

Least square estimator of the parameter of AR-ARCH model in the presence of missing data

Abdelghani HAMAZ¹ Belkais ALTENDJI²

¹ Laboratory of Pure and Applied Mathematics,
Mouloud Mammeri University of Tizi-Ouzou, Algeria

ghani.hamaz@ummto.dz

²Statistics & Risk Management Department

Katholieke University Leuven, Belgium

belaltendji@yahoo.com

ABSTRACT

In this paper we prove the strong consistency and asymptotic normality of the least square estimator of the parameter of autoregressive models, with the noise sequence driven by a ARCH model, in the presence of missing data. Simulation experiments are presented. .

Keywords: AR-ARCH processes; Least square estimation; Missing data; Asymptotic normality.

2000 Mathematics Subject Classification: 62M20, 60G25

1 Introduction

Autoregressive conditionally heteroscedastic (ARCH) models were introduced by Engle, R.F. (1982). Their GARCH (generalized ARCH) extension is due to Bollerslev, T. (1986). In these models, the key concept is the conditional variance, that is, the variance conditional on the past. In the classical GARCH models, the conditional variance is expressed as a linear function of the squared past values of the series. This particular specification allows to capture the main stylized facts characterizing financial series, as described in Francq, C., Zakoian, J.M. (2010). The presence of ARCH type effects in financial and macroeconomic time series is a well-established fact. The combination of the ARCH specification for the conditional variance and the autoregressive specification for the conditional mean has many appealing features, including a better specification of the forecast variance and the possibility of testing the presence of momentum in stock returns in a well-specified model. Recently AR-ARCH type models have

been used as basic "building blocks" for Markov switching and mixture models see, e.g., Lanne, M., Saikkoun, P. (2003). Asymptotic inference in the AR-ARCH model is also considered in Francq, C., Zakoian, J.M. (2004) and Lange, T. *et al.* (2011).

In many practical situations, the values of a time series are not observed at equally spaced times either because some data are missing or because prudent statistical analysis of the data leads to discard some observed suspect values. Missing data in economic time series can be attributed to certain events in the economic or social environment, such as the stock market crash of October 1987, or the events in USA in September 2001. Likewise, it can be considered that holidays or weekends are days when there are no measurements of the variable under consideration, although economic activity continues as a product of political and social phenomena that will have a direct influence on the next value of the studied index. Therefore, major differences can be observed in the values of certain economic indexes after a weekend or a public holiday. On the other hand, in financial markets, transaction-by-transaction data are examples of high frequency data that occur at irregular times.

This paper is organized as follows. In section 2, the model and some technical tools are presented. Strong consistency and asymptotic normality of an preliminary estimator are derived in section 3. In section 4, we use the preliminary estimator to construct an improved estimator. A simulation study is undertaken to assess the performance of the improved estimators.

2 Preliminaries

In this section, we present the model and discuss the application of a law of Large Numbers to functions of the processes, which is a critical tool in the asymptotic inference. We denote by $(\Omega, \mathcal{A}, P)$ the probability space in which all the random variables considered in the following are defined. Consider the following autoregressive model with ARCH(p) errors:

$$\left\{ \begin{array}{l} X_t = \alpha X_{t-1} + \epsilon_t \\ \epsilon_t = \sigma_t \eta_t \\ \sigma_t^2 = \omega + \sum_{i=1}^p \phi_i \epsilon_{t-i}^2, \quad t \in \mathbf{N} \end{array} \right. \quad (2.1)$$

where $\alpha \in \mathbf{R}$, $\omega > 0$, $\phi_i \geq 0$, and $\{\eta_t, t \in \mathbf{N}\}$ are independent and identically distributed (*iid*) random variables with mean zero and variance one, and X_0 is independent of $\{\epsilon_t, t \in \mathbf{N}\}$. Asymptotic inference in the AR-ARCH model is considered in Francq and Zakoian (2004), Ling and McAleer (2003). The parameter vector is denoted $\theta = (\alpha, \phi, w)'$ with $\phi' = (\phi_1, \dots, \phi_p)'$. We assume that $|\alpha| < 1$. In this case, the unique stationary solution is given by the causal

representation $X_t = \sum_{j=1}^{\infty} \psi_j \epsilon_{t-j}$ where the coefficients $\{\psi_j\}$ are determined by the relation:

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{1}{1 - \alpha z}, \quad |z| \leq 1. \quad (2.2)$$

The numerical calculations of the coefficients ψ_j is discussed in Brockweel, P.J., Davis, R.A. (1991).

The proof of the consistency and asymptotic normality of the least square estimator rests on the ergodicity of the underlying time series. Recall that a stationary sequence is said to be ergodic if it satisfies the strong law of large numbers. A standard example of a stationary ergodic sequence of random elements is an *i.i.d.* sequence. General transformations of ergodic sequences remain ergodic as shown by the following result. Its proof can be found, for instance, in Billingsley, P. (1995).

Theorem 2.1. *If $(Z_t)_{t \in \mathbf{Z}}$ is an ergodic strictly stationary sequence and if $(Y_t)_{t \in \mathbf{Z}}$ is defined by*

$$Y_t = f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots),$$

where f is a measurable function from $\mathbf{R}^{\infty}$ to $\mathbf{R}$, then $(Y_t)_{t \in \mathbf{Z}}$ is also an ergodic strictly stationary sequence.

The ergodic theorem for stationary sequences

Theorem 2.2. *If $(Z_t)_{t \in \mathbf{Z}}$ is strictly stationary and ergodic, if f is measurable and if*

$$E |f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots)| < \infty$$

then

$$\frac{1}{n} \sum_{t=1}^n f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots) \rightarrow E f(\cdots, Z_{t-1}, Z_t, Z_{t+1}, \cdots) \quad a.s.$$

In particular, if $(X_t)_{t \in \mathbf{Z}}$ is the causal stationary solution of the AR(1) equation $X_t = \alpha X_{t-1} + u_t$, $|\alpha| < 1$, (u_t) *iid* $(0, \sigma^2)$, then Theorem 1 shows that $(X_t)_{t \in \mathbf{Z}}$, $(X_{t-1}u_t)_{t \in \mathbf{Z}}$ and $(X_{t-1}^2)_{t \in \mathbf{Z}}$ are also ergodic stationary sequences.

Consider the least-squares estimator $\hat{\alpha}_{LS}$ of the parameter α , that is

$$\hat{\alpha}_{LS} = \operatorname{argmin}_{\alpha} \sum_{t=2}^n (X_t - \alpha X_{t-1})^2.$$

From the first-order condition, we obtain

$$\hat{\alpha}_{LS} = \frac{\sum_{t=2}^n X_{t-1} X_t}{\sum_{t=2}^n X_{t-1}^2}.$$

The ergodic theorem shows that the numerator tends almost surely to $\gamma(1) = \text{Cov}(X_t, X_{t-1}) = \alpha\gamma(0)$ and that the denominator tends to $\gamma(0)$. It follows that $\hat{\alpha}_{LS} \rightarrow \alpha$ almost surely as $n \rightarrow \infty$. Note that this result still holds true when the assumption that u_t is a strong white noise is replaced by the assumption that u_t is a semi-strong white noise, or even by the assumption that u_t is an ergodic and stationary weak white noise.

3 The estimation procedure and asymptotic theory

Let (X_t) be the AR(1)-ARCH(p) time series defined by equation (2.1) where $|\alpha| < 1$. Following Bondon, P. and Bahamonde, N. (2012), a time series with missing observations can be regarded as an amplitude-modulated version of the original time series, *i.e.*

$$X_t^* = A_t X_t, \quad (3.1)$$

where X_t is assumed to be defined for all time, A_t is given by

$$A_t = \begin{cases} 1 & \text{if } X_t \text{ is observed} \\ 0 & \text{if } X_t \text{ is missing} \end{cases} \quad (3.2)$$

and X_t^* represents the actually observed value of X_t , with 0 inserted in the series whenever the value of X_t is missing.

In equation (2.1), if we replace X_t by X_t^* , we obtain

$$X_t^* = \alpha X_{t-1}^* + A_t \epsilon_t \quad 1 \leq t \leq n. \quad (3.3)$$

One can obtain a preliminary least square estimator of α as

$$\hat{\alpha}_{LS,0} = \left[\sum_{t=2}^n X_{t-1}^{*2} \right]^{-1} \sum_{t=2}^n X_t^* X_{t-1}^*. \quad (3.4)$$

3.1 Assumptions

To show the consistency and the normality of the $\hat{\alpha}_{LS,0}$, with the presence of the missing data, we base on the following assumptions:

(H1): Processes (A_t) is an independent and identically distributed processes.

(H2): Processes (A_t) and (X_t) are independent.

3.2 The main results

We are in position to state our first main result.

Theorem 3.1. *Let (X_t) be the AR(1)-ARCH(p) processes defined by (2.1) and (A_t) be a $\{0, 1\}$ processes satisfying **(H1)** and **(H2)** and $A_0 \neq 0$, a.s. Then, as $n \rightarrow \infty$,*

$$\hat{\alpha}_{LS,0} \xrightarrow{a.s.} \alpha. \quad (3.5)$$

Proof. The least square estimator is

$$\begin{aligned} \hat{\alpha}_{LS,0} &= \left(\frac{1}{n} \sum_{t=2}^n X_{t-1}^{*2} \right)^{-1} \left(\frac{1}{n} \sum_{t=2}^n X_t^* X_{t-1}^* \right) \\ &= \frac{1}{n} \left(\sum_{t=2}^n A_{t-1}^2 X_{t-1}^2 \right)^{-1} \left(\frac{1}{n} \sum_{t=2}^n A_t X_t A_{t-1} X_{t-1} \right) \end{aligned} \quad (3.6)$$

Since (X_t) is strictly stationary and ergodic and **(H1)** and **(H2)** are satisfied, it follows from Theorem 1 that combined processes (X_t, A_t) is strictly stationary and ergodic. Then $(A_t^2 X_t^2)$ and $(A_t A_{t-1} X_t X_{t-1})$ are stationary and ergodic. The ergodic theorem (Theorem 2) entails that

$$\begin{aligned} \frac{1}{n} \sum_{t=2}^n A_{t-1}^2 X_{t-1}^2 &\xrightarrow{a.s.} E[A_t^2] E[X_t^2] = E[A_0^2] E[X_0^2] \\ \frac{1}{n} \sum_{t=2}^n A_t A_{t-1} X_t X_{t-1} &\xrightarrow{a.s.} E[A_t A_{t-1}] E[X_{t-1} (\alpha X_{t-1} + \epsilon_t)] \end{aligned}$$

It follows from **(H1)** and the causality of (X_t) that

$$\begin{aligned} E[A_t A_{t-1}] E[X_t (\alpha X_{t-1} + \epsilon_t)] &= E[A_t^2] \alpha E[X_{t-1}^2] + E[A_t^2] E[\epsilon_t X_{t-1}] \\ &= E[A_0^2] \alpha E[X_0^2] \end{aligned} \quad (3.7)$$

Hence, (3.5) is established. $\square$

In order to establish the asymptotic normality of $\hat{\alpha}_{LS,0}$, one often needs to recall the following elementary result about difference martingale. We refer to Billingsley, P. (1995) for details.

Lemma 3.2. *If $(\nu_t, \mathcal{F}_t)_t$ is a stationary and ergodic sequence of square integrable martingale increments such that $\sigma_\nu^2 = \text{Var}(\nu_t) \neq 0$, then*

$$\frac{1}{\sqrt{n}} \sum_{t=2}^n \nu_t \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_\nu^2).$$

Theorem 3.3. *Let (X_t) be the AR(1)-ARCH(p) processes defined by (2.1) and (A_t) be a $\{0, 1\}$ processes satisfying **(H1)** and **(H2)** and $A_0 \neq 0$, a.s. Then, as $n \rightarrow \infty$,*

$$\frac{1}{\sqrt{n}} (\hat{\alpha}_{LS,0} - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N} \left(0, \text{Var}(A_0)^2 \{E[X_t^2]\}^{-2} E[\sigma_t^2 X_{t-1}^2] \right), \quad \text{when } E[\epsilon_t^4] < \infty. \quad (3.8)$$

Proof. Recall that

$$\hat{\alpha}_{LS,0} = \left(\frac{1}{n} \sum_{t=2}^n X_{t-1}^{*2} \right)^{-1} \left(\frac{1}{n} \sum_{t=2}^n X_t^* X_{t-1}^* \right)$$

which is equivalent

$$\sqrt{n}(\hat{\alpha}_{LS,0} - \alpha) = \left(\frac{1}{n} \sum_{t=2}^n X_{t-1}^{*2} \right)^{-1} \frac{1}{\sqrt{n}} \sum_{t=2}^n A_t \epsilon_t X_{t-1}^*.$$

If $E(\epsilon_t^4) < \infty$ the sequence $(A_t A_{t-1} \epsilon_t X_{t-1})_t$ is integrable ergodic stationary difference, with the variance

$$\text{Var}(A_t A_{t-1} \epsilon_t X_{t-1}) = \text{Var}(A_t A_{t-1}) E[\sigma_t^2 X_{t-1}^2].$$

Developing the product

$$\sigma_t X_{t-1}^2 = \left(\omega + \sum_{i=1}^p \epsilon_{t-i}^2 \right) \left(\sum_{i=0}^{\infty} \alpha^i \epsilon_{t-i-1} \right)^2,$$

by Lemma 1

$$\frac{1}{\sqrt{n}} \sum_{t=2}^n A_t A_{t-1} \epsilon_t X_{t-1} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \text{Var}(A_t A_{t-1}) E[\sigma_t^2 X_{t-1}^2]).$$

By the previous arguments, we have

$$\frac{1}{n} \sum_{t=2}^n A_{t-1}^2 X_{t-1}^2 \xrightarrow{a.s.} E[A_0]^2 E[X_0^2].$$

Thus, the conclusion follows. $\square$

Remarks

1. Comparing the asymptotic variances of $\hat{\alpha}_{LS,0}$ and $\hat{\alpha}_{LS}$, when (A_t) is a sequence of independent Bernoulli trials with $q = P(A_t = 1)$, we see that the missing data increases the variance with multiplicative factor $\text{Var}(A_0)^2$.
2. As $q \rightarrow 1$, the asymptotic variance of $\hat{\alpha}_{LS,0}$ tends to the asymptotic variance of the $\hat{\alpha}_{LS}$ with no missing data.
3. The asymptotic variance of $\hat{\alpha}_{LS,0}$ increases as q decreases.

4 Performance of the improved estimator by a simulation study

When the values of a time series are not observed at equally spaced times, several methods of replacement of the missing values were developed in the literature. Pourahmadi (1989) proposed a complete solution of estimating the missing values of a stationary time series

obtained by decomposing it into a prediction plus regression problem. He obtained the best linear interpolator (LI) of the missing value with respect to the mean square error.

An improved estimator $\hat{\alpha}_{LS,I}$ of α can be obtained by using the interpolation method of missing data. If we assume that there is no existing two consecutive missing data, the method consists of replacing a missing value X_k by the well-known formula, see Pourahmadi, M. (1989),

$$\hat{X}'_k = \begin{cases} \frac{\alpha}{1 + \alpha^2}(X_{k-1} + X_{k+1}) & \text{if } X_k \text{ is missing} \\ X_k & \text{if } X_k \text{ is observed.} \end{cases} \quad (4.1)$$

The improved estimator $\hat{\alpha}_{LS,I} = \left[\sum_{t=1}^n \hat{X}'_{t-1} \hat{X}'_{t-1} \right]^{-1} \sum_{t=1}^n \hat{X}'_{t-1}$ is obtained by replacing α by the preliminary estimator $\hat{\alpha}_{LS,0}$ in (4.1).

In order to assess the performance of the improved estimators, we examine the finite sample properties. For (A_t) , we take a sequence of independent Bernoulli trials with $q = P\{A_t = 1\}$. All the experiments are based on $1000 \times N$ replications of AR(1)-ARCH(1) processes defined by (1)-(3) where $\eta_t \sim \mathcal{N}(0, 1)$ and $\theta = (0.5, 0.2, 0.4)'$ and then $\theta = (0.5, 0.6, 0.3)'$. Using the notations of the previous sections, we investigate the impacts of varying the sample size q , among $1 - q = 0\%, 5\%, 10\%$ and $n = 50, 100, 250, 500, 1000, 2000$.

Table I: Sample mean (s.e) of the different estimators when $\theta = (0.5, 0.2, 0.4)'$

	$1 - q$	0%	5%	10%
N=50	$\hat{\alpha}_{LS,0}$	0.478(17.10e-3)	0.453(132.22e-3)	0.428(135.65e-3)
	$\hat{\alpha}_{LS,I}$	0.478(17.10e-3)	0.475(130.34e-3)	0.472(132.48e-3)
N=100	$\hat{\alpha}_{LS,0}$	0.488(08.81e-3)	0.463(95.43e-3)	0.437(96.92e-3)
	$\hat{\alpha}_{LS,I}$	0.488(08.81e-03)	0.487(94.28e-3)	0.483(96.18e-3)
N=250	$\hat{\alpha}_{LS,0}$	0.495(03.55e-3)	0.470(60.70e-3)	0.445(61.29e-3)
	$\hat{\alpha}_{LS,I}$	0.495(03.55e-3)	0.494(60.37e-3)	0.492(63.02e-3)
N=500	$\hat{\alpha}_{LS,0}$	0.497(01.73e-3)	0.472(43.33e-3)	0.448(43.55e-3)
	$\hat{\alpha}_{LS,I}$	0.497(01.73e-3)	0.495(43.64e-3)	0.494(47.60e-3)
N=1000	$\hat{\alpha}_{LS,0}$	0.498(00.80e-3)	0.473(30.01e-3)	0.448(31.19e-3)
	$\hat{\alpha}_{LS,I}$	0.498(00.80e-3)	0.497(31.80e-3)	0.494(37.64e-3)
N=2000	$\hat{\alpha}_{LS,0}$	0.499(00.43e-3)	0.474(21.12e-3)	0.449(21.65e-3)
	$\hat{\alpha}_{LS,I}$	0.499(00.43e-3)	0.497(23.89e-3)	0.494(31.18e-3)

Table II: Sample mean (s.e) of the different estimators when $\theta = (0.5, 0.6, 0.3)'$

	$1 - q$	0%	5%	10%
N=50	$\hat{\alpha}_{MC,0}$	0.467(17.19e-3)	0.444(134.10e-3)	0.426(139.03e-3)
	$\hat{\alpha}_{MC,I}$	0.467(17.19e-3)	0.466(130.12e-3)	0.477(131.65e-3)
N=100	$\hat{\alpha}_{MC,0}$	0.487(09.02e-3)	0.462(97.10e-3)	0.442(95.31e-3)
	$\hat{\alpha}_{MC,I}$	0.487(09.02e-3)	0.485(95.41e-3)	0.491(95.87e-3)
N=250	$\hat{\alpha}_{MC,0}$	0.495(03.43e-3)	0.470(59.92e-3)	0.444(61.81e-3)
	$\hat{\alpha}_{MC,I}$	0.495(03.43e-3)	0.494(58.95e-3)	0.492(62.22e-3)
N=500	$\hat{\alpha}_{MC,0}$	0.498(01.17e-3)	0.473(42.82e-3)	0.447(43.52e-3)
	$\hat{\alpha}_{MC,I}$	0.498(01.17e-3)	0.496(43.52e-3)	0.494(47.72e-3)
N=1000	$\hat{\alpha}_{MC,0}$	0.498(00.86e-3)	0.474(30.04e-3)	0.448(30.33e-3)
	$\hat{\alpha}_{MC,I}$	0.498(00.86e-3)	0.497(31.69e-3)	0.495(63.81e-3)
N=2000	$\hat{\alpha}_{MC,0}$	0.499(00.40e-3)	0.474(21.26e-3)	0.449(21.81e-3)
	$\hat{\alpha}_{MC,I}$	0.499(00.40e-3)	0.498(23.73e-3)	0.495(31.32e-3)

From Table I, it can be observed that the biais of $\alpha_{MC,0}$ increases severely as $1 - q$ increases. On the contrary, the biais of the $\alpha_{MC,I}$ is much less sensitive to the percentage of missing observations.

As seen from Table I, the empirical standard errors of $\alpha_{MC,0}$ are slightly better than the standard error of $\hat{\alpha}_{MC,I}$. Moreover, as n increases, the preciseness of $\hat{\alpha}_{MC,I}$ increases. From Table II, it can be observed that the results are similar for another values of (ϕ, ω) . For these reasons, $\hat{\alpha}_{MC,I}$ has better performances than $\hat{\alpha}_{MC,0}$.

Acknowledgments

The authors are grateful to thank an anonymous referee whose careful and thorough reading gave then the opportunity to improve the quality of the paper

References

- [1] Billingsley, P. (1995). *Probability and Measure*. New York: John Wiley & Sons, Inc.
- [2] Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **31** (3), 307–327.
- [3] Bondon, P., Bahamonde, N. (2012). Least square estimation of ARCH modeMC with missing observations. *J. Time Ser. Anal.*, **33**, 380–391.
- [4] Bose, P. and Mukherjee, K. (2003). Estimating the ARCH parameters by solving linear equations. *J. Time Ser. Anal.*, **24**, 127–136.
- [5] Brockweel, P.J., Davis, R.A. (1991). *Time series: Theory and Methods. 2nd ed.* New York: Springer-Verlag.
- [6] Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of Unit Kingdom inflation. *Econometrica*, **50**, 987–1007.
- [7] Francq, C., Zakoian, J.M. (2004). Maximum Likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, **10**(4), 605–637.
- [8] Francq, C., Zakoian, J.M. (2010). *GARCH modeMC, structure, statistical inference and financial applications*. John Wiley & Sons Ltd.
- [9] Lange, T. et al. (2011). Estimation and asymptotic inference in the AR-ARCH model. *Econometric Reviews*, **30:2**, 129–153.
- [10] Lanne, M., Saikkoun, P. (2003). Modeling the U.S. short-term interest rate by mixture autoregressive processes. *Journal of Financial Econometrics*, **1**, 96–125.
- [11] Ling, S., McAleer, M. (2003). Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory*, **19**, 280–310.
- [12] Pourahmadi, M. (1989). Estimation and interpolation of missing values of stationary time series. *Journal of Time Series Analysis*, **10** (2), 149–169.