A bootstrap approximation to a unit root test statistic for heavy-tailed observations

Lajos Horváth
University of Utah

Piotr Kokoszka
Utah State University

June 10, 2002

Abstract

In the context of the AR(1) model with innovations in the domain of attraction of an α-stable law, we develop a residual bootstrap approximation to the distribution of a least-squares estimator of the autoregressive parameter when this parameter is equal to unity. Our procedure requires drawing bootstrap samples of size $m < n$, n being the size of the original sample. We establish the convergence in probability of the bootstrap distribution function assuming that $m \to \infty$ and $m/n \to 0$. An analogous result is established for the partial sum process of the bootstrap noise sequence.

Keywords: bootstrap, heavy tails, stable distrinution, unit root
2000 MSC: 62M10, 62P20 60E07

1 Introduction

We consider the model

$$X_k = \theta X_{k-1} + \varepsilon_k, \quad X_0 = 0, \tag{1.1}$$

where θ is an unknown parameter and the ε_k are a noise sequence. The unit root problem for this model consists in testing $H_0 : \theta = 1$ against $H_1 : |\theta| < 1$; if $\theta = 1$, we say that the process has a unit root. Unit root tests in various settings have been extensively studied in econometric literature, we refer to Chapter 17 of Hamilton (1994) and Philllips and Xiao (1998) for extensive reviews. Essentially all work in this area assumed that the error process, perhaps in a more complex setting than (1.1), is stationary with finite variance. The goal is then to establish the null distribution of a test statistic. Typically the limiting distribution is a complicated functional of a Brownian motion and the parameters describing the process $\{\varepsilon_k\}$, see e.g Wang *et al.* (2002) and Ling and Li (2002) for recent contributions in this direction. Basawa *et al.* (1991), Heimann and Kreiss (1996) and Paparoditis and Politis (2001) used different types of bootstrap procedures to develop unit root tests for which the exact form of the limiting null distribution is not required. These test however also assume that the noise process has finite variance.

The focus of this paper is on the simple model (1.1), but with the infinite variance ε_t; some economic and financial time series are believed not to have finite variance, see Mittnik and Rachev (2000) and references therein. Such a model was considered by Chan and Tran (1989) whose theory was extended in Chapter 15 of Mittnik and Rachev (2000) to more complex models and to t-tests. The approach of Chan and Tran (1989) and Mittnik and Rachev (2000) relies on finding the asymptotic distribution of a test statistic which is shown to depend on the tail index parameter α, as explained in Section 2. The present paper shows that the null distribution of the popular test statistic based on the least-squares estimator of θ can be approximated using residual bootstrap. Our method is similar to those of Basawa *et al.* (1991) and Heimann and Kreiss (1996). As in the latter reference, it requires drawing bootstrap samples of size $m < n$, n being the size of the original sample. The fact that the ususal bootstrap based on resamples of size n is not consistent for infinite variance observations was first realized by Athreya (1987). In the context of estimating the mean, the approach based on resamples of size $m < n$ was developed by Athreya (1986) and Arcones and Giné (1989, 1991).

While we do not claim that our aproach provides the full solution to the unit root problem in the infinite variance case, it is hoped that the rigourous mathematical theory developed here will point out to the potential practical applicability and robustness of procedures of this type.

The paper is organized as follows. I section 2 we state the assumptions, describe the bootstrap procedure and state the main results. Proofs are collected in Section 3.

2 Assumptions and results

We assume the the ε_k in (1.1) satisfy the following assumption:

We impose the following assumption on the ε_k in (1.1):

ASSUMPTION 2.1 The ε_k are in the domain of attraction of a strictly α-stable law with $1 < \alpha < 2$ and $E\varepsilon_k = 0$.

Chan and Tran (1989) showed that if $\theta = 1$ and Assumption 2.1 holds, then

$$n(\hat{\theta}_n - 1) \xrightarrow{d} \xi, \tag{2.1}$$

where

$$\hat{\theta}_n = \frac{\sum\limits_{2\le j\le n} X_{j-1}X_j}{\sum\limits_{2\le j\le n} X_{j-1}^2}$$

is the least-squares estimator of θ and

$$\xi = \frac{\int_0^1 U(t-)dU(t)}{\int_0^1 U^2(t)dt}. \tag{2.2}$$

In (2.2), $U(\cdot)$ is an α-stable Levy process defined by

$$a_n^{-1} \sum_{1\le i\le nt} \varepsilon_i \to U(t) \quad \text{in } D[0,1], \tag{2.3}$$

where

$$a_n = \inf\left\{x : P(|\varepsilon_0| > x) \le n^{-1}\right\} = n^{1/\alpha}L(n) \tag{2.4}$$

for some slowly varying function L, see Section 4.4 of Resnick (1986) for the proof of (2.3).

We propose to use the following bootstrap procedure to approximate the distribution of ξ without estimating the tail index α:

1. Compute the residuals $\varepsilon_1^* = X_1,\ \varepsilon_2^* = X_2 - \hat\theta_n X_1, \ldots, \varepsilon_n^* = X_n - \hat\theta_n X_{n-1}$ and the centered residuals
$$\hat\varepsilon_j = \varepsilon_j^* - \frac{1}{n}\sum_{1\le j\le n} \varepsilon_j^*, \quad j = 1, 2, \ldots, n.$$

2. For a fixed $m \le n$ select with replacement a bootstrap sample $\tilde\varepsilon_1, \ldots, \tilde\varepsilon_m$ from $\hat\varepsilon_1, \ldots, \hat\varepsilon_n$.

3. Construct the bootstrap process
$$\tilde X_0 = 0, \quad \tilde X_i = \hat\theta \tilde X_{i-1} + \tilde\varepsilon_i, \quad 1 \le i \le m.$$
and calculate the estimator
$$\tilde\theta_m = \frac{\sum_{2\le j\le m} \tilde X_{j-1}\tilde X_j}{\sum_{2\le j\le m} \tilde X_{j-1}^2}.$$

4. Approximate the distribution of $n(\hat\theta_n - 1)$ by that of $m(\tilde\theta_m - \hat\theta_n)$.

We impose the following asymptotic conditions on $m = m(n)$:

Assumption 2.2 As $n \to \infty$, $m \to \infty$ and $m/n \to 0$.

In the following we denote $\mathcal{E} = \sigma(\varepsilon_i,\ i \ge 1)$. Note that if $\theta = 1$, then $\mathcal{E} = \sigma(X_i, i \ge 1)$. By $P_{\mathcal{E}}$ and $E_{\mathcal{E}}$ we denote, respectively, the conditional probability and expectation with respect to $\mathcal{E}$.

THEOREM 2.1 *If $\theta = 1$ and Assumptions 2.1 and 2.2 hold, then for every real x*

$$P_{\mathcal{E}}(m(\tilde{\theta}_m - \hat{\theta}_n) \leq x) \xrightarrow{P} P(\xi \leq x). \tag{2.5}$$

Our next theorem is used in the proof of Theorem 2.1, but is also of independent interest.

THEOREM 2.2 *Suppose $\theta = 1$ and Assumptions 2.1 and 2.2 hold. Then, for any bounded continuous functional h on $D[0,1]$*

$$P_{\mathcal{E}}\left(h(a_m^{-1}\sum_{1\leq i\leq m\cdot} \tilde{\varepsilon}_i) \leq x\right) \xrightarrow{P} P\left(h(U(\cdot)) \leq x\right)$$

for all points of continuity of the limit.

REMARK 2.1 Heimann and Kreiss (1996) developed a similar bootstrap approximation which is valid for all real θ in (1.1) under the assumption that the ε_k have finite variance. Their method relies on the observation that for finite variance ε_k the statistic

$$T_n = \left[\sum_{2\leq i\leq n} X_{i-1}^2\right]^{1/2} (\hat{\theta}_n - \theta)$$

has a limiting distribution for all values of θ. This is not the case for the ε_k considered in this paper. If $\theta = 1$, the case we consider, $m^{1/2-1/\alpha}T_n/L(n)$ has a limiting distribution. For other values of θ, different normalizations are required, the only case when the normalizing constants in analogs of (2.1) do not depend on α being $\theta = 1$.

3 Proofs

We begin by introducing some notation. Denoting $\bar{\varepsilon}^* = n^{-1}\sum_{j=1}^n \varepsilon_j^*$, it is easy to see that

$$\hat{\varepsilon}_i = \varepsilon_i + (\theta - \hat{\theta}_n)X_{i-1} - \bar{\varepsilon}^*. \tag{3.1}$$

When we select one of $\hat{\varepsilon}_1, \ldots, \hat{\varepsilon}_n$, we select the corresponding unobservable noise variable $\varepsilon_\cdot$ and the observation $X_\cdot$, and these random variables will be denoted $\underline{\varepsilon}_i$ and $\underline{X}_{i-1}$. This means that

$$\tilde{\varepsilon}_i = \underline{\varepsilon}_i + (\theta - \hat{\theta}_n)\underline{X}_{i-1} - \bar{\varepsilon}^* \tag{3.2}$$

so

$$\frac{1}{a_m}\sum_{1\leq i\leq mt}\tilde{\varepsilon}_i = \frac{1}{a_m}\sum_{1\leq i\leq mt}\underline{\varepsilon}_i + (\theta - \hat{\theta}_n)\frac{1}{a_m}\sum_{1\leq i\leq mt}\underline{X}_{i-1} - \frac{[mt]}{a_m}\bar{\varepsilon}^*. \tag{3.3}$$

PROOF OF THEOREM 2.2: Follows from Lemmas 3.1, 3.2 and 3.3 which establish the necessary asymptotics for the three terms on the right-hand side of (3.3).

LEMMA 3.1 *Under the assumptions of Theorem 2.2, $ma_m^{-1}\bar{\varepsilon}^* \xrightarrow{P} 0$.*

PROOF: Observe that

$$\bar{\varepsilon}^* = \frac{1}{n}\sum_{1\le i\le n}\varepsilon_i + (\theta - \hat{\theta}_n)\frac{1}{n}\sum_{1\le i\le n} X_{i-1}.$$

By (2.3),

$$\left(\frac{1}{a_n}\sum_{1\le i\le n}\varepsilon_i,\ \frac{1}{na_n}\sum_{1\le i\le n} X_{i-1}\right) \xrightarrow{d} \left(U(1),\ \int_0^1 U(t)dt\right) \tag{3.4}$$

and therefore by (3.4) and (2.1)

$$ma_m^{-1}\bar{\varepsilon}^* = O_P(1)\left(\frac{m}{n}\right)^{1-1/\alpha}\frac{L(n)}{L(m)}. \tag{3.5}$$

By Theorem 1.5.6 of Bingham *et al.* (1987) (Potter bounds), for any $\delta > 0$, there is $x_0 = x_0(\delta)$ such that for $m > x_0$, $L(n)/L(m) \le 2(n/m)^{\delta}$. Choosing δ so small that $1 - 1/\alpha - \delta > 0$, we obtain

$$\lim_{n\to\infty}\left(\frac{m}{n}\right)^{1-1/\alpha}\frac{L(n)}{L(m)} \le 2\lim_{n\to\infty}\left(\frac{m}{n}\right)^{1-1/\alpha-\delta} = 0. \tag{3.6}$$

LEMMA 3.2 *Under the assumptions of Theorem 2.2, for any $x > 0$,*

$$P_{\mathcal{E}}\left(\frac{|1-\hat{\theta}_n|}{a_m}\sup_{1\le t\le 1}\left|\sum_{1\le i\le mt}\underline{X}_{i-1}\right| \ge x\right) \xrightarrow{P} 0.$$

PROOF: Let

$$\Delta = \frac{1}{n}\sum_{0\le i\le n-1} X_i.$$

By (2.1) and (2.3), we have

$$\begin{aligned}\frac{|1-\hat{\theta}_n|}{a_m}m\Delta &= O_P(1)\frac{m}{n}\frac{n^{1/\alpha}L(n)}{m^{1/\alpha}L(m)}\frac{1}{nn^{1/\alpha}L(n)}\sum_{0\le i\le n-1} X_i \\ &= O_P(1)\left(\frac{m}{n}\right)^{1-1/\alpha}\frac{L(n)}{L(m)}.\end{aligned}$$

By (3.6), we conclude that

$$a_m^{-1}|1-\hat{\theta}|m\Delta \xrightarrow{P} 0. \tag{3.7}$$

Conditionally on $\mathcal{E}$, $\underline{X}_1 - \Delta, \ldots, \underline{X}_m - \Delta$ are iid random variables with zero mean and

$$E_{\mathcal{E}}(\underline{X}_1 - \Delta)^2 = \frac{1}{n}\sum_{1\le i\le n-1}(X_i - \Delta)^2 = O_P(1)n^{2/\alpha}L^2(n),$$

by (2.3). Using the Marcinkiewicz-Zygmund inequality, c.f. Petrov (1995), p. 82, there is a constant C such that

$$E_{\mathcal{E}} \max_{0\le k\le m-1} \left(\sum_{0\le i\le k} (\underline{X}_i - \Delta) \right)^2 \le C E_{\mathcal{E}} \left(\sum_{0\le k\le m-1} (\underline{X}_1 - \Delta) \right)^2$$

$$= CmE_{\mathcal{E}}(\underline{X}_1 - \Delta)^2 = O_P(1) m n^{2/\alpha} L^2(n).$$

Therefore, by the Markov inequality, we have

$$P_{\mathcal{E}} \left(\frac{|1-\hat{\theta}_n|}{a_m} \max_{0\le k\le m-1} \left| \sum_{0\le i\le k} (\underline{X}_i - \Delta) \right| \ge x \right) = O_P(1) \frac{mn^{2/\alpha} L^2(n)}{n^2 m^{2/\alpha} L^2(m)} \tag{3.8}$$

$$= O_P(1) n^{2(1/\alpha - 1)} L^2(n) m^{1-2/\alpha} L^{-2}(m) \xrightarrow{P} 0.$$

Combining (3.7) and (3.8) completes the proof of the lemma.

Lemma 3.3 *If the assumptions of Theorem 2.2 hold, then*

$$P_{\mathcal{E}} \left(h(a_m^{-1} \sum_{1\le i\le m\cdot} \underline{\varepsilon}_i) \le x \right) \xrightarrow{P} P(h(U(\cdot)) \le x).$$

Proof: Conditionally on $\mathcal{E}$, $\underline{\varepsilon}_1, \dots, \underline{\varepsilon}_m$ are iid with distribution function

$$F_n(x) = \frac{1}{n} \sum_{1\le i\le n} \mathbf{1}\left\{\varepsilon_i \le x\right\}.$$

By Proposition 3.4 of Resnick (1986) it is enough to prove that for some $0 \le p, q \le 1, p+q = 1$ and any $x > 0$

$$m(1 - F_n(xa_m) \xrightarrow{P} px^{-\alpha}; \tag{3.9}$$

$$mF_n(-xa_m) \xrightarrow{P} qx^{-\alpha}. \tag{3.10}$$

Since ε_1 is in the domain of attraction of an α-stable law, its distribution function F satisfies

$$m(1 - F(xa_m)) \to px^{-\alpha}; \tag{3.11}$$

$$mF(-xa_m) \to qx^{-\alpha} \tag{3.12}$$

for some $0 \le p, q \le 1, p + q = 1$.

It is easy to see that

$$\mathrm{Var}[m(1 - F_n(xa_m))] \le m^2 n^{-1} (1 - F(xa_m)) \to 0, \tag{3.13}$$

by (3.11) and the assumption $m/n \to 0$. Since $EF_n(x) = F(x)$, (3.9) follows from (3.11) and (3.13). Relation (3.10) follows from the same argument.

PROOF OF THEOREM 2.1: Let

$$\tilde{S}(i) = \tilde{\varepsilon}_1 + \ldots + \tilde{\varepsilon}_i.$$

Arguing as in Chan and Tran (1989), we conclude from Theorem 2.2 that

$$P_{\mathcal{E}}\left(a_m^{-2}\sum_{2\le i\le m}\tilde{\varepsilon}_i\tilde{S}(i-1)\le x\right)\xrightarrow{P} P\left(\int_0^1 U(t-)dU(t)\le x\right), \tag{3.14}$$

$$P_{\mathcal{E}}\left(m^{-1}a_m^{-2}\sum_{2\le i\le m}\tilde{S}(i-1)^2\le x\right)\xrightarrow{P} P\left(\int_0^1 U^2(t)dt\le x\right). \tag{3.15}$$

Consequently,

$$\frac{m\sum\limits_{2\le i\le m}\tilde{\varepsilon}_i\tilde{S}(i-1)}{\sum\limits_{2\le i\le m}\tilde{S}(i-1)^2}\xrightarrow{d}\xi,\quad \text{in probability}. \tag{3.16}$$

It is easy to see that

$$\tilde{\theta}_m-\hat{\theta}_n=\frac{\sum\limits_{2\le i\le m}\tilde{\varepsilon}_i\tilde{X}_{i-1}}{\sum\limits_{2\le i\le m}\tilde{X}_{i-1}^2},$$

so to complete the proof, we must verify that

$$P_{\mathcal{E}}\left\{\left|\frac{\sum\limits_{2\le i\le m}\tilde{\varepsilon}_i\tilde{X}_{i-1}}{\sum\limits_{2\le i\le m}\tilde{X}_{i-1}^2}-\frac{\sum\limits_{2\le i\le m}\tilde{\varepsilon}_i\tilde{S}(i-1)}{\sum\limits_{2\le i\le m}\tilde{S}(i-1)^2}\right|>\frac{x}{m}\right\}\xrightarrow{P}0. \tag{3.17}$$

Relation (3.17) follows from Lemmas 3.6 and 3.7, which are preceded by Lemmas 3.4 and 3.5.

In the remainder of the proof we show how (3.17) follows from Lemmas 3.6 and 3.7. To lighten the notation, rewrite (3.17) as

$$P_{\mathcal{E}}\left\{\left|\frac{N_X}{D_X}-\frac{N_S}{D_S}\right|>\frac{x}{m}\right\}\xrightarrow{P}0. \tag{3.18}$$

It then suffices to show that

$$P_{\mathcal{E}}\left\{\left|\frac{N_X(D_S-D_X)}{D_XD_S}\right|>\frac{x}{m}\right\}\xrightarrow{P}0 \tag{3.19}$$

and

$$P_{\mathcal{E}}\left\{\left|\frac{N_X-N_S}{D_S}\right|>\frac{x}{m}\right\}\xrightarrow{P}0. \tag{3.20}$$

We focus on (3.19), a similar, in fact shorter, argument applies to (3.20).

Observe that for arbitrarily small $r > 0$ and arbitrarily large C

$$P_{\mathcal{E}}\left\{\left|\frac{N_X(D_S - D_X)}{D_X D_S}\right| > \frac{x}{m}\right\} \tag{3.21}$$

$$\begin{aligned}
&\le P_{\mathcal{E}}\left\{|D_S - D_X| > xm^{-1},\ D_X > rma_m^2,\ D_S > rma_m^2,\ N_X \le Ca_m^2\right\} \\
&\quad + P_{\mathcal{E}}\left\{D_X \le rma_m^2\right\} + P_{\mathcal{E}}\left\{D_S \le rma_m^2\right\} + P_{\mathcal{E}}\left\{N_X > Ca_m^2\right\}. \\
&\le P_{\mathcal{E}}\left\{|D_S - D_X| > xr^2C^{-1}ma_m^2\right\} + P_{\mathcal{E}}\left\{D_X - D_S \le -rma_m^2\right\} \\
&\quad + P_{\mathcal{E}}\left\{D_S \le 2rma_m^2\right\} + P_{\mathcal{E}}\left\{D_S \le rma_m^2\right\} + P_{\mathcal{E}}\left\{N_X > Ca_m^2\right\}.
\end{aligned}$$

By Lemma 3.7, for any positive r and C,

$$P_{\mathcal{E}}\left\{|D_S - D_X| > xr^2C^{-1}ma_m^2\right\} + P_{\mathcal{E}}\left\{D_X - D_S \le -rma_m^2\right\} \xrightarrow{P} 0 \tag{3.22}$$

By (3.15),

$$P_{\mathcal{E}}\left\{D_S \le 2rma_m^2\right\} \xrightarrow{P} \left\{\int_0^1 U^2(t)dt \le 2r\right\}, \tag{3.23}$$

$$P_{\mathcal{E}}\left\{D_S \le rma_m^2\right\} \xrightarrow{P} P\left\{\int_0^1 U^2(t)dt \le r\right\} \tag{3.24}$$

and by (3.14),

$$P_{\mathcal{E}}\left\{N_X > Ca_m^2\right\} \xrightarrow{P} P\left\{\int_0^1 U(t-)dU(t) > C\right\}. \tag{3.25}$$

Combining relations (3.21) – (3.25), we obtain (3.19).

Lemma 3.4 *If $\theta = 1$ and Assumption 2.1 holds, then $E_{\mathcal{E}}\tilde{\varepsilon}_1 = 0$ and $E_{\mathcal{E}}\tilde{\varepsilon}_1^2 = O_P(1)n^{2/\alpha-1}L^2(n)$.*

Proof: The relation $E_{\mathcal{E}}\tilde{\varepsilon}_1 = 0$ follows from the definition. By (3.1),

$$\begin{aligned}
E_{\mathcal{E}}\tilde{\varepsilon}_1^2 &= \frac{1}{n}\sum_{1\le i\le n}[\varepsilon_i + (\theta - \hat{\theta}_n)X_{i-1} - \bar{\varepsilon}^*]^2 \\
&= \frac{1}{n}\sum_{1\le i\le n}\varepsilon_i^2 + \frac{1}{n}(\theta - \hat{\theta}_n)^2\sum_{1\le i\le n}X_{i-1}^2 + (\bar{\varepsilon}^*)^2 \\
&+ \frac{2}{n}(\theta - \hat{\theta}_n)\sum_{1\le i\le n}\varepsilon_i X_{i-1} - \frac{2}{n}\bar{\varepsilon}^*\sum_{1\le i\le n}\varepsilon_i - \frac{2}{n}\bar{\varepsilon}^*(\theta - \hat{\theta}_n)\sum_{1\le i\le n}X_{i-1}.
\end{aligned}$$

Therefore the lemma follows from the following relations:

$$|\theta - \hat{\theta}_n| = O_P(1)n^{-1}, \qquad \sum_{1\le i\le n}\varepsilon_i X_{i-1} = O_P(1)n^{2/\alpha}L^2(n),$$

$$\sum_{1\le i\le n}\varepsilon_i^2 = O_P(1)n^{2/\alpha}L^2(n), \qquad \sum_{1\le i\le n}X_{i-1}^2 = O_P(1)n^{1+2/\alpha}L^2(n),$$

$$\bar{\varepsilon}^* = O_P(1)n^{1/\alpha-1}L(n), \qquad \sum_{1\le i\le n}X_{i-1} = O_P(1)n^{1+1/\alpha}L(n).$$

Lemma 3.5 *If* $n(\hat{\theta}_n - 1) = O_P(1)$, *then*

$$\sum_{1\le i\le m}\sum_{0\le j\le i-1}(\hat{\theta}_n^j - 1)^2 = O_P(1)\frac{m^4}{n^2}.$$

Proof: By the mean value theorem

$$|\hat{\theta}_n^j - 1| \le (1 + |\hat{\theta}_n - 1|)^{j-1} j |\hat{\theta}_n - 1| \le \max_{1\le k\le m}(1 + |\hat{\theta}_n - 1|)^k j |\hat{\theta}_n - 1|.$$

To complete the proof, it remains to verify that

$$\max_{1\le k\le m}(1 + |\hat{\theta}_n - 1|)^k j |\hat{\theta}_n - 1| = O_P(1) j n^{-1}. \tag{3.26}$$

By (2.1), (3.26) reduces to

$$\lim_{C\to\infty}\limsup_{n\to\infty} P\left(\max_{1\le k\le m}\left[1 + |\hat{\theta}_n - 1|^k\right] > C\right) = 0. \tag{3.27}$$

Introduce the events

$$\Omega_{n,C} = \left\{|\hat{\theta}_n - 1| \le n^{-1}C\right\}.$$

By (2.1),

$$\lim_{C\to\infty}\limsup_{n\to\infty} P(\Omega_{n,C}^c) = 0. \tag{3.28}$$

On the event $\Omega_{n,C}$,

$$\max_{1\le k\le m}(1 + |\hat{\theta}_n - 1|)^k \le \left[\left(1 + n^{-1}C\right)^n\right]^{m/n} \to 1$$

because $m/n \to 0$. Therefore

$$\lim_{C\to\infty}\limsup_{n\to\infty} P\left(\max_{1\le k\le m}\left[1 + |\hat{\theta}_n - 1|^k\right] > C,\ \Omega_{n,C}\right) = 0. \tag{3.29}$$

Relations (3.28) and (3.29) yield (3.27), and so the proof of the lemma is complete.

Lemma 3.6 *If the assumptions of Theorem 2.1 hold, then for any* $x > 0$

$$P_{\mathcal{E}}\left(\left|\sum_{2\le i\le m}\tilde{\varepsilon}_i(\tilde{X}_{i-1} - \tilde{S}(i-1))\right| > x a_m^2\right) \xrightarrow{P} 0.$$

PROOF: Observe that

$$\tilde{X}_i = \sum_{0\le j\le i-1} \hat{\theta}_n^j \tilde{\varepsilon}_{i-j}.$$

and that conditionally on $\mathcal{E}$, $\tilde{\varepsilon}_i$ and $\tilde{X}_{i-1}, \tilde{S}(i-1)$ are independent. Therefore

$$E_{\mathcal{E}}\left[\sum_{2\le i\le m} \tilde{\varepsilon}_i(\tilde{X}_{i-1} - \tilde{S}(i-1))\right]^2 = [E_{\mathcal{E}}\tilde{\varepsilon}_1^2] \sum_{2\le i\le m} E_{\mathcal{E}}(\tilde{X}_{i-1} - \tilde{S}(i-1))^2$$

$$= [E_{\mathcal{E}}\tilde{\varepsilon}_1^2] \sum_{2\le i\le m} E_{\mathcal{E}}\left(\sum_{0\le j\le i-1} (\hat{\theta}_n^j - 1)\tilde{\varepsilon}_{i-j}\right)^2 = [E_{\mathcal{E}}\tilde{\varepsilon}_1^2]^2 \sum_{2\le i\le m}\sum_{0\le j\le i-1} (\hat{\theta}_n^j - 1)^2$$

$$= O_P(1) m^4 n^{4(1/\alpha-1)} L^4(n),$$

by Lemmas 3.4 and 3.5. By the Chebyshev inequality we therefore obtain

$$P_{\mathcal{E}}\left(\left|\sum_{2\le i\le m} \tilde{\varepsilon}_i(\tilde{X}_{i-1} - \tilde{S}(i-1))\right| > x m^{2/\alpha} L^2(m)\right) = O_P(1)\left[\left(\frac{m}{n}\right)^{1-1/\alpha} \frac{L(n)}{L(m)}\right]^4. \tag{3.30}$$

The right hand side of (3.30) tends in probability to zero by (3.6).

LEMMA 3.7 *If the assumptions of Theorem 2.1 hold, then for any* $x > 0$

$$P_{\mathcal{E}}\left(\left|\sum_{2\le i\le m}\left(\tilde{X}_{i-1}^2 - \tilde{S}^2(i-1)\right)\right| > x m a_m^2\right) \xrightarrow{P} 0.$$

PROOF: Observe that

$$\tilde{X}_{i-1}^2 - \tilde{S}^2(i-1) = \left(\tilde{X}_{i-1} - \tilde{S}(i-1)\right)^2 + 2\left(\tilde{X}_{i-1} - \tilde{S}(i-1)\right)\tilde{S}(i-1).$$

Using Abel's summation, we obtain

$$\begin{aligned}\left|\tilde{X}_i - \tilde{S}(i)\right| &\le \max_{0\le k\le m}(1 + |\hat{\theta}_n - 1|)^k |\hat{\theta}_n - 1| \sum_{1\le j\le i} |\tilde{S}(j)| \\ &\le (1 + |\hat{\theta}_n - 1|)^m |\hat{\theta}_n - 1| \sum_{1\le j\le i} |\tilde{S}(j)|.\end{aligned}$$

By the weak convergence of $a_m^{-1}\tilde{S}(m\cdot)$ we have

$$P_{\mathcal{E}}\left(\sum_{1\le i\le m-1} |\tilde{S}(i)| \sum_{1\le j\le i} |\tilde{S}(j)| \ge x m^{2+2/\alpha} L^2(m)\right) \xrightarrow{P} P\left(\int_0^1 |U(t)| \int_0^t |U(s)| ds \ge x\right).$$

Since $|\hat{\theta}_n - 1| = O_p(n^{-1})$,

$$P_{\mathcal{E}}\left(\sum_{1\le i\le m-1}\left(\tilde{X}_{i-1} - \tilde{S}(i-1)\right)\tilde{S}(i-1) \ge x\right) \xrightarrow{P} 0.$$

A similar argument applies to $\sum_{2\le i\le m}\left(\tilde{X}_{i-1} - \tilde{S}(i-1)\right)^2$.

References

Arcones, M. and Giné, E. (1989). The bootstrap of the mean with arbitrary bootstrap sample size. *Annals of the Institute Henri Poincaré*, **22,** 457–481.

Arcones, M. and Giné, E. (1991). Additions and corrections to 'the bootstrap of the mean with arbitrary bootstrap sample size'. *Annals of the Institute Henri Poincaré*, **27,** 583–595.

Athreya, K. B. (1986). Bootstrap of the mean in the infinite variance case, ii. Technical Report 86-21. Department of Statistics, Iowa State University.

Athreya, K. B. (1987). Bootstrap of the mean in the infinite variance case. *The Annals of Statistics*, **15,** number 2, 724–731.

Basawa, I. V., Malik, A. K., McCormick, W. P. and Taylor, R. L. (1991). Bootstrap test of significance and sequential bootstrap estimation for unstable first order autoregressive processes. *Commun. Statist. – Theory Meth.*, **20,** 1015–1026.

Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation.* Cambridge University Press.

Chan, N. H. and Tran, L. T. (1989). On the first order autoregressive process with infinite variance. *Econometric Theory*, **5,** 354–362.

Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press, Princeton, NJ.

Heimann, G. and Kreiss, J-P. (1996). Bootstrapping general first order autoregression. *Statistics and Probability Letters*, **30,** 87–98.

Ling, S. and Li, W. K. (2002). Asymptotic inference for unit root processes with GARCH(1,1) errors. *Econometric Theory*, **00,**; fortcoming.

Mittnik, S. and Rachev, S. (2000). *Stable Paretian Models in Finance.* Wiley, New York.

Paparoditis, E. and Politis, D. N. (2001). Residual based block bootstrap for unit root testing. Technical Report. University of California, San Diego. Department of Economics Discussion paper No. 2001-6.

Petrov, V. V. (1995). *Limit Theorems of Probability Theory: Sequences of Independent Random Varianles.* Clarendon Press.

Philllips, P. C. B. and Xiao, Z. (1998). A primer on unit root testing. Technical Report 972. Cowles Foundation for Research in Economics. Available at `http://korora.econ.yale.edu/phillips.htm`.

Resnick, S. I. (1986). Point processes, regular variation and weak convergence. *Adv. Appl. Prob.*, **18,** 66–138.

Wang, Q., Lin, Y-X. and Gulati, C. M. (2002). Asymptotics for general fractionally integerated processes with applications to unit root tests. *Econometric Theory*, **00,**; Forthcoming.