



HAL
open science

Testing for long memory in panel random-coefficient AR(1) data

Remigijus Leipus, Anne Philippe, Vytautė Pilipauskaitė, Donatas Surgailis

► **To cite this version:**

Remigijus Leipus, Anne Philippe, Vytautė Pilipauskaitė, Donatas Surgailis. Testing for long memory in panel random-coefficient AR(1) data. 2017. hal-01622201v1

HAL Id: hal-01622201

<https://hal.science/hal-01622201v1>

Preprint submitted on 25 Oct 2017 (v1), last revised 13 Sep 2019 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Testing for long memory in panel random-coefficient AR(1) data

Remigijus Leipus¹, Anne Philippe², Vytautė Pilipauskaitė¹, Donatas Surgailis¹

October 25, 2017

¹Vilnius University, Faculty of Mathematics and Informatics, Naugarduko 24, 03225 Vilnius, Lithuania

²Université de Nantes, Laboratoire de Mathématiques Jean Leray, 44322 Nantes Cedex 3, France

Abstract

It is well-known that random-coefficient AR(1) process can have long memory depending on the index β of the tail distribution function of the random coefficient, if it is a regularly varying function at unity. We discuss estimation of β from panel data comprising N random-coefficient AR(1) series, each of length T . The estimator of β is constructed as a version of the tail index estimator of Goldie and Smith (1987) applied to sample lag 1 autocorrelations of individual time series. Its asymptotic normality is derived under certain conditions on N , T and some parameters of our statistical model. Based on this result, we construct a statistical procedure to test if the panel random-coefficient AR(1) data exhibit long memory. A simulation study illustrates finite-sample performance of the introduced estimator and testing procedure.

Keywords: random-coefficient autoregression; panel data; tail index estimator; long memory process

2010 MSC: 62G32, 62M10.

1 Introduction

Dynamic panels comprising observations taken at regular time intervals for the same individuals such as households, firms, etc. in a large heterogeneous population, are often described by simple autoregressive models with random parameters. One of the simplest and the most studied models for individual evolution is the random-coefficient AR(1) (RCAR(1)) process

$$X(t) = aX(t-1) + \zeta(t), \quad t \in \mathbb{Z}, \quad (1)$$

where the innovations $\{\zeta(t), t \in \mathbb{Z}\}$ are independent identically distributed (i.i.d.) random variables (r.v.s) with $E\zeta(0) = 0$, $E\zeta^2(0) < \infty$ and the autoregressive coefficient $a \in (0, 1)$ is a r.v., independent of $\{\zeta(t), t \in \mathbb{Z}\}$. If the distribution of a is sufficiently dense near unity, then statistical properties of the individual evolution in (1) and the corresponding panel can differ greatly from those in the case of fixed $a \in (0, 1)$. To be more specific, assume that the AR coefficient a has a density function $g(x)$, $x \in (0, 1)$, satisfying

$$g(x) \sim g_1(1-x)^{\beta-1}, \quad x \rightarrow 1-, \quad (2)$$

for some $\beta > 1$ and $g_1 > 0$. Then a stationary solution of RCAR(1) equation (1) has the following autocovariance function

$$EX(0)X(t) = E \frac{a^t}{1-a^2} \sim \frac{g_1}{2} \Gamma(\beta-1) t^{-(\beta-1)}, \quad t \rightarrow \infty,$$

and exhibits long memory in the sense that $\sum_{t \in \mathbb{Z}} |\text{Cov}(X(0), X(t))| = \infty$ for $\beta \in (1, 2)$. The same long memory property applies to the contemporaneous aggregate of N independent individual evolutions $\{X_i(t), t \in \mathbb{Z}\}$, $i = 1, \dots, N$, of (1) and its Gaussian limit arising as $N \rightarrow \infty$. For Beta distributed a , these facts were first uncovered by Granger [9] and later extended to more general distributions and/or RCAR equations in Gonçalves and Gouriéroux [8], Zaffaroni [23], Celov et al. [3], Oppenheim and Viano [16], Puplinskaitė and Surgailis [20], Philippe et al. [17] and other works. Assumption (2) and the parameter β play a crucial role for statistical (dependence) properties of the panel $\{X_i(t), t = 1, \dots, T, i = 1, \dots, N\}$ as N and T increase, possibly at different rates. Particularly, Pilipauskaitė and Surgailis [18] proved that for $\beta \in (1, 2)$ the distribution of the normalized sample mean $\sum_{i=1}^N \sum_{t=1}^T X_i(t)$ is asymptotically normal if $N/T^\beta \rightarrow \infty$ and β -stable if $N/T^\beta \rightarrow 0$ (in the ‘intermediate’ case $N/T^\beta \rightarrow c \in (0, \infty)$ this limit distribution is more complicated and given by an integral with respect to a certain Poisson random measure). A similar but non-identical trichotomy of the limit distribution of the sample mean for a panel comprising RCAR(1) series driven by common innovations is proved in Pilipauskaitė and Surgailis [19].

In the above context, a natural statistical problem concerns inference on the distribution of the random AR coefficient a , e.g., its cumulative distribution function (c.d.f.) G or the parameter β in (2). Leipus et al. [10], Celov et al. [4] estimated the density g using sample autocovariances of the limit aggregated process. For estimating parameters of G , Robinson [21] used the method of moments. He proved asymptotic normality of the estimators for moments of G based on the panel RCAR(1) data as $N \rightarrow \infty$ for fixed T , under the condition $E(1 - a^2)^{-2} < \infty$ which does not allow for long memory in $\{X(t), t \in \mathbb{Z}\}$. For the parameters of Beta distribution, Beran et al. [1] discussed maximum likelihood estimation based on (truncated) sample lag 1 autocorrelations computed from $\{X_i(1), \dots, X_i(T)\}$, $i = 1, \dots, N$, and proved consistency and asymptotic normality of the introduced estimator as $N, T \rightarrow \infty$. In nonparametric context, Leipus et al. [11] studied the empirical c.d.f. of a based on (truncated) sample lag 1 autocorrelations similarly to [1], and derived its asymptotic properties as $N, T \rightarrow \infty$, including those of a kernel density estimator. Moreover, [11] proposed a nonparametric estimator of moments of G and proved its asymptotic normality as $N, T \rightarrow \infty$. Except for parametric situations, the afore mentioned results do not allow for inferences about the tail parameter β in (2) and testing for the presence or absence of long memory in panel RCAR(1) data.

The present paper discusses in semiparametric context, the estimation of β in (2) from RCAR(1) panel $\{X_i(t), t = 1, \dots, T, i = 1, \dots, N\}$ with finite variance $EX_i^2(t) < \infty$. We use the fact that (2) implies $P(1/(1-a) > y) \sim (g_1/\beta)y^{-\beta}$, $y \rightarrow \infty$, i.e. r.v. $1/(1-a)$ follows a heavy-tailed distribution with index $\beta > 1$. Thus, if a_1, \dots, a_N were observed, β could be estimated by a number of tail index estimators. Given panel data, we estimate unobservable a_i by (truncated) sample lag 1 autocorrelation \tilde{a}_i computed from $\{X_i(1), \dots, X_i(T)\}$ similarly to [1, 11], for each $i = 1, \dots, N$. We then apply to observations $1/(1 - \tilde{a}_1), \dots, 1/(1 - \tilde{a}_N)$ the tail-index estimator introduced by Goldie and Smith [7], also studied by Novak and Utev [13], Novak [14, 15]. The main result of our paper is Theorem 2 giving sufficient conditions for asymptotic normality of the constructed estimator $\tilde{\beta}_N$. These conditions involve β , rates of growth of N, T and a threshold parameter $\delta = \delta_N \rightarrow 0$ whose choice depends on the second-order regularity parameter ν of G , see (3) below, and the $2p$ -moment of innovations. Based on the above asymptotic result, we construct a statistical procedure to test the presence of long memory in the panel, more precisely, the null hypothesis $H_0 : \beta \geq 2$ vs. the long memory alternative $H_1 : \beta \in (1, 2)$.

The paper is organized as follows. In Section 2 we make the assumptions about the statistical (panel) model. We also define the estimator $\tilde{\beta}_N$ based on the panel data and state the main Theorem 2. Section 3

details the choice of the threshold δ_N in terms of other parameters of our RCAR(1) model. In Section 4 a simulation study illustrates finite-sample properties of the estimator $\tilde{\beta}_N$ and the testing procedure for long memory. Proofs can be found in Section 5.

In what follows, C stands for a positive constant whose precise value is unimportant and which may change from line to line. We write \rightarrow_p , \rightarrow_d for the convergence in probability and distribution respectively, whereas $\rightarrow_{D[0,1]}$ denotes the weak convergence in the space $D[0, 1]$ with the uniform metric. Notation $\mathcal{N}(\mu, \sigma^2)$ is used for the normal distribution with mean μ and variance σ^2 .

2 Assumptions and the main results

To derive asymptotic results about estimation of β in the RCAR(1) panel model, condition (2) is strengthened as follows.

(G) a is a $(0, 1)$ -valued r.v. with $G(x) := \mathbb{P}(a \leq x)$, $x \in [0, 1]$. There exists $\epsilon \in (0, 1)$ such that G is continuously differentiable on $(1 - \epsilon, 1)$ with derivative satisfying

$$g(x) = \kappa\beta(1 - x)^{\beta-1}(1 + O((1 - x)^\nu)), \quad x \rightarrow 1-, \quad (3)$$

for some $\beta > 1$, $\nu > 0$ and $\kappa > 0$.

Let $Y := 1/(1 - a)$. Assumption (G) implies that

$$\mathbb{P}(Y > y) = \kappa y^{-\beta}(1 + O(y^{-\nu})), \quad y \rightarrow \infty. \quad (4)$$

Let Y_1, \dots, Y_N be i.i.d. r.v.s with a c.d.f. satisfying (4). To estimate the tail index β in (4), Goldie and Smith [7] introduced the estimator

$$\beta_N := \frac{\sum_{i=1}^N \mathbf{1}(Y_i \geq v)}{\sum_{i=1}^N \mathbf{1}(Y_i \geq v) \ln(Y_i/v)}, \quad (5)$$

where $v = v_N \rightarrow \infty$ is a threshold level, and proved asymptotic normality and other properties of this estimator.

For independent realizations a_1, \dots, a_N of a under assumption (G), we rewrite the tail-index estimator in (5) as

$$\beta_N = \frac{\sum_{i=1}^N \mathbf{1}(a_i > 1 - \delta)}{\sum_{i=1}^N \mathbf{1}(a_i > 1 - \delta) \ln(\delta/(1 - a_i))}, \quad (6)$$

where $\delta := 1/v$ is a threshold close to 0.

Theorem 1. *Assume that a, a_1, \dots, a_N are i.i.d. r.v.s and (G) holds. If $\delta = \delta_N \rightarrow 0$ and $N\delta^\beta \rightarrow \infty$ and $N\delta^{\beta+2\nu} \rightarrow 0$ as $N \rightarrow \infty$, then*

$$\sqrt{N\delta^\beta}(\beta_N - \beta) \rightarrow_d \mathcal{N}(0, \beta^2/\kappa).$$

Theorem 1 is due to Goldie and Smith [7, Thm. 4.3.2]. The proof in [7] uses Lyapunov's CLT conditionally on the number of exceedances over a threshold. Further sufficient conditions for asymptotic normality of β_N were obtained in Novak [13, 14]. In Section 5 we give an alternative proof of Theorem 1 based on the tail empirical process. Our proof has the advantage that it can be more easily adapted to prove asymptotic normality of the estimator $\tilde{\beta}_N$ in (11) of parameter β in the panel RCAR(1) model.

Let $X_i := \{X_i(t), t \in \mathbb{Z}\}$, $i = 1, 2, \dots$, be stationary random-coefficient AR(1) processes

$$X_i(t) = a_i X_i(t-1) + \zeta_i(t), \quad t \in \mathbb{Z}, \quad (7)$$

where innovations $\{\zeta_i(t), t \in \mathbb{Z}\}$ admit the following decomposition:

$$\zeta_i(t) = b_i \eta(t) + c_i \xi_i(t), \quad t \in \mathbb{Z}. \quad (8)$$

Let the following assumptions hold:

(A1) $\eta, \eta(t)$, $t \in \mathbb{Z}$, are i.i.d. with $E\eta = 0$, $E\eta^2 = 1$, $E|\eta|^{2p} < \infty$ for some $p > 1$.

(A2) $\xi, \xi_i(t)$, $t \in \mathbb{Z}$, $i = 1, 2, \dots$, are i.i.d. with $E\xi = 0$, $E\xi^2 = 1$, $E|\xi|^{2p} < \infty$ for the same $p > 1$ as in (A1).

(A3) $(b, c), (b_1, c_1), (b_2, c_2), \dots$, are i.i.d. random vectors with possibly dependent components $b \geq 0, c \geq 0$ satisfying $P(b + c = 0) = 0$ and $E(b^2 + c^2) < \infty$.

(A4) a, a_1, a_2, \dots are i.i.d. satisfying assumption (G).

(A5) $\{\eta(t), t \in \mathbb{Z}\}$, $\{\xi_i(t), t \in \mathbb{Z}\}$, a_i and (b_i, c_i) are mutually independent for each $i = 1, 2, \dots$.

Assumptions (A1)–(A3) about the innovations are very general and allow a uniform treatment of common shock (case $(b, c) = (1, 0)$) and idiosyncratic shock (case $(b, c) = (0, 1)$) situations. Similar assumptions about the innovations are made in [11]. If b or c are random (nonconstant), the innovations $\{\zeta_i(t)\}$ in (8) form a possibly dependent but otherwise uncorrelated stationary process with $E\zeta_i(0) = 0$, $E\zeta_i^2(0) = E(b^2 + c^2)$, $E\zeta_i(0)\zeta_i(t) = 0$, $t \neq 0$. Under assumptions (A1)–(A5), for each $i = 1, 2, \dots$ there exists a unique strictly stationary solution of (7) given by

$$X_i(t) = \sum_{s \leq t} a_i^{t-s} \zeta_i(s), \quad t \in \mathbb{Z},$$

with $EX_i(0) = 0$ and $EX_i^2(0) = E(b^2 + c^2)E(1 - a^2)^{-1} < \infty$, see [11].

From the panel RCAR(1) data $\{X_i(t), t = 1, \dots, T, i = 1, \dots, N\}$ we compute sample lag 1 autocorrelation coefficients

$$\hat{a}_i := \frac{\sum_{t=1}^{T-1} (X_i(t) - \bar{X}_i)(X_i(t+1) - \bar{X}_i)}{\sum_{t=1}^T (X_i(t) - \bar{X}_i)^2}, \quad (9)$$

where $\bar{X}_i := T^{-1} \sum_{t=1}^T X_i(t)$ is the sample mean, $i = 1, \dots, N$. By the Cauchy-Schwarz inequality, the estimator \hat{a}_i in (9) does not exceed 1 in absolute value a.s. Moreover, \hat{a}_i is invariant to shift and scale transformations of $\{X_i(t)\}$ in (7), i.e., we can replace $\{X_i(t)\}$ by $\{\sigma_i X_i(t) + \mu_i\}$ with some (unknown) $\mu_i \in \mathbb{R}$ and $\sigma_i > 0$ for every $i = 1, 2, \dots$

Next, we choose a threshold level $\delta > 0$ and introduce a truncated estimator

$$\tilde{a}_i := \min(\hat{a}_i, 1 - \delta^2) \quad (10)$$

for $i = 1, \dots, N$. We then define the ‘RCAR’ version of the Goldie-Smith estimator in (6) as

$$\tilde{\beta}_N := \frac{\sum_{i=1}^N \mathbf{1}(\tilde{a}_i > 1 - \delta)}{\sum_{i=1}^N \mathbf{1}(\tilde{a}_i > 1 - \delta) \ln(\delta / (1 - \tilde{a}_i))}. \quad (11)$$

In what follows, let $T = T_N$ be a positive integer-valued function of N , such that $\lim_{N \rightarrow \infty} T_N = \infty$. Let also $\delta = \delta_N > 0$ be a function of N such that $\lim_{N \rightarrow \infty} \delta_N = 0$. For ease of presentation we suppress the dependence of T and δ on N .

Theorem 2. Assume (A1)–(A5). Let $N \rightarrow \infty$ so that $N\delta^{\beta+2(\beta\wedge\nu)} \rightarrow 0$ and $N\delta^\beta/(\ln \delta)^4 \rightarrow \infty$ and

$$\sqrt{N\delta^\beta}\gamma \ln \delta \rightarrow 0 \quad \text{if } 1 < p \leq 2, \quad (12)$$

$$\sqrt{N\delta^\beta}((T\delta^\beta)^{-1} \vee \gamma) \ln \delta \rightarrow 0 \quad \text{if } p > 2, \quad (13)$$

where $\gamma = \gamma_N := (T^{(p-1)\wedge(p/2)}\delta^{p+\beta})^{-1/(p+1)}$. Then

$$\sqrt{N\delta^\beta}(\tilde{\beta}_N - \beta) \rightarrow_d \mathcal{N}(0, \beta^2/\kappa).$$

Corollary 3. Set $\tilde{K}_N := \sum_{i=1}^N \mathbf{1}(\tilde{a}_i > 1 - \delta)$. Under assumptions of Theorem 2,

$$\sqrt{\tilde{K}_N}(\tilde{\beta}_N - \beta) \rightarrow_d \mathcal{N}(0, \beta^2).$$

Remark 1. The reason for truncating sample lag 1 autocorrelation \hat{a}_i at a level less than 1 as in (10) is explained in Beran et al. [1]. In principle, in the estimator (11) we could use a different truncation level from $1 - \delta^2$ in (10), however this new level would enter and further complicate conditions (12)–(13).

3 The choice of the threshold in Theorem 2

Let us discuss conditions for the choice of the threshold $\delta = \delta_N$ in Theorem 2. Note that (A4) restricts this result to the case $\beta > 1$. Assume $p \geq 2$ in (A1), (A2) and also that $T = T_N$, $\delta = \delta_N$ increase as

$$T \sim C_1 N^a, \quad \delta \sim C_2 N^{-b}$$

for some $a > 0$, $b > 0$ and $C_1 > 0$, $C_2 > 0$. Then $N\delta^\beta/(\ln \delta)^4 \rightarrow \infty$ is equivalent to

$$\beta < \frac{1}{b}. \quad (14)$$

Condition $N\delta^{\beta+2(\beta\wedge\nu)} \rightarrow 0$ is equivalent to

$$\max \left\{ \frac{1}{3b}, \frac{1}{b} - 2\nu \right\} < \beta, \quad (15)$$

whereas (13) is satisfied if and only if

$$\frac{1 + p(1 - a + 2b)}{b(p - 1)} < \beta < \frac{2a - 1}{b}. \quad (16)$$

Inequalities (14)–(16) can be summarized as

$$\max \left\{ 1, \frac{1}{3b}, \frac{1}{b} - 2\nu, \frac{1 + p(1 - a + 2b)}{b(p - 1)} \right\} < \beta < \min \left\{ \frac{1}{b}, \frac{2a - 1}{b} \right\}. \quad (17)$$

In order that the interval for β in (17) is nonempty, we restrict the set of possible values of the parameters a and b . Particularly, this is the case and (17) holds if

$$\begin{cases} 1 < \beta < \frac{1}{b}, \\ \max \left\{ \frac{1}{3}, \frac{1}{1 + 2\nu} \right\} \leq b < 1, \quad a \geq \frac{(1 + b)(1 + p)}{p}, \quad \nu > 0. \end{cases} \quad (18)$$

Indeed, the upper bound in (17) is obvious since (18) implies $a > 1$ and the lower bound in (17) holds due to

$$\max \left\{ \frac{1}{3b}, \frac{1}{b} - 2\nu, \frac{1 + p(1 - a + 2b)}{b(p - 1)} \right\} \leq 1, \quad a > 1,$$

which follow from (18).

Albeit being only sufficient for Theorem 2, inequalities (18) provide some limitations and recommendations for estimation of β . Note that (18) restricts the range of β to the interval $(1, 3)$ provided the second-order parameter ν in (3) satisfies $\nu \geq 1$ (which roughly means that the density $g(x)$, $x \in (0, 1)$, is well-approximated by power function $C(1 - x)^{\beta-1}$ in the vicinity of $x = 1$). Condition $\beta < 1/b$ in (18) says that for larger values of β the threshold δ should decrease slower with N , or should be taken larger for fixed N , compared to the choice of δ for smaller β . Finally, the lower bound for a in (18) reflects the fact that the panel length $T \sim C_1 N^a$ should grow much faster than N , with exponent $a > 1 + b > 4/3$ in the limiting case $p = \infty$, in other words, the results of the present paper apply to long panels, similarly to [11].

4 Simulation study

The simulation study compares finite-sample performance of the estimators β_N in (6) and $\tilde{\beta}_N$ in (11) based on unobservable AR coefficients a_1, \dots, a_N and their estimates $\hat{a}_1, \dots, \hat{a}_N$ from a simulated panel respectively. We simulate N independent RCAR(1) processes X_i , $i = 1, \dots, N$ of length T , each of them driven by i.i.d. standard normal innovations $\{\zeta_i(t)\} \equiv \{\xi_i(t)\}$ in (8), with random coefficients a_i drawn from the Beta distribution with a density function

$$g(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad x \in [0, 1], \quad (19)$$

where parameters $\alpha = 2$ and $\beta > 1$. Then g satisfies (3) with the same β and $\kappa = (\beta B(\alpha, \beta))^{-1}$ and $\nu = 1$. The simulation procedure is the following:

- (S1) We simulate 10000 panels for each configuration of N , T , β , where $N = 1000$, $T = 1000, 5000, 10000$, $\beta = 1.5, 1.75, 2, 2.25, 2.5, 2.75$.
- (S2) For each simulated $N \times T$ panel we compute $\tilde{\beta}_N$ and β_N for different values of $\delta = 0.01, 0.011, 0.012, \dots, 0.5$. (The computation of β_N uses only the simulated sample a_1, \dots, a_N .)
- (S3) In each case, we choose the optimal threshold δ by a heuristic criterion called the automated Eye-Ball method (see [5]). Figure 1 illustrates this choice in the case of a single simulated panel. The automated choice matches the visual mean of the estimate over all values of δ . The optimal $\tilde{\beta}_N$ produced by this method in Figure 1 is very close to the true value of the parameter.
- (S4) In each experiment, we compute the empirical bias and standard deviation of β_N and $\tilde{\beta}_N$.

The results of our simulation experiments are presented in Table 1. Surprisingly, Table 1 shows a rather good agreement of the (root) mean squared errors between the two very different estimators β_N and $\tilde{\beta}_N$. The bias differences between the estimators for $T = 1000$ are more pronounced but they decrease with T increasing. We conclude that the AR coefficients in these experiments are estimated accurately enough so that we do not see the estimation effect on the behaviour of $\tilde{\beta}_N$ compared to β_N . We note that our experiment

$\beta = 2.5, N = T = 5000$

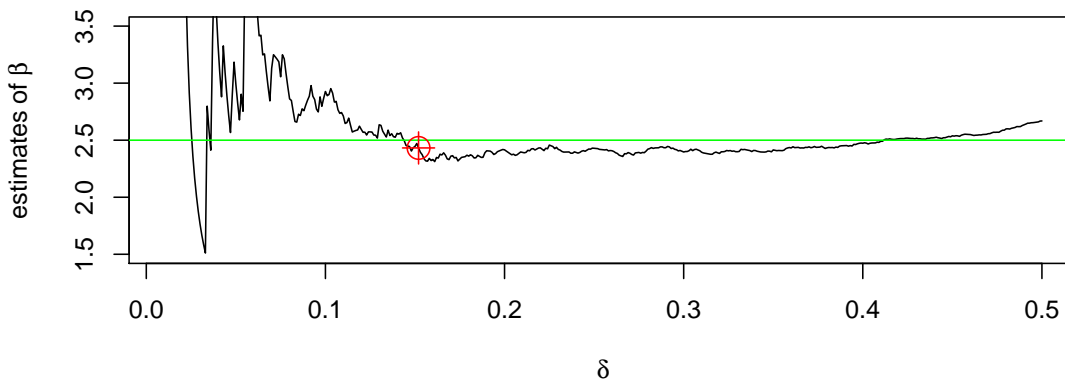


Figure 1: The graph of $\tilde{\beta}_N$ as a function of δ computed for a single simulated panel. The (green) horizontal line indicates the true value of β and the (red) \oplus symbol marks the optimal value of $\tilde{\beta}_N$ selected by the automated Eye-Ball method.

is in agreement with the recommendation at the end of the previous section that the length T of the panel must be large enough with respect to N .

Recall from Introduction that for $\beta \in (1, 2)$ the panel RCAR(1) data exhibit long memory. To test the null hypothesis $H_0 : \beta \geq 2$ vs. the long memory alternative $H_1 : \beta < 2$ we use the following statistic

$$\tilde{Z}_N := \sqrt{\tilde{K}_N}(\tilde{\beta}_N - 2)/\tilde{\beta}_N,$$

where $\tilde{K}_N = \sum_{i=1}^N \mathbf{1}(\tilde{a}_i > 1 - \delta)$, see Corollary 3. According to this corollary, $\tilde{Z}_N \rightarrow_d \mathcal{N}(0, 1)$ for $\beta = 2$ and $\tilde{Z}_N \rightarrow_p +\infty$ for $\beta > 2$, whereas $\tilde{Z}_N \rightarrow_p -\infty$ under $H_1 : \beta < 2$. Hence, for a fixed significance level $\omega \in (0, 1)$ we reject H_0 if $\tilde{Z}_N < z(\omega)$, where $z(\omega)$ is the ω -quantile of the standard normal distribution.

The same limit results hold for the test statistic

$$Z_N := \sqrt{K_N}(\beta_N - 2)/\beta_N,$$

where β_N defined in (6) and $K_N := \sum_{i=1}^N \mathbf{1}(a_i > 1 - \delta)$ are computed from known AR coefficients a_1, \dots, a_N . We compare the empirical performance of \tilde{Z}_N and Z_N from the same simulations as in Table 1, as follows:

- (S5) For each simulated values $\tilde{\beta}_N$ and β_N in (S1)–(S3) we compute the p -value of \tilde{Z}_N and Z_N .
- (S6) The empirical c.d.f.s of computed p -values of \tilde{Z}_N and Z_N for $\beta = 1.5, 1.75, 2, 2.25, 2.5, 2.75$ are plotted in Figure 2. We also provide in Table 2 the empirical probabilities for these statistics to reject H_0 at 5% level of significance.

Table 2 confirms the impression got from Table 1, that for $N = 1000$ and $T \geq 5000$, the empirical probabilities to reject the null hypothesis are similar when using test statistics Z_N and \tilde{Z}_N based on estimators β_N and $\tilde{\beta}_N$ respectively. However, for $T = 1000$ the empirical size of \tilde{Z}_N disagrees with the nominal level.

Figure 2 shows that under the null $\beta = 2$ the empirical probability to reject H_0 is close to nominal almost uniformly in p . Note that for $\beta > 2$, the probability graphs deviate from the straight line in the direction

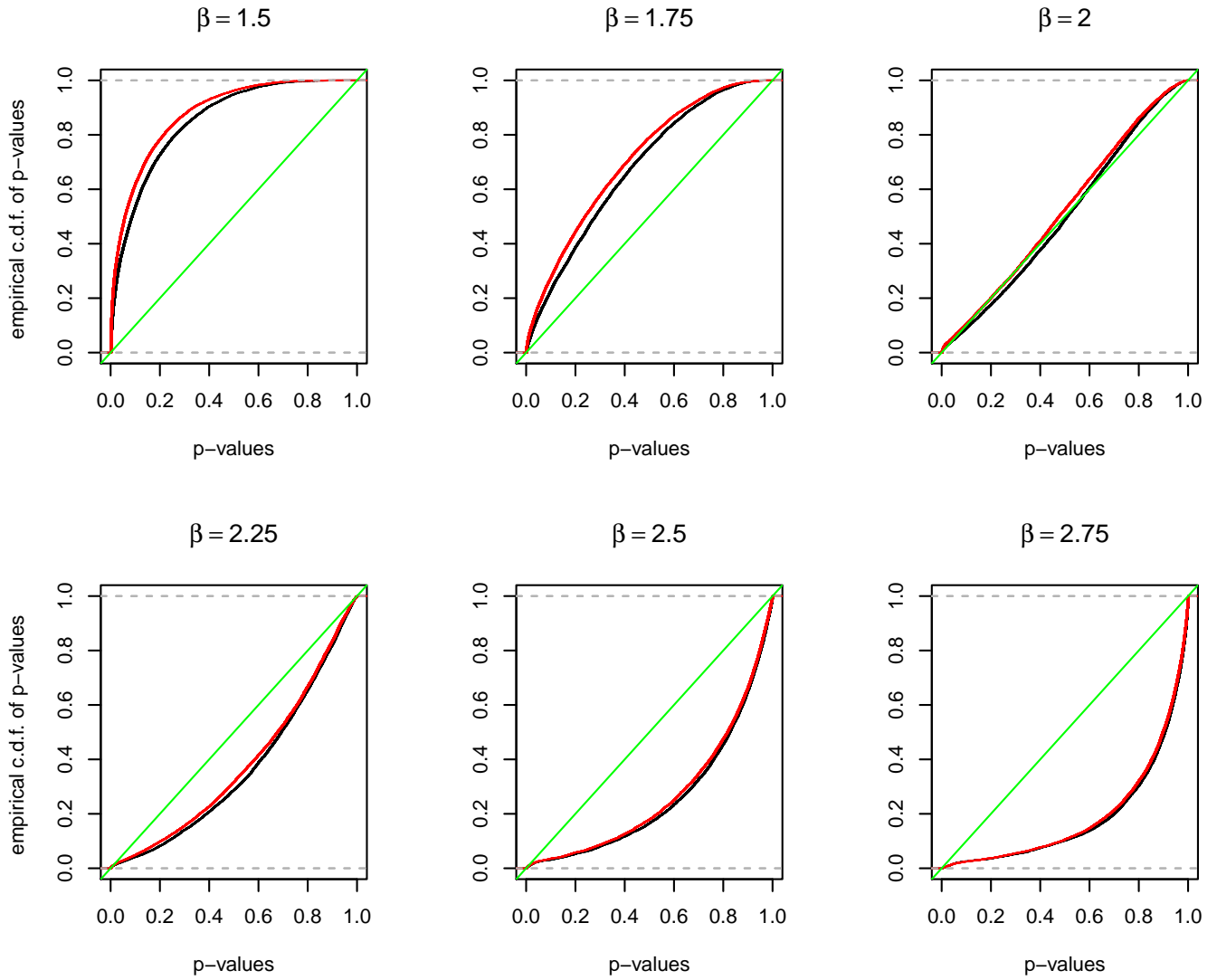


Figure 2: Empirical c.d.f. of p -values of the statistics Z_N (red) and \tilde{Z}_N (black) for testing $\beta \geq 2$ vs. $\beta < 2$. $N = 1000$, $T = 5000$. The AR coefficient has Beta distribution in (19) with parameters $(2, \beta)$. The green line represents the c.d.f. of the uniform distribution on $[0, 1]$. The number of replications of each experiment is 10000.

			$\beta = 1.5$	1.75	2	2.25	2.5	2.75
β_N		bias	0.002	-0.016	-0.044	-0.074	-0.099	-0.141
		sd	0.287	0.340	0.394	0.445	0.490	0.545
		mse	0.082	0.115	0.157	0.204	0.250	0.317
$\tilde{\beta}_N$	$T = 1000$	bias	0.222	0.150	0.077	0.025	-0.015	-0.071
		sd	0.244	0.301	0.356	0.418	0.469	0.519
		mse	0.109	0.113	0.133	0.176	0.22	0.274
$\tilde{\beta}_N$	$T = 5000$	bias	0.059	0.026	-0.006	-0.044	-0.084	-0.123
		sd	0.267	0.325	0.381	0.432	0.502	0.537
		mse	0.074	0.106	0.145	0.188	0.259	0.304
$\tilde{\beta}_N$	$T = 10000$	bias	0.033	0.011	-0.023	-0.052	-0.079	-0.126
		sd	0.276	0.332	0.391	0.445	0.489	0.545
		mse	0.077	0.110	0.154	0.200	0.245	0.313

Table 1: Empirical performance of β_N and $\tilde{\beta}_N$ for an i.i.d. sample a_1, \dots, a_N and for a simulated RCAR(1) $N \times T$ panel respectively with $N = 1000$. The random AR coefficient a_i is Beta distributed according to (19) with parameters $(2, \beta)$. The number of replications of each experiment is 10000.

		$\beta = 1.5$	1.75	2	2.25	2.5	2.75
Z_N		0.465	0.173	0.063	0.031	0.021	0.018
\tilde{Z}_N	$T = 1000$	0.174	0.052	0.023	0.014	0.013	0.010
	$T = 5000$	0.386	0.137	0.048	0.024	0.024	0.016
	$T = 10000$	0.419	0.148	0.055	0.027	0.019	0.017

Table 2: Empirical probability to reject $H_0 : \beta \geq 2$ at significance level 5%, using Z_N and \tilde{Z}_N with $N = 1000$. The AR coefficient is Beta distributed with parameters $(2, \beta)$. The number of replications of each experiment is 10000.

of zero, in agreement with our asymptotic results. The graphs in Figure 2 corresponding to $\beta = 1.5$ and $\beta = 1.75$ illustrate the power of the two tests and their consistency. Again, for $T = 5000$ the graphs for Z_N and \tilde{Z}_N seem to be very close.

5 Proofs

Notation. In what follows, let $G_N(x) := N^{-1} \sum_{i=1}^N \mathbf{1}(a_i \leq x)$, where a_1, \dots, a_N are i.i.d. with $G(x) := P(a_1 \leq x)$, $x \in [0, 1]$. Let $\hat{G}_N(x) := N^{-1} \sum_{i=1}^N \mathbf{1}(\hat{a}_i \leq x)$, where $\hat{a}_1, \dots, \hat{a}_N$ defined by (9) have a common c.d.f. $\hat{G}(x) := P(\hat{a}_1 \leq x)$, $x \in [0, 1]$.

The following result of [11] will be useful in the sequel.

Proposition 4 (Leipus et al. [11]). *Under assumptions (A1)–(A5), for all $\varepsilon \in (0, 1)$ and $T \geq 1$, it holds*

$$P(|\hat{a}_1 - a_1| > \varepsilon) \leq C(T^{-((p-1) \wedge (p/2))} \varepsilon^{-p} + T^{-1})$$

with $C > 0$ independent of ε, T .

The proof of Theorem 2 uses Proposition 5.

Proposition 5. *Let assumptions (A1)–(A5) hold. Let $N \rightarrow \infty$ so that $N\delta^\beta/(\ln \delta)^4 \rightarrow \infty$ and (12), (13) hold. Then*

$$|\ln \delta|(N\delta^{-\beta})^{1/2} \sup_{x \in [1-\delta, 1]} |\widehat{G}_N(x) - G_N(x)| = o_p(1).$$

Proof. We follow the proof of Theorem 3.1 in [11]. For $x \in [1 - \delta, 1]$, write

$$\widehat{G}_N(x) - G_N(x) = \frac{1}{N} \sum_{i=1}^N (\mathbf{1}(a_i + \widehat{\rho}_i \leq x) - \mathbf{1}(a_i \leq x)) = D'_N(x) - D''_N(x),$$

where $\widehat{\rho}_i := \widehat{a}_i - a_i$, $i = 1, \dots, N$, and

$$D'_N(x) := \frac{1}{N} \sum_{i=1}^N \mathbf{1}(x < a_i \leq x - \widehat{\rho}_i, \widehat{\rho}_i \leq 0),$$

$$D''_N(x) := \frac{1}{N} \sum_{i=1}^N \mathbf{1}(x - \widehat{\rho}_i < a_i \leq x, \widehat{\rho}_i > 0).$$

For all $\gamma > 0$, we have

$$D'_N(x) \leq \frac{1}{N} \sum_{i=1}^N \mathbf{1}(x < a_i \leq x + \gamma\delta) + \frac{1}{N} \sum_{i=1}^N \mathbf{1}(|\widehat{\rho}_i| > \gamma\delta) =: I'_N(x) + I''_N.$$

(Note that I''_N does not depend on x .) Choose $\gamma = \gamma_N := (T^{(p-1) \wedge (p/2)} \delta^{p+\beta})^{-1/(p+1)} = o(1)$ in view of (12), (13) and $N\delta^\beta/(\ln \delta)^4 \rightarrow \infty$. Then Proposition 4 yields

$$\mathbf{E}I''_N = \mathbf{P}(|\widehat{\rho}_1| > \gamma\delta) \leq C(T^{-(p-1) \wedge (p/2)}(\gamma\delta)^{-p} + T^{-1})$$

and thus $|\ln \delta|(N\delta^{-\beta})^{1/2} \mathbf{E}I''_N = o(1)$. Next,

$$\begin{aligned} |\ln \delta|(N\delta^{-\beta})^{1/2} I'_N(x) &= |\ln \delta|(N\delta^{-\beta})^{1/2} [G_N(x + \gamma\delta) - G(x + \gamma\delta) - G_N(x) + G(x)] \\ &\quad + |\ln \delta|(N\delta^{-\beta})^{1/2} [G(x + \gamma\delta) - G(x)]. \end{aligned} \quad (20)$$

For the same $\gamma = o(1)$, relation (3) implies

$$\Gamma_N := \sup_{x \in [1-\delta, 1]} |G(x + \gamma\delta) - G(x)| \leq C \sup_{x \in [1-\delta, 1]} \int_x^{x+\gamma\delta} (1-u)^{\beta-1} du \leq C\delta^{\beta-1}\gamma\delta = C\gamma\delta^\beta,$$

hence, by (12), (13), $|\ln \delta|(N\delta^{-\beta})^{1/2} \Gamma_N = O(|\ln \delta|(N\delta^\beta)^{1/2}\gamma) = o(1)$, whereas the first term on the r.h.s. of (20) vanishes in the uniform metric in probability, because $\gamma\delta(\ln \delta)^4 \rightarrow 0$ and $N\delta^\beta/(\ln \delta)^4 \rightarrow \infty$, see Lemma 6. Since $D''_N(x)$ is analogous to $D'_N(x)$, this proves the proposition. \square

Define $U_N(x) := (N\delta^{-\beta})^{1/2}(G_N(x) - G(x))$, $x \in [0, 1]$, and the modulus of continuity of U_N restricted on $[1 - \delta, 1]$:

$$\omega_N(h) := \sup_{\substack{1-\delta \leq x \leq y \leq 1, \\ y-x \leq h}} |U_N(y) - U_N(x)|, \quad h > 0.$$

Lemma 6. Assume that (G) holds. Then for all $h > 0$ and all $\varepsilon > 0$,

$$\varepsilon^4 \mathbf{P}(\omega_N(h) > \varepsilon) \leq C(h + (N\delta^\beta)^{-1}),$$

where C is a constant independent of $h, \varepsilon, \delta, N$.

Proof. Let $1 - \delta \leq x \leq y \leq z \leq 1$. By [2, p. 150, (14.9), (14.10)],

$$\begin{aligned} \mathbf{E}|U_N(y) - U_N(x)|^2 |U_N(z) - U_N(y)|^2 &\leq 3\delta^{-2\beta} \mathbf{P}(a \in (x, y]) \mathbf{P}(a \in (y, z]) \\ &\leq 3\delta^{-2\beta} \mathbf{P}(a \in (x, z])^2. \end{aligned}$$

Similarly, considerations of the 4th central moment of a binomial variable lead to

$$\mathbf{E}|U_N(y) - U_N(x)|^4 \leq 3\delta^{-2\beta} \mathbf{P}(a \in (x, y])^2 + (N\delta^{2\beta})^{-1} \mathbf{P}(a \in (x, y]).$$

Now fix $m \geq 1$ and split $[1 - \delta, 1] = \cup_{i=1}^m \Delta_i$, where $\Delta_i = [1 - \delta i/m, 1 - \delta(i-1)/m]$. Therefore, by [22, p. 49, Lemma 1], for all $m \geq 1$ and all $\varepsilon > 0$ we have

$$\begin{aligned} \varepsilon^4 \mathbf{P}(\omega_N(1/m) \geq 6\varepsilon) &\leq 3(K+1)\delta^{-2\beta} \mathbf{P}(a \in [1 - \delta, 1]) \max_{1 \leq i \leq m} \mathbf{P}(a \in \Delta_i) \\ &\quad + (N\delta^{2\beta})^{-1} \mathbf{P}(a \in [1 - \delta, 1]), \end{aligned}$$

where K is some universal constant independent of $m, \varepsilon, \delta, N$. Finally, (G) implies $\mathbf{P}(a \in [1 - \delta, 1]) \leq C\delta^\beta$ and $\max_{1 \leq i \leq m} \mathbf{P}(a \in \Delta_i) \leq C\delta^\beta/m$, which proves the lemma. \square

Proof of Theorem 1. We rewrite the estimator in (6) as

$$\beta_N = \frac{1 - G_N(1 - \delta)}{\int_{1-\delta}^1 \ln(\delta/(1-x)) dG_N(x)} = \frac{1 - G_N(1 - \delta)}{\int_{1-\delta}^1 (1 - G_N(x)) \frac{dx}{1-x}} = \frac{1 - G_N(1 - \delta)}{\int_0^\delta (1 - G_N(1-x)) \frac{dx}{x}}.$$

Next, we decompose $\beta_N - \beta = D^{-1} \sum_{i=1}^4 I_i$, where

$$\begin{aligned} I_1 &:= \beta \int_0^\delta (G_N(1-x) - G(1-x)) \frac{dx}{x}, & I_2 &:= -(G_N(1-\delta) - G(1-\delta)), \\ I_3 &:= -\beta \int_0^\delta (1 - \kappa x^\beta - G(1-x)) \frac{dx}{x}, & I_4 &:= 1 - \kappa \delta^\beta - G(1-\delta) \end{aligned} \quad (21)$$

and

$$D := \int_0^\delta (1 - G_N(1-x)) \frac{dx}{x} = \frac{1}{\beta} (\kappa \delta^\beta - I_1 - I_3). \quad (22)$$

According to the assumptions $(N\delta^\beta)^{1/2} \delta^\nu \rightarrow 0$ and (G), we get $(N\delta^{-\beta})^{1/2} I_4 \rightarrow 0$ and $(N\delta^{-\beta})^{1/2} I_3 \rightarrow 0$.

From the tail empirical process theory, see e.g. [6, Thm. 1], [12, (1.1)–(1.3)], we have that

$$(N\delta^{-\beta})^{1/2} (G_N(1-x\delta) - G(1-x\delta)) \rightarrow_{D[0,1]} \kappa^{1/2} B(x^\beta), \quad (23)$$

where $\{B(x), x \in [0, 1]\}$ is a standard Brownian motion. Therefore, we can expect that

$$(N\delta^{-\beta})^{1/2} (I_1 + I_2) \rightarrow_d \kappa^{1/2} \left(\beta \int_0^1 B(x^\beta) \frac{dx}{x} - B(1) \right). \quad (24)$$

The main technical point to prove (24) is to justify the application of the invariance principle (23) to the integral $(N\delta^{-\beta})^{1/2}I_1$, which is not a continuous functional in the uniform topology on the whole space $D[0, 1]$. For $\varepsilon > 0$, we split $I_1 := \beta(I_0^\varepsilon + I_\varepsilon^1)$, where

$$I_0^\varepsilon := \int_0^\varepsilon (G_N(1 - \delta x) - G(1 - \delta x)) \frac{dx}{x}, \quad I_\varepsilon^1 := \int_\varepsilon^1 (G_N(1 - \delta x) - G(1 - \delta x)) \frac{dx}{x}.$$

By (23), $(N\delta^{-\beta})^{1/2}I_\varepsilon^1 \rightarrow_d \kappa^{1/2} \int_\varepsilon^1 B(x^\beta) \frac{dx}{x}$, where $\mathbb{E}|\int_\varepsilon^1 B(x^\beta) \frac{dx}{x} - \int_0^1 B(x^\beta) \frac{dx}{x}|^2 \rightarrow 0$ as $\varepsilon \rightarrow 0$. Hence, (24) follows from

$$\lim_{\varepsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E}|(N\delta^{-\beta})^{1/2}I_0^\varepsilon|^2 = 0. \quad (25)$$

In the i.i.d. case $\mathbb{E}|I_0^\varepsilon|^2 = \int_0^\varepsilon \int_0^\varepsilon \text{Cov}(G_N(1 - \delta x), G_N(1 - \delta y)) \frac{dx dy}{xy}$, where

$$\text{Cov}(G_N(x), G_N(y)) = N^{-1}G(x \wedge y)(1 - G(x \vee y)) \leq N^{-1}(1 - G(x \vee y)),$$

and

$$\mathbb{E}|I_0^\varepsilon|^2 \leq \frac{C}{N} \int_0^\varepsilon \frac{dx}{x} \int_0^x (1 - G(1 - \delta y)) \frac{dy}{y} \leq \frac{C}{N} \int_0^\varepsilon \frac{dx}{x} \int_0^x (\delta y)^\beta \frac{dy}{y} = \frac{C}{N\delta^{-\beta}} \int_0^\varepsilon x^{\beta-1} dx = \frac{C\varepsilon^\beta}{N\delta^{-\beta}}, \quad (26)$$

proving (25) and hence (24) too.

Finally, we obtain $\delta^{-\beta}D \rightarrow_p \kappa/\beta$ in view of $(N\delta^{-\beta})^{1/2}(I_1 + I_3) = O_p(1)$ and $N\delta^\beta \rightarrow \infty$.

We conclude that

$$(N\delta^\beta)^{1/2}(\beta_N - \beta) \rightarrow_d \frac{\beta}{\kappa^{1/2}} \left(\beta \int_0^1 B(x^\beta) \frac{dx}{x} - B(1) \right) =: W. \quad (27)$$

Clearly, W follows a normal distribution with zero mean and variance

$$\mathbb{E}W^2 = \frac{\beta^2}{\kappa} \left(2\beta^2 \int_0^1 \frac{dx}{x} \int_0^x y^{\beta-1} dy - 2\beta \int_0^1 x^{\beta-1} dx + 1 \right) = \frac{\beta^2}{\kappa},$$

which agrees with the one in [7]. The proof is complete. \square

Proof of Theorem 2. Rewrite

$$\tilde{\beta}_N = \frac{1 - \widehat{G}_N(1 - \delta)}{\int_{\delta^2}^\delta (1 - \widehat{G}_N(1 - x)) \frac{dx}{x}}.$$

Split $\tilde{\beta}_N - \beta = \widetilde{D}^{-1}(\sum_{i=1}^4 I_i + \sum_{i=1}^4 R_i)$, where I_i , $i = 1, \dots, 4$, are defined in (21) and

$$\begin{aligned} R_1 &:= \beta \int_{\delta^2}^\delta (\widehat{G}_N(1 - x) - G_N(1 - x)) \frac{dx}{x}, & R_2 &:= G_N(1 - \delta) - \widehat{G}_N(1 - \delta), \\ R_3 &:= \beta \int_0^{\delta^2} (G(1 - x) - G_N(1 - x)) \frac{dx}{x}, & R_4 &:= \beta \int_0^{\delta^2} (1 - G(1 - x)) \frac{dx}{x} \end{aligned}$$

and

$$\widetilde{D} := \int_{\delta^2}^\delta (1 - \widehat{G}_N(1 - x)) \frac{dx}{x} = D - \frac{1}{\beta}(R_1 + R_3 + R_4)$$

with D given by (22). By Proposition 5, $(N\delta^{-\beta})^{1/2}R_2 = o_p(1)$ and $(N\delta^{-\beta})^{1/2}R_1 = o_p(1)$. In view of (26), we have $\mathbb{E}|(N\delta^{-\beta})^{1/2}R_3|^2 \leq C\delta^{2\beta} = o(1)$ and so $(N\delta^{-\beta})^{1/2}R_3 = o_p(1)$. Finally, $(N\delta^{-\beta})^{1/2}R_4 = o(1)$ as $N\delta^{3\beta} \rightarrow 0$. \square

Proof of Corollary 3. Let $K_N = \sum_{i=1}^N \mathbf{1}(a_i > 1 - \delta)$. By Proposition 5, we have $(N\delta^\beta)^{-1}(\tilde{K}_N - K_N) = o_p(1)$. Since $\text{Var}(K_N) \leq N(1 - G(1 - \delta))$ and $N(1 - G(1 - \delta)) \rightarrow \infty$, Markov's inequality yields

$$\frac{K_N}{N(1 - G(1 - \delta))} \rightarrow_p 1,$$

consequently, $(N\delta^\beta)^{-1}K_N \rightarrow_p \kappa$. We conclude that $(N\delta^\beta)^{-1}\tilde{K}_N \rightarrow_p \kappa$. □

References

- [1] Beran, J., Schützner, M. and Ghosh, S. (2010) From short to long memory: Aggregation and estimation. *Comput. Stat. Data Anal.* **54**, 2432–2442.
- [2] Billingsley, P. (1999) *Convergence of Probability Measures*, 2nd ed. Wiley, New York.
- [3] Celov, D., Leipus, R. and Philippe, A. (2007) Time series aggregation, disaggregation and long memory. *Lithuanian Math. J.* **47**, 379–393.
- [4] Celov, D., Leipus, R. and Philippe, A. (2010) Asymptotic normality of the mixture density estimator in a disaggregation scheme. *J. Nonparametric Statist.* **22**, 425–442.
- [5] Danielson, J., Ergun L.M., de Haan, L. and de Vries, C.G (2016+) Tail index estimation: Quantile driven threshold selection. Available at <http://dx.doi.org/10.2139/ssrn.2717478>.
- [6] Einmahl, J.H.J. (1990) The empirical distribution function as a tail estimator. *Stat. Neerl.* **44**, 79–82.
- [7] Goldie, C.M., Smith, R.L. (1987) Slow variation with remainder: theory and applications. *Quart. J. Math. Oxford* **38**, 45–71.
- [8] Gonçalves, E. and Gouriéroux, C. (1988) Aggrégation de processus autoregressifs d'ordre 1. *Annales d'Economie et de Statistique* **12**, 127–149.
- [9] Granger, C.W.J. (1980) Long memory relationship and the aggregation of dynamic models. *J. Econometrics* **14**, 227–238.
- [10] Leipus, R., Oppenheim, G., Philippe, A. and Viano, M.-C. (2006) Orthogonal series density estimation in a disaggregation scheme. *J. Statist. Plan. Inf.* **136**, 2547–2571.
- [11] Leipus, R., Philippe, A., Pilipauskaitė, V. and Surgailis, D. (2017). Nonparametric estimation of the distribution of the autoregressive coefficient from panel random-coefficient AR(1) data. *J. Multiv. Anal.* **153**, 121–135.
- [12] Mason, D.M. (1988) A strong invariance theorem for the tail empirical process *Ann. Inst. Henri Poincaré* **24**, 491–506.
- [13] Novak, S.Y. and Utev, S. (1990). On the asymptotic distribution of the ratio of sums of random variables. *Siberian Math. J.* **31**, 781–788.
- [14] Novak S.Y. (1992) Inference about the Pareto-type distribution. In: *Trans. 11th Prague Conf. Inform Theory Statist. Decis. Func. Random Processes*. Prague: Academia, **B**, 251–258.
- [15] Novak, S.Y. (2002) Inference on heavy tails for dependent data. *Siberian Adv. Math.* **12**, 73–96.
- [16] Oppenheim, G. and Viano, M.-C. (2004) Aggregation of random parameters Ornstein-Uhlenbeck or AR processes: some convergence results. *J. Time Ser. Anal.* **25**, 335–350.
- [17] Philippe, A., Puplinskaitė, D. and Surgailis, D. (2014) Contemporaneous aggregation of triangular array of random-coefficient AR(1) processes. *J. Time Series Anal.* **35**, 16–39.
- [18] Pilipauskaitė, V. and Surgailis, D. (2014) Joint temporal and contemporaneous aggregation of random-coefficient AR(1) processes. *Stochastic Process. Appl.* **124**, 1011–1035.

- [19] Pilipauskaitė, V. and Surgailis, D. (2015) Joint aggregation of random-coefficient AR(1) processes with common innovations. *Statist. Probab. Lett.* **101**, 73–82.
- [20] Puplinskaitė, D. and Surgailis, D. (2010) Aggregation of random coefficient AR(1) process with infinite variance and idiosyncratic innovations. *Adv. Appl. Probab.* **42**, 509–527.
- [21] Robinson, P.M. (1978) Statistical inference for a random coefficient autoregressive model. *Scand. J. Stat.*, **5**, 163–168.
- [22] Shorack, G.R., Wellner, J.A. (1986) *Empirical Processes with Applications to Statistics*. Wiley, New York.
- [23] Zaffaroni, P. (2004) Contemporaneous aggregation of linear dynamic models in large economies. *J. Econometrics* **120**, 75–102.