INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2008-07

Estimating ARCH Models when the Coefficients are Allowed to be Equal to Zero

C. FRANCQ[1]
J.-M. ZAKOÏAN[2]

[1] Université Lille III, GREMARS-EQUIPPE, BP 60149, 59653 Villeneuve d'Ascq Cédex, France.
Tél. : 33 (0) 3 20 41 64 87 (francq@univ-lille3.fr)
[2] CREST and GREMARS-EQUIPPE, 15 Boulevard Gabriel Péri, 92245 Malakoff Cédex, France.
Tél. : 33 (0) 1 41 17 78 25 (zakoian@ensae.fr).

Estimating ARCH Models When the Coefficients are Allowed to be Equal to Zero

Christian Francq, Jean-Michel Zakoïan
University Lille 3 and CREST

Abstract: In order to be consistent with volatility processes, the autoregressive conditional heteroskedastic (ARCH) models are constrained to have non-negative coefficients. The estimators incorporating these constraints possess non standard asymptotic distributions when the true parameter has zero coefficients. This situation, where the parameter is on the boundary of the parameter space, must be considered to derive the critical values of tests that one or several ARCH coefficients are equal to zero. In this paper we compare the asymptotic theoretical properties, as well as the finite sample behavior, of the main estimation methods in this framework.

Keywords: ARCH model, Boundary of the parameter space, Conditional heteroskedasticity, Quasi Maximum Likelihood Estimation, Non-normal asymptotic distribution.

Résumé : Afin de préserver les propriétés d'un processus de volatilité, les coefficients des modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) sont contraints à être positifs. Les estimateurs tenant compte de ces contraintes ont des distributions asymptotiques non standard quand la vraie valeur du paramètre a des coefficients nuls. Dans cette situation, où la vraie valeur est sur la frontière de l'espace des paramètres, les valeurs critiques des tests de nullité d'un ou plusieurs coefficients ARCH sont également différentes du cas standard. Nous comparons dans ce cadre les propriétés asymptotiques théoriques, ainsi que le comportement en échantillon fini, des principales méthodes d'estimation.

Mots clés: modèle ARCH, paramètre sur le bord, hétéroscédasticité conditionnelle, estimateur du quasi maximum de vraisemblance, distribution asymptotique non normale.

1 Introduction

Least squares (LS) and quasi-maximum likelihood (QML) procedures are arguably the most widely-used estimations methods for ARCH models, and were already considered in the seminal paper by Engle (1982). The LS estimator (LSE) has the advantage of being a closed-form estimator that can be easily implemented and does not require the use of optimization procedures, but has the disadvantage of being generally much less efficient than the QMLE. The quasi-generalized least squares estimator (QGLSE) improves the efficiency of the LSE but remains user-friendly. Deriving the asymptotic properties of these estimators is not a trivial task. Berkes, Horváth, and Kokoszka (2003) is the first reference in which the asymptotic properties of the QMLE of ARCH and the generalized ARCH (GARCH) models were captured in a mathematically rigorous way under weak conditions (see also Francq & Zakoïan, 2004 and Straumann, 2005 where several technical assumptions made in Berkes et al., 2003 are relaxed).

For an estimator to be asymptotically normal (AN), a crucial assumption is that the true parameter must belong to the interior of the parameter space. This requirement, made by the above-mentioned papers, is not satisfied when the ARCH coefficients are constrained to be positive by the estimation procedure and when some components of the true ARCH parameter are equal to zero. Following Chernoff (1954) or Andrews (1999) who studied in general frameworks the asymptotic distribution of estimators when the parameter is on a boundary of the parameter space, Jordan (2003) and Francq and Zakoïan (2007) studied the ARCH and GARCH QMLE when the parameter is allowed to have zero components. This framework is particularly relevant for hypothesis testing problems, which often put the parameter on the boundary of the parameter space under the null. Tests of the significance of the coefficients and tests of conditional homoscedasticity constitute typical situations where we have to study the estimators when the parameter is at the boundary.

In this paper we compare the asymptotic behaviour of the LSE, QGLSE and QMLE of ARCH models, when the true parameter may have zero coefficients. We also consider constrained and truncated versions of the LSE and QGLSE. We limit ourselves to ARCH models because the LSE and QGLSE lose their main practical advantage (namely the fact that they do not need any numerical optimization procedure) in the GARCH framework.

2 Constrained and unconstrained estimators

Consider the standard ARCH(q) model given by the equations

$$\begin{cases} \epsilon_t = \sqrt{h_t}\eta_t \\ \sigma_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i}\epsilon_{t-i}^2, \quad \forall t \in \mathbb{Z} \end{cases} \tag{1}$$

where the noise sequence (η_t) is independent and identically distributed (iid) with mean 0 and variance 1, the distribution of η_t^2 is not degenerated[1], and

$$\theta_0 := (\omega_0, \alpha_{01}, \dots, \alpha_{0q})$$

[1] *i.e.* η_t^2 is not almost surely equal to 1 (this assumption entails the identifiability of the ARCH coefficients and the invertibility of the Fisher information matrix).

is a parameter vector which satisfies the positivity constraints

$$\omega_0 > 0, \quad \alpha_{0i} \geq 0 \;\; (i = 1, \ldots, q).$$

These constraints are sufficient (and are also necessary when η_t has a positive density) to ensure the positivity of the volatility process $\sigma_t^2 = \sigma_t^2(\theta_0)$.

2.1 Quasi-maximum likelihood estimator

The QML estimation procedure requires the computation of the logarithm of $\sigma_t^2(\theta) = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2$ at any point $\theta = (\omega, \alpha_1, \ldots, \alpha_q)$ of the parameter space Θ. If $\alpha_i < 0$ for some i, the volatility $\sigma_t^2(\theta)$ is likely to take negative values[2] and the QML procedure fails. Assuming that the parameter space Θ is a compact subset of $[0, \infty)^{q+1}$ that bounds the first component away from zero, one can compute the quasi-likelihood criterion

$$\mathbf{l}_n(\theta) = n^{-1} \sum_{t=q+1}^{n} \ell_t(\theta), \qquad \ell_t(\theta) = \frac{\epsilon_t^2}{\sigma_t^2(\theta)} + \log \sigma_t^2(\theta).$$

The QMLE is then defined as any measurable solution of

$$\hat{\theta}_n^{QML} = \arg\min_{\theta \in \Theta} \mathbf{l}_n(\theta).$$

Assume $\theta_0 \in \Theta$. We have strong consistency under the sole assumption that (ϵ_t) is a strictly stationary solution to (1) such that ϵ_t is measurable with respect to $\{\eta_u, u \leq t\}$ (this assumption[3] is maintained throughout the paper). We have AN

$$\sqrt{n}(\hat{\theta}_n^{QML} - \theta_0) \xrightarrow{d} \mathcal{N}\left\{0, \;\; (E\eta_1^4 - 1)J^{-1}\right\}, \quad J = E\left\{\frac{1}{\sigma_{q+1}^4} \frac{\sigma_{q+1}^2}{\partial \theta} \frac{\sigma_{q+1}^2}{\partial \theta'}(\theta_0)\right\}, \qquad (2)$$

under the additional assumption that

$$\theta_0 \in \overset{\circ}{\Theta}, \qquad (3)$$

where $\overset{\circ}{\Theta}$ denotes the interior of the parameter space Θ (see Berkes et al., 2003, Francq & Zakoïan, 2004 and Straumann, 2005). When some of its components are equal to zero, the parameter θ_0, which is constrained to have nonnegative components, belongs to the boundary of the parameter space and (3) is not satisfied.

The following elementary example shows that the asymptotic distribution of the QMLE cannot be Gaussian when (3) is not satisfied.

[2] If $\alpha_i < 0$ and η_t has a positive density over $\mathbb{R}$, one can show that, almost surely, $\sigma_t^2(\theta) < 0$ for infinitely many t.

[3] Jensen and Rahbek (2004) showed that the coefficient α_{01} of an ARCH(1) model can be consistently estimated without any stationarity condition, but the strict stationarity condition is required for the estimation of ω_0 (this point is often misunderstood in the recent literature).

Example 2.1 *Due to the positivity constraints, the QMLE of an ARCH(1) model satisfies* $\hat{\alpha}_n \geq 0$ *almost surely, for all* n. *When the DGP is a white noise, then* $\alpha_{01} = 0$ *and with probability one*

$$\sqrt{n}(\hat{\alpha}_n - \alpha_{01}) = \sqrt{n}\hat{\alpha}_n \geq 0, \qquad \forall n.$$

In this case $\sqrt{n}(\hat{\alpha}_n - \alpha_{01})$ *cannot converge in law to any non-degenerate Gaussian distribution* $\mathcal{N}(m, s^2)$ *with* $s^2 > 0$. *Indeed*

$$\lim_{n\to\infty} P\left\{\sqrt{n}(\hat{\alpha}_n - \alpha_{01}) < 0\right\} = 0 \quad \textit{whereas} \quad P\left\{\mathcal{N}(m, s^2) < 0\right\} > 0.$$

For the same reason, when the true value of a general GARCH parameter has zero components, the asymptotic distribution cannot be Gaussian, for the QMLE or for any other estimator which takes into account the positivity constraints.

2.2 Least squares estimators

The LSE is an alternative estimator based on the AR(q) representation for ϵ_t^2

$$\epsilon_t^2 = \omega_0 + \sum_{i=1}^{q} \alpha_{0i}\epsilon_{t-i}^2 + u_t, \tag{4}$$

where $u_t = \epsilon_t^2 - h_t = (\eta_t^2 - 1)h_t$. When $\sum_{i=1}^q \alpha_{0i} < 1$ we have $E\epsilon_1^2 < \infty$. Denoting by $\mathcal{F}_{t-1}$ the σ-field generated by $\{\eta_u, u < t\}$, the sequence $(u_t, \mathcal{F}_{t-1})_t$ is then a martingale difference. Introducing the vector $Z'_{t-1} = \left(1, \epsilon_{t-1}^2, \ldots, \epsilon_{t-q}^2\right)$, we get from (4)

$$Y = X\theta_0 + U$$

where

$$X = \begin{pmatrix} Z'_{n-1} \\ \vdots \\ Z'_q \end{pmatrix}, \quad Y = \begin{pmatrix} \epsilon_n^2 \\ \vdots \\ \epsilon_{q+1}^2 \end{pmatrix}, \quad U = \begin{pmatrix} u_n \\ \vdots \\ u_{q+1} \end{pmatrix}.$$

2.2.1 Unconstrained LSE

With probability one, it can be shown that the matrix $X'X$ is non-singular for large enough n. The LSE of θ_0 is thus given by

$$\hat{\theta}_n^{LS} = (X'X)^{-1}X'Y. \tag{5}$$

If $E(\epsilon_1^4) < +\infty$, the LSE can be shown to be strongly consistent (see Bose & Mukherjee, 2003). If, in addition $E(\epsilon_1^8) < +\infty$ the estimator is AN (see also Bose & Mukherjee, 2003) and

$$\sqrt{n}(\hat{\theta}_n^{LS} - \theta_0) \xrightarrow{d} \mathcal{N}\left\{0, \ (E\eta_1^4 - 1)A^{-1}BA^{-1}\right\}, \tag{6}$$

where

$$A = E(Z_q Z'_q), \qquad B = E(\sigma_{q+1}^4 Z_q Z'_q)$$

are non-singular matrices.

2.2.2 Constrained LSE

Contrary to the QMLE, the computation of the LSE does not require positivity constraints. Note that one or several components of the $\hat{\theta}_n^{LS}$, as defined by (5), can be negative. This is a serious practical problem because we know that ARCH models with negative coefficients are not viable and can produce negative predictions of the volatility. This is why it is worth considering the constrained LSE (CLSE) defined by

$$\hat{\theta}_n^{cLS} = \arg \min_{\theta \in [0,\infty[^{q+1}} \|Y - X\theta\|^2.$$

When X has rank $q+1$, the constrained estimator $\hat{\theta}_n^{cLS}$ is the orthogonal projection of $\hat{\theta}_n^{LS}$ on $[0,+\infty[^{q+1}$ with respect to the metric $X'X$:

$$\hat{\theta}_n^{cLS} = \arg \min_{\theta \in [0,+\infty[^{q+1}} (\hat{\theta}_n^{LS} - \theta)' X'X (\hat{\theta}_n^{LS} - \theta). \tag{7}$$

The following proposition states that, when θ_0 belongs to the interior of the parameter space, the asymptotic behaviors of the constrained and unconstrained LSE are the same.

Proposition 2.1 *Under the assumption $E\epsilon_1^4 < \infty$ we have $\hat{\theta}_n^{cLS} \to \theta_0$ almost surely. Under the additional assumption $\theta_0 \in (0,+\infty[^{q+1}$, with probability one we have $\hat{\theta}_n^{cLS} = \hat{\theta}_n^{LS}$ for n large enough, and thus when $E(\epsilon_1^8) < +\infty$*

$$\sqrt{n}(\hat{\theta}_n^{cLS} - \theta_0) \xrightarrow{d} \mathcal{N}\left\{0, \ (E\eta_1^4 - 1)A^{-1}BA^{-1}\right\}. \tag{8}$$

Example 2.1 shows that the AN (8) does not hold when some ARCH coefficients are equal to zero.

2.2.3 Truncated LSE

Since all the ARCH coefficients must be positive, a naive approach could be to replace any negative component of the LSE $\hat{\theta}_n^{LS} = (\hat{\theta}_{n,1}^{LS}, \dots, \hat{\theta}_{n,q+1}^{LS})'$ by zero. This leads to the truncated LSE (TLSE) defined by

$$\hat{\theta}_n^{tLS} = \left(\hat{\theta}_{n,1}^{tLS}, \dots, \hat{\theta}_{n,q+1}^{tLS}\right)', \qquad \hat{\theta}_{n,i}^{tLS} = \hat{\theta}_{n,i}^{LS} 1_{\{\hat{\theta}_{n,i}^{LS} \geq 0\}} \quad i = 1, \dots, q+1.$$

Defining the vector $\mathbf{1}_{\{\hat{\theta}_n^{LS} \geq 0\}} = \left(1_{\{\hat{\theta}_{n,1}^{LS} \geq 0\}}, \dots, 1_{\{\hat{\theta}_{n,q+1}^{LS} \geq 0\}}\right)'$ and using the Hadamard product $\odot$, the truncated estimator can be written as $\hat{\theta}_n^{tLS} = \hat{\theta}_n^{LS} \odot \mathbf{1}_{\{\hat{\theta}_n^{LS} \geq 0\}}$. This estimator is simpler to implement than the CLSE and the following proposition shows that its asymptotic properties are the same as those of the constrained and unconstrained LSE when θ_0 is not on the boundary of the parameter space.

Proposition 2.2 *Proposition 2.1 remains valid when $\hat{\theta}_n^{cLS}$ is replaced by $\hat{\theta}_n^{tLS}$.*

2.2.4 Quasi-generalized least squares estimator

For linear regression models with heteroscedastic observations, it is well known that the (ordinary) LSE is outperformed by the QGLSE (see *e.g.* Hamilton, 1994 Chapter 8). In the ARCH framework the QGLSE is defined by

$$\hat{\theta}_n^{QG} = (X'\hat{\Omega}X)^{-1}X'\hat{\Omega}Y,$$

where X is supposed to have full rank $q+1$, and $\hat{\Omega}$ is a non singular consistent estimator of $\Omega = \mathrm{Diag}(\sigma_n^{-4}, \dots, \sigma_{1+q}^{-4})$. If a first-step estimator $\hat{\theta}_n = (\hat{\omega}, \hat{\alpha}_1, \dots, \hat{\alpha}_1)'$ is available, the matrix $\hat{\Omega}$ can be obtained by replacing σ_t^2 by $\hat{\omega} + \sum_{i=1}^q \hat{\alpha}_i \epsilon_{t-i}^2$ in Ω. In order to be sure that $\hat{\Omega}$ is well defined and invertible, one can employ the truncated LSE $\hat{\theta}_n = \hat{\theta}_n^{tLS}$. Then the two-stage least squares estimator $\hat{\theta}_n^{QG}$ is consistent and asymptotically normal

$$\sqrt{n}(\hat{\theta}_n^{QG} - \theta_0) \xrightarrow{d} \mathcal{N}\left\{0, \ (E\eta_1^4 - 1)J^{-1}\right\}, \quad J = E(\sigma_{q+1}^{-4} Z_q Z_q'), \tag{9}$$

under the moment assumption $E\epsilon_1^4 < \infty$ when all the ARCH coefficients are strictly positive, and under a slightly stronger moment assumption in the general case (see Bose & Mukherjee, 2003 and Gouriéroux, 1997).

Obviously one can define constrained and truncated versions of the QGLSE by

$$\begin{aligned}\hat{\theta}_n^{cQG} &= \arg\min_{\theta \in [0,\infty[^{q+1}} (Y - X\theta)'\hat{\Omega}(Y - X\theta) \\ &= \arg\min_{\theta \in [0,+\infty[^{q+1}} (\hat{\theta}_n^{QG} - \theta)'X'\hat{\Omega}X(\hat{\theta}_n^{QG} - \theta)\end{aligned}$$

and

$$\hat{\theta}_n^{tQG} = \hat{\theta}_n^{QG} \odot \mathbf{1}_{\{\hat{\theta}_n^{QG} \geq 0\}}$$

Proposition 2.3 *Under the assumptions $E\epsilon_t^4 < \infty$ and $\theta_0 \in (0, +\infty[^{q+1}$, the three estimators $\hat{\theta}_n^{QG}$, $\hat{\theta}_n^{cQG}$ and $\hat{\theta}_n^{tQG}$ converge almost surely to θ_0 and have the same asymptotic distribution given by (9), as $n \to \infty$.*

3 Conditions for AN of the estimators and comparison of the asymptotic variances

In view of (2), (6) and (9), the following lemma shows the well known result that, under assumptions ensuring AN, the LSE and its variants (the CLSE and TLSE) are less efficient than the QMLE and the (unconstrained, constrained and truncated) QGLSE.

Lemma 3.1 *Under the assumption $E\epsilon_t^8 < \infty$,*

$$A^{-1}BA^{-1} - J^{-1}$$

is positive semi-definite.

Note however that the conditions required to obtain AN are not the same for the different estimators. In particular the computation of the QMLE requires positivity constraints[4], contrary to the LSE and QGLSE. On the other hand the LSE and its extensions require moment conditions, whereas the QMLE requires only the strict stationarity condition. For an ARCH(1) model the strict stationarity condition is $\alpha_{01} < \exp\{-E \log \eta_t^2\}$ and the second-order stationarity requires the much stronger condition $\alpha_{01} < 1$, and the condition $E\epsilon_t^{2m} < \infty$ is equivalent to $\alpha^m E\eta_t^{2m} < 1$ for all $m \in \{1, 2, \ldots, \}$. The absence of moment conditions is an important advantage for the QMLE over the other estimators because the ARCH models are often fitted to financial series showing evidence of fat tails.

The following table summarizes the constraints on the different estimators in the simple ARCH(1) case, when η_t follows a standard Gaussian $\mathcal{N}(0, 1)$ or Student distributions normalized in such a way that $E\eta_1^2 = 1$ (St_ν stands for a normalized Student distribution with ν degrees of freedom). Note that, as shown by a trivial extension of Example 2.1 to general constrained estimators, the value $\alpha_{01} = 0$ is not allowed for the AN of the QMLE, CLSE and CQGLSE. The next section gives the asymptotic distribution of these estimators when θ_0 belongs to the boundary of $[0, \infty[^{q+1}$.

Table 1: Conditions ensuring asymptotic normality for estimators of an ARCH(1) model with coefficient (ω_0, α_{01}), when the iid noise η_t follows a standard Gaussian or normalized Student distributions.

	QMLE	LSE and QGLSE[a]	Constrained and truncated LSE and QGLSE
Normal	$\alpha_{01} \in (0, 3.562)$	$\alpha_{01} \in [0, 0.312)$	$\alpha_{01} \in (0, 0.312)$
St_3	$\alpha_{01} \in (0, 7.389)$	$\alpha_{01} \in \emptyset$	$\alpha_{01} \in \emptyset$
St_5	$\alpha_{01} \in (0, 4.797)$	$\alpha_{01} \in \emptyset$	$\alpha_{01} \in \emptyset$
St_9	$\alpha_{01} \in (0, 4.082)$	$\alpha_{01} \in [0, 0.143)$	$\alpha_{01} \in (0, 0.143)$

[a]For the proof of the AN of this estimator, a technical additional assumption (see Equation (8) in Bose & Mukherjee, 2003) is required. This technical assumption is satisfied, in particular, when $\alpha_{01} > 0$ or when $E\epsilon_t^6 < \infty$.

4 Asymptotic distribution of the estimators when the parameter is on the boundary

The parameter θ_0 is allowed to contains zero components, but we exclude the situation where θ_0 attains the upper boundary of the parameter space. Under this assumption the

[4]In a recent paper Iglesias and Linton (2007) propose to approximate the distribution of constrained estimators using Edgeworth expansions of unrestricted estimators. This approach fails for the QMLE of ARCH models because, the volatility being constrained to be positive to compute the QML, the unrestricted QMLE does not exist.

set $\sqrt{n}(\Theta - \theta_0)$ converges to the so-called local parameter space Λ defined by

$$\Lambda = \Lambda(\theta_0) = \Lambda_1 \times \cdots \times \Lambda_{q+1},$$

where $\Lambda_1 = \mathbf{R}$, and, for $i = 2, \ldots, q+1$,

$$\Lambda_i = \mathbf{R} \quad \text{if} \quad \theta_{0i} \neq 0 \qquad \text{and} \qquad \Lambda_i = [0, \infty) \quad \text{if} \quad \theta_{0i} = 0.$$

In view of the positivity constraints, the random vector $\sqrt{n}(\hat{\theta}_n^{QML} - \theta)$ belongs to Λ with probability one. Following Chernoff (1954) or Andrews (1999) who studied boundary problems in very general frameworks, Francq and Zakoian (2007) gave conditions under which

$$\sqrt{n}(\hat{\theta}_n^{QML} - \theta_0) \xrightarrow{d} \lambda^{\Lambda} := \arg\inf_{\lambda \in \Lambda} \{\lambda - Z\}' J \{\lambda - Z\}, \tag{10}$$

with

$$Z \sim \mathcal{N}\left\{0, (E\eta_1^4 - 1)J^{-1}\right\}, \quad J = E\left(\sigma_{q+1}^{-4} Z_q Z_q'\right).$$

In the ARCH framework these conditions reduce to the moment condition $E\epsilon_t^6 < \infty$.[5] When (3) holds true, we have $\Lambda = \mathbf{R}^{q+1}$ and we retrieve the standard result because $\lambda^{\Lambda} = Z \sim \mathcal{N}\{0, (E\eta_1^4 - 1)J^{-1}\}$. When θ_0 is on the boundary, the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n^{QML} - \theta_0)$ is more complex than a Gaussian. This is the law of the projection of the Gaussian vector Z on the convex cone Λ. The asymptotic distributions of the constrained LSE and QGLSE are of the same type.

Proposition 4.1 *When (6) holds (*i.e. *when $E\epsilon_1^8 < \infty$) we have*

$$\sqrt{n}(\hat{\theta}_n^{cLS} - \theta_0) \xrightarrow{d} \arg\inf_{\lambda \in \Lambda} \left\{\lambda - \tilde{Z}\right\}' A \left\{\lambda - \tilde{Z}\right\},$$

*with $\tilde{Z} \sim \mathcal{N}\{0, (E\eta_1^4 - 1)A^{-1}BA^{-1}\}$. When (9) holds (*i.e. *when $E\epsilon_1^4 < \infty$) we have $\sqrt{n}(\hat{\theta}_n^{cQG} - \theta_0) \xrightarrow{d} \lambda^{\Lambda}$.*

The asymptotic distributions of the truncated estimator is simply the truncation of asymptotic distribution of the unrestricted estimators.

Proposition 4.2 *With the notation $\tilde{Z}$ introduced in Proposition 4.1, when (6) holds (*i.e. *when $E\epsilon_1^8 < \infty$) we have $\sqrt{n}(\hat{\theta}_n^{tLS} - \theta_0) \xrightarrow{d} \tilde{Z} \odot \mathbf{1}_{\{\tilde{Z} \geq 0 \text{ or } \theta_0 > 0\}}$. When (9) holds (*i.e. *when $E\epsilon_1^4 < \infty$) we have $\sqrt{n}(\hat{\theta}_n^{tQG} - \theta_0) \xrightarrow{d} Z \odot \mathbf{1}_{\{Z \geq 0 \text{ or } \theta_0 > 0\}}$.*

We use $\mathrm{MSE}^{QML} = \text{trace}\left\{E\left(\lambda^{\Lambda}\lambda^{\Lambda'}\right)\right\}$ as a scalar measure for the asymptotic accuracy of the QMLE, and define similarly the MSE of the other estimators. Because we do not have an explicit expression for the matrix J, it seems difficult to compute and compare the MSE of all the estimators in the general setting. Comparison is however possible on the following example.

[5] The moment assumption $E\epsilon_t^4 < \infty$ seems necessary (for the existence of J) and is maybe sufficient, but a stronger moment assumption is required in the proof given by Francq and Zakoïan (2007).

4.1 Comparing the accuracy of the estimators under conditional homoscedasticity

Consider an ARCH(q) in which $\alpha_{01} = \cdots = \alpha_{0q} = 0$. This framework is encountered in conditional homoscedasticity tests. In this case the local parameter space is $\Lambda = \mathbb{R} \times (0, \infty)^q$ and the information matrix J has a simple expression. A straightforward computation, available from the authors under request, yields

$$\begin{aligned} \mathrm{MSE}^{QML} &= \mathrm{MSE}^{CLS} = \mathrm{MSE}^{CQGLS} \\ &= (E\eta_1^4 + q - 1)\omega_0^2 - q\omega_0^2/2 + q(q-1)\omega_0^2/2\pi + q/2, \qquad (11) \\ \mathrm{MSE}^{TLS} &= \mathrm{MSE}^{TQGLS} = (E\eta_1^4 + q - 1)\omega_0^2 + q/2, \qquad (12) \\ \mathrm{MSE}^{LS} &= \mathrm{MSE}^{QGLS} = (E\eta_1^4 + q - 1)\omega_0^2 + q. \qquad (13) \end{aligned}$$

It is interesting to note that

$$\mathrm{MSE}^{QML} < \mathrm{MSE}^{TLS} < \mathrm{MSE}^{LS}$$

when $q \leq 4$,

$$\mathrm{MSE}^{TLS} < \mathrm{MSE}^{QML} < \mathrm{MSE}^{LS}$$

when $5 \leq q < (\pi + 1) + \pi/\omega_0^2$, and

$$\mathrm{MS}^{TLS} < \mathrm{MS}^{LSE} < \mathrm{MS}^{QML}$$

when $q > (\pi + 1) + \pi/\omega_0^2$. That the QMLE (which is actually the maximum likelihood estimator (MLE) when η_t is gaussian) might be dominated by another estimator seems quite surprising, at least at first sight. Our interpretation of this interesting phenomenon is the following. According to Le Cam's theory on convergence of local experiments (see *e.g.* van der Vaart, 1998) our problem is closely related to the problem of estimating $m \in \Lambda$ from one observation $X \sim \mathcal{N}(m, 2J^{-1})$, assuming for simplicity that $\eta_t \sim \mathcal{N}(0, 1)$. The form of J^{-1} being very simple, it can be shown that the MLE of m is explicitly given by $\hat{m}^{ML} = \left(X_1 + \omega_0 \sum_{i=2}^{q+1} X_i^-, X_2^+, \ldots, X_{q+1}^+\right)'$. It is easy to see that this MLE estimator is less efficient than $\hat{m}^{tLS} = \left(X_1, X_2^+, \ldots, X_{q+1}^+\right)'$ when $m = (m_1, 0, \ldots, 0)'$ and $q > 4$.

4.2 Monte Carlo results

Table 2 summarizes the output of Monte Carlo experiments. The empirical MSE's are generally close to the asymptotic MSE's obtained from (11)-(13) (given in the rows $n = \infty$). The smallest MSE's are displayed in bold type. We note that the (Q)MLE can be outperformed by simpler estimators, in finite samples and also asymptotically when $q = 6$. The TLSE, although particularly simple to implement, performs remarkably well in the framework of Table 2. Other simulation experiments, not reported here, reveal that, as expected, the QMLE and (C/T)QGLSE are much more accurate than the other estimators in the presence of conditionally heteroscedastic data.

From the asymptotic theory, as well as from Table 2 and other numerical experiments not presented here, we draw the conclusion that i) the QMLE is generally superior to the other estimators in terms of accuracy when the data show evidence of conditional heteroscedasticity and/or heavy tail distribution, ii) the (Q)MLE can be outperformed by simpler estimators, such as the TLSE, when the true value of the parameters stay at the boundary of the parameter space. Detailed proofs of the results given in this paper are available from the authors.

Table 2: Empirical MSE and asymptotic MSE for estimators of an ARCH(q) model when the data generated process is a $\mathcal{N}(0, 0.2^2)$ iid sequence. The number of replications is $N = 1,\ 000$.

q	n	QMLE	LSE	QGLSE	CLSE	CQGLSE	TLSE	TQGLSE
1	100	0.74	1.07	1.11	**0.58**	0.61	0.60	0.64
	1000	0.59	1.14	1.12	0.59	**0.57**	0.61	0.59
	∞	**0.60**	1.12	1.12	**0.60**	**0.60**	0.62	0.62
3	100	2.30	3.30	3.29	**1.47**	1.59	1.59	1.67
	1000	1.76	3.26	3.18	1.73	**1.66**	1.77	1.70
	∞	**1.68**	3.20	3.20	**1.68**	**1.68**	1.70	1.70
6	100	5.27	6.94	6.70	**3.13**	3.27	3.31	3.38
	1000	3.51	6.35	6.19	3.34	**3.24**	3.34	**3.24**
	5000	3.42	6.71	6.68	3.37	3.35	3.35	**3.33**
	∞	3.39	6.32	6.32	3.39	3.39	**3.32**	**3.32**

Appendix

A Computing the Constrained LSE with linear projections

In view of (7), the vector $X\hat{\theta}_n^{cLS}$ coincides with one of the $2^{q+1}-1$ projections of Y on the linear subspaces generated by the column $i_1, \dots, i_k$ of X, with $k = 1, \dots, q+1$. We now give a result which can be helpful to compute $\hat{\theta}_n^{cLS}$ without considering all the $2^{q+1} - 1$ projections. It is clear that $\hat{\theta}_n^{cLS} = \hat{\theta}_n^{LS}$ when $\hat{\theta}_n^{LS} \in [0, +\infty[^{q+1}$. Suppose that one of the components of $\hat{\theta}_n^{LS}$ is strictly negative, for instance the last one. Let

$$X = (X^{(1)}, X^{(2)}), \quad X^{(2)} = \begin{pmatrix} \epsilon_{n-q}^2 & \cdots & \epsilon_1^2 \end{pmatrix}',$$

and

$$\hat{\theta}_n^{LS} = (X'X)^{-1} X'Y = \begin{pmatrix} \hat{\theta}_n^{(1)} \\ \hat{\alpha}_q \end{pmatrix}, \quad \tilde{\theta}_n = \begin{pmatrix} \tilde{\theta}_n^{(1)} \\ 0 \end{pmatrix} = \begin{pmatrix} \left(X^{(1)'}X^{(1)}\right)^{-1} X^{(1)'}Y \\ 0 \end{pmatrix}.$$

Note that $\tilde{\theta}_n^{(1)}$ is the LSE of the ARCH($q-1$) model.

Proposition A.1 *Assume that X has rank $q+1$ and $\hat{\alpha}_q < 0$. Then*

$$\tilde{\theta}_n^{(1)} \in [0, +\infty[^q \qquad \textit{if and only if} \qquad \hat{\theta}_n^{cLS} = \tilde{\theta}_n.$$

Proof of Proposition A.1. Let $P^{(1)} = X^{(1)} \left(X^{(1)'} X^{(1)}\right)^{-1} X^{(1)'}$ be the projection matrix on the columns of $X^{(1)}$ and let $M^{(1)} = I_n - P^{(1)}$. One can verify that

$$\left(\hat{\theta}_n^{LS'} - \tilde{\theta}_n'\right) X'X = \left(0, Y' M^{(1)} X^{(2)}\right).$$

Let $e_{q+1} = (0, \dots, 0, 1)' \in \mathbb{R}^{q+1}$. Since $\hat{\theta}_n \hat{\theta}_n^{LS} e_{q+1} < 0$, we have $(\hat{\theta}_n^{LS'} - \tilde{\theta}_n') e_{q+1} < 0$, which can be written as

$$\left(\hat{\theta}_n^{LS'} - \tilde{\theta}_n'\right) X'X \left(X'X\right)^{-1} e_{q+1} < 0,$$

or

$$Y' M^{(1)} X^{(2)} \{\left(X'X\right)^{-1}\}_{q+1, q+1} < 0.$$

Thus $Y' M^{(1)} X^{(2)} < 0$. It follows that for all $\theta = (\theta^{(1)'}, \theta^{(2)})' \in [0, +\infty[^{q+1}$ we have

$$\begin{aligned}
\left\langle \hat{\theta}_n^{LS} - \tilde{\theta}_n, \tilde{\theta}_n - \theta \right\rangle_{X'X} &= \left(\hat{\theta}_n^{LS} - \tilde{\theta}_n\right)' X'X \left(\tilde{\theta}_n^{LS} - \theta\right) \\
&= \left(0, Y' M^{(1)} X^{(2)}\right) \begin{pmatrix} \tilde{\theta}_n^{(1)} - \theta^{(1)} \\ -\theta^{(2)} \end{pmatrix} \\
&= -\theta^{(2)} Y' M^{(1)} X^{(2)} \geq 0.
\end{aligned}$$

When $\tilde{\theta}_n \in [0, +\infty[^{q+1}$, this vector is thus the projection of $\hat{\theta}_n^{LS}$ on the convex set $[0, +\infty[^{q+1}$ with respect to the metric $X'X$.

B Proofs

Some of the proofs of this appendix are direct adaptations of well-known results on regression models (see *e.g.* Gouriéroux & Monfort, 1989), and are thus given for the reader convenience.

Proof of Proposition 2.1. Using the Minkowski inequality and (7), we have

$$\left\|\hat{\theta}_n^{cLS} - \theta_0\right\|_{X'X/n} \leq \left\|\hat{\theta}_n^{cLS} - \hat{\theta}_n^{LS}\right\|_{X'X/n} + \left\|\hat{\theta}_n^{LS} - \theta_0\right\|_{X'X/n} \leq 2 \left\|\hat{\theta}_n^{LS} - \theta_0\right\|_{X'X/n}.$$

Note that $X'X/n$ converge to the positive definite matrix A. Thus the consistency of $\hat{\theta}_n^{cLS}$ follows from that of $\hat{\theta}_n^{LS}$. If $\hat{\theta}_n^{LS} \to \theta_0 \in (0, \infty)^{q+1}$ then, for n large enough, $\hat{\theta}_n^{LS} \in (0, \infty)^{q+1}$ and, in view (7), $\hat{\theta}_n^{cLS} = \hat{\theta}_n^{LS}$. Thus the two estimators have the same asymptotic distribution.

Proof of Proposition 2.2. Because all the components of θ_0 are positive, $\|\hat{\theta}_n^{tLS} - \theta_0\| \leq \|\hat{\theta}_n^{LS} - \theta_0\|$, and the consistency of the LSE entails that of the truncated estimator. If $\hat{\theta}_n^{LS} \to \theta_0 \in (0, \infty)^{q+1}$ then, for n large enough, all the components of $\hat{\theta}_n^{LS}$ are positive and we thus have $\hat{\theta}_n^{tLS} = \hat{\theta}_n^{LS}$. In this case the two estimators have the same asymptotic distribution.

Proof of Proposition 2.3. The proof follows from the arguments given in the proofs of Propositions 2.1 and 2.2.

Proof of Lemma 3.1. Letting $D = \sigma_t^2 A^{-1} Z_{t-1} - \sigma_t^{-2} J^{-1} Z_{t-1}$, we have

$$\begin{aligned} E(DD') &= A^{-1}E(\sigma_t^4 Z_{t-1}Z_{t-1}')A^{-1} + J^{-1}E(\sigma_t^{-4} Z_{t-1}Z_{t-1}')J^{-1} \\ &\quad - A^{-1}E(Z_{t-1}Z_{t-1}')J^{-1} - J^{-1}E(Z_{t-1}Z_{t-1}')A^{-1} \\ &= A^{-1}BA^{-1} - J^{-1} \end{aligned}$$

and the result follows.

Proof of Proposition 4.1. In view of (7)

$$\begin{aligned} \hat{\theta}_n^{cLS} &= \arg \min_{\theta \in [0,+\infty[^{q+1}} \|\sqrt{n}(\hat{\theta}_n^{LS} - \theta)\|_{n^{-1}X'X} \\ &= \arg \min_{\theta \in [0,+\infty[^{q+1}} \|\sqrt{n}(\hat{\theta}_n^{LS} - \theta_0) - \sqrt{n}(\theta - \theta_0)\|_{n^{-1}X'X}. \end{aligned}$$

Noting that $\theta \in [0, +\infty[^{q+1}$ if and only if $\sqrt{n}(\theta - \theta_0) \in \Lambda$, we obtain

$$\sqrt{n}(\hat{\theta}_n^{cLS} - \theta_0) = \arg \min_{\lambda \in \Lambda} \|\tilde{Z}_n - \lambda\|_{n^{-1}X'X},$$

with $\tilde{Z}_n = \sqrt{n}(\hat{\theta}_n^{LS} - \theta_0)$. Since $(\tilde{Z}_n, n^{-1}X'X) \xrightarrow{d} (\tilde{Z}, A)$, we conclude by the continuous mapping theorem.

Proof of Proposition 4.2. Because $\hat{\theta}_n^{tLS} \to \theta_0$ with probability one, we have

$$\hat{\theta}_n^{tLS} \odot \mathbf{1}_{\{\theta_0 > 0\}} = \hat{\theta}_n^{LS} \odot \mathbf{1}_{\{\theta_0 > 0\}}$$

for n large enough. We also have $\hat{\theta}_n^{tLS} \odot \mathbf{1}_{\{\theta_0=0\}} = \hat{\theta}_n^{LS} \odot \mathbf{1}_{\{\theta_n^{LS}>0\}} \odot \mathbf{1}_{\{\theta_0=0\}}$. Thus

$$\sqrt{n}\left(\hat{\theta}_n^{tLS} - \theta_0\right) = \sqrt{n}\left(\hat{\theta}_n^{LS} - \theta_0\right) \odot \mathbf{1}_{\{\theta_0>0\}} + \sqrt{n}\left(\hat{\theta}_n^{LS} - \theta_0\right) \odot \mathbf{1}_{\{\theta_0=0\}} \odot \mathbf{1}_{\{\theta_n^{LS}>0\}}$$

for n large enough. Since

$$\sqrt{n}(\hat{\theta}_n^{LS} - \theta_0) \xrightarrow{d} \tilde{Z} \quad \text{and} \quad \mathbf{1}_{\{\theta_0=0\}} \odot \mathbf{1}_{\{\theta_n^{LS}>0\}} = \mathbf{1}_{\{\theta_0=0\}} \odot \mathbf{1}_{\{\theta_n^{LS}-\theta_0>0\}},$$

the continuous mapping theorem entails $\sqrt{n}(\hat{\theta}_n^{tLS} - \theta_0) \xrightarrow{d} \tilde{Z} \odot \mathbf{1}_{\{\tilde{Z}\geq 0 \text{ or } \theta_0>0\}}$. The second convergence of the proposition is obtained by the same arguments.

Proof of (11)-(13). For an ARCH(q) model with $\alpha_{01} = \cdots = \alpha_{0q} = 0$ we have

$$(E\eta_1^4 - 1)J^{-1} = \begin{pmatrix} (E\eta_1^4 + q - 1)\omega_0^2 & -\omega_0 \ \dots \ -\omega_0 \\ -\omega_0 & \\ \vdots & I_q \\ -\omega_0 & \end{pmatrix}.$$

The local parameter space is $\Lambda = \mathbb{R} \times (0, \infty)^q$ and the vector λ^Λ can be interpreted as the orthogonal projection of Z onto Λ for the scalar product $< x, y >_J = x' J y$. Since Λ is convex, λ^Λ is uniquely defined. It is clear that $\lambda^\Lambda = Z$ when $Z \in \Lambda$. When $Z \notin \Lambda$ the solution λ^Λ coincides with the $<,>_J$-orthogonal projection of Z on a linear subspace defined by $K_i \lambda = 0$, where K_i is one of the $2^q - 1$ matrix obtained by cancelling 0 or up to $q-1$ rows of the matrix $K = (0_q, I_q)$ (see Francq & Zakoïan, 2007). These projections are defined by

$$\lambda_{K_i} = P_i Z, \qquad \text{where} \qquad P_i = I_{d_1+d_2} - J^{-1} K_i' \left(K_i J^{-1} K_i' \right)^{-1} K_i.$$

The form of J^{-1} implies that $K_i J^{-1} K_i' = (E\eta_1^4 - 1)^{-1} I_{n_i}$ and $J^{-1} K_i' K_i = (E\eta_1^4 - 1)^{-1} K_i' K_i$ where n_i is the number of rows of K_i. The solution λ^Λ is finally given by the projection which belongs to Λ and is the closest to Z according to the metric J. Thus it can be shown that

$$\lambda^\Lambda = \begin{pmatrix} Z_1 + \omega_0 \sum_{i=2}^{q+1} Z_i^- \\ Z_2^+ \\ \vdots \\ Z_{q+1}^+ \end{pmatrix}, \quad Z \sim \mathcal{N} \left\{ 0, \begin{pmatrix} (E\eta_1^4 + q - 1)\omega_0^2 & -\omega_0 & \cdots & -\omega_0 \\ -\omega_0 & & & \\ \vdots & & I_q & \\ -\omega_0 & & & \end{pmatrix} \right\},$$

where $z^+ = z 1_{\{z>0\}}$ and $z^- = z 1_{\{z<0\}}$. By a symmetry argument we have $E Z_1 Z_i^- = E Z_1 Z_i^+ = E Z_1 Z_i / 2 = -\omega_0/2$ for $i \geq 2$. We also have $E Z_i^+ = -E Z_i^- = 1/\sqrt{2\pi}$ and $E \left(Z_i^+ \right)^2 = E \left(Z_i^- \right)^2 = 1/2$ for $i \geq 2$. Using also the independence of $Z_2, \ldots, Z_{q+1}$, we obtain

$$\begin{aligned} \text{MSE}^{QML} &= E \left(Z_1 + \omega_0 \sum_{i=2}^{q+1} Z_i^- \right)^2 + q E \left(Z_2^+ \right)^2 \\ &= E Z_1^2 + q \omega_0^2 E \left(Z_2^- \right)^2 + 2 q \omega_0 E Z_1 Z_2^- + q(q-1) \omega_0^2 \left(E Z_2^- \right)^2 + q E \left(Z_2^+ \right)^2 \\ &= (E\eta_1^4 + q - 1)\omega_0^2 - q\omega_0^2/2 + q(q-1)\omega_0^2/2\pi + q/2. \end{aligned}$$

We now compute the MSE of the LSE. First note that $A = \omega_0^2 J$ and $B = \omega_0^2 A$. Thus the asymptotic distribution of the LSE is that of Z, and

$$\text{MSE}^{LS} = (E\eta_1^4 + q - 1)\omega_0^2 + q.$$

Let us consider the constrained LSE. Because $A = \omega_0^2 J$, we have

$$P_i = I_{d_1+d_2} - J^{-1} K_i' \left(K_i J^{-1} K_i' \right)^{-1} K_i = I_{d_1+d_2} - A^{-1} K_i' \left(K_i A^{-1} K_i' \right)^{-1} K_i,$$

which entails that

$$\text{MSE}^{CLS} = \text{MSE}^{QML}.$$

The law of $\tilde{Z}$ introduced in Proposition 4.1 being equal to that of Z, we have

$$\text{MSE}^{TLS} = E Z_1^2 + \sum_{i=2}^{q+1} E \left(Z_i^+ \right)^2 = (E\eta_1^4 + q - 1)\omega_0^2 + q/2.$$

B.1 Additional Monte Carlo experiments

The process simulated in Table 3 is an ARCH(1) with $\alpha_{01} \neq 1$. When the fitted model is an ARCH(q) with $q = 1$ then the parameter belongs to the interior of the parameter space. When $\alpha_{01} = 0.2$ the moment condition $E\epsilon_t^8 < \infty$ is satisfied, and the AN given in Propositions 2.1, 2.2 and 2.3 thus hold. As expected from Lemma 3.1, when n is large the LSE, CLSE and TLSE are less efficient than the QMLE, QGLSE, CQGLSE and TQGLSE. In agreement with Proposition 2.3-type results, for n large, the unconstrained, constrained and truncated versions of the LSE are the same, and the same is true for the QGLSE. Note however that for n small and α small, the QMLE can be dominated by other estimators. When $\alpha_{01} = 0.4$ the moment condition $E\epsilon_t^4 < \infty$ is satisfied, so the (C/T)LSE remain consistent and the (C/T)QGLSE remain consistent and AN. As expected, the behavior of the QMLE and QGLSE are very similar. When $\alpha_{01} = 0.6$ or $\alpha_{01} = 0.8$, only the QMLE works well. The behavior of the QGLSE is particularly bad when $\alpha_{01} = 0.8$, but this estimator works nicely when $\alpha_{01} = 0.6$ and $n = 1,000$, which is a little bit surprising because the consistency condition $\alpha_{01} < \sqrt{1/3} = 0.577$ is not satisfied.

When the fitted model is an ARCH(q) with $q > 1$ then the parameter is at the boundary of the parameter space. We are now in a situation where the theoretical comparison of the asymptotic MSE of the different estimators is not available. The QMLE is always the best, except when n and α_{01} are both small, in which case the QMLE can be beaten by truncated or constrained estimators.

The last set of Monte Carlo experiments is aimed to verify that the following inequalities hold for large sample sizes, and see if they hold for smaller sample sizes:

$$\mathrm{MSE}^{CLS} < \mathrm{MSE}^{TLS} < \mathrm{MSE}^{LS}$$

when $q \leq 4$,

$$\mathrm{MSE}^{TLS} < \mathrm{MSE}^{CLS} < \mathrm{MSE}^{LS}$$

when $5 \leq q < (\pi + 1) + \pi/\omega_0^2$, and

$$\mathrm{MSE}^{TLS} < \mathrm{MSE}^{LS} < \mathrm{MSE}^{CLS}$$

when $q > (\pi + 1) + \pi/\omega_0^2$. We known that, asymptotically, $\mathrm{MSE}^{CLS} = \mathrm{MSE}^{QML}$, but we have not computed the empirical MSE^{QML} because we performed simulation experiments involving ARCH(q) models with large q and large sample sizes n, and the QMLE is time consuming in this case, compared to the LSE. From Table 4 it is seen that the LSE is always dominated, either by the CLSE or the TLSE. That the TLSE (which can be considered as a naive estimator, without much theoretical support) can sometimes be more accurate than the CLSE (which is asymptotically equivalent to the MLE in the framework of Tables 2 and 4) is a surprising result of the present paper.

Table 3: Empirical MSE and asymptotic MSE for estimators of an ARCH(q) model when the data generated process is an ARCH(1) model with $\eta_t \sim \mathcal{N}(0,1)$, $\omega_0 = 0.2$ and $\alpha_{01} = 0.2, 0.4, 0.6, 0.8$. The number of replications is $N = 1,\ 000$.

q	n	α_{01}	QMLE	LSE	QGLSE	CLSE	CQGLSE	TLSE	TQGLSE
1	100	0.2	2.21	2.28	2.35	**1.95**	2.02	1.96	2.03
	1000		**2.53**	4.17	**2.53**	4.17	**2.53**	4.17	**2.53**
	100	0.4	**3.76**	4.51	3.99	4.39	3.87	4.39	3.88
	1000		**3.62**	10.47	3.79	10.47	3.79	10.47	3.79
	100	0.6	**5.09**	14.68	26.72	14.53	26.67	14.62	26.68
	1000		**4.66**	37.54	5.43	37.54	5.43	37.54	5.43
	100	0.8	**6.32**	1126.8	31.20	$1.1\ 10^3$	31.00	$1.1\ 10^3$	31.04
	1000		**5.74**	532.1	$3.6\ 10^6$	531.9	$3.6\ 10^6$	531.9	$3.6\ 10^6$
2	100	0.2	2.87	3.79	3.42	**2.31**	2.43	2.65	2.55
	1000		3.05	5.66	3.52	4.78	**3.04**	5.00	3.09
	100	0.4	4.45	6.83	6.45	4.93	**4.31**	5.48	4.39
	1000		**4.10**	16.27	4.68	12.10	4.33	14.00	4.37
	100	0.6	**5.81**	15.39	8.37	12.90	6.63	13.40	6.70
	1000		**5.05**	52.16	9.06	44.11	8.81	46.08	8.85
	100	0.8	**7.07**	363.7	86.59	357.0	$4.5\ 10^5$	360.2	83.85
	1000		**6.03**	412.0	14.48	383.4	13.84	401.5	14.17

Table 4: Empirical MSE and asymptotic MSE for estimators of an ARCH(q) model when the data generated process is a $\mathcal{N}(0, 9)$ iid sequence. The number of replications is $N = 1,\ 000$.

q	n	LSE	CLSE	TLSE
3	100	51.68	**32.95**	49.98
	1 000	50.04	**40.21**	48.55
	10 000	46.23	**38.94**	44.58
	∞	48.00	**41.59**	46.50
5	100	82.01	**52.80**	79.06
	1000	70.33	**65.30**	67.84
	10 000	65.84	65.55	**63.10**
	∞	68.00	71.65	**65.50**
7	100	117.74	**85.94**	113.56
	1000	95.24	104.02	**91.80**
	10 000	87.48	106.21	**83.71**
	∞	88.00	113.16	**84.50**

Corresponding Author's Address:

Jean-Michel Zakoïan
EQUIPPE-GREMARS and CREST
15 Bd Gabriel Péri
92245 Malakoff Cedex France.

E-mail: zakoian@ensae.fr

References

Andrews, D. (1999). Estimation when a parameter is on a boundary. *Econometrica*, *67*, 1341-1384.

Berkes, I., Horváth, L., & Kokoszka, P. (2003). GARCH processes: structure and estimation. *Bernoulli*, *9*, 201-227.

Bose, A., & Mukherjee, K. (2003). Estimating the arch parameters by solving linear equations. *Journal of Time Series Analysis*, *24*, 127-136.

Chernoff, H. (1954). On the distribution of the likelihood ratio. *Annals of Mathematical Statistics*, *54*, 573-578.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of the united kingdom inflation. *Econometrica*, *50*, 987-1007.

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

Francq, C., & Zakoïan, J.-M. (2007). Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications*, *117*, 1265-1284.

Gouriéroux, C. (1997). *Arch models and financial applications*. New York: Springer-Verlag.

Gouriéroux, C., & Monfort, A. (1989). *Statistique et modèles économétriques*. Paris: Economica.

Hamilton, J. (1994). *Time series analysis*. Princeton, New Jersey: Princeton University Press.

Iglesias, E., & Linton, O. (2007). Higher order asymptotic theory when a parameter is on a boundary with an application to GARCH models. *Econometric Theory*, *23*, 1136-1161.

Jensen, S., & Rahbek, A. (2004). Asymptotic normality of the QMLE estimator of ARCH in the nonstationary case. *Econometrica*, *72*, 641-646.

Jordan, H. (2003). *Asymptotic properties of ARCH*(p) *quasi maximum likelihood estimators under weak conditions*. University of Vienna: PhD Thesis.

Straumann, D. (2005). *Estimation in conditionally heteroscedastic time series models*. Lecture Notes in Statistics, Springer.

van der Vaart, A. (1998). *Asymptotic statistics*. Cambridge University Press.