sum_{t=1}^{n} \tau_t M_t \left\| \hat{\psi} - \psi_0 \right\| \to 0$ a.s. and (A.7) is shown.

This result and the ergodic theorem show that $\hat{Q}_n(\beta) \to Q_\infty(\beta) := E\tau_t \nu_t^2(\psi_0, \beta) \in \mathbb{R}^+ \cup \{+\infty\}$, a.s. and uniformly in a neighborhood of $\beta$, as $n \to \infty$. Since $\tau_t \nu_t(\psi_0, \beta_0) = \tau_t (1 + \sum_{i=1}^{p} b_i u_{t-i}^2) \left( \epsilon_{t-i}^2 - \sigma_0^2 \right)$ and $\tau_t \left\{ \nu_t(\psi_0, \beta) - \nu_t(\psi_0, \beta_0) \right\} = \tau_t \left\{ \sigma_t^2(\psi_0, \beta_0) - \sigma_t^2(\psi_0, \beta) \right\} \in \mathcal{F}_{t-1}$ are orthogonal (when $Q_\infty(\beta)$ is finite, which is the case at $\beta = \beta_0$ in view of the moment condition A11), it can be shown...
that under the identifiability condition \( A4 \), \( Q_\infty(\beta) > Q_\infty(\beta_0) \) when \( \beta \neq \beta_0 \). The consistency follows from standard arguments.

Under \( A6 \), the derivative of the criterion defined in (4.3) vanishes at \( \hat{\beta} = \hat{\beta}_{WLS} \), for sufficiently large \( n \). A Taylor expansion at the order 1 of the derivative around \( \beta_0 \) yields

\[
0 = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\partial \nu_t^2(\hat{\psi}, \beta_0)}{\partial \beta} + \left( \frac{1}{n} \sum_{t=1}^{n} \tau_t \frac{\partial^2 \nu_t^2(\hat{\psi}, \beta_0)}{\partial \beta \partial \beta'} + R_n \right) \sqrt{n} \left( \hat{\beta} - \beta_0 \right),
\]

where the element of the matrix \( R_n \) are of the form

\[
R_n(i, j) = \frac{1}{n} \sum_{t=1}^{n} \tau_t \left\{ \frac{\partial^2 \nu_t^2(\hat{\psi}, \beta_0)}{\partial \beta_i \partial \beta_j} - \frac{\partial^2 \nu_t^2(\hat{\psi}, \beta_0)}{\partial \beta_i \partial \beta_j} \right\}
\]

for some \( \beta^* \) between \( \hat{\beta} \) and \( \beta_0 \). In view of the consistency result, the moment condition \( E \sup_{\theta \in V(\theta_0)} \left\| \frac{\partial^2 \nu_t^2(\theta)}{\partial \theta \partial \theta'} \right\| < \infty \), and the continuity of the derivative, \( R_n(i, j) \to 0 \) a.s. Similar arguments and a Taylor expansion around \( \psi_0 \) yields

\[
0 = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \tau_t \frac{\partial \nu_t^2(\psi_0, \beta_0)}{\partial \beta} + \frac{1}{n} \sum_{t=1}^{n} \tau_t \frac{\partial^2 \nu_t^2(\psi_0, \beta_0)}{\partial \beta \partial \psi'} \sqrt{n} (\hat{\psi} - \psi_0)
\]
\[+ o_P(1) + \left( \frac{1}{n} \sum_{t=1}^{n} \tau_t \frac{\partial^2 \nu_t^2(\psi_0, \beta_0)}{\partial \beta \partial \beta'} + o_P(1) \right) \sqrt{n} \left( \hat{\beta} - \beta_0 \right).
\]

Applying the CLT of Billingsley (1961) to the square integrable stationary martingale difference \( \{(\tau_t \nu_t \partial \nu_t(\psi_0, \beta_0)) / \partial \beta', \omega_t u_t X_t' \}, \mathcal{F}_t \), we obtain

\[
\left( \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \tau_t \frac{\partial \nu_t^2(\psi_0, \beta_0)}{\partial \beta} \right) \sim N \left( 0, \frac{2}{\sqrt{n}} \sum_{t=1}^{n} \tau_t \nu_t \frac{\partial^2 \nu_t^2(\psi_0, \beta_0)}{\partial \beta} \right)
\]

\[
\Rightarrow \left( Z_{\beta} \right) \sim N \left( 0, \left( \begin{array}{cc} 4B_\beta & -2B_\beta A^{-1}_\psi \\ -2A^{-1}_\psi B_\beta & A^{-1}_\psi B_\beta A^{-1}_\psi \end{array} \right) \right)
\]

Applying the ergodic theorem we have a.s.

\[
\frac{1}{n} \sum_{t=1}^{n} \tau_t \frac{\partial^2 \nu_t^2(\psi_0, \beta_0)}{\partial \beta \partial \psi'} \to -2A_{\beta \psi}, \quad \frac{1}{n} \sum_{t=1}^{n} \tau_t \frac{\partial^2 \nu_t^2(\psi_0, \beta_0)}{\partial \beta \partial \beta'} \to 2A_{\beta}.
\]
By arguments already given $A_\beta$ is invertible. Thus

$$\sqrt{n} \begin{pmatrix} \hat{\beta} - \beta_0 \\ \hat{\psi} - \psi_0 \end{pmatrix} \xrightarrow{\mathcal{L}} \begin{pmatrix} (-2A_\beta)^{-1}(Z_\beta - 2A_\beta \psi Z_\psi) \\ Z_\psi \end{pmatrix}$$

and the proof follows.

### A.5 Proof of (5.2)

A Taylor expansion at the order 1 around $\theta_0$ yields

$$0_{q+1} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \tau_t \frac{\partial^2 \nu_t^2(\hat{\psi}, \hat{b}, \hat{\sigma}_\epsilon^2)}{\partial \beta \partial \beta'} \sqrt{n} \begin{pmatrix} \hat{b} \\ \hat{\sigma}_\epsilon^2 - \hat{\sigma}_\epsilon^{2c} \end{pmatrix} + o_P(1).$$

(A.8)

Notice that the last component of the first term in the right-hand side is null. It follows that

$$\sqrt{n}(\hat{\sigma}_\epsilon^2 - \hat{\sigma}_\epsilon^{2c}) = -\frac{1}{A_\sigma} A_{sb} \sqrt{n} \hat{b} + o_P(1).$$

Now using the first $q$ components of (A.8) we get $\Delta_n^c = A_* \sqrt{n} \hat{b} + o_P(1)$, from which the convergence in (5.2) follows.

### A.6 Proof of Lemma 6.1.

We start by establishing a lemma which will be used to show Lemma 6.1. Let, for $0 \leq \ell < n$,

$$\gamma(\ell) = \frac{1}{n} \sum_{t=1}^{n-\ell} u_t u_{t+\ell} \quad \text{and} \quad \rho(\ell) = \frac{\gamma(\ell)}{\gamma(0)}$$

denote the white noise “empirical” autocovariances and autocorrelations. Let $\gamma_m = (\gamma(1), \ldots, \gamma(m))'$ and $\rho_m = (\rho(1), \ldots, \rho(m))'$.
Lemma A.1 Under the assumptions of Theorem 4.1, \( \sqrt{n}(\hat{\psi} - \psi_0, \gamma_m)' \xrightarrow{\mathcal{L}} \mathcal{N}(0, \Sigma_{\hat{\psi}, \gamma_m}) \) when \( p > 0 \), where
\[
\Sigma_{\hat{\psi}, \gamma_m} = \begin{pmatrix}
A_{\psi}^{-1}B_{\psi}A_{\psi}^{-1} & A_{\psi}^{-1}E(\omega_t u_t^2 X_t U_t') \\
E(\omega_t u_t^2 U_t X_t')A_{\psi}^{-1} & E(u_t^2 U_t U_t')
\end{pmatrix}.
\]

Proof. From the proof of Theorem 4.1, we have
\[
\sqrt{n}(\hat{\psi} - \psi_0) = A_{\psi}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \omega_t u_t X_t + o_P(1).
\]
We have
\[
\sqrt{n}\gamma_m = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} u_t U_t.
\]
Applying the CLT of Billingsley (1961) to the square integrable stationary martingale difference \( \{(\omega_t u_t X_t', u_t U_t'), \mathcal{F}_t\} \), Lemma A.1 is proved.

Now, in view of Francq et al. (2004, proof of Theorem 2) we have
\[
\dot{\gamma}_m := (\dot{\gamma}(1), \ldots, \dot{\gamma}(m))' = \gamma_m - \sigma_u^2 \Lambda' (\hat{\psi} - \psi_0) + o_P(1/n).
\]
Hence, by Lemma A.1, the asymptotic distribution of \( \sqrt{n}\gamma_m \) is normal, with mean zero and covariance matrix
\[
\text{Var}_a(s)(\sqrt{n}\gamma_m) = \text{Var}_a(s)(\sqrt{n}\gamma_m) + \sigma_u^4 \Lambda' \text{Var}_a(s)(\sqrt{n}\hat{\psi})\Lambda
\]
\[
- \sigma_u^2 \Lambda' \text{Cov}_a(s)(\sqrt{n}\hat{\psi}, \sqrt{n}\gamma_m) - \sigma_u^2 \text{Cov}_a(s)(\sqrt{n}\gamma_m, \sqrt{n}\hat{\psi})\Lambda.
\]
Finally, we have
\[
\hat{\rho}_m = \dot{\gamma}_m / \sigma_u^2 + O_P(1/n),
\]
from which Lemma 6.1 straightforwardly follows.
A.7 Proof of Theorem 6.2.

To show Theorem 6.2 we establish an intermediate result which is the analog of Lemma A.1. We set
\[ \gamma_{\epsilon^2}(\ell) = \frac{1}{n} \sum_{t=\ell+1}^{n} (\epsilon_t^2 - \sigma_t^2)(\epsilon_{t-\ell}^2 - \sigma_{t-\ell}^2) \quad \text{and} \quad \rho_{\epsilon^2}(\ell) = \frac{\gamma_{\epsilon^2}(\ell)}{\gamma_{\epsilon^2}(0)} \]
for \(0 \leq \ell < n\). Let
\[ \gamma_{m}^{\epsilon^2} = (\gamma_{\epsilon^2}(1), \ldots, \gamma_{\epsilon^2}(m))' \quad \text{and} \quad \rho_{m}^{\epsilon^2} = (\rho_{\epsilon^2}(1), \ldots, \rho_{\epsilon^2}(m))'. \]
Write \( \hat{\theta} = (\hat{\psi}_{WLS}', \hat{\beta}'_{WLS})' \).

Lemma A.2 Under the assumptions of Theorem 4.2, when \(p + q \neq 0\),
\[ \sqrt{n} \left( \frac{\hat{\theta} - \theta_0}{\gamma_{m}^{\epsilon^2}} \right) \xrightarrow{D} \mathcal{N} \left( 0, \Sigma_{\hat{\theta}, \gamma_{m}^{\epsilon^2}} : = \begin{pmatrix} \Sigma_{WLS} & S \\ S' & E(V_{t}V_{t}') \end{pmatrix} \right) \].

Proof. The proof is written for \(pq \neq 0\), but can be straightforwardly modified when \(p = 0\) or \(q = 0\). From the proof of Theorem 4.2, we have
\[ \sqrt{n}(\hat{\theta} - \theta_0) = \begin{pmatrix} A_{\psi}^{-1} & 0 \\ A_{\beta}^{-1}A_{\beta_0}A_{\psi}^{-1} & A_{\beta}^{-1} \end{pmatrix} \left( \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \omega_t u_t X_t \right) + o_P(1). \]

Noting that
\[ \sqrt{n}\gamma_{m}^{\epsilon^2} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} V_t, \]
and applying the CLT of Billingsley (1961) to the square integrable stationary martingale difference \( \left\{ (\omega_t u_t X_t, \tau_t \nu_t \frac{\partial \sigma_t^2(\omega_0, \beta_0)}{\partial \beta}, V_{t}')', \mathcal{F}_t \right\} \), Lemma A.2 is proved.

Now remark that Assumptions A13 and A14 entail the existence \(\Lambda^{\epsilon^2}\). Consider for simplicity the case of an AR(0)-BL(1), then
\[ E\|\epsilon_t \frac{\partial \epsilon_t}{\partial b}(\theta_0)\|^2 = E \left( \frac{u_t^2 u_{t-1}}{(1 + b_0 u_{t-1})^2} \right)^2 \leq \frac{E u_1^6}{e^4} < \infty. \]

In the general case, one can similarly check that \(E\|\epsilon_t \frac{\partial \epsilon_t}{\partial b}\|^2 < \infty\), from which the existence of \(\lambda_{\epsilon}^2 = 2E\epsilon_t \frac{\partial \epsilon_t}{\partial b}(\epsilon_t^2 - \sigma_t^2)(\theta_0)\), and thus of \(\Lambda^{\epsilon^2}\), follow. The existence of \(S\) is a consequence of A9-A12.
Replacing $\epsilon^2$ by $\sigma^2_\epsilon$ in $\hat{\gamma}_{i2}(\ell)$, we define

$$\tilde{\gamma}_{i2}(\ell) = \frac{1}{n} \sum_{t=\ell+1}^n (\epsilon_t^2 - \sigma^2_\epsilon)(\epsilon_{t-\ell}^2 - \sigma^2_\epsilon), \quad \ell = 0, \ldots, n - 1.$$  

We similarly define $\tilde{\rho}_{i2}(\ell)$, $\tilde{\gamma}_m$ and $\tilde{\rho}_m$. It is easy to check that $\tilde{\gamma}_{i2}(\ell) - \hat{\gamma}_{i2}(\ell) = o_p(1)$. Consequently $\sqrt{n}\tilde{\gamma}_m$ and $\sqrt{n}\hat{\gamma}_m$ have the same asymptotic distribution, when existing. The same is true for $\sqrt{n}\tilde{\rho}_m$ and $\sqrt{n}\hat{\rho}_m$.

Note that $\tilde{\gamma}_{i2}(\ell)$ is a function of $\hat{\theta}$ which takes the value $\gamma_{i2}(\ell)$ at the point $\theta_0$.

Assumption A 13 entails that $\tilde{\gamma}_{i2}(\ell)$ is well defined, and even derivable, when $n$ is large enough for $\hat{\theta} \in V(\theta_0)$. Moreover, the ergodic theorem entails that a.s.

$$\frac{\partial \tilde{\gamma}_{i2}(\ell)}{\partial \theta}(\theta_0) = \frac{1}{n} \sum_{t=\ell+1}^n (\epsilon_t^2 - \sigma^2_\epsilon) \frac{\partial \epsilon_t^2}{\partial \theta}(\theta_0) + \frac{2}{n} \sum_{t=\ell+1}^n \epsilon_t \frac{\partial \epsilon_t}{\partial \theta}(\epsilon_{t-\ell}^2 - \sigma^2_\epsilon)(\theta_0) \to \lambda_\ell^2$$

for $\ell > 0$. A Taylor expansion then gives

$$\tilde{\gamma}_m^{\ell^2} := (\tilde{\gamma}_{i2}(1), \ldots, \tilde{\gamma}_{i2}(m))^\prime = \gamma_m^{\ell^2} + \Lambda^{\ell^2}(\hat{\theta} - \theta_0) + O_p(1/n).$$

It follows from Lemma A.2 that $\sqrt{n}\tilde{\gamma}_m^{\ell^2}$ converges in law to a gaussian distribution with mean zero and covariance matrix

$$E(V_tV_t') + \Lambda^{\ell^2} \Sigma WLS \Lambda^{\ell^2} + S' \Lambda^{\ell^2} + \Lambda^{\ell^2} S.$$  

Since

$$\hat{\gamma}_{i2}(0) \rightarrow \text{Var} \epsilon_t^2 = \sigma^4_\epsilon(\mu_4 - 1),$$

and

$$E(V_tV_t') = \sigma^8_\epsilon(\mu_4 - 1)^2 I_m,$$

the first result of Theorem 6.2 follows. In the case $q = 0$, the vector $(\partial \epsilon_t/\partial \theta)(\theta_0)$ belongs to $\mathcal{F}_{t-1}$, which implies $\lambda_\ell^2 = 0$. The simplification of the asymptotic variance when $q = 0$ follows. The last result is obvious.
References


Inconsistency of the QMLE and asymptotic normality of the weighted LSE for a class of conditionally heteroscedastic models: complementary results

A Technical details

A.1 Proof of Remark 4.2

The result stated in the remark can be viewed as a version of the Gauss-Markov theorem. Let

\[ D_t = A_{\psi}^{-1} \omega_t \sigma_t(\theta_0) X_t - A_{\psi}^{-1} \sigma_t(\theta_0) X_t, \]

where \( A_{\psi}^* = E(\sigma_t^{-2}(\theta_0) X_t X_t') \). Note that \( A_{\psi}^{-1} \) is the asymptotic variance of the WLSE based on the weights \( \sigma_t^{-2}(\theta_0) \). The result then follows from

\[ ED_t D_t' = A_{\psi}^{-1} B_{\psi} A_{\psi}^{-1} - A_{\psi}^* -1. \]

A.2 Remark on Assumption (6.8)

We now show that (6.8) is satisfied when the distribution of \( \epsilon_t \) admits a density \( f \) with respect to the Lebesgue measure. To simplify the notation, consider the LARCH(1) case \( u_t = (1 + b_{01} u_{t-1}) \epsilon_t \), the arguments being the same in the general LARCH(\( q \)) case. Assumption (6.8) being always satisfied when \( b_{01} = 0 \), one can assume that \( b_{01} \neq 0 \) and \( \epsilon_t \) has a density. Lemma A.1 below entails that \( P(u_t = c) = 0 \) for all \( c \neq 0 \).

Assumption (6.8) follows since, by stationarity, \( P(1 + b_{01} u_{t-1} = 0) = P(u_t = -1/b_{01}) = 0. \)

Lemma A.1 If \( X \) and \( Y \) are two independent random variables and \( Y \) has a density \( f \) respect to the Lebesgue measure \( \lambda \), then for every \( \varepsilon \) there exists a \( \delta \) such that

\[ P(XY \in A) < \varepsilon \quad \text{if} \quad \lambda(A) < \delta \quad \text{and} \quad 0 \notin A. \]  

(A.1)

Proof. For all Borel set \( A \) which does not contain 0, we define the set \( \frac{1}{x} A = \{ y \in \mathbb{R} : xy \in A \} \). The independence of \( X \) and \( Y \), the dominated convergence theorem and the absolute continuity of \( P_Y \) entail that for every \( \varepsilon \) there exists a \( \delta \) such that

\[ P(XY \in A) = \int P(Y \in \frac{1}{x} A) dP_X(x) < \varepsilon \quad \text{if} \quad \lambda(A) < \delta. \]

\[ \square \]

A.3 Estimating the asymptotic covariance matrices of the estimators and test statistics

The asymptotic variance \( \Sigma \) of the QMLE, given in Theorem 2.1, could be easily estimated under the assumptions of this theorem, but, in view of the inconsistency result given in Proposition 3.1, this is not of interest. So we focus on the WLSE.

The matrices \( A_{\psi} \) and \( B_{\psi} \) defined in Theorem 4.1 can be consistently estimated by the matrices \( \hat{A}_{\psi} \) and \( \hat{B}_{\psi} \) defined by (6.5). On order to define estimates for the other matrices involved in \( \Sigma_{WLS} \), we set

\[ \hat{Y}_t = \begin{pmatrix} 2 \hat{\sigma}_t^2 \left( 1 + \sum_{i=1}^q \hat{b}_i \hat{u}_{t-i} \right) \hat{u}_{t-1} \\ \vdots \\ 2 \hat{\sigma}_t^2 \left( 1 + \sum_{i=1}^q \hat{b}_i \hat{u}_{t-i} \right) \hat{u}_{t-q} \\ \left( 1 + \sum_{i=1}^q \hat{b}_i \hat{u}_{t-i} \right)^2 \end{pmatrix} \]
for $t = p + q + 1, \ldots, n$. In view of Theorem 4.2, at first sight it could seen natural to estimate the matrices $A_\beta$ and $B_\beta$ by

$$\hat{A}_\beta = \frac{1}{n} \sum_{t=r \wedge (p+q)+1}^{n+1} \tau_t \hat{Y}_t \hat{Y}_t', \quad \hat{B}_\beta = \frac{\hat{\mu}_4 - 1}{n} \sum_{t=r \wedge (p+q)+1}^{n+1} \tau^2_t \sigma^4_t (\hat{\theta}) \hat{Y}_t \hat{Y}_t',$$

where $\hat{\mu}_4 = n^{-1} \sum_{t=p+q+1}^{n} \hat{e}_t^4 / \hat{\sigma}^4_t$ and $\hat{e}_t$ is defined as in (6.7) by $\hat{e}_t = \hat{u}_t / (1 + \sum_{i=1}^{q} \hat{b}_i \hat{u}_{t-i})$. The estimator of $\hat{A}_\beta$ is indeed reasonable, but the estimators $\hat{B}_\beta$ and $\hat{\mu}_4$ have a bad behavior when $1 + \sum_{i=1}^{q} \hat{b}_i \hat{u}_{t-i}$ is close to zero for some $t$, which likely to occur when Assumption A13 cannot be made. It is therefore safer to use the estimator

$$\tilde{B}_\beta = \frac{1}{n} \sum_{t=r \wedge (p+q)+1}^{n+1} \tau^2_t \tilde{\nu}_t^2 (\hat{\theta}) \hat{Y}_t \hat{Y}_t', \quad \tilde{\nu}_t = \hat{u}_t^2 - \sigma^2_t (\hat{\theta}).$$

We can then define an estimator of $\mu_4$ by

$$\hat{\mu}_4 = 1 + \frac{\| \tilde{B}_\beta \|}{n^{-1} \sum_{t=r \wedge (p+q)+1}^{n+1} \tau^2_t \sigma^4_t (\hat{\theta}) \hat{Y}_t \hat{Y}_t'}. $$

Under the symmetry assumption $E \epsilon^3_t = 0$, which has been made for the numerical illustrations of the present paper, it can be seen that $B_{3\psi} = 0$ and $B_{3\psi} = 0$. Then the asymptotic variance of the WLS can be estimated by

$$\hat{\Sigma}_{WLS} = \begin{pmatrix} \hat{\Sigma}_{WLS} & 0 \\ 0 & \hat{\Sigma}_{WLS} \end{pmatrix}, \quad \hat{\Sigma}_{WLS} = \hat{A}_\psi^{-1} \hat{B}_\psi \hat{A}_\psi^{-1}, \quad \hat{\Sigma}_{WLS} = \hat{A}_\psi^{-1} \hat{B}_\psi \hat{A}_\psi^{-1}. $$

When the assumption $E \epsilon^3_t = 0$ is relaxed, the matrices $B_{3\psi}$ and $B_{3\psi}$ must be estimated. This can be done in an obvious way, setting

$$\hat{Z}_t = -2 \hat{u}_t X_t + 2 \hat{\sigma}^2_t \left( 1 + \sum_{i=1}^{q} \hat{b}_i \hat{u}_{t-i} \right) \sum_{i=1}^{q} \hat{b}_i X_{t-i}$$

for $t = p + q + 1, \ldots, n$. In view of (6.4), it is then natural to propose an estimator of $\Sigma_{\mu_m}$ of the form

$$\hat{\Sigma}_{\mu_m} = \frac{1}{n \hat{\sigma}_u^2} \sum_{t=m+p+1}^{n} Y_t Y_t',$$

where $n \hat{\sigma}_u^2 = n^{-1} \sum_{t=p+1}^{n} \hat{u}_t^2$. There is however no guarantee that this estimator be positive definite. To avoid the problem, it is preferable to define an estimator of $\Sigma_{\mu_m}$ by

$$\hat{\Sigma}_{\mu_m} = \frac{1}{n \hat{\sigma}_u^2} \sum_{t=m+p+1}^{n} Y_t Y_t,$$

where

$$Y_t = \hat{u}_t \hat{U}_t - \hat{\sigma}_u^2 \hat{u}_t \hat{U}_t \hat{A}_\psi^{-1} X_t.$$
and, under A13 and A14,

\[
\lambda_t^2 := 2 \mathbb{E}_t \frac{\partial \ell_t}{\partial \theta} (\ell_{t-\ell} - \sigma^2_{\theta})(\theta_0)
\]

\[
= 2\sigma^2_0 \mathbb{E}_t \frac{1}{\left(1 + \sum_{i=1}^{q} b_{0i} u_{t-i}\right)} \left( \sum_{i=1}^{q} b_{0i} X_t - \frac{-u_{t-i}}{0} \right) (\ell_{t-\ell} - \sigma^2_{\theta}).
\]

Thus, one can propose the estimator

\[
\hat{\Lambda}^2 = \left(\hat{\lambda}_1^2, \ldots, \hat{\lambda}_m^2\right) \text{'}, \quad \hat{\lambda}_t^2 = \frac{2\sigma^2_0}{n} \sum_{t=\ell+p+q+\ell+1}^{n} \frac{\hat{\epsilon}_t}{u_t} \left( \sum_{i=1}^{q} \hat{b}_i X_{t-i} \right) (\hat{\epsilon}_{t-\ell} - \hat{\sigma}^2_{0e}) I_{\{u_t \neq 0\}}.
\]

B Implementation in R

The programs given in this section are written in the R language (see http://cran.r-project.org/).

B.1 Auxiliary routines

The function `estimARCHq.qml(omega0, alpha0, u)` computes the QMLE of an ARCH(q) model for the series \( u \), with initial values \( \omega_0 \) and \( \alpha_0 \).

```r
estimARCHq.qml <- function(omega0, alpha0, u) {
  q <- length(alpha0); valinit <- c(omega0, alpha0)
  res <- nlminb(valinit, objf.arch.qml, lower = c(0.05, rep(0.00, q)), upper = c(rep(Inf, q+1)), u = u)
  res$par
}

objf.arch.qml <- function(x, u) {
  q <- length(x) - 1; omega <- x[1]; alpha <- x[2:(q+1)]; n <- length(u);
  sigma2 <- as.numeric(n)
  for (t in (q+1):n) sigma2[t] <- omega + sum(alpha[1:q] * (u[(t-1):(t-q)]^2))
  qml <- mean(log(sigma2[(q+1):n]) + u[(q+1):n]^2 / sigma2[(q+1):n])
  qml
}
```

The function `phi.star.ARinv(psi, lagmax=m)` computes the coefficients \( \psi^*_1, \ldots, \psi^*_m \) defined by

\[
\left(1 - \sum_{i=1}^{p} \psi_i z^i\right)^{-1} = 1 + \sum_{i=1}^{\infty} \psi^*_i z^i, \quad |z| \leq 1,
\]

when the zeros of the polynomial \( 1 - \sum_{i=1}^{p} \psi_i z^i \) are outside the unit circle.

```r
phi.star.ARinv <- function(psi, lagmax=50) {
  p <- length(psi); psi.star <- rep(0, lagmax); psi.star[1] <- psi[1]
  if(p>1) for(h in 2:p) psi.star[h] <- psi[h] + sum(psi[(h-1):1] * psi.star[1:(h-1)])
  if(lagmax>p) for(h in (p+1):lagmax) psi.star[h] <- sum(psi[p:1] * psi.star[(h-p):(h-1)])
  psi.star
}
```

B.2 Weighted Least Squares Estimator (WLSE)

B.2.1 WLSE of the AR parameter

First consider the implementation of the AR WLSE defined in Theorem 4.1. The observation \( x_1, \ldots, x_n \) are stored in the vector \( x \). The function `WLSE.g.AR(psi.init, p, q, x)` computes the WLSE of the autoregressive parameter \( (\psi_1, \ldots, \psi_p) \). The weights are the inverse of the volatility of an ARCH\((p+q)\) model fitted to linear innovations, obtained from an initial value \( \text{psi.init} \) of the AR parameter.

```r
WLSE.g.AR <- function(psi.init, p, q, x) {
  n <- length(x); if(p>0) X <- matrix(nrow=(n-p), ncol=p); u <- rep(as.numeric(NA), n)
  if(p>0) for(t in 1:p) u[t] <- x[t] - sum(psi.init[1:(t-1)] * x[1:(t-1)])
  # WLSE for the AR part, with weights 1/sig^2 where sig^2 is the volatility of an ARCH(r) model
  # WLSE.g.AR <- function(psi.init, p, q, x) {
  #   n <- length(x); if(p>0) X <- matrix(nrow=(n-p), ncol=p); u <- rep(as.numeric(NA), n)
  #   if(p>0) for(t in 1:p) u[t] <- x[t] - sum(psi.init[1:(t-1)] * x[1:(t-1)])
  #   ...
  # }
}
```
for (t in (p+1):n) u[t]<-x[t]-sum(psi.init[1:p]*x[(t-1):(t-p)]) # u contains the linear innovations
omega.init<-var(u[(p+1):n]) # initial value for the ARCH intercept

r<-p+q # order of the ARCH
if(r>0 & 0.1*r<1) omega.init<-var(u[(p+1):n])*(1-0.1*r) # initial values for the ARCH coefficients
omega<-arch.estim[1] # estimated values of the intercept
omegainv<-rep(omega,n) # omegainv contains the inverse of the weights
if(r>0) for (t in (r+1):n) omegainv[t]<-omegainv[t]+sum(alpha[1:r]*u[(t-1):(t-r)]^2)

psi<-c() else for (j in 1:p) X[1:(n-p),j]<-x[(n-j):(p+1-j)]/sqrt(omegainv[n:(p+1)])
 psi<-solve(t(X)%*%X,t(X)%*%(x[n:(p+1)])/(n*omegainv[n:(p+1)])) # psi is the WLSE
res<-x # linear innovations induced from the WLSE
if(p>0) {for (t in 1:p) res[t]<-x[t]-sum(psi[1:(t-1)]*x[(t-1):1])}
for (t in (p+1):n) res[t]<-x[t]-sum(psi[1:p]*x[(t-1):(t-p)])

A.psi<-matrix(0,nrow=p,ncol=p)
B.psi<-matrix(0,nrow=p,ncol=p)
for (t in (n:(p+1))) {A.psi<-A.psi+ c(x[(t-1):(t-p)])%*%t(c(x[(t-1):(t-p)])/(n*omegainv[t:n]))}
S.psi<-solve(A.psi)%*%B.psi%*%solve(A.psi) # estimes of the WLSE asymptotic variance

list(psi=psi,res=res,A.psi=A.psi,B.psi=B.psi,Sigma.psi=S.psi,omegainv=omegainv)

B.2.2 WLSE of the AR-LARCH model
The function WLSE.g.ARLARCH() computes the WLSE of all the coefficients of the AR(p) LARCH(q) model. An estimates of the information matrices \(A_\beta, B_\beta, A_\psi, B_\psi\) and of the asymptotic variances \(\Sigma_{WLS,\beta}\) and \(\Sigma_{WLS,\psi}\) are also computed. Estimated standard deviations for the estimates are deduced.

WLS.e.g.AR.LARCH<- function(phi.init,b.init,sig.init,x,quantil=0.01){
  p<-length(phi.init); q<-length(b.init); n<-length(x)
  result<-WLSE.g.AR(phi.init,p,q,x) # estimation of the AR part
  u<-result$res # linear innovations induced from the WLSE
  sig.psi<-sqrt(diag(result$Sigma.psi)/n) # estimated sdv for the WLSE of the AR coefs
  valinit<-c(b.init,sig.init)
  res <- nlminb(valinit,objf.wlse, lower=c(rep(-Inf,q),0),
               upper=c(rep(Inf,q),Inf),x=x,u=u) # estimation of the LARCH coefficients
  beta<-res$par; b <- beta[1:q]; sig2 <- beta[1:(q+1)]
  h<-rep(sig2,n); s<-rep(1,n)
  if(q>0) for (t in (q+p+1):n) s[t]<-1+sum(b[1:q]*u[(t-1):(t-q)]); h[t]<-sig2*(s[t])^2
  epsilon<-u; tol<-as.numeric(quantile(abs(s),quantil))
  if(q>0) for (t in (q+p+1):n) {if(abs(s[t])>=tol) epsilon[t]<-u[t]/s[t]}
  Y.hat<-matrix(0,nrow=(q+1),ncol=n)
  for (t in (q+p+1):n) {Y.hat[1:q,t]<-2*sig2*s[t]*u[(t-1):(t-q)]; Y.hat[(q+1),t]<-s[t]**2}
  A.beta<-matrix(0,nrow=(q+1),ncol=(q+1)); B.beta<-matrix(0,nrow=q+1,ncol=q+1)
  for (t in (p+q+1):n){A.beta<-A.beta+tau[t]*Y.hat[1:q,t]*t(Y.hat[1:q,t])}
  if(q>0) for (t in (q+p+1):n) {A.beta<-A.beta+tau[t]*Y.hat[1:q,t]}
  for (t in (q+p+1):n) {B.beta<-B.beta+tau[t]*Y.hat[1:q,t]}
  B.beta<-B.beta*var(epsilon^2)/(sig2**2); mu4<- var(epsilon^2)/(sig2**2)+1
  Sigma.beta<-solve(A.beta)%*%B.beta%*%solve(A.beta)

  objf.wlse <- function(para,p,q,tau,epsilon){
    if(q>0) b <- para[1:q]
    sig2 <- para[q+1]; n <- length(u); h<-rep(sig2,n)
    if(q>0) for (t in (q+p+1):n) h[t]<-sig2*(1+sum(b[1:q]*u[(t-1):(t-q)]))
    obj <- mean(tau*(p+q+1):n)^2+mu4*(h[(p+q+1):n])**2
  }
  list(psi=psi,res=res,A.beta=A.beta,B.beta=B.beta, Sigma.beta=Sigma.beta, Sigma.beta.sig=Sigma.beta)
}
B.3 Testing for LARCH effect

The following function performs a score test of the null hypothesis of conditional homoscedasticity against that of a LARCH\((q)\) model with coefficient \(b \neq 0\).

```r
########################################################################
# score test for conditional homoscedasticity (against LARCH effect) #
# score.LARCH<- function(u,tau,q) {
# n<-length(u); sig2c<-sum(tau*u^2)/sum(tau)
# U<-matrix(0,nrow=n,ncol=q); V<-rep(0,n)
# for(t in (q+1):n){U[t,1:q]<-tau[t]*u[(t-1):(t-q)]; V[t]<-u[t]^2-sig2c}
# score.stat<-n*as.numeric(t(V)%*%U%*%solve(t(U)%*%U)%*%t(U)%*%V/t(V)%*%V)
# pval<-1-pchisq(score.stat,df=q); list(stat=score.stat,pval=pval) }

B.4 Checking the adequacy of the AR\((p)\)-LARCH\((q)\) model

B.4.1 Checking the adequacy of the AR equation

The following function performs the portmanteau test defined in Remark 6.3. It uses the function \(\phi.star\) defined in section B.1, and also the function \(\rho(u,h)\) which computes the autocorrelation of the vector \(u\) at lag \(h\).

```r
########################################################################
# portmanteau test for the adequacy of the AR equation #
# Portmanteau.AR<- function(m,x,psi,u,A.psi,B.psi,Sigma.psi,omega){
# n<-length(u); p<-length(psi); r<-max(p,m); Upsilon<-matrix(0,nrow=m,ncol=n)
# s2u<-mean(u^2); r<-max(p+q,p+m)+1; if(p>0) Ainv<-solve(A.psi); Sigma.rho<-matrix(0,nrow=m,ncol=m)
# if(p>0) Lambda<-matrix(0,nrow=p,ncol=m); if(p>0 & m>0)phi.star<-phi.star.ARinv(psi,lagmax=(m-1))
# if(p>0) for(i in 1:p){if(m>i) Lambda[i,i]<-1; if(m==i) Lambda[i,(i+1):m]<-phi.star[1:(m-i)]}
# if(p>0) for (t in r:n) Upsilon[,t]<-u[t]*u[(t-1):(t-m)]-s2u*omega[t]*u[t]*t(Lambda)%*%Ainv%*%x[(t-1):(t-p)]
# if(p<=0) for (t in r:n) Upsilon[,t]<-u[t]*u[(t-1):(t-m)]
# for (t in r:n) Sigma.rho<-Sigma.rho+Upsilon[,t]%*%t(Upsilon[,t]) /n
# Sigma.rho<-Sigma.rho/s2u^2; rho.m <- rep(as.numeric(NA),m)
# for(h in 1:m) rho.m[h]<-rho(u,h); Q.m<-n*(rho.m)%*%solve(Sigma.rho)%*%rho.m
# pval<-1-pchisq(Q.m,df=m); list(stat=Q.m,pval=pval)

B.4.2 Checking the adequacy of the LARCH equation

The following function performs the portmanteau test defined in Remark 6.6.

```r
########################################################################
# portmanteau test for the adequacy of the LARCH equation #
# Portmanteau.LARCH<- function(m,x,u,psi,b,sig2,mu4,Sigma.psi,Sigma.beta,epsilon,tol){
# n<-length(u); p<-length(psi); q<-length(b); Sigma.rho2<-diag(rep(1,m))
# Sigma.WLS<-matrix(0,nrow=(p+q+1),ncol=(p+q+1))
# if(p>0)Sigma.WLS<-cbind(Sigma.psi,matrix(0,nrow=p,ncol=(q+1))),
# if(q>0) {Lambda2<-matrix(0,nrow=m,ncol=(p+q+1))
# X<-matrix(0,nrow=(p+q+1),ncol=n); Y<-matrix(0,nrow=n,ncol=m)
# h<-rep(as.numeric(NA),n); s<-rep(0,n)
# for (t in (q+p+1):n) {s[t]<-1+sum(b[1:q]*u[(t-1):(t-q)]); h[t]<-sig2*s[t]^2}
# if(abs(s[t])>tol) {if(p>0) for (i in 1:q) X[1:p,t]<-X[1:p,t]+b[i]*x[(t-i-1):(t-i-p)]
# X[(p+1):(p+q),t]<- -u[(t-1):(t-m)]; X[t]<-sig2*X[t]/s[t] } }
# for (t in (m+1):n) {Y[t]<-epsilon[(t-1):(t-m)]^2- sig2; Lambda2<-Lambda2+2*Y[t]^%*%X[t]/n }
# if(q>0) Sigma.rho2<-Lambda2+2*Y[t]^%*%X[t]/n
# rho2.m <- rep(as.numeric(NA),m)
# for(h in 1:m) rho2.m[h]<-rho(epsilon^2,h); Q.m<-n*(rho2.m)%*%solve(Sigma.rho2)%*%rho2.m
# pval<-1-pchisq(Q.m,df=m); list(stat=Q.m,pval=pval)
```
Table 6: WLSE for a simulation of length \( n = 5,000 \) of the AR(1)-LARCH(5) model (B.1).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \psi_1 )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
<th>( b_3 )</th>
<th>( b_4 )</th>
<th>( b_5 )</th>
<th>( \sigma^2_e )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.25</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.25</td>
<td>1</td>
</tr>
<tr>
<td>Estimate</td>
<td>0.255</td>
<td>0.009</td>
<td>-0.004</td>
<td>0.002</td>
<td>0.011</td>
<td>0.245</td>
<td>1.008</td>
</tr>
<tr>
<td>(Standard Deviation)</td>
<td>(0.014)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.016)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>( t )-ratio</td>
<td>18.78</td>
<td>0.71</td>
<td>-0.34</td>
<td>0.19</td>
<td>0.87</td>
<td>15.67</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: As Table 6, but the model is an AR(2)-LARCH(6), whereas the DGP is still the AR(1)-LARCH(5) model (B.1).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \psi_1 )</th>
<th>( \psi_2 )</th>
<th>( b_1 )</th>
<th>( b_2 )</th>
<th>( b_3 )</th>
<th>( b_4 )</th>
<th>( b_5 )</th>
<th>( b_6 )</th>
<th>( \sigma^2_e )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.25</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.25</td>
<td>0.005</td>
<td>1.005</td>
</tr>
<tr>
<td>Estimate</td>
<td>0.253</td>
<td>0.030</td>
<td>-0.015</td>
<td>-0.009</td>
<td>0.010</td>
<td>-0.008</td>
<td>0.247</td>
<td>-0.013</td>
<td>1.005</td>
</tr>
<tr>
<td>(Std)</td>
<td>(0.014)</td>
<td>(0.014)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.012)</td>
<td>(0.016)</td>
<td>(0.012)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>( t )-ratio</td>
<td>18.06</td>
<td>2.19</td>
<td>-1.26</td>
<td>-0.71</td>
<td>0.85</td>
<td>-0.67</td>
<td>15.55</td>
<td>-1.04</td>
<td></td>
</tr>
</tbody>
</table>

B.5 Application of the R programs

We simulated a trajectory of length \( n = 5,000 \) of the AR(1)-LARCH(5) model

\[
\begin{align*}
\{ x_t = 0.25x_{t-1} + u_t, \\
u_t = (1 + 0.25u_{t-5})\epsilon_t, \\
\epsilon_t \text{ iid } \mathcal{N}(0, 1).
\end{align*}
\]

(B.1)

Table 6 displays the WLSE, when the fitted model corresponds to the DGP. Table 7 displays the WLSE when the estimated model is an AR(2)-LARCH(6). From this table, one can see that, as expected, the null hypothesis \( H_0 : b_6 = 0 \) can not be rejected. The conclusion is less clear concerning the null hypothesis \( H_0 : \psi_2 = 0 \). Table 8 presents portmanteau tests based on a quadratic form of the first \( m \) autocorrelations of the linear residuals \( \hat{u}_t \), for an AR(1) model and for an AR(0) model (i.e. \( \hat{u}_t = u_t = x_t \)). As expected the AR(0) is rejected and the AR(1) is not rejected. Table 9 clearly shows that the hypothesis of conditional homoscedasticity must be rejected against that of a LARCH(\( q \)) for \( q \geq 5 \). The portmanteau tests of Table 10 indicate significant autocorrelations for the squares of AR(1)-LARCH(\( q \)) residuals when \( q < 5 \). As expected the portmanteau tests based on the AR(1)-LARCH(5) residuals do not reject the adequacy of the model, which is indeed the data generating model.

Table 8: Portmanteau test of adequacy of the AR(1) model and of the AR(0) model (i.e. no linear part) for the AR part of the AR-LARCH model.

<table>
<thead>
<tr>
<th>( m )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{Q}^{\hat{u}}_m )</td>
<td>&gt; 280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p )-value</td>
<td>( \simeq 0 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Portmanteau test for the AR(1) model
| \( \bar{Q}^{\hat{u}}_m \) | 1.03 | 1.18 | 4.7 | 6.87 | 7.01 | 7.36 | 7.47 | 7.94 | 7.95 | 9.51 | 9.52 | 9.52 |
| \( p \)-value | 0.31 | 0.55 | 0.2 | 0.14 | 0.22 | 0.29 | 0.38 | 0.44 | 0.54 | 0.48 | 0.57 | 0.66 |
Table 9: Test of conditional homoscedasticity against a LARCH\textit{(q)} model for the linear innovations.

<table>
<thead>
<tr>
<th>$m$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_n$</td>
<td>0.16</td>
<td>0.68</td>
<td>0.71</td>
<td>1.65</td>
<td>439.99</td>
<td>440.43</td>
<td>440.27</td>
<td>440.23</td>
<td>440.67</td>
<td>445.27</td>
</tr>
<tr>
<td>$p$-value</td>
<td>0.693</td>
<td>0.711</td>
<td>0.872</td>
<td>0.799</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 10: Portmanteau test of adequacy of different AR(1)-LARCH\textit{(q)} models.

<table>
<thead>
<tr>
<th>$m$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
</table>
| Portmanteau test for the LARCH(0) model \textit{(i.e. absence of LARCH part)} | \begin{tabular}{l|l|l|l|l|l|l|l|l|l|l|l} 
 $\tilde{Q}_m^2$ & 0 & 0.36 & 0.56 & 2.97 & 16.35 & 19.09 & 19.22 & 19.53 & 21.01 & 21.94 & 25.64 & 26.92 \ 
 $p$-value & 0.975 & 0.836 & 0.906 & 0.564 & 0.006 & 0.004 & 0.008 & 0.012 & 0.013 & 0.015 & 0.007 & 0.008 \\
\end{tabular} | \begin{tabular}{l|l|l|l|l|l|l|l|l|l|l|l} 
 | 0 & 0.35 & 0.54 & 2.61 & 16.05 & 18.97 & 19.11 & 19.46 & 20.92 & 21.84 & 25.57 & 26.88 \ 
 $p$-value & 0.995 & 0.84 & 0.91 & 0.625 & 0.007 & 0.004 & 0.008 & 0.013 & 0.013 & 0.016 & 0.008 & 0.008 \\
\end{tabular} | \begin{tabular}{l|l|l|l|l|l|l|l|l|l|l|l} 
 | 0 & 0.01 & 0.22 & 0.27 & 2.16 & 15.24 & 18.19 & 18.43 & 18.74 & 19.89 & 20.75 & 24.65 & 25.97 \ 
 $p$-value & 0.928 & 0.894 & 0.966 & 0.706 & 0.009 & 0.006 & 0.01 & 0.016 & 0.019 & 0.023 & 0.01 & 0.011 \\
\end{tabular} | \begin{tabular}{l|l|l|l|l|l|l|l|l|l|l|l} 
 | 0 & 0.01 & 0.26 & 0.31 & 2.3 & 15.4 & 18.31 & 18.52 & 18.83 & 19.99 & 20.86 & 24.73 & 26.05 \ 
 $p$-value & 0.929 & 0.878 & 0.958 & 0.681 & 0.009 & 0.006 & 0.01 & 0.016 & 0.018 & 0.022 & 0.01 & 0.011 \\
\end{tabular} | \begin{tabular}{l|l|l|l|l|l|l|l|l|l|l|l} 
 | 0 & 0.01 & 0.38 & 0.44 & 2.48 & 15.61 & 18.6 & 18.8 & 19.1 & 20.3 & 21.18 & 24.81 & 26.08 \ 
 $p$-value & 0.909 & 0.829 & 0.933 & 0.648 & 0.008 & 0.005 & 0.009 & 0.014 & 0.016 & 0.02 & 0.01 & 0.01 \ 
\end{tabular} | \begin{tabular}{l|l|l|l|l|l|l|l|l|l|l|l} 
 | 0 & 0.03 & 0.72 & 1.96 & 3.8 & 4.31 & 7.78 & 7.95 & 8.02 & 8.32 & 9.03 & 10.78 & 12.61 \ 
 $p$-value & 0.869 & 0.698 & 0.582 & 0.434 & 0.506 & 0.255 & 0.337 & 0.432 & 0.502 & 0.529 & 0.462 & 0.398 \\
\end{tabular} |
C Complementary simulation experiments

We first consider an extension of the AR(1)-LARCH(1) model considered in Section 7.1:

\[ x_t = \psi_{01} x_{t-1} + u_t, \quad u_t = (1 + b_{01} u_{t-1}) \epsilon_t, \quad \epsilon_t \text{iid } \mathcal{N}(0, \sigma_0^2), \quad \sigma_0 > 0 \]  

(C.1)

with uniform and gaussian distributions for the error term. A simulation of length 100 of \(^{(u_t)}\), with gaussian innovations, is displayed in Figure 3. Volatility clustering can be noticed, as well as the absence of significant correlations and, on the contrary, the presence of significant autocorrelations for the squares.

![Figure 3: Simulation of Model (7.1) (or Model (C.1)) with \( \psi_{01} = 0, b_{01} = 0.9 \) and \( \epsilon_t \sim \mathcal{N}(0,1) \) (left panel), empirical autocorrelations (middle panel) and empirical autocorrelations of the squares (right panel).](image)

We first study the properties of the QMLE. The true parameter is taken to be \( \phi_{01} = 0.9, b_{01} = -0.5 \) and \( \epsilon_t \sim U(-0.5,0.5) \) (thus \( \sigma_0^2 = 1/12 \)). The number of replications is 500. Figure 4 displays boxplots and Q-Q plots of the estimation errors when the sample size is \( n = 100 \). It is seen that the biases are very small but that values very far from the true value can be obtained, particularly for the ARCH coefficient \( b_{01} \). Moreover, the Q-Q plots indicate important departures from the normality, especially for the AR coefficient \( \psi_{01} \). As \( n \) increases from 100 to 1000, the imprecision on the coefficient \( b_{01} \) becomes smaller and no significant departure from the asymptotic normality is noticed, see Figure 5. Similar results are obtained for the WLSE and the results displayed in Figure 6 indicate that the choice of the weights does not have dramatic effects on the bias and accuracy of the estimators.

From these experiments, it could seem that QML is a reasonable estimation procedure for this model. This is in fact the case because the error distribution has a sufficiently small compact support. Even if the parameter space is not specified in the numerical procedure, values which would entail cancelation of the volatility are not considered by the algorithm. We now investigate the properties of the QMLE and WLSE when the errors distribution is gaussian. Figure 7 shows that the performance of the QMLE is very bad in this case. Both bias and accuracy are disastrous. This confirms our discussion in Section 3. On the contrary, the behavior of the WLSE remains satisfactory whatever the choice of the weights. It is seen that constant weights are not the most appropriate in this case.

C.1 Comparison of different version of the WLSE

Table 11 is the same than Table 2, but for simulations is \( n = 1000 \) instead of \( n = 100 \). The conclusion is similar: the WLSE base on an ARCH proxy of the volatility seems to be superior to the 3 other versions of the WLSE, specially in terms of RMSE. The difference seems even more important for \( n = 1,000 \) than for \( n = 100 \).

C.2 Empirical distribution of the portmanteau tests

The asymptotic validity of the portmanteau tests of adequacy of the AR equation in shown in Theorem 6.1 under very mild assumptions. To prove the asymptotic validity of the portmanteau tests of adequacy of the LARCH equation, Theorem 6.2 required the two restrictive assumptions \( A_{13} \) and \( A_{14} \). The simulation experiments of this section aim to see whether these assumptions are indeed necessary or not in practice. We simulated \( N = 500 \) independent trajectories of length \( n = 100 \) of the AR(1)-LARCH(1) model considered in Section 7.1:

\[
\begin{cases}
  x_t = 0.9 x_{t-1} + u_t, \\
  u_t = (1 + b u_{t-1}) \epsilon_t, \\
  \epsilon_t \text{iid } \mathcal{N}(0,1),
\end{cases}
\]  

(C.2)
Figure 4: Boxplots (left-top panel) and Q-Q plots (other panels) of the 500 QML estimation errors. The data are generated from Model (C.1) with $\psi_0 = 0.9, b_{01} = -0.5$ and $\epsilon_t \sim \mathcal{U}(-0.5, 0.5)$. The sample size is $n = 100$.

For different values of $b$. For each of the $N$ trajectories, we computed the test statistics $\tilde{Q}_{m}^{\psi}$ and $\tilde{Q}_{m}^{\sigma^2}$ for $m = 6$. The asymptotic distribution of these two statistics is the $\chi^2_6$. Figures 8 and 9 show that the empirical distributions of the portmanteau test statistics are close to the $\chi^2_6$, except for the $\tilde{Q}_{m}^{\sigma^2}$ statistics when $b$ is large (i.e. $b = 0.6$ and $b = 0.9$).
(a) Parameter estimations

(b) Normal Q–Q Plot of $\hat{\psi}$

(c) Normal Q–Q Plot of $\hat{b}$

(b) Normal Q–Q Plot of $\hat{\sigma^2}$

Figure 5: Same as Figure 4 but for $n = 1,000$.

Figure 6: Boxplots of 500 estimation errors, for the QMLE (left-top panel) and the WLSE with different choices of the weights (other panels). The data are generated as in Figure 4.
Table 11: As Table 2, but for simulations is \( n = 1000 \) (Comparison of four different versions of the WLS estimator. The DGP is an AR(1)-LARCH(1) process with a gaussian iid noise \( \epsilon_t \). The number of replications is \( N = 500 \) and the length of the simulations is \( n = 1000 \)).

<table>
<thead>
<tr>
<th></th>
<th>LSE</th>
<th>WLS</th>
<th>WLS^{\text{HR}}</th>
<th>WLS^{\text{L}}</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi_{01} = 0.9 )</td>
<td>-0.004</td>
<td>0.020</td>
<td>0.001</td>
<td>0.014</td>
</tr>
<tr>
<td>( \beta_{03} = -0.62 )</td>
<td>0.016</td>
<td>0.822</td>
<td>-0.007</td>
<td>0.064</td>
</tr>
<tr>
<td>( \sigma^2_{\epsilon_{03}} = 1 )</td>
<td>0.106</td>
<td>0.271</td>
<td>-0.004</td>
<td>0.067</td>
</tr>
<tr>
<td>( \psi_{01} = 0.9 )</td>
<td>-0.004</td>
<td>0.021</td>
<td>0.003</td>
<td>0.014</td>
</tr>
<tr>
<td>( \beta_{03} = -0.63 )</td>
<td>-0.044</td>
<td>1.522</td>
<td>-0.007</td>
<td>0.064</td>
</tr>
<tr>
<td>( \sigma^2_{\epsilon_{03}} = 1 )</td>
<td>0.115</td>
<td>0.300</td>
<td>-0.004</td>
<td>0.067</td>
</tr>
<tr>
<td>( \psi_{01} = 0.9 )</td>
<td>-0.007</td>
<td>0.027</td>
<td>0.001</td>
<td>0.013</td>
</tr>
<tr>
<td>( \beta_{03} = -0.75 )</td>
<td>0.440</td>
<td>4.999</td>
<td>-0.008</td>
<td>0.068</td>
</tr>
<tr>
<td>( \sigma^2_{\epsilon_{03}} = 1 )</td>
<td>0.310</td>
<td>1.319</td>
<td>-0.006</td>
<td>0.072</td>
</tr>
<tr>
<td>( \psi_{01} = 0.9 )</td>
<td>-0.019</td>
<td>0.064</td>
<td>0.001</td>
<td>0.012</td>
</tr>
<tr>
<td>( \beta_{03} = -0.99 )</td>
<td>1.997</td>
<td>8.935</td>
<td>-0.008</td>
<td>0.075</td>
</tr>
<tr>
<td>( \sigma^2_{\epsilon_{03}} = 1 )</td>
<td>7.656</td>
<td>34.533</td>
<td>-0.007</td>
<td>0.085</td>
</tr>
<tr>
<td>( \psi_{01} = 0.9 )</td>
<td>-0.033</td>
<td>0.095</td>
<td>0.001</td>
<td>0.015</td>
</tr>
<tr>
<td>( \beta_{03} = -1.1 )</td>
<td>1.323</td>
<td>6.825</td>
<td>-0.011</td>
<td>0.081</td>
</tr>
<tr>
<td>( \sigma^2_{\epsilon_{03}} = 1 )</td>
<td>122.564</td>
<td>1221.361</td>
<td>-0.011</td>
<td>0.093</td>
</tr>
</tbody>
</table>

Figure 7: Same as in Figure 6 but with \( \epsilon_t \sim \mathcal{N}(0, 1) \).
Figure 8: Distributions of the portmanteau test statistics $\hat{Q}_6^u$ for the adequacy of the AR(1) part (left panels) and of $\hat{Q}_6^{\epsilon^2}$ for the adequacy of the LARCH(1) part (right panels). The kernel density estimators are based on $N = 500$ replications of size $n = 100$ of the AR(1)-LARCH(1) model (C.2) for $b = 0$, $b = 0.2$ or $b = 0.4$. The density of the $\chi^2_6$ is plotted in dotted line.
Figure 9: As Figure 8, but for $b = 0.3$, $b = 0.6$ and $b = 0.9$. 
Figure 10: Trajectory $t \rightarrow 1 + \hat{b} \hat{u}_{t-1}$ of the rescaled volatility. The $\tilde{Q}_m^2$-based portmanteau tests are likely to perform badly when this trajectory frequently approaches the horizontal line.
Table 12: Portmanteau test of adequacy of the AR(1) model for the linear dynamics of nine stock market returns.

<table>
<thead>
<tr>
<th>m</th>
<th>CAC $\tilde{Q}_m^u$</th>
<th>p-value</th>
<th>DAX $Q_m^u$</th>
<th>p-value</th>
<th>Shanghai $\tilde{Q}_m^u$</th>
<th>p-value</th>
<th>DJA $Q_m^u$</th>
<th>p-value</th>
<th>DJT $\tilde{Q}_m^u$</th>
<th>p-value</th>
<th>FTSE $\tilde{Q}_m^u$</th>
<th>p-value</th>
<th>Nasdaq $Q_m^u$</th>
<th>p-value</th>
<th>Nikkei $\tilde{Q}_m^u$</th>
<th>p-value</th>
<th>SP 500 $\tilde{Q}_m^u$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.508</td>
<td>2.7</td>
<td>0.111</td>
<td>1.9</td>
<td>0.164</td>
<td>2.9</td>
<td>0.61</td>
<td>0.3</td>
<td>0.069</td>
<td>3.3</td>
<td>0.069</td>
<td>0.95</td>
<td>0.008</td>
<td>6.4</td>
<td>0.012</td>
<td>3.3</td>
<td>0.069</td>
</tr>
<tr>
<td>2</td>
<td>0.9</td>
<td>0.642</td>
<td>2.7</td>
<td>0.146</td>
<td>2.5</td>
<td>0.258</td>
<td>2.5</td>
<td>0.72</td>
<td>9.5</td>
<td>0.191</td>
<td>3.9</td>
<td>0.191</td>
<td>9.6</td>
<td>0.022</td>
<td>14.3</td>
<td>0.001</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>3</td>
<td>5.8</td>
<td>0.124</td>
<td>5.8</td>
<td>0.119</td>
<td>2.6</td>
<td>0.022</td>
<td>9.6</td>
<td>0.001</td>
<td>14.4</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>9.8</td>
<td>0.024</td>
<td>14.4</td>
<td>0.002</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>4</td>
<td>6.5</td>
<td>0.162</td>
<td>8.3</td>
<td>0.059</td>
<td>2.6</td>
<td>0.045</td>
<td>9.8</td>
<td>0.002</td>
<td>14.5</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>10.1</td>
<td>0.047</td>
<td>14.5</td>
<td>0.013</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>5</td>
<td>11.3</td>
<td>0.046</td>
<td>9.5</td>
<td>0.092</td>
<td>3.4</td>
<td>0.072</td>
<td>10.1</td>
<td>0.002</td>
<td>14.5</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>11.2</td>
<td>0.074</td>
<td>14.5</td>
<td>0.013</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>6</td>
<td>12</td>
<td>0.062</td>
<td>9.5</td>
<td>0.082</td>
<td>5</td>
<td>0.074</td>
<td>11.2</td>
<td>0.002</td>
<td>14.5</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>11.4</td>
<td>0.102</td>
<td>14.5</td>
<td>0.013</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>7</td>
<td>13.3</td>
<td>0.066</td>
<td>12.7</td>
<td>0.122</td>
<td>3.4</td>
<td>0.102</td>
<td>12.9</td>
<td>0.002</td>
<td>14.5</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>13.6</td>
<td>0.143</td>
<td>14.5</td>
<td>0.013</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>8</td>
<td>13.4</td>
<td>0.097</td>
<td>12.9</td>
<td>0.17</td>
<td>5</td>
<td>0.143</td>
<td>13.6</td>
<td>0.002</td>
<td>14.5</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>13.7</td>
<td>0.154</td>
<td>14.5</td>
<td>0.013</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>9</td>
<td>13.4</td>
<td>0.143</td>
<td>13.6</td>
<td>0.228</td>
<td>3.4</td>
<td>0.154</td>
<td>13.7</td>
<td>0.002</td>
<td>14.5</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>13.9</td>
<td>0.212</td>
<td>14.5</td>
<td>0.013</td>
<td>4.8</td>
<td>0.191</td>
</tr>
<tr>
<td>10</td>
<td>13.7</td>
<td>0.188</td>
<td>13.7</td>
<td>0.284</td>
<td>5</td>
<td>0.212</td>
<td>13.9</td>
<td>0.002</td>
<td>14.5</td>
<td>0.191</td>
<td>4.8</td>
<td>0.191</td>
<td>13.9</td>
<td>0.212</td>
<td>14.5</td>
<td>0.013</td>
<td>4.8</td>
<td>0.191</td>
</tr>
</tbody>
</table>

D  Complementary results for the application to the stock market indices

Figure 11 displays the returns of the nine stock market indices used in the empirical application. Figure 12 shows the autocorrelation function (ACF), and Figure 13 shows the ACF of the squares of the series. For each series, the squares are much more autocorrelated than the initial returns, which is very standard for such financial series.

Table 12 of the present section and Table 5 of Section 7.2 display portmanteau tests for the adequacy of an AR(1) model, or of an AR(0) model, for the linear part. Very often the AR(0) is not rejected, meaning that no linear part in needed, which is in accordance to the standard economic theory of efficient markets.

Based on the portmanteau tests defined in Theorem 6.2, Table 13 clearly rejects the AR(1)-LARCH(1) model, for all the stock market returns. In view of Table 14, the LARCH(5) model is also frequently rejected. Note however that, in view of the empirical results of Section C.2, the results provided by these portmanteau tests must be interpreted with caution. In view of Figure 14, the assumption $A_{13}$ required in Theorem 6.2 is not plausible for most of the indices, except for the Shanghai and the Nikkei. For these two indices, the empirical study conducted in Section C.2 indicates that the portmanteau tests based on $\tilde{Q}_m^2$ are likely to perform reasonably well. Even for these two indices, Table 14 indicates that the portmanteau tests reject the LARCH(5) model. This leads to think that for the non linear part of the model, a LARCH($q$) with a small order $q$, is not perfectly adequate. To solve the problem one can consider two types of extensions: (i) introducing a parametrization of the $b_i$ coefficients in (1.1), for instance of the form $b_i = ci^d$, and allowing for $q = \infty$, (ii) adding a persistence term of the form $\beta \sigma_{t-1}$ to the volatility. Such extensions are left for future work.
Table 13: Portmanteau tests of adequacy of the AR(1)-LARCH(1) model for nine stock market returns.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{Q}_m^2$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC</td>
<td>0.7</td>
<td>0.906</td>
<td>244.9</td>
<td>274.3</td>
<td>343.2</td>
<td>410.5</td>
<td>448.3</td>
<td>514.9</td>
<td>577.3</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.412</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Shanghai</td>
<td>0</td>
<td>2</td>
<td>35.9</td>
<td>49.4</td>
<td>69.2</td>
<td>69.9</td>
<td>74.6</td>
<td>76.3</td>
<td>81</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.879</td>
<td>0.362</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DAX</td>
<td>1.5</td>
<td>176.1</td>
<td>227.2</td>
<td>280.4</td>
<td>351.7</td>
<td>419.8</td>
<td>447.9</td>
<td>562.4</td>
<td>576.9</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.228</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DJA</td>
<td>1.2</td>
<td>35.9</td>
<td>113.2</td>
<td>158.4</td>
<td>207</td>
<td>245.5</td>
<td>280.8</td>
<td>296.6</td>
<td>320.6</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.268</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DJT</td>
<td>0.7</td>
<td>22.7</td>
<td>74.7</td>
<td>81.6</td>
<td>98.7</td>
<td>108.5</td>
<td>114.4</td>
<td>156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.988</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>FTSE</td>
<td>0.1</td>
<td>105.8</td>
<td>207.7</td>
<td>238.4</td>
<td>312.5</td>
<td>344.8</td>
<td>358</td>
<td>468.6</td>
<td>516.9</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.737</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nasdaq</td>
<td>0</td>
<td>4.3</td>
<td>30.7</td>
<td>56.4</td>
<td>109.9</td>
<td>115.6</td>
<td>116.4</td>
<td>119.5</td>
<td>131</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.871</td>
<td>0.117</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nikkei</td>
<td>0.5</td>
<td>38.1</td>
<td>78.8</td>
<td>103.1</td>
<td>152.9</td>
<td>201.5</td>
<td>252.2</td>
<td>270.3</td>
<td>284.9</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.495</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SP 500</td>
<td>4.9</td>
<td>51.7</td>
<td>94.1</td>
<td>131.5</td>
<td>188.9</td>
<td>210.9</td>
<td>252</td>
<td>281.2</td>
<td>311.4</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.027</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Table 14: Portmanteau tests for the adequacy of the LARCH(5) model for nine stock market returns.

<table>
<thead>
<tr>
<th></th>
<th>$\hat{Q}_m^2$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAC</td>
<td>47.1</td>
<td>69.3</td>
<td>76</td>
<td>97.7</td>
<td>113.5</td>
<td>137.2</td>
<td>165.7</td>
<td>213.4</td>
<td>236.9</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.145</td>
<td>0.045</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Shanghai</td>
<td>2.1</td>
<td>6.2</td>
<td>22.2</td>
<td>23.9</td>
<td>24.1</td>
<td>24.5</td>
<td>25</td>
<td>26.2</td>
<td>27.5</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.014</td>
<td>0.001</td>
<td>0.003</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.001</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DAX</td>
<td>6.1</td>
<td>13.1</td>
<td>14.2</td>
<td>15.2</td>
<td>17.2</td>
<td>19.2</td>
<td>21</td>
<td>25.1</td>
<td>31.4</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.014</td>
<td>0.001</td>
<td>0.003</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.004</td>
<td>0.001</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DJA</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>13.9</td>
<td>30.8</td>
<td>31.4</td>
<td>39</td>
<td>42.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.567</td>
<td>0.841</td>
<td>0.937</td>
<td>0.008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>DJT</td>
<td>22.4</td>
<td>26.7</td>
<td>33.3</td>
<td>37.4</td>
<td>43.4</td>
<td>57</td>
<td>78.9</td>
<td>90.2</td>
<td>110.8</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>FTSE</td>
<td>3</td>
<td>5.5</td>
<td>5.5</td>
<td>7.5</td>
<td>8.2</td>
<td>19.3</td>
<td>31.7</td>
<td>52.1</td>
<td>53.3</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.085</td>
<td>0.065</td>
<td>0.141</td>
<td>0.11</td>
<td>0.143</td>
<td>0.004</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nasdaq</td>
<td>0.997</td>
<td>0.994</td>
<td>0.999</td>
<td>1</td>
<td>1</td>
<td>0.999</td>
<td>0.992</td>
<td>0.997</td>
<td>0.998</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Nikkei</td>
<td>17.4</td>
<td>25.6</td>
<td>48.4</td>
<td>72.8</td>
<td>98.9</td>
<td>159.6</td>
<td>203</td>
<td>216.3</td>
<td>230.8</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SP 500</td>
<td>0.1</td>
<td>0.2</td>
<td>1.2</td>
<td>3.9</td>
<td>9</td>
<td>16.6</td>
<td>25.6</td>
<td>34.1</td>
<td>40.2</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.701</td>
<td>0.924</td>
<td>0.757</td>
<td>0.422</td>
<td>0.111</td>
<td>0.011</td>
<td>0.001</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
Figure 11: Nine stock market indices.
Figure 12: ACF of nine stock market indices.
Figure 13: ACF of nine stock market indices.
Figure 14: Comparison between the estimated volatility of the stock market indices and zero (the horizontal line).