LEAST SQUARES ESTIMATION OF ARCH MODELS WITH MISSING OBSERVATIONS

Pascal Bondon
CNRS UMR 8506, Université Paris XI, France.

and

Natalia Bahamonde
Instituto de Estadística, Pontificia Universidad Católica de Valparaíso, Chile.

January 19, 2012

Abstract

A least squares estimator for ARCH models in the presence of missing data is proposed. Strong consistency and asymptotic normality are derived. Monte Carlo simulation results are analysed and an application to real data of a Chilean stock index is reported.

Keywords: ARCH models, missing observations, conditional heteroscedasticity, least squares estimation, martingale central limit theorem.

AMS 2010 subject classification: Primary 62M10, secondary 62F12.

This research was partially supported by grants ECOS/CONICYT C10E03 and FONDECYT 3080009.

1 Introduction

Autoregressive conditionally heteroscedastic (ARCH) type modeling, introduced by Engle (1982), is a very active area of research in econometrics since the past two decades. These models are often used in financial econometrics because their characteristics are close to the observed features of empirical financial data and capture various stylized facts. ARCH models have gained popularity thanks to their easy applicability and flexibility to allow various extensions that better fit the empirical facts of financial data.

The most commonly used estimation procedure for an ARCH model is the quasi maximum likelihood estimation (QMLE) whose asymptotic properties have been extensively studied, see for instance Berkes et al. (2003), Francq and Zakoïan (2004), and Straumann (2005). Alternatively, other estimation methods based on AR and ARMA representations of the squared ARCH process are available. In this way, Giraitis and Robinson (2001) propose a Whittle estimator for some ARCH processes including the Bollerslev (1986) GARCH model, Bose and Mukherjee (2003) advocate a two–stage least squares estimator (LSE) for the Engle (1982) ARCH model, and Kristensen and Linton (2006) investigate an estimator based on fourth-order moments for a GARCH(1,1) model. The estimators proposed by Bose and Mukherjee (2003) and Kristensen and Linton (2006) have closed-form expressions, and therefore have the advantage over the QMLE that they do not require the use of any numerical optimisation procedure. Furthermore, the estimator presented by Bose and Mukherjee (2003) has the same asymptotic efficiency as that of the QMLE, and a finite number of Newton-Raphson iterations using the estimator studied by Kristensen and Linton (2006) as starting point yields also the same asymptotic efficiency as that of the QMLE.

Most of the work on time series assume that the observations are consecutive and equally spaced. In real data sets, however, this is not unusual to find a large number of missing observations or irregularly observed data. A survey of methodologies for handling missing-data problems with emphasis on likelihood methods and the expectation–maximization (EM) algorithm is presented by Little and Rubin (2002). The monograph edited by Parzen (1983) concentrates on the analysis of irregularly observed time series and contains theoretical and practical contributions.

Following Parzen (1963), a time series with missing observations can be regarded as an amplitude modulated version of the original time series, i.e.,

$$X_t^\star = a_t X_t, \tag{1}$$

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

$$a_t = \begin{cases} 1 & \text{if } X_t \text{ is observed,} \\ 0 & \text{if } X_t \text{ is missing,} \end{cases} \tag{2}$$

and $X_t^\star$ represents the actually observed value of X_t, with 0 inserted in the series whenever the value of X_t is missing. In practice, missing values may occur regularly or randomly. Jones (1962) and Parzen (1963) consider the case of periodic sampling where the observed data consist of repeated groups of A consecutive observations followed by B missed observations. On the other hand, Scheinok (1965) and Bloomfield (1970) investigate the case where X_t is observed when a "success" is achieved on a Bernoulli trial. These four references concentrate on non-parametric spectral analysis of time series with missing values, while a parametric approach is considered by Dunsmuir and Robinson (1981b). In the time domain, asymptotic properties of non-parametric estimators of autocovariances and autocorrelations of amplitude modulated time series are established by Dunsmuir and Robinson (1981a) and Yajima and Nishino (1999). These results can be used to build Yule-Walker type estimators for an AR process with missing observations. It is worth noticing that the asymptotic distributions derived by Dunsmuir and Robinson (1981a) and Yajima and Nishino (1999) apply to the case of linear stationary processes (X_t) with conditionally homoscedastic innovations (ϵ_t), i.e., (X_t) satisfies

$$X_t = \sum_{j=0}^{\infty} \beta_j \epsilon_{t-j}, \qquad \sum_{j=0}^{\infty} \beta_j^2 < \infty,$$

where $\mathrm{E}(\epsilon_t \mid \mathcal{F}_{t-1}^{\epsilon}) = 0$ and $\mathrm{E}(\epsilon_t^2 \mid \mathcal{F}_{t-1}^{\epsilon}) = \sigma^2$, $\mathcal{F}_t^{\epsilon}$ being the σ-algebra generated by $(\epsilon_j)_{j \le t}$. Therefore, these results do not apply to an ARCH process, nor to its square. The estimation of ARMA models for non consecutively observed time series has received some attention. Sakai (1980) obtains asymptotic results of a Yule-Walker type estimator for an AR process observed under periodic sampling. Dunsmuir (1981) considers also the case of periodic sampling and derives the strong consistency and asymptotic normality of Gaussian MLE of an ARMA model. From a practical point of view, Jones (1980) shows that the Gaussian likelihood of an ARMA process with missing values can be easily calculated using a state space representation. From a theoretical viewpoint, consistency and asymptotic normality of MLE seem difficult to establish under general sampling schemes. For a Gaussian time series, Dunsmuir (1983) derives asymptotic normality of the one-step Newton-Raphson estimator initialized with a consistent estimator. Without the assumption of normality, Reinsel and Wincek (1987) derive asymptotic normality of the one-step Newton-Raphson estimator and a weighted LSE, both initialized with a consistent estimator, for a first order AR process. Lastly, under mild assumptions on the sampling schemes, Shin and Sarkar (1995) establish the weak consistency and asymptotic normality of Gaussian MLE of a first order AR process.

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.

The purpose of this paper is to contribute to the treatment of missing observations in financial time series. To the best of our knowledge, the problem of estimating an ARCH model in the presence of missing values has not yet been addressed. Both practical and theoretical aspects deserve attention since the state space model and Kalman filter approach in Jones (1980) cannot apply to compute the likelihood of an ARCH process with missing values, and the estimation results mentioned above do not include the case of conditionally heteroscedastic time series. Here, we propose a LSE for ARCH time series affected by random missing values inspired by the estimator proposed by Bose and Mukherjee (2003) and which has a closed-form expression. Strong consistency and asymptotic normality are derived in Section 2. In Section 3, the behaviour of the estimator for finite samples is analysed via Monte Carlo simulations, and is compared to a Yule-Walker estimator and to some estimators based on a complete data set obtained after filling the missing observations by imputation procedures. An application to real data of a Chilean stock index is also reported.

2 Model and Estimator

We denote by $(\Omega, \mathcal{A}, \mathsf{P})$ the probability space on which all the random variables considered in the following are defined. Let (X_t) be the ARCH(p) time series defined by the equation

$$X_t = \sigma_t(\alpha)\epsilon_t, \tag{3}$$

where (ϵ_t) is a sequence of iid random variables with $\mathrm{E}\epsilon_0 = 0$ and $\mathrm{E}\epsilon_0^2 = 1$, and $(\sigma_t(\alpha))$ is a non negative process satisfying the difference equation

$$\sigma_t^2(\alpha) = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2. \tag{4}$$

Here $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_p)'$ where the parameters α_i are nonnegative and α_p is positive. According to (Giraitis et al., 2000, Theorem 2.1), there exists a strictly stationary and ergodic solution to (3)–(4) with $\mathrm{E}X_0^2 < \infty$ when $\sum_{i=1}^{p} \alpha_i < 1$. Moreover, such a solution with finite second-order moment is a unique nonanticipative solution in the sense that X_t is independent of the $(\epsilon_s)_{s>t}$. If $\alpha_0 = 0$, this unique solution is $X_t = 0$. In all the following, we assume that $\alpha_0 > 0$ and

$\sum_{i=1}^{p} \alpha_i < 1$. Conditions for the existence of higher-order moments of (X_t) are given by Milhøj (1985) and He and Teräsvirta (1999).

Let $Y_t = X_t^2$,

$$Z_t = (1, Y_{t-1}, \ldots, Y_{t-p})',$$

and $\eta_t = \epsilon_t^2 - 1$. Then (4) is equivalent to $\sigma_t^2(\alpha) = Z_t'\alpha$ and squaring (3) gives

$$Y_t = Z_t'\alpha + \sigma_t^2(\alpha)\eta_t, \tag{5}$$

where $\mathrm{E}(\sigma_t^2(\alpha)\eta_t \mid \mathcal{F}_{t-1}^X) = \sigma_t^2(\alpha)\,\mathrm{E}(\eta_t \mid \mathcal{F}_{t-1}^X) = 0$, $\mathcal{F}_t^X$ being the σ-algebra generated by $(X_j)_{j\le t}$. Equation (5) shows that (Y_t) follows an autoregressive model of order p whose innovation is a martingale difference sequence. However, it is worth noticing that the innovation is conditionally heteroscedastic. Using (5), Bose and Mukherjee (2003) have proposed a two-stage LSE of parameter α. Our purpose is to generalize their estimator to the case where the series (X_t) is only partially observed.

We express observed data $(X_t^\star)_{1\le t\le n}$ by (1) where (a_t) represents the state of observation defined by (2). Throughout we make the two following assumptions :

(A1) Process (a_t) is strictly stationary and weakly mixing.

(A2) Processes (a_t) and (X_t) are independent.

Let $A_t = \prod_{i=0}^{p} a_{t-i}$, and

$$Y_t^\star = A_t Y_t = A_t Z_t'\alpha + A_t\sigma_t^2(\alpha)\eta_t = Z_t^{\star\prime}\alpha + \sigma_t^{\star 2}(\alpha)\eta_t, \qquad p+1 \le t \le n. \tag{6}$$

Equation (6) coincides with (5) when $p+1$ consecutive data are observed. Ignoring the randomness of $\sigma_t^{\star 2}(\alpha)$ and also the presence of α in it, one can obtain a preliminary LSE of α as

$$\widehat{\alpha}_{\mathrm{pr}} = \left[\sum_{t=p+1}^{n} Z_t^\star Z_t^{\star\prime}\right]^{-1} \sum_{t=p+1}^{n} Z_t^\star Y_t^\star. \tag{7}$$

The consistency and asymptotic normality of $\widehat{\alpha}_{\mathrm{pr}}$ are given in the following lemma.

Lemma 1. Let (X_t) be the ARCH(p) process defined by (3)–(4) and (a_t) be a $\{0,1\}$ process satisfying (A1), (A2) and $A_0 \neq 0$ a.s. Then, as $n \to \infty$,

(i) $\widehat{\alpha}_{\mathrm{pr}} \xrightarrow{a.s.} \alpha$ when $\mathrm{E}X_0^4 < \infty$,

(ii) $n^{1/2}(\widehat{\alpha}_{\mathrm{pr}} - \alpha) \xrightarrow{d} \mathrm{N}\left[0, (\mathrm{E}A_0)^{-1}\,\mathrm{Var}(\epsilon_0^2)\{\mathrm{E}Z_0Z_0'\}^{-1}\,\mathrm{E}\{(\alpha'Z_0)^2 Z_0Z_0'\}\{\mathrm{E}Z_0Z_0'\}^{-1}\right]$ when $\mathrm{E}X_0^8 < \infty$.

Like Bose and Mukherjee (2003), we use $\widehat{\alpha}_{\rm pr}$ to construct an improved estimator $\widehat{\alpha}$ of α as follows. Dividing (6) by $\sigma_t^2(\alpha)$, we get

$$\frac{Y_t^\star}{\sigma_t^2(\alpha)} = \frac{Z_t^{\star\prime}}{\sigma_t^2(\alpha)}\alpha + A_t\eta_t.$$

In this expression, the errors $A_t\eta_t$ are homoscedastic. Replacing α by $\widehat{\alpha}_{\rm pr}$ in $\sigma_t^2(\alpha)$, we estimate α by the ordinary LSE $\widehat{\alpha}$ given by

$$\widehat{\alpha} = \left[\sum_{t=p+1}^{n} \frac{Z_t^\star Z_t^{\star\prime}}{\sigma_t^4(\widehat{\alpha}_{\rm pr})}\right]^{-1} \sum_{t=p+1}^{n} \frac{Z_t^\star Y_t^\star}{\sigma_t^4(\widehat{\alpha}_{\rm pr})}. \tag{8}$$

Observe that each term in (8) can be calculated since when $A_t = 0$, $Z_t^\star = 0$, and when $A_t = 1$, data $X_{t-1}, \ldots, X_{t-p}$ are observed and $\sigma_t^2(\widehat{\alpha}_{\rm pr})$ is obtained by (4). The following theorem gives the consistency and asymptotic normality of $\widehat{\alpha}$.

Theorem 1. Let (X_t) be the ARCH(p) process defined by (3)–(4) where $\alpha_i > 0$ for $i = 0, \ldots, p$, and (a_t) be a $\{0,1\}$ process satisfying (A1), (A2) and $A_0 \neq 0$ a.s. Then, as $n \to \infty$,

(i) $\widehat{\alpha} \xrightarrow{a.s.} \alpha$ when $\mathrm{E}X_0^6 < \infty$,

(ii) $n^{1/2}(\widehat{\alpha} - \alpha) \xrightarrow{d} \mathrm{N}\left[0, (\mathrm{E}A_0)^{-1}\,\mathrm{Var}(\epsilon_0^2)\left\{\mathrm{E}\{(\alpha' Z_0)^{-2} Z_0 Z_0'\}\right\}^{-1}\right]$ when $\mathrm{E}X_0^8 < \infty$.

Remark 1. Comparing the asymptotic variances of $\widehat{\alpha}_{\rm pr}$ and $\widehat{\alpha}$ with the corresponding ones given by Bose and Mukherjee (2003), we see that the missing data increases the variances with the multiplicative factor $(\mathrm{E}A_0)^{-1}$. In particular, the asymptotic variance of $\widehat{\alpha}$ is equal to the asymptotic variance of the QMLE with no missing data multiplied by $(\mathrm{E}A_0)^{-1}$. When (a_t) is a sequence of independent Bernoulli trails with $q = \mathsf{P}\{a_t = 1\}$, $(\mathrm{E}A_0)^{-1} = q^{-(p+1)}$. From this expression we draw the two following conclusions :

(i) As $q \to 1$, the asymptotic variance of $\widehat{\alpha}$ tends to the asymptotic variance of the QMLE with no missing data.

(ii) The asymptotic variance of $\widehat{\alpha}$ increases as q decreases.

Remark 2. Another estimator of α can be obtained by applying the Yule-Walker method proposed by Dunsmuir and Robinson (1981a) to the autoregressive model (5). Indeed, it follows from (5) that

$$\alpha_0 = \left(1 - \sum_{i=1}^{p} \alpha_i\right) \mathrm{E}Y_0, \tag{9}$$

and $(\alpha_1, \ldots, \alpha_p)$ is given by the unique solution to the Yule-Walker equation

$$\begin{pmatrix} \gamma_Y(0) & \gamma_Y(1) & \cdots & \gamma_Y(p-1) \\ \gamma_Y(1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \gamma_Y(1) \\ \gamma_Y(p-1) & \cdots & \gamma_Y(1) & \gamma_Y(0) \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix} = \begin{pmatrix} \gamma_Y(1) \\ \gamma_Y(2) \\ \vdots \\ \gamma_Y(p) \end{pmatrix}, \tag{10}$$

where $\gamma_Y(k) = \mathrm{Cov}(Y_0, Y_k)$. Estimates $(\widehat{\alpha}_{\mathrm{yw},1}, \ldots, \widehat{\alpha}_{\mathrm{yw},p})$ can be obtained by replacing $\gamma_Y(k)$ by an appropriate estimate in (10). Following Dunsmuir and Robinson (1981a), we estimate $\gamma_Y(k)$ by

$$\widehat{\gamma}_Y(k) = \frac{\sum_{t=1}^{n-k} a_t a_{t+k}(Y_t - \widehat{\mu}_Y)(Y_{t+k} - \widehat{\mu}_Y)}{\sum_{t=1}^{n-k} a_t a_{t+k}},$$

where $\widehat{\mu}_Y = \sum_{t=1}^{n} a_t Y_t / \sum_{t=1}^{n} a_t$. Then $\widehat{\alpha}_{\mathrm{yw},0}$ is obtained by replacing in the right hand side of (9) α_i by $\widehat{\alpha}_{\mathrm{yw},i}$ and $\mathrm{E}Y_0$ by $\widehat{\mu}_Y$. The finite sample behavior of the Yule-Walker estimator $\widehat{\alpha}_{\mathrm{yw}} = (\widehat{\alpha}_{\mathrm{yw},0}, \widehat{\alpha}_{\mathrm{yw},1}, \ldots, \widehat{\alpha}_{\mathrm{yw},p})'$ is presented in Section 3.1.

Remark 3. Lemma 1 and Theorem 1 hold for any amplitude modulating sequence (a_t) satisfying (A1), (A2), $A_0 \neq 0$ a.s. and $\mathrm{E}A_0^4 < \infty$. In this case, Lemma 1(ii) and Theorem 1(ii) should be replaced by (15) and (27), respectively.

3 Numerical results

3.1 Simulations

Here we analyse the finite sample behavior of estimator $\widehat{\alpha}$ in (8) by Monte Carlo simulations. For (a_t), we take a sequence of independent Bernoulli trails with $q = \mathsf{P}\{a_t = 1\}$. Different sample sizes n and probabilities $1 - q$ of missing observation are considered. All the experiments are based on $1000 \times n$ replications of an ARCH(1) process defined by (3)–(4) where (ϵ_t) is a sequence of iid zero-mean Gaussian random variables with unit variance and $\alpha = (0.3, 0.5)'$. The sample mean and standard error (s.e.) of $\widehat{\alpha}$ are compared to those of $\widehat{\alpha}_{\mathrm{yw}}$ and some estimators based on a complete data set obtained after filling the missing observations by imputation procedures. Different imputation methods can be applied and we consider the following estimators.

- $\widehat{\alpha}_{\mathrm{b}}$, $\widehat{\alpha}_{\mathrm{m}}$, $\widehat{\alpha}_{\mathrm{p}}$ and $\widehat{\alpha}_{\mathrm{p}}^{\mathrm{n}}$ are the two–stage LSE of Bose and Mukherjee (2003) obtained respectively when each missing value is (a) deleted and the observed data are bound together, (b) replaced by the sample mean of the observed values, (c) replaced by the first non missing value prior to it, (d) interpolated by the half sum of the first non missing value prior to it and the first non missing value next to it.

- $\widetilde{\alpha}_{\mathrm{b}}$, $\widetilde{\alpha}_{\mathrm{m}}$, $\widetilde{\alpha}_{\mathrm{p}}$ and $\widetilde{\alpha}_{\mathrm{p}}^{\mathrm{n}}$ are defined like $\widehat{\alpha}_{\mathrm{b}}$, $\widehat{\alpha}_{\mathrm{m}}$, $\widehat{\alpha}_{\mathrm{p}}$ and $\widehat{\alpha}_{\mathrm{p}}^{\mathrm{n}}$, respectively, except that the QMLE is calculated in place of the LSE of Bose and Mukherjee (2003).

The tables below show the results for $n = 500, 1000, 2000$ and $1-q = 0\%, 5\%, 10\%, 15\%, 20\%$.

$1-q$	0%	5%	10%	15%	20%
$\widehat{\alpha}$	0.305 (3.161e-2) 0.483 (9.618e-2)	0.305 (3.895e-2) 0.483 (4.728e-1)	0.305 (4.422e-2) 0.480 (3.792e-1)	0.306 (8.414e-2) 0.477 (5.255e-1)	0.307 (5.794e-2) 0.475 (5.984e-1)
$\widehat{\alpha}_{\mathrm{yw}}$	0.356 (6.816e-2) 0.394 (1.121e-1)	0.357 (7.115e-2) 0.392 (1.202e-1)	0.358 (7.706e-2) 0.390 (1.290e-1)	0.359 (8.437e-2) 0.388 (1.378e-1)	0.359 (9.480e-2) 0.386 (1.478e-1)
$\widehat{\alpha}_{\mathrm{b}}$	0.305 (3.161e-2) 0.483 (9.618e-2)	0.312 (3.400e-2) 0.470 (1.375e-1)	0.321 (3.659e-2) 0.455 (1.089e-1)	0.330 (3.997e-2) 0.439 (1.645e-1)	0.340 (4.272e-2) 0.423 (1.145e-1)
$\widetilde{\alpha}_{\mathrm{b}}$	0.302 (3.011e-2) 0.491 (8.944e-2)	0.310 (3.223e-2) 0.478 (9.317e-2)	0.318 (3.455e-2) 0.464 (9.712e-2)	0.327 (3.716e-2) 0.449 (1.017e-1)	0.336 (3.999e-2) 0.434 (1.064e-1)
$\widehat{\alpha}_{\mathrm{m}}$	0.305 (3.161e-2) 0.483 (9.618e-2)	0.318 (4.133e-2) 0.423 (1.054e-1)	0.325 (4.760e-2) 0.372 (7.926e-1)	0.328 (5.199e-2) 0.329 (1.184e-1)	0.325 (5.297e-2) 0.289 (4.015e-1)
$\widetilde{\alpha}_{\mathrm{m}}$	0.302 (3.011e-2) 0.491 (8.944e-2)	0.317 (3.976e-2) 0.427 (9.019e-2)	0.324 (4.532e-2) 0.377 (8.939e-2)	0.326 (4.897e-2) 0.335 (8.854e-2)	0.323 (5.026e-2) 0.299 (8.772e-2)
$\widehat{\alpha}_{\mathrm{p}}$	0.305 (3.161e-2) 0.483 (9.618e-2)	0.296 (3.191e-2) 0.498 (1.018e-1)	0.288 (3.324e-2) 0.512 (1.315e-1)	0.279 (3.273e-2) 0.527 (9.721e-2)	0.270 (3.326e-2) 0.543 (9.686e-2)
$\widetilde{\alpha}_{\mathrm{p}}$	0.302 (3.011e-2) 0.491 (8.944e-2)	0.294 (3.061e-2) 0.504 (8.851e-2)	0.286 (3.113e-2) 0.519 (8.766e-2)	0.277 (3.170e-2) 0.534 (8.715e-2)	0.268 (3.230e-2) 0.550 (8.676e-2)
$\widehat{\alpha}_{\mathrm{p}}^{\mathrm{n}}$	0.305 (3.161e-2) 0.483 (9.618e-2)	0.294 (3.180e-2) 0.492 (1.199e-1)	0.282 (3.226e-2) 0.501 (9.841e-2)	0.270 (3.222e-2) 0.512 (9.741e-2)	0.258 (3.216e-2) 0.524 (1.396e-1)
$\widetilde{\alpha}_{\mathrm{p}}^{\mathrm{n}}$	0.302 (3.011e-2) 0.491 (8.944e-2)	0.291 (3.058e-2) 0.500 (9.054e-2)	0.279 (3.110e-2) 0.512 (9.179e-2)	0.267 (3.113e-2) 0.525 (9.309e-2)	0.254 (3.123e-2) 0.540 (9.447e-2)

Table 1: Sample mean (s.e.) of the different estimators when $n = 500$.

We see that $\widehat{\alpha}$, $\widehat{\alpha}_{\mathrm{yw}}$, $\widehat{\alpha}_{\mathrm{b}}$, $\widetilde{\alpha}_{\mathrm{b}}$, $\widehat{\alpha}_{\mathrm{m}}$ and $\widetilde{\alpha}_{\mathrm{m}}$ tend to overestimate α_0 and underestimate α_1, while $\widehat{\alpha}_{\mathrm{p}}$, $\widetilde{\alpha}_{\mathrm{p}}$, $\widehat{\alpha}_{\mathrm{p}}^{\mathrm{n}}$ and $\widetilde{\alpha}_{\mathrm{p}}^{\mathrm{n}}$ tend to underestimate α_0 and overestimate α_1. For all estimators, the s.e. of the estimator of α_1 is larger than the s.e. of the estimator of α_0.

Not surprisingly, $\widehat{\alpha}_{\mathrm{yw}}$ is not a good estimator of α even for a complete data set. For a given imputation procedure, the two–stage LSE and the QMLE behave similarly. For all imputation based estimation methods, the bias increases severely (in absolute value) as $1-q$ increases, while the s.e. remains almost the same.

$1-q$	0%	5%	10%	15%	20%
$\widehat{\alpha}$	0.302 (2.192e-2) 0.492 (6.630e-2)	0.302 (2.925e-2) 0.492 (4.897e-1)	0.303 (3.670e-2) 0.491 (7.868e-1)	0.303 (4.445e-2) 0.489 (4.640e-1)	0.304 (5.144e-2) 0.487 (7.792e-1)
$\widehat{\alpha}_{\mathrm{yw}}$	0.346 (5.433e-2) 0.415 (9.763e-2)	0.347 (5.976e-2) 0.413 (1.053e-1)	0.348 (6.489e-2) 0.412 (1.136e-1)	0.349 (7.286e-2) 0.410 (1.223e-1)	0.349 (7.747e-2) 0.408 (1.316e-1)
$\widehat{\alpha}_{\mathrm{b}}$	0.302 (2.192e-2) 0.492 (6.630e-2)	0.310 (2.382e-2) 0.478 (7.156e-2)	0.318 (2.546e-2) 0.464 (7.489e-2)	0.327 (2.748e-2) 0.449 (7.880e-2)	0.336 (2.991e-2) 0.434 (1.003e-1)
$\widetilde{\alpha}_{\mathrm{b}}$	0.301 (2.114e-2) 0.495 (6.283e-2)	0.309 (2.264e-2) 0.482 (6.567e-2)	0.317 (2.429e-2) 0.469 (6.852e-2)	0.326 (2.616e-2) 0.454 (7.167e-2)	0.335 (2.816e-2) 0.439 (7.503e-2)
$\widehat{\alpha}_{\mathrm{m}}$	0.302 (2.192e-2) 0.492 (6.630e-2)	0.317 (2.999e-2) 0.430 (7.087e-2)	0.324 (3.384e-2) 0.377 (8.374e-2)	0.327 (3.989e-2) 0.332 (7.733e-2)	0.325 (3.833e-2) 0.294 (7.618e-2)
$\widetilde{\alpha}_{\mathrm{m}}$	0.301 (2.114e-2) 0.495 (6.283e-2)	0.317 (2.906e-2) 0.430 (6.391e-2)	0.324 (3.233e-2) 0.380 (6.339e-2)	0.326 (3.818e-2) 0.338 (6.279e-2)	0.323 (3.637e-2) 0.302 (6.194e-2)
$\widehat{\alpha}_{\mathrm{p}}$	0.302 (2.192e-2) 0.492 (6.630e-2)	0.294 (2.232e-2) 0.505 (7.344e-2)	0.286 (2.244e-2) 0.519 (6.590e-2)	0.277 (2.297e-2) 0.534 (6.893e-2)	0.268 (2.367e-2) 0.550 (7.774e-2)
$\widetilde{\alpha}_{\mathrm{p}}$	0.301 (2.114e-2) 0.495 (6.283e-2)	0.293 (2.151e-2) 0.509 (6.228e-2)	0.285 (2.191e-2) 0.523 (6.194e-2)	0.276 (2.239e-2) 0.537 (6.153e-2)	0.267 (2.284e-2) 0.553 (6.118e-2)
$\widehat{\alpha}_{\mathrm{p}}^{\mathrm{n}}$	0.302 (2.192e-2) 0.492 (6.630e-2)	0.291 (2.225e-2) 0.500 (9.561e-2)	0.280 (2.250e-2) 0.509 (6.714e-2)	0.268 (2.241e-2) 0.520 (7.911e-2)	0.256 (2.274e-2) 0.533 (7.335e-2)
$\widetilde{\alpha}_{\mathrm{p}}^{\mathrm{n}}$	0.301 (2.114e-2) 0.495 (6.283e-2)	0.290 (2.150e-2) 0.505 (6.381e-2)	0.278 (2.174e-2) 0.516 (6.469e-2)	0.266 (2.193e-2) 0.529 (6.562e-2)	0.253 (2.204e-2) 0.544 (6.651e-2)

Table 2: Sample mean (s.e.) of the different estimators when $n = 1000$.

$1-q$	0%	5%	10%	15%	20%
$\widehat{\alpha}$	0.301 (1.525e-2)	0.301 (2.433e-2)	0.301 (3.348e-2)	0.302 (4.721e-2)	0.302 (4.102e-2)
	0.496 (4.903e-2)	0.496 (3.211e-1)	0.494 (3.504e-1)	0.494 (6.005e-1)	0.495 (9.730e-1)
$\widehat{\alpha}_{\mathrm{yw}}$	0.338 (4.612e-2)	0.339 (5.154e-2)	0.340 (5.644e-2)	0.340 (6.210e-2)	0.341 (7.228e-2)
	0.431 (8.581e-2)	0.430 (9.313e-2)	0.428 (1.009e-1)	0.427 (1.090e-1)	0.426 (1.177e-1)
$\widehat{\alpha}_{\mathrm{b}}$	0.301 (1.525e-2)	0.309 (1.644e-2)	0.317 (1.774e-2)	0.326 (1.922e-2)	0.335 (2.079e-2)
	0.496 (4.903e-2)	0.483 (4.873e-2)	0.469 (5.126e-2)	0.454 (5.582e-2)	0.439 (5.610e-2)
$\widetilde{\alpha}_{\mathrm{b}}$	0.301 (1.491e-2)	0.308 (1.597e-2)	0.316 (1.713e-2)	0.325 (1.844e-2)	0.334 (1.986e-2)
	0.498 (4.441e-2)	0.485 (4.630e-2)	0.471 (4.839e-2)	0.457 (5.064e-2)	0.442 (5.295e-2)
$\widehat{\alpha}_{\mathrm{m}}$	0.301 (1.525e-2)	0.316 (2.134e-2)	0.324 (2.447e-2)	0.327 (2.669e-2)	0.325 (3.225e-2)
	0.496 (4.903e-2)	0.432 (8.287e-2)	0.379 (5.236e-2)	0.333 (5.455e-2)	0.294 (6.168e-2)
$\widetilde{\alpha}_{\mathrm{m}}$	0.301 (1.491e-2)	0.316 (2.082e-2)	0.324 (2.346e-2)	0.326 (2.535e-2)	0.323 (3.111e-2)
	0.498 (4.441e-2)	0.432 (4.530e-2)	0.381 (4.508e-2)	0.339 (4.454e-2)	0.303 (4.399e-2)
$\widehat{\alpha}_{\mathrm{p}}$	0.301 (1.525e-2)	0.293 (1.555e-2)	0.285 (1.582e-2)	0.277 (1.610e-2)	0.267 (1.632e-2)
	0.496 (4.903e-2)	0.509 (4.701e-2)	0.523 (4.580e-2)	0.537 (4.563e-2)	0.553 (4.471e-2)
$\widetilde{\alpha}_{\mathrm{p}}$	0.301 (1.491e-2)	0.293 (1.518e-2)	0.285 (1.547e-2)	0.276 (1.580e-2)	0.267 (1.612e-2)
	0.498 (4.441e-2)	0.511 (4.402e-2)	0.524 (4.374e-2)	0.539 (4.345e-2)	0.554 (4.334e-2)
$\widehat{\alpha}_{\mathrm{p}}^{\mathrm{n}}$	0.301 (1.525e-2)	0.290 (1.584e-2)	0.279 (1.567e-2)	0.267 (1.577e-2)	0.255 (1.580e-2)
	0.496 (4.903e-2)	0.504 (4.755e-2)	0.514 (4.834e-2)	0.525 (4.799e-2)	0.537 (4.744e-2)
$\widetilde{\alpha}_{\mathrm{p}}^{\mathrm{n}}$	0.301 (1.491e-2)	0.289 (1.516e-2)	0.278 (1.536e-2)	0.265 (1.547e-2)	0.253 (1.555e-2)
	0.498 (4.441e-2)	0.507 (4.502e-2)	0.518 (4.569e-2)	0.531 (4.636e-2)	0.545 (4.695e-2)

Table 3: Sample mean (s.e.) of the different estimators when $n = 2000$.

On the contrary, the bias of $\widehat{\alpha}$ is much less sensitive to the percentage of missing observations and is much smaller (in absolute value) than the bias of any other estimator, while the s.e. of the different estimators have the same magnitude (except the s.e. of $\widehat{\alpha}_1$ and $\widehat{\alpha}_{\text{yw},1}$ which are larger). Moreover, as n increases, the preciseness of $\widehat{\alpha}$ increases. For these reasons, $\widehat{\alpha}$ presents better performances than the other estimators considered here.

3.2 A real time series example

We consider the daily results (P_t) of Chile's IPSA stock index from January 3, 1994 through December 30, 2004. The IPSA is an index composed of the 40 most heavily traded stocks and revised quarterly. We ignore the unequal spacing of the data resulting from the five-day working week. The series contains 2869 days of which 127 are public holidays on which the Chilean stock exchange is closed. We consider that the data of these 127 days are missing. The percentage of missing data in the original time series (P_t) is therefore equal to 4.427%. Our data of interest are the daily log returns (R_t) defined by $R_t = \ln P_t - \ln P_{t-1}$. The percentage of missing data in the series (R_t) is equal to 8.438%. Figure 1 shows the 100 last values of R_t of which five blocks of two consecutive data are missing. The Q-Q plot shows departure from the Gaussian assumption with a typical value of the kurtosis for a financial time series. The sample autocorrelation function (ACF) obtained with the estimator proposed by Parzen (1961) and further analysed by Dunsmuir and Robinson (1981a), and the sample partial autocorrelation function (PACF) suggest that (R_t) follows an MA(1) or an AR(1) process.

Fitting all the ARMA(p, q) models with $1 \leq p + q \leq 2$ to (R_t), the AR(1) model has the smallest AIC value. We retain this model which is

$$R_t = 0.220 R_{t-1} + X_t,$$

where the s.e. of the AR parameter is 1.947e-2, showing that this parameter is significant at the 5% level. The percentage of missing data in the residuals (X_t) is the same as in (R_t), i.e., 8.438%. The kurtosis of (X_t) is 8.958. Figure 2 shows the sample ACF and PACF of some functions of (X_t). The sample ACF of (X_t) shows that (X_t) behaves like a white noise process, while the sample ACF of $(|X_t|)$ and (X_t^2) clearly show the existence of conditional heteroscedasticity. The sample PACF of (X_t^2) indicates that an ARCH(p) model with $1 \leq p \leq 10$ might be appropriate for (X_t).

We choose to fit an ARCH(1) model and an ARCH(3) model to (X_t) using (8), and we compare the abilities of these two models to predict the volatility of (X_t). For $p = 1$, the model is

$$\sigma_t^2 = 8.675\text{e-}5 + 0.332 X_{t-1}^2,$$

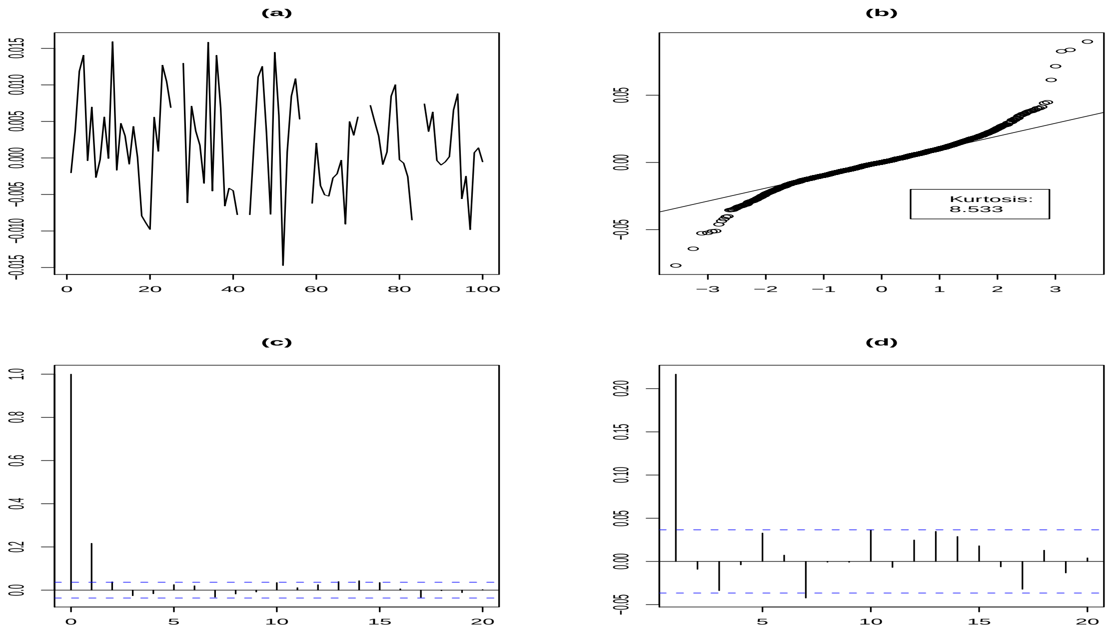

Figure 1: Daily log returns for IPSA stock from January 3, 1994 to December 30, 2004; (a) 100 last data, (b) Normal Q-Q plot, (c) Sample ACF, (d) Sample PACF.

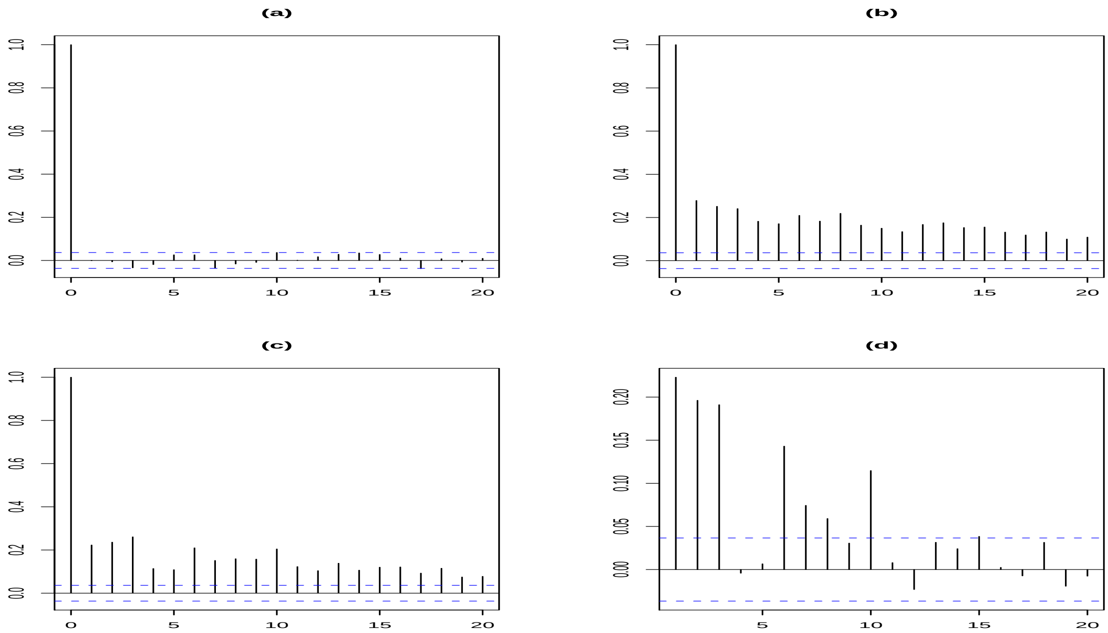

Figure 2: Innovations series from AR(1) fit on daily log returns for IPSA stock; (a), (b) and (c) show the sample ACF of the residuals, their absolute values, and squared values, respectively, and (d) is the sample PACF of the squared residuals.

where the s.e. of the parameters are 5.663e-6 and 5.597e-2, respectively. Then the parameters are significant at the 5% level. For $p = 3$, the model is

$$\sigma_t^2 = 5.837\text{e-}5 + 0.203X_{t-1}^2 + 0.176X_{t-2}^2 + 0.181X_{t-3}^2,$$

where the s.e. of the parameters are 4.814e-6, 3.993e-2, 4.074e-2, 6.465e-2, respectively. Again, the parameters are significant at the 5% level. Figures 3 and 4 plot the residuals of the ARCH(1) model and the ARCH(3) model, respectively, with the sample ACF of some functions of these residuals. If the sample ACF show that the residuals behave like white noise processes in both cases, the sample ACF of the squared and absolute values of the residuals exhibits more heteroscedasticity in the case of the ARCH(1) model than in the case of the ARCH(3) model. The kurtosis of the residuals of the ARCH(1) model is 6.010, while the residuals of the ARCH(3) model are less leptokurtic with a kurtosis equal to 4.646. Figures 5 and 6 plot the 90% (based on normal quantiles) one step prediction intervals for the 100 last data of the daily log returns for IPSA stock by fitting an AR(1)-ARCH(1) model and an AR(1)-ARCH(3) model, respectively. In both figures, the constant prediction intervals obtained with a non-heteroscedastic AR(1) model are plotted and we show clearly that the heteroscedastic models perform better. Moreover, the prediction intervals are more precise with an AR(1)-ARCH(3) model than with an AR(1)-ARCH(1) model.

Appendix

Proof of Lemma 1. From (6) and (7),

$$\widehat{\alpha}_{\text{pr}} - \alpha = \left[n^{-1} \sum_{t=p+1}^{n} A_t^2 Z_t Z_t' \right]^{-1} n^{-1} \sum_{t=p+1}^{n} U_t, \tag{11}$$

where $U_t = A_t^2(\alpha' Z_t) Z_t \eta_t$. Since (X_t) is strictly stationary and ergodic and (A1) and (A2) are satisfied, it follows from Hannan (1973) that the combined process (X_t, a_t) is strictly stationary and ergodic. Then $(A_t^2 Z_t Z_t')$ and (U_t) are strictly stationary and ergodic. When $\mathrm{E}X_0^4 < \infty$ and $\mathrm{E}A_0^2 < \infty$, the elements of $\mathrm{E}|A_t^2 Z_t Z_t'|$ and $\mathrm{E}|U_t|$ are finite, and it follows from the pointwise ergodic theorem for stationary sequences (Stout, 1974, Theorem 3.5.7) that

$$n^{-1} \sum_{t=p+1}^{n} A_t^2 Z_t Z_t' \xrightarrow{a.s.} \mathrm{E}A_0^2\, \mathrm{E}Z_0 Z_0', \tag{12}$$

$$n^{-1} \sum_{t=p+1}^{n} U_t \xrightarrow{a.s.} \mathrm{E}A_0^2\, \mathrm{E}\{(\alpha' Z_0) Z_0\}\, \mathrm{E}\eta_0 = 0, \tag{13}$$

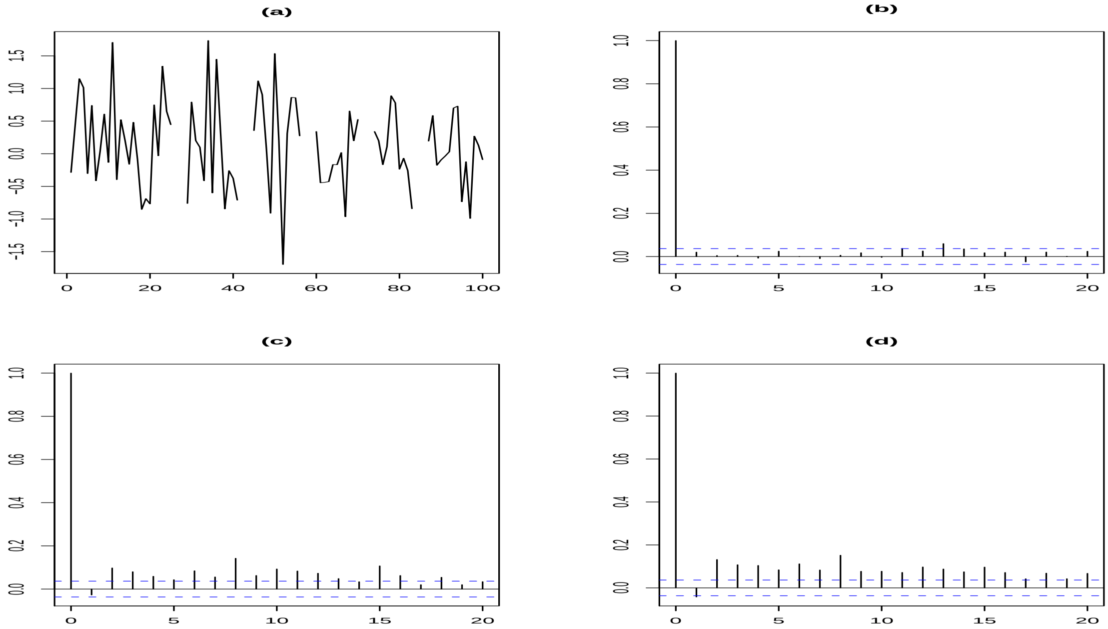

Figure 3: Residuals series from ARCH(1) fit on the innovations series from AR(1) fit on daily log returns for IPSA stock; (a) 100 last data, (b), (c) and (d) show the sample ACF of the residuals, their squared values, and absolute values, respectively.

as $n \to \infty$. When $A_0 \neq 0$ a.s., $\mathrm{E}A_0^2\, \mathrm{E}Z_0 Z_0'$ is invertible, and (i) follows from (11), (12) and (13). Let $\mathcal{F}_t^U$ be the σ-algebra generated by the random variables $(U_j)_{j\leq t}$ and $\mathcal{G}_t$ be the σ-algebra generated by the random variables $\{(X_j)_{j\leq t}, (a_j)_{j\leq t+1}\}$. We have $\mathrm{E}(U_t \mid \mathcal{G}_{t-1}) = A_t^2(\alpha' Z_t) Z_t\, \mathrm{E}(\eta_t \mid \mathcal{G}_{t-1})$. It results from (A2) that the random variables $\{\epsilon_t, (X_j)_{j\leq t-1}\}$ and $(a_j)_{j\leq t}$ are mutually independent. Furthermore, ϵ_t is independent of $(X_j)_{j\leq t-1}$. Then ϵ_t is independent of $\mathcal{G}_{t-1}$ and $\mathrm{E}(\eta_t \mid \mathcal{G}_{t-1}) = \mathrm{E}\eta_t = 0$, which implies that $\mathrm{E}(U_t \mid \mathcal{G}_{t-1}) = 0$. Since $\mathcal{F}_t^U \subset \mathcal{G}_t$, $\mathrm{E}(U_t \mid \mathcal{F}_{t-1}^U) = \mathrm{E}(\mathrm{E}(U_t \mid \mathcal{G}_{t-1}) \mid \mathcal{F}_{t-1}^U) = 0$, and (U_t) is a martingale difference sequence. According to the central limit theorem of Billingsley (1961) and Ibragimov (1963) for stationary ergodic martingale differences, if the elements of $W = \mathrm{E}U_0 U_0'$ are finite, then

$$n^{-1/2} \sum_{t=p+1}^{n} U_t \xrightarrow{d} \mathrm{N}(0, W) \tag{14}$$

as $n \to \infty$. When $\mathrm{E}X_0^8 < \infty$, the elements of $\mathrm{E}\{(\alpha' Z_0)^2 Z_0 Z_0'\}$ are finite. Moreover, for any $m > 0$, if $\mathrm{E}|X_0|^m < \infty$, then $\mathrm{E}|\epsilon_0|^m < \infty$. Indeed, according to (4), $\sigma_t^2(\alpha) \geq \alpha_0 > 0$ and it follows from (3) that $|\epsilon_t|^m \leq \alpha_0^{-m/2} |X_t|^m$. Therefore, when $\mathrm{E}X_0^8 < \infty$ and $\mathrm{E}A_0^4 < \infty$, the elements of W are finite and $W = \mathrm{E}A_0^4\, \mathrm{E}\{(\alpha' Z_0)^2 Z_0 Z_0'\}\, \mathrm{Var}(\epsilon_0^2)$. Combining (12), (14) and (11),

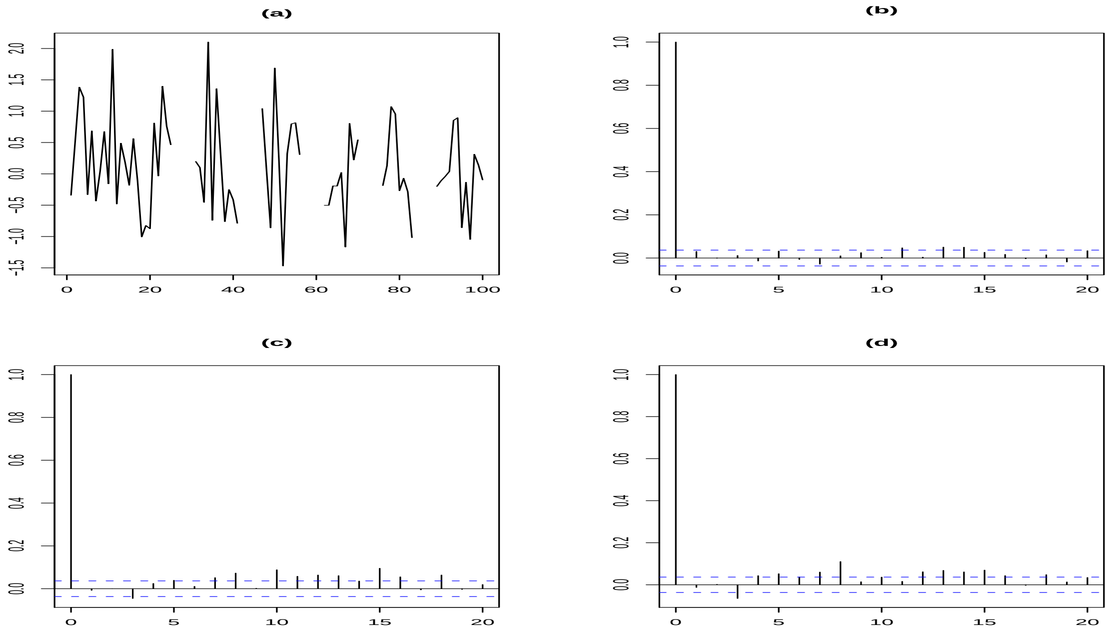

Figure 4: Residuals series from ARCH(3) fit on the innovations series from AR(1) fit on daily log returns for IPSA stock; (a) 100 last data, (b), (c) and (d) show the sample ACF of the residuals, their squared values, and absolute values, respectively.

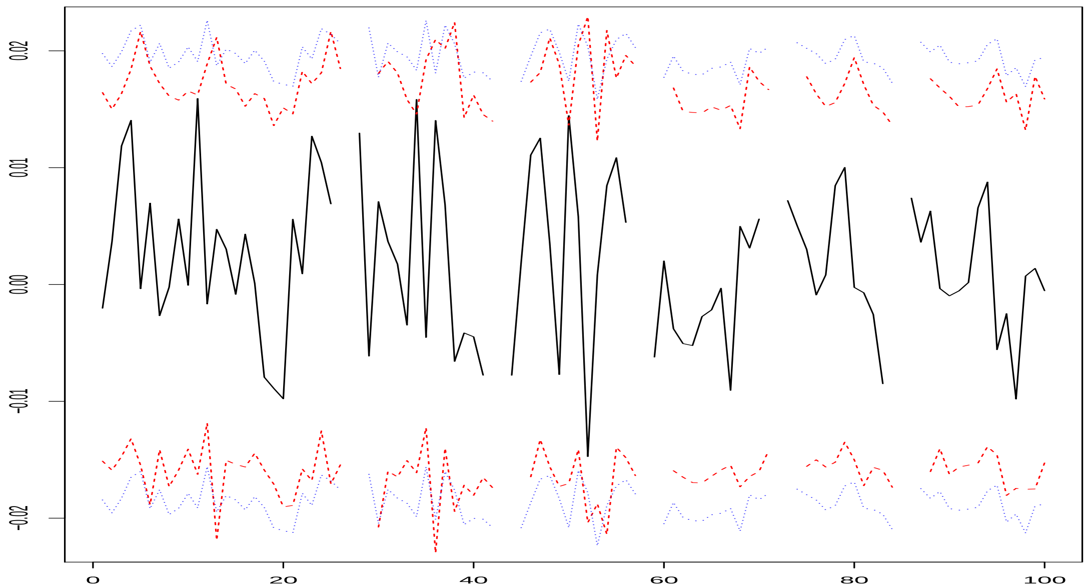

Figure 5: One step prediction intervals for the 100 last data of the daily log returns for IPSA stock; AR(1) model : dotted lines; AR(1)-ARCH(1) model : dashed lines.

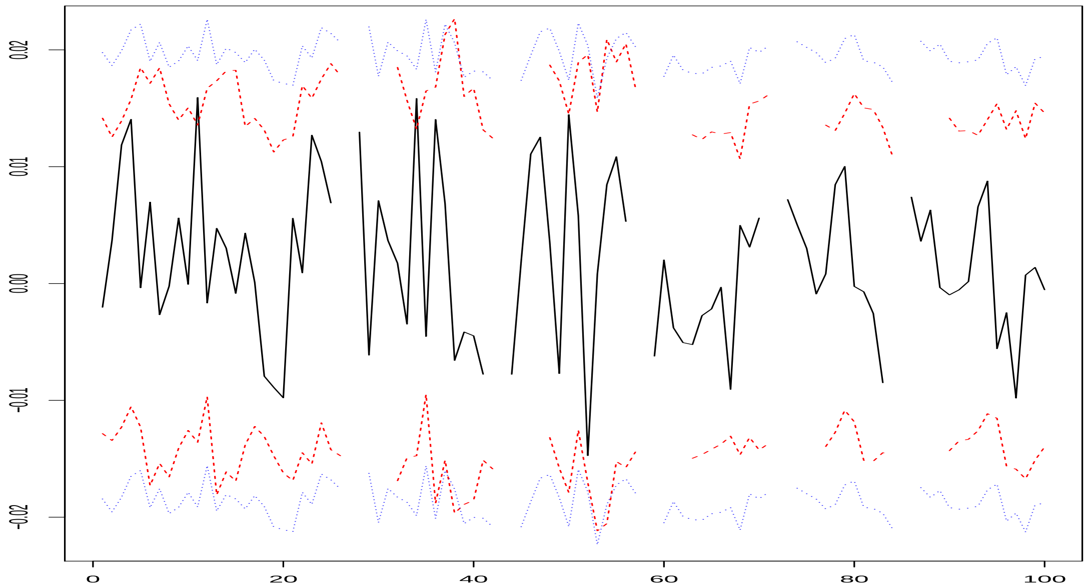

Figure 6: One step prediction intervals for the 100 last data of the daily log returns for IPSA stock; AR(1) model : dotted lines; AR(1)-ARCH(3) model : dashed lines.

we get

$$n^{1/2}(\widehat{\alpha}_{\text{pr}} - \alpha) \xrightarrow{d} \mathrm{N}\left[0, \frac{\mathrm{E}A_0^4}{\{\mathrm{E}A_0^2\}^2} \operatorname{Var}(\epsilon_0^2)\{\mathrm{E}Z_0Z_0'\}^{-1}\,\mathrm{E}\{(\alpha' Z_0)^2 Z_0 Z_0'\}\{\mathrm{E}Z_0Z_0'\}^{-1}\right], \tag{15}$$

which is equivalent to (ii) when (a_t) is a $\{0,1\}$ process.

Proof of Theorem 1. From (6) and (8),

$$\widehat{\alpha} - \alpha = \left[n^{-1} \sum_{t=p+1}^{n} \frac{A_t^2 Z_t Z_t'}{\sigma_t^4(\widehat{\alpha}_{\text{pr}})}\right]^{-1} n^{-1} \sum_{t=p+1}^{n} \frac{A_t^2 \sigma_t^2(\alpha) Z_t \eta_t}{\sigma_t^4(\widehat{\alpha}_{\text{pr}})}. \tag{16}$$

To establish (i), we show that

$$n^{-1} \sum_{t=p+1}^{n} \left[\frac{1}{\sigma_t^4(\widehat{\alpha}_{\text{pr}})} - \frac{1}{\sigma_t^4(\alpha)}\right] A_t^2 Z_t Z_t' \xrightarrow{a.s.} 0, \tag{17}$$

$$n^{-1} \sum_{t=p+1}^{n} \left[\frac{1}{\sigma_t^4(\widehat{\alpha}_{\text{pr}})} - \frac{1}{\sigma_t^4(\alpha)}\right] A_t^2 \sigma_t^2(\alpha) Z_t \eta_t \xrightarrow{a.s.} 0, \tag{18}$$

$$n^{-1} \sum_{t=p+1}^{n} \frac{A_t^2 Z_t Z_t'}{\sigma_t^4(\alpha)} \xrightarrow{a.s.} \mathrm{E}A_0^2\,\mathrm{E}\{(\alpha' Z_0)^{-2} Z_0 Z_0'\}, \tag{19}$$

$$n^{-1} \sum_{t=p+1}^{n} \frac{A_t^2 Z_t \eta_t}{\sigma_t^2(\alpha)} \xrightarrow{a.s.} \mathrm{E}A_0^2\,\mathrm{E}\{(\alpha' Z_0)^{-1} Z_0\}\,\mathrm{E}\eta_0 = 0, \tag{20}$$

as $n \to \infty$. Following Bose and Mukherjee (2003), we use the Taylor expansions for $u, v > 0$,

$$\frac{1}{u^2} - \frac{1}{v^2} = -2\frac{(u-v)}{\chi^3}, \tag{21}$$

$$= -2\frac{(u-v)}{v^3} + 3\frac{(u-v)^2}{\xi^4}, \tag{22}$$

where χ and ξ satisfy $0 < 1/\chi, 1/\xi \leq 1/v + 1/u$. Assume that $\mathrm{E}X_0^4 < \infty$ and $0 < \mathrm{E}A_0^2 < \infty$. According to Lemma 1(i), $\widehat{\alpha}_{\mathrm{pr}} \xrightarrow{a.s.} \alpha$, i.e., there exists an event $E \in \mathcal{A}$ with probability one such that $\widehat{\alpha}_{\mathrm{pr}} \to \alpha$ as $n \to \infty$ for all outcomes in E. In the following, consider a fixed outcome in E. For all $\nu > 0$, there exists n_0 such that $|\widehat{\alpha}_{\mathrm{pr},i} - \alpha_i| \leq \nu$ for all $n \geq n_0$ and for all $i \in \{0, \ldots, p\}$. Take $\nu = \frac{1}{2}\min_{i\in\{0,\ldots,p\}} \alpha_i$. Then $0 < \frac{1}{2}\alpha_i \leq \alpha_i - \nu \leq \widehat{\alpha}_{\mathrm{pr},i}$ for all $n \geq n_0$ and for all $i \in \{0, \ldots, p\}$. Therefore, for all $n \geq n_0$ and for all $t \in \mathbb{Z}$, $0 < \frac{1}{2}\alpha_0 \leq \frac{1}{2}Z_t'\alpha \leq Z_t'\widehat{\alpha}_{\mathrm{pr}}$, i.e., $0 < \frac{1}{2}\alpha_0 \leq \frac{1}{2}\sigma_t^2(\alpha) \leq \sigma_t^2(\widehat{\alpha}_{\mathrm{pr}})$. Using (21) and (22) with $u = \sigma_t^2(\widehat{\alpha}_{\mathrm{pr}})$ and $v = \sigma_t^2(\alpha)$, we deduce that the corresponding intermediate points $\chi_{t,n}$ and $\xi_{t,n}$ satisfy $0 < 1/\chi_{t,n}, 1/\xi_{t,n} \leq (1 + v/u)/v \leq 3/\alpha_0$ for all $n \geq n_0$ and for all $t \in \mathbb{Z}$. Let $\delta_{n,i}$ and $Z_{t,i}$ for $i = 0, \ldots, p$ be the components of $\widehat{\alpha}_{\mathrm{pr}} - \alpha$ and Z_t, respectively. For proving (17), observe that for all $n \geq n_0$,

$$n^{-1}\sum_{t=p+1}^{n}\left[\frac{1}{\sigma_t^4(\widehat{\alpha}_{\mathrm{pr}})} - \frac{1}{\sigma_t^4(\alpha)}\right]A_t^2 Z_t Z_t' = -2\sum_{i=0}^{p}\delta_{n,i}\, n^{-1}\sum_{t=p+1}^{n}\frac{A_t^2 Z_{t,i} Z_t Z_t'}{\chi_{t,n}^3}, \tag{23}$$

and a typical (k,l) entry inside the summation with respect to "t" in the right hand side of (23) satisfies

$$0 \leq \frac{A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}}{\chi_{t,n}^3} \leq \left(\frac{3}{\alpha_0}\right)^3 A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}.$$

Process $(A_t^2 Z_{t,i} Z_{t,k} Z_{t,l})$ is strictly stationary and ergodic with a finite mean when $\mathrm{E}X_0^6 < \infty$ and $\mathrm{E}A_0^2 < \infty$. Thus, $n^{-1}\sum_{t=p+1}^{n} A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}$ converges a.s., and $n^{-1}\sum_{t=p+1}^{n}\frac{A_t^2 Z_{t,i} Z_{t,k} Z_{t,l}}{\chi_{t,n}^3}$ is bounded a.s. Since $\delta_{n,i} \xrightarrow{a.s.} 0$ for $i = 0, \ldots, p$, (17) follows from (23). To prove (18), note that for all $n \geq n_0$,

$$n^{-1}\sum_{t=p+1}^{n}\left[\frac{1}{\sigma_t^4(\widehat{\alpha}_{\mathrm{pr}})} - \frac{1}{\sigma_t^4(\alpha)}\right]A_t^2\sigma_t^2(\alpha) Z_t \eta_t = -2\sum_{i=0}^{p}\delta_{n,i}\, n^{-1}\sum_{t=p+1}^{n}\frac{A_t^2\sigma_t^2(\alpha) Z_{t,i} Z_t \eta_t}{\chi_{t,n}^3}, \tag{24}$$

where

$$n^{-1}\left|\sum_{t=p+1}^{n}\frac{A_t^2\sigma_t^2(\alpha) Z_{t,i} Z_{t,k}\eta_t}{\chi_{t,n}^3}\right| \leq \left(\frac{3}{\alpha_0}\right)^3 n^{-1}\sum_{t=p+1}^{n} A_t^2\sigma_t^2(\alpha) Z_{t,i} Z_{t,k}|\eta_t|, \tag{25}$$

and process $(A_t^2\sigma_t^2(\alpha) Z_{t,i} Z_{t,k}|\eta_t|)$ is strictly stationary and ergodic with a finite mean when $\mathrm{E}X_0^6 < \infty$ and $\mathrm{E}A_0^2 < \infty$. Then the right hand side of (25) converges a.s., and the left hand side is bounded a.s. Thus, (18) follows from (24). (19) and (20) follow from the pointwise ergodic theorem for the stationary sequences $(A_t^2\sigma_t^{-4}(\alpha)Z_tZ_t')$ and $(A_t^2\sigma_t^{-2}(\alpha)Z_t\eta_t)$, respectively, when $\mathrm{E}X_0^4 < \infty$ and $\mathrm{E}A_0^2 < \infty$. To prove (ii), we show that

$$n^{-1/2}\sum_{t=p+1}^{n}\frac{A_t^2\sigma_t^2(\alpha)Z_t\eta_t}{\sigma_t^4(\widehat{\alpha}_{\mathrm{pr}})}\xrightarrow{d}\mathrm{N}[0,\mathrm{E}A_0^4\,\mathrm{Var}(\epsilon_0^2)\,\mathrm{E}\{(\alpha'Z_0)^{-2}Z_0Z_0'\}] \tag{26}$$

when $\mathrm{E}X_0^8 < \infty$ and $0 < \mathrm{E}A_0^4 < \infty$. Then it will follow from (16), (17), (19) and (26) that

$$n^{1/2}(\widehat{\alpha}-\alpha)\xrightarrow{d}\mathrm{N}\left[0,\frac{\mathrm{E}A_0^4}{\{\mathrm{E}A_0^2\}^2}\,\mathrm{Var}(\epsilon_0^2)\left\{\mathrm{E}\{(\alpha'Z_0)^{-2}Z_0Z_0'\}\right\}^{-1}\right], \tag{27}$$

which is equivalent to (ii) when (a_t) is a $\{0,1\}$ process. To establish (26), we show that

$$n^{-1/2}\sum_{t=p+1}^{n}\left[\frac{1}{\sigma_t^4(\widehat{\alpha}_{\mathrm{pr}})}-\frac{1}{\sigma_t^4(\alpha)}\right]A_t^2\sigma_t^2(\alpha)Z_t\eta_t\xrightarrow{p}0, \tag{28}$$

$$n^{-1/2}\sum_{t=p+1}^{n}\frac{A_t^2Z_t\eta_t}{\sigma_t^2(\alpha)}\xrightarrow{d}\mathrm{N}[0,\mathrm{E}A_0^4\,\mathrm{Var}(\epsilon_0^2)\,\mathrm{E}\{(\alpha'Z_0)^{-2}Z_0Z_0'\}]. \tag{29}$$

For proving (28) note that by (22), for all $n \geq n_0$,

$$\begin{aligned}n^{-1/2}\sum_{t=p+1}^{n}\left[\frac{1}{\sigma_t^4(\widehat{\alpha}_{\mathrm{pr}})}-\frac{1}{\sigma_t^4(\alpha)}\right]A_t^2\sigma_t^2(\alpha)Z_t\eta_t &= -2\sum_{i=0}^{p}(n^{1/2}\delta_{n,i})\,n^{-1}\sum_{t=p+1}^{n}\frac{A_t^2Z_{t,i}Z_t\eta_t}{\sigma_t^4(\alpha)}\\ +3\sum_{i,j=0}^{p}(n^{1/2}\delta_{n,i})(n^{1/2}\delta_{n,j})\,n^{-3/2}\sum_{t=p+1}^{n}\frac{A_t^2\sigma_t^2(\alpha)Z_{t,i}Z_{t,j}Z_t\eta_t}{\xi_{t,n}^4} &= -2T_1+3T_2,\ \text{say.}\end{aligned}$$

According to Lemma 1(ii), $n^{1/2}\delta_{n,i} = O_p(1)$. The pointwise ergodic theorem for the stationary sequence $\frac{A_t^2Z_{t,i}Z_t\eta_t}{\sigma_t^4(\alpha)}$ implies that $n^{-1}\sum_{t=p+1}^{n}\frac{A_t^2Z_{t,i}Z_t\eta_t}{\sigma_t^4(\alpha)}\xrightarrow{a.s.}\mathrm{E}A_0^2\,\mathrm{E}\{(\alpha'Z_0)^{-2}Z_{0,i}Z_0\}\,\mathrm{E}\eta_0 = 0$, and then $T_1\xrightarrow{p}0$. Now, for all $n \geq n_0$,

$$n^{-1}\left|\sum_{t=p+1}^{n}\frac{A_t^2\sigma_t^2(\alpha)Z_{t,i}Z_{t,j}Z_{t,k}\eta_t}{\xi_{t,n}^4}\right| \leq \left(\frac{3}{\alpha_0}\right)^4 n^{-1}\sum_{t=p+1}^{n}A_t^2\sigma_t^2(\alpha)Z_{t,i}Z_{t,j}Z_{t,k}|\eta_t|, \tag{30}$$

and process $(A_t^2\sigma_t^2(\alpha)Z_{t,i}Z_{t,j}Z_{t,k}|\eta_t|)$ is strictly stationary and ergodic with a finite mean when $\mathrm{E}X_0^8 < \infty$ and $\mathrm{E}A_0^2 < \infty$. Then the right hand side of (30) converges a.s., and the left hand side is bounded a.s. Thus, $n^{-3/2}\sum_{t=p+1}^{n}\frac{A_t^2\sigma_t^2(\alpha)Z_{t,i}Z_{t,j}Z_t\eta_t}{\xi_{t,n}^4}\xrightarrow{a.s.}0$, and $T_2\xrightarrow{p}0$. Finally, as in the proof of Lemma 1 for process (U_t), we see that process $(A_t^2\sigma_t^{-2}(\alpha)Z_t\eta_t)$ is a martingale difference sequence, and (29) follows from the central limit theorem for stationary ergodic martingale differences.

References

I. Berkes, L. Horváth, and P. Kokoszka. GARCH processes: structure and estimation. *Bernoulli*, 9(2):201–227, 2003.

P. Billingsley. The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.*, 12:788–792, 1961.

P. Bloomfield. Spectral analysis with randomly missing observations. *J. Roy. Statist. Soc. Ser. B*, 32:369–380, 1970.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *J. Econometrics*, 31(3): 307–327, 1986.

A. Bose and K. Mukherjee. Estimating the ARCH parameters by solving linear equations. *J. Time Ser. Anal.*, 24(2):127–136, 2003.

W. Dunsmuir. Estimation for stationary time series when data are irregularly spaced or missing. In David F. Findley, editor, *Applied time series analysis II*, pages 609–649. Academic Press, New York, 1981.

W. Dunsmuir. A central limit theorem for estimation in Gaussian stationary time series observed at unequally spaced times. *Stochastic Process. Appl.*, 14(3):279–295, 1983.

W. Dunsmuir and P. M. Robinson. Asymptotic theory for time series containing missing and amplitude modulated observations. *Sankhyā Ser. A*, 43(3):260–281, 1981a.

W. Dunsmuir and P. M. Robinson. Parametric estimators for stationary time series with missing observations. *Adv. in Appl. Probab.*, 13(1):129–146, 1981b.

R. F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

C. Francq and J.-M. Zakoïan. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637, 2004.

L. Giraitis and P. M. Robinson. Whittle estimation of ARCH models. *Econometric Theory*, 17 (3):608–631, 2001.

L. Giraitis, P. Kokoszka, and R. Leipus. Stationary ARCH models: dependence structure and central limit theorem. *Econometric Theory*, 16(1):3–22, 2000.

E. J. Hannan. Central limit theorems for time series regression. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 26:157–170, 1973.

C. He and T. Teräsvirta. Fourth moment structure of the GARCH(p, q) process. *Econometric Theory*, 15(6):824–846, 1999.

I. A. Ibragimov. A central limit theorem for a class of dependent random variables. *Theory Probab. Appl.*, 8:83–89, 1963.

R. H. Jones. Spectral analysis with regularly missed observations. *Ann. Math. Statist.*, 33: 455–461, 1962.

R. H. Jones. Maximum likelihood fitting of ARMA models to time series with missing observations. *Technometrics*, 22(3):389–395, 1980.

D. Kristensen and O Linton. A closed-form estimator for the GARCH$(1,1)$ model. *Econometric Theory*, 22(2):323–337, 2006.

R. J. A. Little and D. B. Rubin. *Statistical analysis with missing data*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2002.

A. Milhøj. The moment structure of ARCH processes. *Scand. J. Statist.*, 12(4):281–292, 1985.

E. Parzen. Spectral analysis of asymptotically stationary time series. *Bull. Inst. Internat. Statist.*, 33rd Session, Paris, 1961.

E. Parzen. On spectral analysis with missing observations and amplitude modulation. *Sankhyā Ser. A*, 25:383–392, 1963.

E. Parzen, editor. *Proceedings of Time Series Analysis of Irregularly Observed Data*, volume 25 of *Lect. Notes in Statistics*. Springer Verlag, New York, 1983.

G. C. Reinsel and M. A. Wincek. Asymptotic distribution of parameter estimators for nonconsecutively observed time series. *Biometrika*, 74(1):115–124, 1987.

H. Sakai. Fitting autoregression with regularly missed observations. *Ann. Inst. Statist. Math.*, 32(3):393–400, 1980.

P. A. Scheinok. Spectral analysis with randomly missed observations: The binomial case. *Ann. Math. Statist.*, 36:971–977, 1965.

D. W. Shin and S. Sarkar. Estimation for stationary AR(1) models with nonconsecutively observed samples. *Sankhyā Ser. A*, 57(2):287–298, 1995.

W. F. Stout. *Almost sure convergence*. Academic Press, New York, 1974.

D. Straumann. *Estimation in conditionally heteroscedastic time series models*, volume 181 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin, 2005.

Y. Yajima and H. Nishino. Estimation of the autocorrelation function of a stationary time series with missing observations. *Sankhyā Ser. A*, 61(2):189–207, 1999.