

# Project Report on Resampling in Time Series Models

Submitted by: Abhishek Bhattacharya

Project Supervisor: Arup Bose

## Abstract

This project revolves around studying estimators for parameters in different Time Series models and studying their asymptotic properties. We introduce various bootstrap techniques for the estimators obtained. Our special emphasis is on Weighted Bootstrap. We establish the consistency of this scheme in a AR model and its variations. Numerical calculations lend further support to our consistency results. Next we analyze ARCH models, and study various estimators used for different error distributions. We also present resampling techniques for estimating the distribution of the estimators. Finally by simulating data, we analyze the numerical properties of the estimators.

## 1 Bootstrap in AR(1) model

Let  $X_t$  be a stationary AR(1) process, that is,

$$\begin{aligned} X_t &= \theta X_{t-1} + Z_t \text{ for } t = 1, 2, \dots \\ Z_t &\text{ iid } (0, \sigma^2); \quad EZ_t^4 < \infty; \quad |\theta| < 1. \end{aligned} \quad (1)$$

We have assumed  $\sigma$  to be known, and  $\theta$  is the unknown parameter of interest. Then the Least Squares estimate for  $\theta$  (which is approximately the MLE in case of normal errors) is given by

$$\hat{\theta}_n = \frac{\sum_{t=2}^n X_t X_{t-1}}{\sum_{t=2}^n X_{t-1}^2}$$

Then it can be established that

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} N(0, (1 - \theta^2)) \quad (2)$$

Let us introduce two particular bootstrap techniques specially used to estimate the distribution of  $\hat{\theta}_n$  from a realization of model (1).

**(a) Residual Bootstrap** Let  $\tilde{Z}_t = X_t - \hat{\theta}_n X_{t-1}$ ,  $t = 2, 3, \dots, n$  and let  $\hat{Z}_t$  be the standardized version of  $\tilde{Z}_t$  such that  $\frac{1}{n-1} \sum \hat{Z}_t = 0$  and  $\frac{1}{n-1} \sum \hat{Z}_t^2 = 1$ . Now we draw  $Z_t^*$ ,  $t = 1, 2, \dots, N$  with replacement from  $\hat{Z}_t$  and define

$$\begin{aligned} X_1^* &= Z_1^* \\ X_t^* &= \hat{\theta}_n X_{t-1}^* + Z_t^* \text{ , } t = 2, \dots, N. \end{aligned}$$

and form the statistic

$$\hat{\theta}_n^* = \frac{\sum_{t=2}^n X_t^* X_{t-1}^*}{\sum_{t=2}^n (X_{t-1}^*)^2} \quad (3)$$

Then (3) forms an estimator of  $\hat{\theta}_n$  and is called the **Residual Bootstrap estimator**. We repeat the simulation process several times to estimate the distribution of  $\hat{\theta}_n^*$ .

**(b) Weighted Bootstrap** Alternatively we define our resampling estimator

$$\hat{\theta}_n^* = \frac{\sum_{t=2}^n w_{nt} X_t X_{t-1}}{\sum_{t=2}^n w_{nt} (X_{t-1})^2} \quad (4)$$

where  $\{w_{nt}; 1 \leq t \leq n, n \geq 1\}$  is a triangular sequence of random variables, independent of  $\{X_t\}$ . These are the so called ‘‘Bootstrap weights’’, and the estimator (4) is the **Weighted Bootstrap Estimator**.

## 1.1 A Bootstrap Central limit theorem

Under suitable conditions on the weights to be stated below, we establish the distributional consistency of the Weighted Bootstrap Estimator,  $\hat{\theta}_n^*$  defined in (4). To establish consistency, we will prove a Bootstrap CLT for which we will need the following established results:

**Result 1 (P-W theorem; see Praestgaard and Wellner(1993))** *Let  $\{c_{nj}; j = 1, 2, \dots, n; n \geq 1\}$  be a triangular array of constants, and let  $\{U_{nj} \ j = 1, 2, \dots, n; n \geq 1\}$  be a triangular array of row exchangeable random variables such that as  $n \rightarrow \infty$ ,*

1.  $\frac{1}{n} \sum_{j=1}^n c_{nj} \rightarrow 0$
2.  $\frac{1}{n} \sum_{j=1}^n c_{nj}^2 \rightarrow \tau^2$
3.  $\frac{1}{n} \max_{1 \leq j \leq n} c_{nj}^2 \rightarrow 0$
4.  $E(U_{nj}) = 0 \quad j = 1, 2, \dots, n. \quad n \geq 1$
5.  $E(U_{nj}^2) = 1 \quad j = 1, 2, \dots, n. \quad n \geq 1$
6.  $\frac{1}{n} \sum_{j=1}^n U_{nj}^2 \xrightarrow{P} 1$
7.  $\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \sqrt{E(U_{nj}^2 I_{\{|U_{nj}| > k\}})} = 0$

Then under the above conditions,

$$\frac{1}{\sqrt{n}} \sum_{j=1}^n c_{nj} U_{nj} \xrightarrow{d} N(0, \tau^2) \quad (5)$$

Result (1) can be generalized by taking  $\{c_{nj}\}$  random variables, independent of  $\{U_{nj}\}$  and the conditions (1), (2) and (3) replaced by convergence in probability. In that case conclusion (5) is replaced by

$$P \left[ \frac{1}{\sqrt{n}} \sum_{j=1}^n c_{nj} U_{nj} \in C \mid \{c_{nj}; j = 1, \dots, n; n \geq 1\} \right] - P[Y \in C] = o_P(1) \quad (6)$$

where  $Y \sim N(0, \tau^2)$  and  $C \in \mathcal{B}(\mathbb{R})$  such that  $P(Y \in \partial C) = 0$ .

**Result 2** *Let  $\{X_1, X_2, \dots, X_n\}$  be the realization of the stationary AR(1) process (1). Then  $\frac{1}{n} \sum_{t=1}^{n-k} X_t^a Z_{t+k}^b \xrightarrow{a.s.} E(X_t^a Z_{t+k}^b)$  whenever  $E Z_t^{\max(a,b)} < \infty \quad \forall a, b, k \in \mathbb{Z}^+; a, b \geq 0; k > 0$ . This can be established using the Martingale SLLN; see Hall and Heyde 1980.*

Let us use the notations  $P_B, E_B, V_B$  to respectively denote probabilities, expectations and variances with respect to the distribution of the weights, conditioned on the given data  $\{X_1, \dots, X_n\}$ . The weights are assumed to be row exchangeable. We henceforth drop the first suffix in the weights  $w_{ni}$  and denote it by  $w_i$ . Let  $\sigma_n^2 = V_B(w_i), W_i = \sigma_n^{-1}(w_i - 1)$ . The following conditions on the row exchangeable weights are assumed:

- A1.**  $E_B(w_1) = 1$
- A2.**  $0 < k < \sigma_n^2 = o(n)$
- A3.**  $c_{1n} = Cov(w_1, w_2) = O(n^{-1})$
- A4.** Conditions of Result(1) hold with  $U_{nj} = W_{nj}$ .

**Theorem 1** Under the conditions (A1)-(A4) on the weights,

$$P_B \left[ \sqrt{n} \sigma_n^{-1} (\hat{\theta}_n^* - \hat{\theta}_n) \leq x | X_1, \dots, X_n \right] - P[Y \leq x] = o_P(1) \quad \forall x \in \mathbb{R} \quad (7)$$

where  $Y \sim N(0, (1 - \theta^2))$ .

**Proof** Note that

$$\begin{aligned} \hat{\theta}_n^* &= \frac{\sum_{t=2}^n w_t X_t X_{t-1}}{\sum_{t=2}^n w_t X_{t-1}^2} \\ &= \frac{\sum_{t=2}^n w_t X_{t-1} (\theta X_{t-1} + Z_t)}{\sum_{t=2}^n w_t X_{t-1}^2} \\ &= \theta + \frac{\sum w_t X_{t-1} Z_t}{\sum w_t X_{t-1}^2} \end{aligned}$$

Similarly

$$\hat{\theta}_n = \frac{\sum X_t X_{t-1}}{\sum X_{t-1}^2} = \theta + \frac{\sum X_{t-1} Z_t}{\sum X_{t-1}^2}$$

Hence

$$\begin{aligned} \hat{\theta}_n^* - \hat{\theta}_n &= \frac{\sum w_t X_{t-1} Z_t}{\sum w_t X_{t-1}^2} - \frac{\sum X_{t-1} Z_t}{\sum X_{t-1}^2} \\ &= \frac{\sum w_t X_{t-1} Z_t}{\sum w_t X_{t-1}^2} - \frac{\sum X_{t-1} Z_t}{\sum w_t X_{t-1}^2} + \frac{\sum X_{t-1} Z_t}{\sum w_t X_{t-1}^2} - \frac{\sum X_{t-1} Z_t}{\sum X_{t-1}^2} \\ &= \frac{\sum (w_t - 1) X_{t-1} Z_t}{\sum w_t X_{t-1}^2} - \frac{\sum X_{t-1} Z_t \sum (w_t - 1) X_{t-1}^2}{\sum X_{t-1}^2 \sum w_t X_{t-1}^2} \end{aligned}$$

Now using Result (2),

$$\frac{\sum X_{t-1} Z_t}{n} \xrightarrow{a.s.} E(X_{t-1} Z_t) = 0 \quad (8)$$

$$\begin{aligned} \frac{\sum X_{t-1}^2 Z_t^2}{n} &\xrightarrow{a.s.} E(X_{t-1}^2 Z_t^2) \\ &= \sigma^4 (1 - \theta^2)^{-1} \end{aligned} \quad (9)$$

**Claim 1.** For  $\tau^2 = \sigma^4 (1 - \theta^2)^{-1}$ ,

$$P_B \left[ \frac{1}{\sqrt{n}} \sum_{t=2}^n W_t X_{t-1} Z_t \leq x \middle| X_1, \dots, X_n \right] \xrightarrow{P} \Phi\left(\frac{x}{\tau}\right) \quad \forall x \in \mathbb{R}$$

To see this let us verify the conditions of Result(1) with  $c_{nj} = X_j Z_{j+1}$  and  $U_{nj} = W_j$  for  $j = 1, \dots, n-1$ .

1.  $\frac{1}{n} \sum_{t=2}^n X_{t-1} Z_t \xrightarrow{P} 0$

Follows from (8).

2.  $\frac{1}{n} \sum_{t=2}^n X_{t-1}^2 Z_t^2 \xrightarrow{P} \sigma^4 (1 - \theta^2)^{-1} (= \tau^2)$

Follows from (9).

3.  $n^{-1} \max(X_{t-1}^2 Z_t^2) \xrightarrow{P} 0$

**Proof** Let  $Y_t = X_{t-1}^2 Z_t^2 = X_t^2 X_{t-1}^2 - 2\theta X_t X_{t-1}^3 + \theta^2 X_{t-1}^4$

Then given  $\epsilon > 0$ ,

$$\begin{aligned} P(n^{-1} \max Y_t > \epsilon) &= P(\max Y_t > n\epsilon) \\ &\leq \sum_{t=1}^n P(Y_t > n\epsilon) \leq \sum \frac{EY_t^2}{n^2 \epsilon^2} = \frac{1}{n\epsilon^2} EY_t^2 \longrightarrow 0 \end{aligned}$$

as  $EY_t^2 = E(X_{t-1}^4 Z_t^4) < \infty$

Conditions (4), (5), (6) and (7) follow from definition and condition on the weights. This proves the claim.

Hence for  $\tau^2 = \sigma^4(1 - \theta^2)^{-1}$

$$P \left[ \frac{1}{\sqrt{n}} \sum_{t=2}^n W_t X_{t-1} Z_t \leq x \middle| X_1, \dots, X_n \right] \xrightarrow{P} \Phi \left( \frac{x}{\tau} \right) \quad \forall x \in \mathbb{R} \quad (10)$$

**Claim 2.** With  $c = \sigma^2(1 - \theta^2)^{-1}$ ,

$$P_B \left[ \left| \frac{1}{n} \sum_{t=2}^n w_t X_{t-1}^2 - c \right| > \epsilon \right] \xrightarrow{P} 0 \quad \forall \epsilon > 0$$

**Proof**

$$\begin{aligned} E_B \left( \frac{1}{n} \sum w_t X_{t-1}^2 \right) &= \frac{1}{n} \sum X_{t-1}^2 \\ V_B \left( \sum w_t X_{t-1}^2 \right) &= \sum X_{t-1}^4 \sigma_n^2 + \sum_{s \neq t} X_{t-1}^2 X_{s-1}^2 \text{Cov}(w_t, w_s) \\ &= \sigma_n^2 \sum X_{t-1}^4 + c_{1n} \sum_{s \neq t} X_{t-1}^2 X_{s-1}^2 \end{aligned}$$

Therefore

$$\begin{aligned} V_B \left( \frac{1}{n} \sum w_t X_{t-1}^2 \right) &= \frac{\sigma_n^2}{n^2} \sum X_{t-1}^4 + \frac{c_{1n}}{n^2} \sum_{s \neq t} X_{t-1}^2 X_{s-1}^2 \\ \frac{1}{n} \sigma_n^2 &\rightarrow 0 \\ \frac{1}{n} \sum X_{t-1}^4 &\xrightarrow{a.s.} E(X_t^4) \end{aligned} \quad (11)$$

Hence the first term in (11)  $\xrightarrow{a.s.} 0$

Also

$$\frac{1}{n^2} \sum_{s \neq t} X_{t-1}^2 X_{s-1}^2 \leq \left( \frac{\sum X_t^2}{n} \right)^2 \xrightarrow{a.s.} (EX_t^2)^2$$

Hence  $\frac{1}{n^2} \sum_{s \neq t} X_{t-1}^2 X_{s-1}^2$  is bounded a.s., and as  $c_{1n} \rightarrow 0$ , the second term in (11) also  $\rightarrow 0$  a.s.

This shows that  $V_B \left( \frac{1}{n} \sum w_t X_{t-1}^2 \right) \rightarrow 0$  a.s.

Hence  $\frac{1}{n} \sum w_t X_{t-1}^2 - \frac{1}{n} \sum X_{t-1}^2 \xrightarrow{P_B} 0$  a.s.

Using Result (2),  $\frac{1}{n} \sum X_{t-1}^2 \xrightarrow{a.s.} E(X_t^2) = \sigma^2(1 - \theta^2)^{-1}$

This implies,  $\frac{1}{n} \sum w_t X_{t-1}^2 \xrightarrow{P_B} \sigma^2(1 - \theta^2)^{-1}$  a.s.

This proves Claim 2.

In fact we have proved that, with  $c = \sigma^2(1 - \theta^2)^{-1}$

$$P_B \left[ \left| \frac{1}{n} \sum_{t=2}^n w_t X_{t-1}^2 - c \right| > \epsilon \right] \xrightarrow{a.s.} 0 \quad \forall \epsilon > 0. \quad (12)$$

Now

$$\begin{aligned} & \sqrt{n} \sigma_n^{-1} (\hat{\theta}_n^* - \hat{\theta}_n) \\ &= \sqrt{n} \sigma_n^{-1} \frac{\sum (w_t - 1) X_{t-1} Z_t}{\sum w_t X_{t-1}^2} - \sqrt{n} \sigma_n^{-1} \frac{\sum X_{t-1} Z_t \sum (w_t - 1) X_{t-1}^2}{\sum X_{t-1}^2 \sum w_t X_{t-1}^2} \\ &= \frac{\sum W_t X_{t-1} Z_t / \sqrt{n}}{\sum w_t X_{t-1}^2 / n} - \sqrt{n} (\hat{\theta}_n - \theta) \sigma_n^{-1} \frac{\sum (w_t - 1) X_{t-1}^2 / n}{\sum w_t X_{t-1}^2 / n} \\ &= T_1 - T_2 \quad (\text{say}) \end{aligned} \quad (13)$$

Then from (10) and (12),  $P_B(T_1 \leq x) - P(T \leq x) = o_P(1)$ , where,

$$T \sim \frac{1}{\sigma^2(1 - \theta^2)^{-1}} N(0, \sigma^4(1 - \theta^2)^{-1}) = N(0, (1 - \theta^2)) \quad (14)$$

**Claim 3.** Define  $A \equiv \sqrt{n}(\hat{\theta}_n - \theta) \sigma_n^{-1} \frac{1}{n} \sum (w_t - 1) X_{t-1}^2$ .

Then  $\forall \epsilon > 0$ ,  $P_B(|A| > \epsilon) \xrightarrow{P} 0$ .

**Proof** Note that,

$$E_B(A) = 0 \quad (15)$$

$$V_B(A) = \frac{n}{\sigma_n^2} (\hat{\theta}_n - \theta)^2 \left[ \frac{\sigma_n^2}{n^2} \sum X_{t-1}^4 + \frac{c_{1n}}{n^2} \sum_{s \neq t} X_{s-1}^2 X_{t-1}^2 \right] \quad (16)$$

$$= (\hat{\theta}_n - \theta)^2 \frac{\sum X_{t-1}^4}{n} + \frac{nc_{1n}}{\sigma_n^2} (\hat{\theta}_n - \theta)^2 \frac{\sum_s \sum_{s \neq t} X_{s-1}^2 X_{t-1}^2}{n^2} \quad (17)$$

$$= A_1 + A_2 \quad (\text{say}) \quad (18)$$

$\frac{\sum X_{t-1}^4}{n}$  converges a.s., and from (2),  $(\hat{\theta}_n - \theta) \xrightarrow{P} 0$ , as a result,  $A_1 \xrightarrow{P} 0$ .

Moreover  $\frac{\sum_{s \neq t} X_{s-1}^2 X_{t-1}^2}{n^2}$  is bounded a.s.,  $nc_{1n}$  is bounded and  $\sigma_n^2$  is bounded away from 0. As a result  $A_2 \xrightarrow{P} 0$ .

Combining,  $V_B(A) \xrightarrow{P} 0$ .

Hence

$$P_B(|A| > \epsilon) \leq \frac{V_B(A)}{\epsilon^2} \xrightarrow{P} 0$$

$$\text{Now } T_2 = \frac{A}{\sum w_t X_{t-1}^2 / n}.$$

From (12), we have,  $\sum w_t X_{t-1}^2 / n$  is bounded away from zero in  $P_B$  a.s., which means that,  $\forall \epsilon > 0$ ,

$$P_B(|T_2| > \epsilon) = o_P(1) \quad (19)$$

Hence from (13), (14) and (19), we have,

$$\mathbf{P}_B[\sqrt{n}\sigma_n^{-1}(\hat{\theta}_n^* - \hat{\theta}_n) \leq \mathbf{x}] - \mathbf{P}[Y \leq \mathbf{x}] = o_P(\mathbf{1}) \quad \forall \mathbf{x} \in \mathbb{R} \quad (20)$$

where  $Y \sim N(0, (1 - \theta^2))$  and this was what was to be proved.

## 1.2 Least Absolute Deviations Estimator

Another estimator of  $\theta_0$  can be the LAD estimator, that is,

$$\hat{\theta}_2 = \arg \min_{\theta} \frac{1}{n} \sum_2^n \left| X_t - \theta X_{t-1} \right|$$

Now we reparametrize the model (1) in such a way that the median of  $Z_t$ , instead of the mean is equal to 0, while  $VZ_t = \sigma^2$  remains unchanged.

## 1.3 Distributional Consistency of the LAD estimator

Under the following assumptions we establish the asymptotic normality of  $\hat{\theta}_2$ .

**A1.** CDF of  $Z_t$ ,  $F$  has a pdf  $f$ , which is continuous at zero.

**A2.**  $|F(x) - F(0) - xf(0)| \leq c|x|^{1+\alpha}$  in a neighborhood of zero, say  $|x| \leq M$ , where  $c, \alpha, M > 0$ .

To do so we use the the following result on random convex functions.

**Result 3 (See Niemire (1992))** *Suppose that  $h_n(a)$ ,  $a \in R^d$  is a sequence of random convex functions which converge in probability to  $h(a)$  for every fixed  $a$ . Then this convergence is uniform on any compact set containing  $a$ .*

**Theorem 2** *Under the conditions (A1)-(A2),  $\sqrt{n}(\hat{\theta}_2 - \theta_0) \xrightarrow{d} N(0, \frac{1}{4f^2(0)EX_t^2})$  as  $n \rightarrow \infty$ .*

**Proof** Define

$$\begin{aligned} f(X_t, \theta) &= (|X_t - \theta X_{t-1}| - |X_t|) \\ g(X_t, \theta) &= X_{t-1}[2I(Z_t(\theta) \leq 0) - 1] \\ \text{where } Z_t(\theta) &= X_t - \theta X_{t-1} \text{ for } t = 2, \dots, n. \\ Y_t(a) &= f(X_t, \theta_0 + n^{-1/2}a) - f(X_t, \theta_0) - n^{-1/2}ag(X_t, \theta_0) \\ &= |Z_t - n^{-1/2}aX_{t-1}| - |Z_t| - n^{-1/2}aX_{t-1}[2I(Z_t \leq 0) - 1] \text{ for } a \in \mathbb{R}. \end{aligned}$$

Also define

$$\begin{aligned} Q_n(\theta) &= \sum f(X_t, \theta) \\ U_n &= \sum g(X_t, \theta_0) \\ V_n &= \sum Y_t(a) = Q_n(\theta_0 + n^{-1/2}a) - Q_n(\theta_0) - n^{-1/2}aU_n \end{aligned}$$

**Step1**  $\sum_{t=2}^n Y_t(a) \xrightarrow{P} a^2 f(0) EX_1^2$

**Step1.1**  $\sum(Y_t - E(Y_t|\mathcal{A}_{t-1})) \xrightarrow{P} 0$

$$\begin{aligned} E(Y_t - E(Y_t|\mathcal{A}_{t-1})) &= 0 \\ V(\sum Y_t - E(Y_t|\mathcal{A}_{t-1})) &= \sum V(Y_t - E(Y_t|\mathcal{A}_{t-1})) \leq \sum V(Y_t) \leq \sum EY_t^2 \end{aligned}$$

By convexity of  $f$ ,

$$0 \leq Y_t(a) \leq n^{-1/2}a[g(X_t, \theta_0 + n^{-1/2}a) - g(X_t, \theta_0)]$$

Therefore

$$\begin{aligned} E(Y_t^2) &\leq \frac{a^2}{n}E[g(X_t, \theta_0 + n^{-1/2}a) - g(X_t, \theta_0)]^2 \\ &= 4\frac{a^2}{n}EX_{t-1}^2[I(Z_t - n^{-1/2}aX_{t-1} \leq 0) - I(Z_t \leq 0)]^2 \end{aligned}$$

Now

$$\sum EY_t^2 = nEY_2^2 \leq 4aEX_1^2[I(Z_2 - n^{-1/2}aX_1 \leq 0) - I(Z_2 \leq 0)]^2$$

which tends to zero using DCT. Therefore

$$V(\sum(Y_t - E(Y_t|\mathcal{A}_{t-1}))) \rightarrow 0$$

This establishes Step 1.1.

**Step1.2**  $\sum E(Y_t|\mathcal{A}_{t-1}) - a^2f(0)EX_1^2 \xrightarrow{P} 0$

$$\begin{aligned} E(Y_t|\mathcal{A}_{t-1}) &= E(|Z_t - n^{-1/2}aX_{t-1}||\mathcal{A}_{t-1}) - E|Z_t| \\ &= \int (|z - n^{-1/2}aX_{t-1}| - |z|)dF(z) \end{aligned}$$

Using the representation,

$$|x - \theta| - |x| = \theta[2I(x \leq 0) - 1] + 2 \int_0^\theta [I(x \leq s) - I(x \leq 0)]ds$$

we have

$$|z - n^{-1/2}aX_{t-1}| - |z| = n^{-1/2}aX_{t-1}[2I(z \leq 0) - 1] + 2 \int_0^{n^{-1/2}aX_{t-1}} [I(z \leq s) - I(z \leq 0)]ds$$

Therefore

$$\begin{aligned} E(Y_t|\mathcal{A}_{t-1}) &= n^{-1/2}aX_{t-1} \int [2I(z \leq 0) - 1]dFz + 2 \int \int_0^{n^{-1/2}aX_{t-1}} [I(z \leq s) - I(z \leq 0)]dsdFz \\ &= 2 \int_0^{n^{-1/2}aX_{t-1}} [F(s) - F(0)]ds \end{aligned} \quad (22)$$

$$= 2n^{-1/2}X_{t-1} \int_0^a [F(n^{-1/2}X_{t-1}x) - F(0)]dx \quad (23)$$

Under assumption A2,

$$\begin{aligned} F(n^{-1/2}X_{t-1}x) - F(0) &= n^{-1/2}X_{t-1}xf(0) + R_{nt}(x) \\ \text{where } |R_{nt}(x)| &\leq cn^{-(1+\alpha)/2}|X_{t-1}|^{1+\alpha}|x|^{1+\alpha} \\ \text{whenever } n^{-1/2}|X_{t-1}||x| &\leq M \end{aligned}$$

Hence

$$\begin{aligned}
E(Y_t|\mathcal{A}_{t-1}) &= 2n^{-1/2}X_{t-1} \int_0^a [n^{-1/2}X_{t-1}xf(0) + R_{nt}(x)]dx \\
&= \frac{1}{n}X_{t-1}^2a^2f(0) + 2n^{-1/2}X_{t-1} \int_0^a R_{nt}(x)dx \\
\sum E(Y_t|\mathcal{A}_{t-1}) &= a^2f(0)\frac{1}{n} \sum X_{t-1}^2 + \frac{2}{n} \sum X_{t-1} \int_0^a \sqrt{n}R_{nt}(x)dx \\
&= I_1 + I_2 \text{ (say)}
\end{aligned}$$

Then  $I_1 \xrightarrow{P} a^2f(0)EX_1^2$ .

Remains to show  $I_2 \xrightarrow{P} 0$ . To show this, let us assume:

1.  $\max_{1 \leq t \leq n} n^{-1/2}|X_{t-1}| \xrightarrow{P} 0$
2.  $\frac{1}{n^{1+\alpha/2}} \sum |X_{t-1}|^{2+\alpha} \xrightarrow{P} 0$

Hence given  $\epsilon > 0$ ,

$$\begin{aligned}
P(\max n^{-1/2}|X_{t-1}| \leq M/|a|) &\rightarrow 1 \\
\text{and } P\left(\frac{c}{n^{1+\alpha/2}} \sum |X_{t-1}|^{2+\alpha} < \epsilon\right) &\rightarrow 1
\end{aligned}$$

Let  $A_n$  be the set where  $\max n^{-1/2}|X_{t-1}| \leq M/|a|$  and  $\frac{c}{n^{1+\alpha/2}} \sum |X_{t-1}|^{2+\alpha} < \epsilon$ .

Then  $\exists N$  such that  $P(A_n) > 1 - \epsilon \forall n \geq N$ . Then on  $A_n$ ,  $|R_{nt}| \leq cn^{-\alpha/2}|X_{t-1}|^{1+\alpha}$ , and hence

$$\begin{aligned}
|I_2| &\leq \frac{2}{n} \sum |X_{t-1}| \int_0^a cn^{-\alpha/2}|X_{t-1}|^{1+\alpha} \\
&\leq \frac{c}{n^{1+\alpha/2}} \sum |X_{t-1}|^{2+\alpha} \\
&< \epsilon
\end{aligned}$$

ie  $P(|I_2| < \epsilon) \rightarrow 1 \forall \epsilon > 0$ . In otherwords  $I_2 \xrightarrow{P} 0$ . This completes Step1.2 and hence Step1. In other words,

$$Q_n(\theta_0 + n^{-1/2}a) - Q_n(\theta_0) - n^{-1/2}aU_n - a^2f(0)EX_1^2 \xrightarrow{P} 0 \quad (24)$$

Due to convexity of  $Q_n$ , the convergence in (24) is uniform on any compact set by Result3. Thus  $\forall \epsilon > 0$ , and  $M > 0$ , for  $n$  sufficiently large, we have

$$P \left[ \sup_{|a| \leq M} \left| Q_n(\theta_0 + n^{-1/2}a) - Q_n(\theta_0) - n^{-1/2}aU_n - a^2f(0)EX_1^2 \right| < \epsilon \right] \geq 1 - \epsilon/2$$

Call

$$\begin{aligned}
A_n(a) &= Q_n(\theta_0 + n^{-1/2}a) - Q_n(\theta_0) , \\
B_n(a) &= n^{-1/2}aU_n + a^2f(0)EX_1^2
\end{aligned}$$

and their minimizers  $a_n$  and  $b_n$  respectively. Then

$$\begin{aligned}
a_n &= \sqrt{n}(\hat{\theta}_2 - \theta_0) \text{ and} \\
b_n &= -(2f(0)EX_1^2)^{-1}n^{-1/2}U_n
\end{aligned}$$

The minimum value of  $B_n$ ,

$$B_n(b_n) = -n^{-1}(4f(0)EX_1^2)^{-1}U_n^2$$

Note that  $b_n$  is bounded in probability. Hence there exists  $M > 0$  such that

$$P \left[ |-(2f(0)EX_1^2)^{-1}n^{-1/2}U_n| < M - 1 \right] \geq 1 - \epsilon/2$$

Let  $A$  be the set where,

$$\begin{aligned} & \sup_{|a| \leq M} |A_n(a) - B_n(a)| < \epsilon \\ \text{and} \quad & |-(2f(0)EX_1^2)^{-1}n^{-1/2}U_n| < M - 1 \end{aligned}$$

Then  $P(A) > 1 - \epsilon$ . On  $A$ ,

$$A_n(b_n) < B_n(b_n) + \epsilon \tag{25}$$

Consider the value of  $A_n$  on the sphere  $S_n = \{a : |a - b_n| = k\epsilon^{1/2}\}$  where  $k$  will be chosen later. By choosing  $\epsilon$  sufficiently small, we have  $|a| \leq M \forall a \in S_n$ . Hence

$$A_n(a) > B_n(a) - \epsilon \forall a \in S_n. \tag{26}$$

Once we chose  $k = 2(2f(0)EX_1^2)^{-1/2}$ ,

$$B_n(a) > B_n(b_n) + 2\epsilon \forall a \in S_n \tag{27}$$

Comparing the bounds (25) and (26), we have  $A_n(a) > A_n(b_n)$  whenever  $a \in S_n$ . If  $|a_n - b_n| > k\epsilon^{1/2}$ , by convexity of  $A_n$ , there exists  $a_n^*$  on  $S_n$  such that  $A_n(a_n^*) \leq A_n(b_n)$  which cannot be the case. Therefore  $|a_n - b_n| < k\epsilon^{1/2}$  on  $A$ . Since this holds with probability atleast  $1 - \epsilon$  and  $\epsilon$  is arbitrary, this proves that  $|a_n - b_n| \xrightarrow{P} 0$ . In other words,

$$\sqrt{n}(\hat{\theta}_2 - \theta_0) = -n^{-1/2}(2f(0)EX_1^2)^{-1}U_n + o_P(1) \tag{28}$$

**Step 2**  $n^{-1/2}U_n \xrightarrow{d} N(0, EX_1^2)$

$$\begin{aligned} U_n &= \sum_{t=2}^n X_{t-1} [2I(Z_t \leq 0) - 1] \\ &= \sum_{t=2}^n Y_t \text{ (say)} \end{aligned}$$

Then note that  $U_n$  is a 0-mean martingale with finite variance increments. Hence to prove Step2, we use the Martingale CLT. Write

$$\begin{aligned} S_n^2 &= \sum_{t=2}^n E(Y_t^2 | \mathcal{A}_{t-1}) = \sum_{t=2}^n X_{t-1}^2 \\ \text{and } s_n^2 &= ES_n^2 = (n-1)EX_1^2 \end{aligned}$$

Then we need to to verify:

1.  $\frac{S_n^2}{s_n^2} \xrightarrow{P} 1$

This follows from Result 2.

2.  $s_n^{-2} \sum_{t=2}^n E(Y_t^2 I(|Y_t| \geq \epsilon s_n)) \longrightarrow 0$  as  $n \rightarrow \infty \forall \epsilon > 0$ .

To see this, note that

$$\begin{aligned} L.H.S. &= \frac{1}{EX_1^2} E(X_1^2 I \left( \frac{|X_1|}{\sqrt{EX_1^2}} \geq \epsilon \sqrt{n-1} \right)) \\ &\longrightarrow 0 \text{ as } EX_1^2 < \infty \end{aligned}$$

Hence using Result 4, we have  $\frac{U_n}{s_n} \xrightarrow{d} N(0, 1)$ , which proves Step2. Combining Step2 and equation(28), we get,

$$\sqrt{n}(\hat{\theta}_2 - \theta_0) \xrightarrow{d} N\left(0, \frac{1}{4f^2(0)EX_1^2}\right)$$

and this was what was to be proved. Finally it remains to verify:

1.  $\max_{2 \leq t \leq n} n^{-1/2}|X_{t-1}| \xrightarrow{P} 0$

Proof: Given  $\epsilon$  positive,

$$\begin{aligned} P(\max_t n^{-1/2}|X_{t-1}| > \epsilon) &\leq \sum_{t=1}^{n-1} P(|X_t| > \epsilon\sqrt{n}) \\ &= (n-1)P(|X_1| > \epsilon\sqrt{n}) \\ &= (n-1) \int I(|X_1| > \epsilon\sqrt{n})dP \\ &\leq (n-1) \int \frac{|X_1|^2}{\epsilon^2 n} I(|X_1| > \epsilon\sqrt{n})dP \\ &= \frac{1}{\epsilon^2} \int |X_1|^2 I(|X_1| > \epsilon\sqrt{n})dP \\ &\rightarrow 0 \text{ as } E|X_1|^2 < \infty \end{aligned}$$

2.  $\frac{1}{n^{1+\alpha/2}} \sum_{t=2}^n |X_{t-1}|^{2+\alpha} \xrightarrow{P} 0$

Proof:

$$\begin{aligned} \frac{1}{n^{1+\alpha/2}} \sum_{t=2}^n |X_{t-1}|^{2+\alpha} &\leq \frac{\max_{1 \leq t \leq n-1} |X_t|^\alpha}{n^{1+\alpha/2}} \frac{1}{n} \sum X_{t-1}^2 \\ &\leq \left(\frac{\max |X_t|}{\sqrt{n}}\right)^\alpha \frac{1}{n} \sum X_{t-1}^2 \\ &\xrightarrow{P} 0 \end{aligned}$$

This follows from (1) and the fact that  $\frac{1}{n} \sum X_{t-1}^2$  is bounded in probability, since  $EX_1^2 < \infty$ . This completes the proof.

#### 1.4 WBS for LAD estimators

Now we define the weighted bootstrap estimators,  $\hat{\theta}_2^*$  of  $\hat{\theta}_2$  as the minimizers of

$$Q_{nB}(\theta) = \sum_{t=2}^n w_{nt} |X_t - \theta X_{t-1}| \tag{29}$$

In the next section, we deduce the consistency of this bootstrap procedure.

#### 1.5 Consistency of the Weighted Bootstrap technique

Now we prove that the Weighted Bootstrap estimator of  $\hat{\theta}_2$  is asymptotically normal with the same asymptotic distribution. In particular WB provides a consistent resampling scheme to estimate the LAD estimator.

**Theorem 3** Let  $\hat{\theta}_2^*$  be the weighted bootstrap estimator of  $\hat{\theta}_2$  as defined in (29). Suppose the bootstrap weights satisfy conditions (A1)-(A4). Also assume that  $n^{-1/2}\sigma_n \max_t |X_t| \xrightarrow{P} 0$ . Then

$$\sup_{x \in \mathbb{R}} \left| P \left[ \sqrt{n}\sigma_n^{-1}(\hat{\theta}_2^* - \hat{\theta}_2) \leq x \mid X_1, \dots, X_n \right] - P[Y \leq x] \right| = o_P(1) \quad (30)$$

where  $Y \sim N\left(0, \frac{1}{4f^2(0)EX_1^2}\right)$ .

**Proof** Define

$$\begin{aligned} U_{nt}(a) &= f(X_t, \theta_0 + n^{-1/2}\sigma_n a) - f(X_t, \theta_0) - n^{-1/2}\sigma_n a g(X_t, \theta_0) \\ U_{nBt}(a) &= w_{nt} U_{nt}(a) \\ S_{nB} &= \sum W_{nt} g(X_t, \theta_0) \\ S_{nw} &= \sum w_{nt} g(X_t, \theta_0) \\ S_n &= \sum g(X_t, \theta_0) \\ H &= 2f(0)EX_1^2 \end{aligned}$$

Then

$$\begin{aligned} E_B U_{nBt} &= U_{nt} \text{ and} \\ \sum U_{nBt}(a) &= Q_{nB}(\theta_0 + n^{-1/2}\sigma_n a) - Q_{nB}(\theta_0) - n^{-1/2}\sigma_n a S_{nw} \end{aligned}$$

**Step1.** We show  $\sqrt{n}\sigma_n^{-1}(\hat{\theta}_2^* - \hat{\theta}_2) = -n^{-1/2}H^{-1}S_{nB} + r_{nB}$  s.t. given  $\epsilon > 0$ ,  $P_B[|r_{nB}| > \epsilon] = o_P(1)$ . To show this, choose  $k = 3H^{-1/2}$  and  $\epsilon$  small enough such that  $k^2\epsilon < 1$  and  $M$  a sufficiently large constant. Let  $\mathcal{A}$  be the set where

$$\begin{aligned} \sup_{|a| \leq M} \sigma_n^2 \left| Q_{nB}(\theta_0 + n^{-1/2}\sigma_n a) - Q_{nB}(\theta_0) - n^{-1/2}\sigma_n a S_{nw} - \frac{\sigma_n^2}{2} a^2 H \right| &< \epsilon \\ \text{and } \left| n^{-1/2}\sigma_n^{-1} H^{-1} S_{nw} \right| &< M - 1 \end{aligned}$$

Then due to convexity of  $Q_{nB}$ , arguing as in the proof of Theorem2, on  $\mathcal{A}$  we have,

$$\begin{aligned} \sqrt{n}\sigma_n^{-1}(\hat{\theta}_2^* - \theta_0) &= -n^{-1/2}\sigma_n^{-1} H^{-1} S_{nW} + r_{nB} \\ \text{s.t. } |r_{nB}| &< k\epsilon^{1/2} \end{aligned}$$

If we show  $1 - P_B[\mathcal{A}] = o_P(1)$ , then

$$P_B[|r_{nB}| > \delta] = o_P(1) \quad \forall \delta > 0$$

Also from equation(28);  $\sqrt{n}(\hat{\theta}_2 - \theta_0) = -n^{-1/2}H^{-1}S_n + o_P(1)$ .

Therefore  $\sqrt{n}\sigma_n^{-1}(\hat{\theta}_2^* - \hat{\theta}_2) = -n^{-1/2}H^{-1}S_{nB} + r_{nB2}$  s.t. given  $\epsilon > 0$ ,  $P_B[|r_{nB2}| > \epsilon] = o_P(1)$ . This will complete Step1.

Hence it remains to show,  $1 - P_B[\mathcal{A}] = o_P(1)$

To show this we show,

$$\forall \quad M > 0, P_B \left[ \sup_{|a| \leq M} \sigma_n^{-2} \left| \sum U_{nBt}(a) - \frac{\sigma_n^2}{2} a^2 H \right| > \epsilon \right] = o_P(1) \quad (31)$$

$$\text{and there exists } M > 0 \text{ s.t. } P_B \left[ |\sigma_n^{-1} n^{-1/2} H^{-1} S_{nw}| \geq M \right] = o_P(1) \quad (32)$$

To show (31), note that,

$$\begin{aligned}
& P_B \left[ \sup_{|a| \leq M\sigma_n^{-2}} \left| \sum U_{nBt}(a) - \frac{\sigma_n^2}{2} a^2 H \right| > \epsilon \right] \\
& \leq \sum_j P_B[\sigma_n^{-1} \left| \sum_t W_t U_t(b_j) \right| > \epsilon/2] + \sum_j I(\sigma_n^{-2} \left| \sum X_t(b_j) - \sigma_n^2 b_j^2 H/2 \right| > \epsilon/2) \\
& \leq \sigma_n^{-2} \sum_j k \sum_t U_t^2(b_j) + \sum_j I(\sigma_n^{-2} \left| \sum_t U_t(b_j) - \sigma_n^2 b_j^2 H/2 \right| > \epsilon/2)
\end{aligned}$$

As a result, we need to show for fixed  $b$ ,

$$\sigma_n^{-2} \sum_t U_{nt}^2(b) = o_P(1) \quad (33)$$

$$\text{and } \sigma_n^{-2} \left[ \sum_t U_{nt}(b) - \sigma_n^2 b^2 H/2 \right] = o_P(1) \quad (34)$$

To see (33),

$$\begin{aligned}
\sigma_n^{-2} \sum_t E U_{nt}^2(b) &= n \sigma_n^{-2} E U_1^2(b) \\
&\leq n \sigma_n^{-2} E [f(X_1, \theta_0 + n^{-1/2} \sigma_n b) - f(X_1, \theta_0) - n^{-1/2} \sigma_n b g(X_1, \theta_0)]^2 \\
&\leq E b^2 [g(X_1, \theta_0 + n^{-1/2} \sigma_n b) - g(X_1, \theta_0)]^2 \\
&\rightarrow 0
\end{aligned}$$

This proves (33).

To prove (34) note that,

$$\begin{aligned}
& \sigma_n^{-2} \left[ \sum U_t(b) - \sigma_n^2 b^2 H/2 \right] = \sigma_n^{-2} \left[ \sum [U_t(b) - E(U_t(b)|\mathcal{A}_{t-1})] + \sum E(U_t(b)|\mathcal{A}_{t-1}) - \sigma_n^2 b^2 H/2 \right] \\
& E \sigma_n^{-2} \sum (U_t(b) - E(U_t(b)|\mathcal{A}_{t-1})) = 0 \\
& V[\sigma_n^{-2} \sum (U_t(b) - E(U_t(b)|\mathcal{A}_{t-1}))] = \sigma_n^{-4} \sum V(U_t - E(U_t|\mathcal{A}_{t-1})) \\
& \leq \sigma_n^{-4} \sum V(U_t(b)) \\
& \leq k_1^{-1} \sigma_n^{-2} \sum E(U_t^2(b)) \quad (\sigma_n^2 > k_1) \\
& = n k_1^{-1} \sigma_n^{-2} E(U_1(b))^2 \\
& \leq k_1^{-1} \sigma_n^{-2} \sigma_n^2 b^2 E[g(X_1, \theta_0 + n^{-1/2} \sigma_n b) - g(X_1, \theta_0)]^2 \\
& = \frac{1}{k_1} E[g(X_1, \theta_0 + n^{-1/2} \sigma_n b) - g(X_1, \theta_0)]^2 \\
& \rightarrow 0
\end{aligned}$$

Hence

$$\sigma_n^{-2} \left[ \sum (U_1(b) - E(U_1(b)|\mathcal{A}_{t-1})) \right] \xrightarrow{P} 0 \quad (35)$$

$$\sigma_n^{-2} \sum E(U_t | \mathcal{A}_{t-1}) = \sigma_n^{-2} \sum 2n^{-1/2} X_{t-1} \int_0^{\sigma_n b} [F(n^{-1/2} \sigma_n X_{t-1} x) - F(0)] dx \quad (36)$$

$$= 2n^{-1/2} \sigma_n^{-1} \sum X_{t-1} \int_0^b [F(n^{-1/2} \sigma_n X_{t-1} x) - F(0)] dx \quad (37)$$

$$= 2n^{-1/2} \sigma_n^{-1} \sum X_{t-1} \left[ n^{-1/2} \sigma_n \frac{b^2}{2} X_{t-1} f(0) + R_{nt} \right] \quad (38)$$

$$\text{where } |R_{nt}| \leq c |n^{-1/2} \sigma_n X_{t-1}|^{1+\alpha} = c (n^{-1/2} \sigma_n)^{1+\alpha} |X_{t-1}|^{1+\alpha} \quad (39)$$

$$= 2 \frac{b^2}{2} f(0) \frac{1}{n} \sum X_{t-1}^2 + 2n^{-1/2} \sigma_n^{-1} \sum X_{t-1} R_{nt} \quad (40)$$

$$= I_1 + I_2 \quad (\text{say}) \quad (41)$$

Here (36) follows from (23), and (39) from assumption A2 on  $F$  and the assumption  $n^{-1/2} \sigma_n \max_t |X_t| \xrightarrow{P} 0$ .

$$\text{Now } I_1 \xrightarrow{P} b^2 f(0) E X_1^2 = b^2 H/2 \quad (42)$$

$$\text{and } |I_2| \leq c (n^{-1/2} \sigma_n^{-1}) (n^{-1/2} \sigma_n)^{1+\alpha} \sum |X_{t-1}|^{2+\alpha} \quad (43)$$

$$= \frac{c \sigma_n^\alpha}{n^{1+\alpha/2}} \sum |X_{t-1}|^{2+\alpha} \quad (44)$$

$$\xrightarrow{P} 0 \quad (45)$$

In this case (45) follows from (44) if we show  $\frac{\sigma_n^\alpha}{n^{1+\alpha/2}} \sum |X_{t-1}|^{2+\alpha} \xrightarrow{P} 0 \forall \alpha > 0$ . To see this, note that

$$\begin{aligned} \frac{\sigma_n^\alpha}{n^{1+\alpha/2}} \sum |X_{t-1}|^{2+\alpha} &\leq \frac{\sigma_n^\alpha \max_t |X_t|^\alpha}{n^{\alpha/2}} \frac{1}{n} \sum X_t^2 \\ &\leq \left( \frac{\sigma_n}{\sqrt{n}} \max_{1 \leq t \leq n} |X_t| \right)^\alpha \frac{1}{n} \sum X_t^2 \\ &\xrightarrow{P} 0 \end{aligned}$$

Combining (42) and (45), we have  $\sigma_n^{-2} \sum E(U_t | \mathcal{A}_{t-1}) \xrightarrow{P} b^2 H/2$ . In o.w.,

$$\sigma_n^{-2} \left[ \sum E(U_t | \mathcal{A}_{t-1}) - \sigma_n^2 b^2 H/2 \right] \xrightarrow{P} 0 \quad (46)$$

Adding (35) and (46), we prove (34). And from (33) and (34) we deduce (31).

$$\begin{aligned} \text{Now } P_B \left[ |\sigma_n^{-1} n^{-1/2} H^{-1} S_{nw}| \geq M \right] \\ &\leq \frac{\sigma_n^{-2} n^{-1} H^{-2}}{M^2} E_B \left[ \sum w_t g(X_t, \theta_0) \right]^2 \\ &\leq \frac{K_1}{M^2 n} \sum g(X_t, \theta_0)^2 + \frac{K_2}{M^2} \left[ \frac{\sum g(X_t, \theta_0)}{\sqrt{n}} \right]^2 \\ &\xrightarrow{P} 0 \end{aligned}$$

if  $M$  is chosen sufficiently large. This proves (32). (31) and (32) together show  $1 - P_B[\mathcal{A}] = o_P(1)$ . This completes step 1.

**Step2.**  $P_B(n^{-1/2}S_{nB} \leq x) - P(Y \leq x) = o_P(1)$ , where  $Y \sim N(0, EX_1^2)$   
 To show this we use Result1.

$$\begin{aligned} S_{nB} &= \sum W_{nt}g(X_t, \theta_0) \\ &= \sum W_{nt}X_{t-1}[2I(Z_t \leq 0) - 1] \end{aligned}$$

Hence we need to show:

1.  $\frac{1}{n} \sum X_{t-1}[2I(Z_t \leq 0) - 1] \xrightarrow{P} 0$
2.  $\frac{1}{n} \sum X_{t-1}^2 \xrightarrow{P} EX_1^2$
3.  $\frac{1}{n} \max_t X_{t-1}^2 \xrightarrow{P} 0$

All these follow from Step2 in the proof of Theorem2.

This completes Step2, and combining with Step1, we get

$$\begin{aligned} P_B\left(\sqrt{n}\sigma_n^{-1}(\hat{\theta}_2^* - \hat{\theta}_2) \leq x\right) - P(Y \leq x) &= o_P(1) \quad \forall x \in \mathbb{R} \\ \text{where } Y &\sim N\left(0, \frac{1}{4f^2(0)EX_1^2}\right) \end{aligned}$$

Using continuity of the normal distribution, we complete the proof.

## 1.6 Special choices for $\mathbf{w}$ .

With  $(w_1, \dots, w_n) \sim \text{Mult}(n, \frac{1}{n}, \dots, \frac{1}{n})$  we get the Paired Bootstrap estimator. This is same as resampling w.r. from  $(X_{t-1}, X_t)$ ,  $t = 1, 2, \dots, n$ . Other choices of  $\{w_i\}$ 's yield the m-out-of-n Bootstrap and their variations. In particular lets check the conditions on the weights in two particular cases.

**Case 1.**  $(\mathbf{w}_1, \dots, \mathbf{w}_n) \sim \text{Mult}(\mathbf{n}, \frac{1}{\mathbf{n}}, \dots, \frac{1}{\mathbf{n}})$

Clearly the weights are exchangeable. Let us verify assumptions (A1)-(A4) on the weights in this case.

**A1.**  $E_B(w_1) = 1$

Obvious in this case.

**A2.**  $0 < k < \sigma_n^2 = o(n)$

$\sigma_n^2 = 1 - \frac{1}{n}$  which clearly satisfies the above condition.

**A3.**  $c_{1n} = O(n^{-1})$

$c_{1n} = -\frac{1}{n}$  which is as above.

**A4.**  $\{W_i\}$  satisfy conditions of P-W theorem.

To show this, we have to verify conditions (6) and (7) of Result (1) with  $U_{nj} = W_j$ .

**Condition(6)**  $\frac{1}{n} \sum W_t^2 \xrightarrow{P} 1$

$$W_t = \sqrt{\frac{n}{n-1}}(w_t - 1)$$

Therefore,

$$\frac{1}{n} \sum W_t^2 = \frac{n}{n-1} \frac{1}{n} \sum (w_t - 1)^2$$

$$V_B(\sum (w_t - 1)^2) = nV_B(w_1 - 1)^2 + n(n-1)\text{Cov}_B((w_1 - 1)^2, (w_2 - 1)^2)$$

Write  $w_1 = \sum_{i=1}^n u_i$  and  $w_2 = \sum_{i=1}^n v_i$ , where  $\{u_i, v_i\}_{i=1}^n$  are iid with the joint distribution of  $(u_i, v_i)$  given by

$$(u_i, v_i) = \begin{cases} (1, 0) & \text{w.p. } 1/n \\ (0, 1) & \text{w.p. } 1/n \\ (1, 0) & \text{w.p. } 1 - 2/n \end{cases}$$

$$\begin{aligned} V_B(w_1 - 1)^2 &= E_B(w_1 - 1)^4 - V_B^2(w_1) \\ &= E_B(w_1 - 1)^4 - \frac{1}{n^2}(1 - \frac{1}{n})^2 \end{aligned}$$

$(w_1 - 1) = \sum_{i=1}^n (u_i - p)$  where  $p = \frac{1}{n}$ ,  $q = 1 - p$ . Hence

$$\begin{aligned} E_B(w_1 - 1)^4 &= E\left( \sum (u_i - p)^4 + 3 \sum_{i \neq j} (u_i - p)^2 (u_j - p)^2 \right) \\ &= nE(u_1 - p)^4 + 3n(n-1)p^2q^2 \\ &= n(pq^4 + p^4q) + 3n(n-1)p^2q^2 \end{aligned}$$

Simplifying

$$= (1 - \frac{1}{n})(4 - \frac{9}{n} + \frac{6}{n^2} + \frac{2}{n^3}) \quad (47)$$

Therefore

$$V_B(w_1 - 1)^2 = (1 - \frac{1}{n})(3 - \frac{9}{n} + \frac{7}{n^2} + \frac{2}{n^3}) \rightarrow 3 \quad (48)$$

$$\begin{aligned} (w_1 - 1)^2 (w_2 - 1)^2 &= [\sum (u_i - p)]^2 [\sum (v_i - p)]^2 \\ &= [\sum (u_i - p)^2 + \sum_{i \neq j} \sum (u_i - p)(u_j - p)] \\ &\quad \times [\sum (v_i - p)^2 + \sum_{i \neq j} \sum (v_i - p)(v_j - p)] \end{aligned}$$

$$\begin{aligned} &E_B(w_1 - 1)^2 (w_2 - 1)^2 \\ &= E\left[ \sum_i (u_i - p)^2 (v_i - p)^2 + \sum_{i \neq j} \sum (u_i - p)^2 (v_j - p)^2 \right. \\ &\quad \left. + \sum_{i \neq j} \sum (u_i - p)(u_j - p)(v_i - p)(v_j - p) \right] \\ &= nE(u_1 - p)^2 (v_1 - p)^2 + n(n-1)V(u_1)V(v_1) \\ &\quad + n(n-1)\text{Cov}(u_1, v_1)\text{Cov}(u_2, v_2) \\ &= n(2pp^2q^2 + p^4(1-2p)) + n(n-1)p^2q^2 - n(n-1)p^4 \\ &= \frac{2}{n^2}(1 - \frac{1}{n})^2 + \frac{1}{n^3}(1 - \frac{2}{n}) \\ &\quad + n(n-1)\frac{1}{n^2}(1 - \frac{1}{n})^2 - n(n-1)\frac{1}{n^4} \\ &= 1 - \frac{3}{n} + \frac{4}{n^2} - \frac{3}{n^3} \end{aligned}$$

$$\begin{aligned}
& Cov_B((w_1 - 1)^2, (w_2 - 1)^2) \\
&= E_B(w_1 - 1)^2(w_2 - 1)^2 - (1 - \frac{1}{n})^2 \\
&= 1 - \frac{3}{n} + \frac{4}{n^2} - \frac{3}{n^3} \\
&\quad - (1 - \frac{2}{n} + \frac{1}{n^2}) \\
&= -\frac{1}{n} + \frac{3}{n^2} - \frac{3}{n^3} \\
&\rightarrow 0
\end{aligned}$$

Therefore

$$\begin{aligned}
V_B(\frac{1}{n} \sum (w_t - 1)^2) &= \frac{1}{n} V_B(w_1 - 1)^2 + (1 - \frac{1}{n}) Cov_B((w_1 - 1)^2, (w_2 - 1)^2) \\
&\rightarrow 0 \\
V_B(\frac{1}{n} \sum W_t^2) &= (\frac{n}{n-1})^2 V_B(\frac{1}{n} \sum (w_t - 1)^2) \rightarrow 0
\end{aligned} \tag{49}$$

$$E_B(\frac{1}{n} \sum W_t^2) = \frac{n}{n-1} E_B(w_1 - 1)^2 = 1 \tag{50}$$

Hence from (49) and (50),

$$\frac{1}{n} \sum W_t^2 \xrightarrow{P_B} 1$$

This proves condition (6).

$$\mathbf{Condition(7)} \quad \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \sqrt{E(W_t^2 I_{(|W_t| > k)})} = 0$$

$$\begin{aligned}
E(W_t^2 I_{|W_t| > k}) &= \frac{1}{\sigma_n^2} E[(w_t - 1)^2 I_{(|w_t - 1| > k\sigma_n)}] \\
&\leq \frac{1}{\sigma_n^2} [E(w_t - 1)^4]^{\frac{1}{2}} [P(|w_t - 1| > k\sigma_n)]^{\frac{1}{2}} \\
&\leq \frac{1}{\sigma_n^2} (M_{n4}) (\frac{\sigma_n^2}{k^2 \sigma_n^2})^{\frac{1}{2}} \\
&= \frac{1}{k} (\frac{M_{n4}}{\sigma_n^4})^{\frac{1}{2}}
\end{aligned}$$

where  $M_{n4} = E(w_t - 1)^4$ . Therefore

$$\begin{aligned}
& \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \sqrt{E(W_t^2 I_{(|W_t| > k)})} \\
&\leq \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{\sqrt{k}} (\frac{M_{n4}}{\sigma_n^4})^{\frac{1}{4}} = 0
\end{aligned}$$

as both  $M_{n4}$  and  $\sigma_n^4$  are bounded (follows from (47)).

**Case 2.**  $(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n)$  iid  $(\mathbf{1}, \sigma^2)$

Again we need to establish (A4), that is, verify conditions 6) and 7) in Result(1).

Condition 6) follows from WLLN.

To verify condition 7), note that since distribution of  $(w_1, w_2, \dots, w_n)$  is independent of  $n$ ,

$$\begin{aligned} & \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \sqrt{E(W_t^2 I_{(|W_t| > k)})} \\ &= \lim_{k \rightarrow \infty} \sqrt{E(W_t^2 I_{(|W_t| > k)})} = 0 \end{aligned}$$

since  $EW_t^2 < \infty$ .

**Remark 1.** Result 2 is true even when the process is nonstationary. This follows from the fact that, given observations  $\{X_t\}$  from the AR process,  $X_t = \theta X_{t-1} + Z_t$ ,  $|\theta| < 1$ ; we can get a stationary solution of the above process, say  $\{Y_t\}$ , such that  $\frac{1}{n} \sum X_t^a Z_{t+k}^b \xrightarrow{a.s.} E(Y_t^a Z_{t+k}^b)$ . As a consequence, Theorem 1) holds even without the assumption of stationarity, which is assumed throughout its proof.

## 2 Bootstrap in Heteroscedastic AR(1) model

Now we introduce heteroscedasticity in the model (1), and study the Weighted Bootstrap estimator. Consider the following model:

$$X_t = \theta_0 X_{t-1} + Z_t; Z_t = \tau_t \epsilon_t \quad t = 1, 2, \dots, n. \quad |\theta_0| < 1 \quad (51)$$

$$X_0 \sim F_0 \text{ with all moments finite.} \quad (52)$$

where  $\theta_0, \tau_t > 0$  are constants,  $\epsilon_t \sim iid(0, 1)$ , and  $\epsilon_t$  is independent of  $\{X_{t-k}, k \geq 1\}$  for all  $t$ .

### 2.1 Estimation

Based on observations  $X_1, X_2, \dots, X_n$  we discuss various methods for estimating  $\theta$  in the model. Listed below are four types of estimators.

**(a) Weighted Least Squares Estimator** Assuming  $\{\tau_t\}$  to be known, consider the following estimator for  $\theta_0$ :

$$\hat{\theta}_1 = \operatorname{argmin}_{\theta} \frac{1}{n} \sum_{t=2}^n \frac{1}{\tau_t^2} (X_t - \theta X_{t-1})^2 \quad (53)$$

$$= \frac{\sum_{t=2}^n \frac{1}{\tau_t^2} X_t X_{t-1}}{\sum_{t=2}^n \frac{1}{\tau_t^2} X_{t-1}^2} \quad (54)$$

If  $\epsilon_t$  in model(51) is normal, (54) turns out to be the (Gaussian) maximum likelihood estimators.

**(b) Least Squares Estimator** In general  $\{\tau_t\}$  are unknown and are non-estimable. Hence we may consider the general least squares estimators, ie,

$$\hat{\theta}_2 = \frac{\sum_{t=2}^n X_t X_{t-1}}{\sum_{t=2}^n X_{t-1}^2} \quad (55)$$

This turns out to be the same as (54) if the  $\{\tau_i\}$  are all equal, that is the model is homoscedastic.

**(c) Weighted Least Absolute Deviations Estimator** The estimators (54) and (55) are  $L_2$ -estimators. It is well known that  $L_1$ -estimators are more robust with respect to heavy-tailed distributions than  $L_2$ -estimators. This motivates the study of various LAD estimators for  $\theta_0$ . Now we reparametrize model(51) in such a way that the median of  $\epsilon_t$ , instead of the mean equals 0 while  $V\epsilon_t = 1$  remains unchanged. Our first absolute deviation estimator takes the form

$$\hat{\theta}_3 = \operatorname{argmin}_{\theta} \sum_{t=2}^n \frac{1}{\tau_t} |X_t - \theta X_{t-1}| \quad (56)$$

This is motivated by the fact that  $\hat{\theta}_3$  turns out to be the maximum likelihood estimator when the errors have double-exponential distribution.

**Least absolute deviations estimator** Estimator(56) uses the fact that  $\tau_t$  are known. In case they are not our absolute deviation estimator takes the form

$$\hat{\theta}_4 = \operatorname{argmin}_{\theta} \sum_{t=2}^n |X_t - \theta X_{t-1}| \quad (57)$$

In the next section we discuss the asymptotic properties of the listed estimators.

## 2.2 Consistency of estimation in heteroscedastic AR(1) process

In this section, we establish the distributional consistency of each of the four estimators discussed in the earlier section. To do so, we will use some established results, the first one being the following Martingale Central Limit theorem:

**Result 4 (Martingale C.L.T.; see Hall and Heyde 1980)** *Let  $\{S_n, \mathcal{F}_n\}$  denote a zero-mean martingale whose increments have finite variance. Write  $S_n = \sum_{i=1}^n X_i$ ,  $V_n^2 = \sum_{i=1}^n E(X_i^2 | \mathcal{F}_{i-1})$  and  $s_n^2 = EV_n^2 = ES_n^2$ . If*

$$s_n^{-2} V_n^2 \xrightarrow{P} 1 \text{ and} \\ s_n^{-2} \sum_{i=1}^n E(X_i^2 I(|X_i| \geq \epsilon s_n)) \longrightarrow 0 \text{ as } n \rightarrow \infty \quad \forall \epsilon > 0.$$

Then  $\frac{S_n}{s_n} \xrightarrow{d} N(0, 1)$ .

Another result we will need is the following one on convergence of a weighted sum of iid random variables.

**Result 5** *Let  $X_1, X_2, \dots, X_n$  be a sequence of iid mean zero random variables, and  $\{c_{in} | i = 1, \dots, n\}$  a triangular sequence of bounded constants. Then  $\frac{1}{n} \sum_{i=1}^n c_{in} X_i \xrightarrow{P} 0$*

### 2.2.1 Distributional consistency of $\hat{\theta}_1$

**Theorem 4** *Define  $\frac{s_n^2}{n} = \frac{1}{n} \sum_{t=2, n} \tau_t^{-2} E X_{t-1}^2$ . Assume that*

- A1.**  $\frac{\tau_i}{\tau_j} \leq M_2 \quad \forall 1 \leq i < j \leq n$
- A2.**  $\frac{1}{n} \sum_{1 \leq i < j \leq n} \theta_0^{2(j-i)} \left(\frac{\tau_i}{\tau_j}\right)^2 \geq M_1 > 0$

$$\mathbf{A3.} \quad \frac{1}{n} \sum_{1 \leq i < j \leq n} \theta_0^{2(j-i)} \left( \frac{\tau_i}{\tau_j} \right)^2 \longrightarrow \rho^2$$

Then under assumptions (A1-A2),  $s_n(\hat{\theta}_1 - \theta_0) \xrightarrow{d} N(0, 1)$

Further if we assume (A3), we have,  $\sqrt{n}(\hat{\theta}_1 - \theta_0) \xrightarrow{d} N(0, \theta_0^2/\rho^2)$  as  $n \rightarrow \infty$

$$\mathbf{Proof} \quad \sqrt{n}(\hat{\theta}_1 - \theta_0) = \sqrt{n} \frac{\sum \tau_t^{-2} X_{t-1} Z_t}{\sum \tau_t^{-2} X_{t-1}^2}$$

**Step1.**  $\frac{1}{\sqrt{n}} \sum \tau_t^{-2} X_{t-1} Z_t$  is asymptotically normal

Let  $S_n = \sum_{t=2}^n \tau_t^{-2} X_{t-1} Z_t$ .

Note that

$$X_t = \theta_0^t X_0 + \sum_{k=1}^t \theta_0^{t-k} Z_k \quad \forall t \geq 1 \quad (58)$$

$$\text{Hence } E(X_t^2) = \theta_0^{2t} E X_0^2 + \sum_{k=1}^t \theta_0^{2t-2k} \tau_k^2 \quad (59)$$

Hence  $S_n$  is a 0 mean  $\mathcal{A}_n$  measurable martingale with increments having finite variance, where  $\mathcal{A}_t = \sigma(X_0, \epsilon_1, \epsilon_2, \dots, \epsilon_t)$ ;  $t = 1, 2, \dots, n$ . This follows from the fact that  $E(X_t^2)$  is finite, and  $E(X_{t-1} Z_t | \mathcal{A}_{t-1}) = 0$ .

To establish the asymptotic normality of  $S_n$ , we use Result (4). Let

$$\begin{aligned} V_n^2 &= \sum_{t=2}^n E(\tau_t^{-4} X_{t-1}^2 Z_t^2 | \mathcal{A}_{t-1}) \\ &= \sum_{t=2}^n \tau_t^{-2} X_{t-1}^2 \end{aligned}$$

Then to accomplish Step1, we need to show

$$\frac{V_n^2}{s_n^2} \xrightarrow{P} 1 \quad (60)$$

$$\frac{1}{s_n^2} \sum_{t=2, n} E [(\tau_t^{-2} X_{t-1} Z_t)^2 I(\tau_t^{-2} |X_{t-1} Z_t| \geq \epsilon s_n)] \longrightarrow 0 \quad (61)$$

To prove (60), note that

$$\begin{aligned} & \frac{V_n^2}{n} - \frac{s_n^2}{n} \\ &= \frac{1}{n} \sum_{t=1, n-1} \tau_{t+1}^{-2} [X_t^2 - E X_t^2] \\ &= \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=0}^t \frac{\tau_k^2}{\tau_{t+1}^2} \theta_0^{2(t-k)} (\epsilon_k^2 - 1) + \frac{2}{n} \sum_{t=1, n-1} \sum_{0 \leq i < j \leq t} \frac{\tau_i \tau_j}{\tau_{t+1}^2} \theta_0^{2t-i-j} \epsilon_i \epsilon_j \\ &= \frac{1}{n} \sum_{k=0}^{n-1} (\epsilon_k^2 - 1) \left( \sum_{t=k}^{n-1} \frac{\tau_k^2}{\tau_{t+1}^2} \theta_0^{2(t-k)} \right) + \frac{2}{n} \sum_{0 \leq i < j \leq n-1} \epsilon_i \epsilon_j \left( \sum_{t=j}^{n-1} \frac{\tau_i \tau_j}{\tau_{t+1}^2} \theta_0^{2t-i-j} \right) \\ &= T_1 + 2T_2 \text{ (say)} \end{aligned}$$

Using assumption (A1) and Resut5, we have  $T_1 \xrightarrow{P} 0$ .

$$\begin{aligned} ET_2 &= 0 \\ VT_2 &= \frac{1}{n^2} \sum_{0 \leq i < j \leq n} \sum_{t=j}^n \theta_0^{2t-i-j} \frac{\tau_i \tau_j}{\tau_{t+1}^2} \\ &\leq \frac{M_2}{n^2} \sum_{t=1}^{n-1} \left( \sum_{k=0}^t |\theta_0|^{t-k} \right)^2 \rightarrow 0 \end{aligned}$$

Hence  $T_2 \xrightarrow{P} 0$ .

Combining,  $\frac{V_n^2}{n} - \frac{s_n^2}{n} \xrightarrow{P} 0$

Also  $\frac{s_n^2}{n} = \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=0}^t \frac{\tau_k^2}{\tau_{t+1}^2} \theta_0^{2(t-k)}$ .

Using assumption(A2),  $\frac{s_n^2}{n}$  is bounded below. Therefore  $\frac{V_n^2}{s_n^2} \xrightarrow{P} 1$  and this proves (60).

Remains to show (61), ie  $\frac{1}{s_n^2} \sum_{t=1}^{n-1} E [\tau_{t+1}^{-2} X_t^2 \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \rightarrow 0$

$$\begin{aligned} |X_t| &\leq \sum_{k=0}^t \tau_k |\theta_0|^{t-k} |\epsilon_k| \\ \frac{X_t^2}{\tau_{t+1}^2} &\leq \sum_{k=0}^t \frac{\tau_k^2}{\tau_{t+1}^2} |\theta_0|^{2(t-k)} \epsilon_k^2 + 2 \sum_{0 \leq i < j \leq t} \frac{\tau_i \tau_j}{\tau_{t+1}^2} |\theta_0|^{2t-i-j} \epsilon_i \epsilon_j \end{aligned}$$

Hence for  $1 \leq t \leq n$ ,

$$\begin{aligned} &E [\tau_{t+1}^{-2} X_t^2 \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \\ &\leq \sum_{k=0}^t \frac{\tau_k^2}{\tau_{t+1}^2} |\theta_0|^{2(t-k)} E [\epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] + 2 \sum_{0 \leq i < j \leq t} \frac{\tau_i \tau_j}{\tau_{t+1}^2} |\theta_0|^{2t-i-j} E [\epsilon_i \epsilon_j \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \\ &\leq A_1 \max_{0 \leq k \leq t} E [\epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] + A_2 \max_{0 \leq i < j \leq t} E [\epsilon_i \epsilon_j \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \\ &\leq A \max_{0 \leq k \leq t} E [\epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \end{aligned}$$

$\frac{s_n}{\sqrt{n}}$  is bounded below by say,  $M > 0$ . Hence for a fixed  $k_0$ ,  $0 \leq k_0 \leq t$ ,

$$\begin{aligned}
& E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( \frac{|X_t \epsilon_{t+1}|}{\tau_{t+1}} \geq \epsilon s_n \right) \right] \\
\leq & E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( \frac{|X_t \epsilon_{t+1}|}{\tau_{t+1}} \geq \epsilon M \sqrt{n} \right) \right] \\
\leq & E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( |\epsilon_{t+1}| \sum_{k=0}^t \frac{\tau_k}{\tau_{t+1}} |\epsilon_k| \geq \epsilon M \sqrt{n} \right) \right] \\
\leq & E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( |\epsilon_{t+1}| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( \sum_{k=0}^t \frac{\tau_k}{\tau_{t+1}} |\theta_0|^{t-k} |\epsilon_k| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] \\
= & E \left[ \epsilon_{t+1}^2 I \left( |\epsilon_{t+1}| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( \sum_{k=0}^t \frac{\tau_k}{\tau_{t+1}} |\theta_0|^{t-k} |\epsilon_k| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] \\
\leq & E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_1 n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( \sum_{k=0}^t |\theta_0|^{t-k} |\epsilon_k| \geq c_2 n^{1/4} \right) \right] \\
\leq & E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_1 n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( |\epsilon_{k_0}| \geq \frac{c_2}{2} n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( \sum_{k \neq k_0} |\theta_0|^{t-k} |\epsilon_k| \geq \frac{c_2}{2} n^{1/4} \right) \right] \\
\leq & E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_1 n^{1/4} \right) \right] + E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq \frac{c_2}{2} n^{1/4} \right) \right] + P \left[ \sum_{k \neq k_0} |\theta_0|^{t-k} |\epsilon_k| \geq \frac{c_2}{2} n^{1/4} \right] \\
\leq & E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_3 n^{1/4} \right) \right] + \frac{c_4}{n^{1/4}}
\end{aligned}$$

Hence  $\max_{0 \leq k \leq t} E \left[ \epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1}^{-1} |X_t \epsilon_{t+1}| \geq \epsilon s_n) \right] \rightarrow 0$ . Using the fact that  $\frac{s_n^2}{n}$  is bounded below, this proves (61). Using Result4, from (60) and (61) we deduce that  $\frac{s_n}{s_n} \xrightarrow{d} N(0, 1)$ , ie

$$\frac{1}{s_n} \sum_{t=2}^n \tau_t^{-2} X_{t-1} Z_t \xrightarrow{d} N(0, 1) \tag{62}$$

**Step2.**

$$\frac{\sum_{t=2}^n \tau_t^{-2} X_{t-1}^2}{s_n^2} \xrightarrow{P} 1 \tag{63}$$

This follows from (60).

From (62) and (63) we deduce,

$$s_n \frac{\sum \tau_t^{-2} X_{t-1} Z_t}{\sum \tau_t^{-2} X_{t-1}^2} \xrightarrow{d} N(0, 1) \tag{64}$$

$$ie s_n (\hat{\theta}_1 - \theta_0) \xrightarrow{d} N(0, 1) \tag{65}$$

$$\frac{s_n^2}{n} = \frac{1}{n} \sum_{t=2}^n \tau_t^{-2} E X_{t-1}^2 \tag{66}$$

$$= \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=1}^t \frac{\tau_k^2}{\tau_{t+1}^2} \theta_0^{2(t-k)} \tag{67}$$

Hence if we assume (A3), we have  $\frac{s_n^2}{n} \rightarrow \frac{\rho^2}{\theta_0^2}$ , and then,

$$\sqrt{n}(\hat{\theta}_1 - \theta_0) \xrightarrow{d} \mathbf{N}(0, \frac{\theta_0^2}{\rho^2})$$

This completes the proof.

**Remark** Assumptions (A1) and (A2) are satisfied if  $\{\tau_t\}$ 's are bounded, or more generally if they are of the same order, ie there exists constants  $c_1, c_2 > 0$  and  $\alpha \geq 0$  such that  $c_1 t^\alpha \leq \tau_t^2 \leq c_2 t^\alpha$  for  $1 \leq t \leq n$ .

### 2.2.2 Distributional consistency of $\hat{\theta}_2$

**Theorem 5** Define  $s_n^2 = \sum_{t=2}^n \tau_t^2 E(X_{t-1}^2)$ . Suppose  $\{\tau_t\}$ 's satisfy the following assumptions.

**A1.**  $M_1 \leq \tau_t \leq M_2$ ;  $t = 1, 2, \dots, n$

**A2.**  $\frac{\sum \tau_t^2}{n} \rightarrow \tau^2 > 0$

**A3.**  $\frac{1}{n} \sum_{1 \leq i < j \leq n} \tau_i^2 \tau_j^2 \theta_0^{2(j-i)} \rightarrow \rho^2$

Then under assumptions(A1) and (A2),  $\left(\frac{s_n}{\sqrt{n}}\right)^{-1} \sqrt{n}(\hat{\theta}_2 - \theta_0) \xrightarrow{d} N\left(0, \frac{(1-\theta_0^2)^2}{\tau^4}\right)$ .

Further if (A3) holds,  $\sqrt{n}(\hat{\theta}_2 - \theta_0) \xrightarrow{d} N\left(0, \frac{\rho^2(1-\theta_0^2)^2}{\tau^4\theta_0^2}\right)$ .

**Proof**  $\sqrt{n}(\hat{\theta}_2 - \theta_0) = \sqrt{n} \frac{\sum X_{t-1} Z_t}{\sum X_{t-1}^2}$

**Step1.**  $\frac{1}{\sqrt{n}} \sum X_{t-1} Z_t$  is asymptotically normal.

Let  $S_n = \sum_{t=2}^n X_{t-1} Z_t$ .

Then  $S_n$  is a 0 mean  $\mathcal{A}_n$  measurable martingale with increments having finite variance, where  $\mathcal{A}_t = \sigma(X_0, \epsilon_1, \epsilon_2, \dots, \epsilon_t)$ ;  $t = 1, 2, \dots, n$ . This follows from the fact that  $E(X_t^2)$  is finite, and  $E(X_{t-1} Z_t | \mathcal{A}_{t-1}) = 0$ .

To establish the asymptotic normality of  $S_n$ , we use Result (4). Let

$$\begin{aligned} V_n^2 &= \sum_{t=2}^n E(X_{t-1}^2 Z_t^2 | \mathcal{A}_{t-1}) \\ &= \sum_{t=2}^n \tau_t^2 X_{t-1}^2 \end{aligned}$$

Then to accomplish Step1, we need to show

$$\frac{V_n^2}{s_n^2} \xrightarrow{P} 1 \tag{68}$$

$$\frac{1}{s_n^2} \sum_{t=2}^n E[(X_{t-1} Z_t)^2 I(|X_{t-1} Z_t| \geq \epsilon s_n)] \rightarrow 0 \tag{69}$$

Using the expressions for  $X_t$  and  $E X_t^2$  from equations(58) and (59), we have

$$\begin{aligned}
& \frac{V_n^2}{n} - \frac{s_n^2}{n} \\
&= \frac{1}{n} \sum_{t=1}^{n-1} \tau_{t+1}^2 [X_t^2 - EX_t^2] \\
&= \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=0}^t \tau_{t+1}^2 \tau_k^2 \theta_0^{2(t-k)} (\epsilon_k^2 - 1) + \frac{2}{n} \sum_{t=1}^{n-1} \sum_{0 \leq i < j \leq t} \tau_i \tau_j \tau_{t+1}^2 \theta_0^{2t-i-j} \epsilon_i \epsilon_j \\
&= \frac{1}{n} \sum_{k=0}^{n-1} (\epsilon_k^2 - 1) \left( \sum_{t=k}^{n-1} \tau_k^2 \tau_{t+1}^2 \theta_0^{2(t-k)} \right) + \frac{2}{n} \sum_{0 \leq i < j \leq n-1} \epsilon_i \epsilon_j \left( \sum_{t=j}^{n-1} \tau_i \tau_j \tau_{t+1}^2 \theta_0^{2t-i-j} \right) \\
&= T_1 + 2T_2 \text{ (say)}
\end{aligned}$$

Using assumption (A1) and Resut5, we have  $T_1 \xrightarrow{P} 0$ .

$$ET_2 = 0 \tag{70}$$

$$VT_2 = \frac{1}{n^2} \sum_{0 \leq i < j \leq n-1} \left( \sum_{t=j}^{n-1} \theta_0^{2t-i-j} \tau_i \tau_j \tau_{t+1}^2 \right)^2 \tag{71}$$

$$\leq \frac{c}{n^2} \sum_{0 \leq i < j \leq n-1} \sum_{t=j}^{n-1} |\theta_0|^{2t-i-j} \tag{72}$$

$$\leq \frac{c}{n^2} \sum_{t=1}^{n-1} \left( \sum_{k=0}^t |\theta_0|^{t-k} \right)^2 \longrightarrow 0 \tag{73}$$

Here  $c$  is some positive constant. (72) follows from (71) using the fact that  $\sum_{t=j}^{n-1} \theta_0^{2t-i-j} \tau_i \tau_j \tau_{t+1}^2$  is bounded which inturn follows from assumption(A1).

Hence  $T_2 \xrightarrow{P} 0$ .

Combining,  $\frac{V_n^2}{n} - \frac{s_n^2}{n} \xrightarrow{P} 0$

Also  $\frac{s_n^2}{n} = \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=0}^t \frac{\tau_k^2}{\tau_{t+1}^2} \theta_0^{2(t-k)}$ .

Again using assumption(A1),  $\frac{s_n^2}{n}$  is bounded below. Therefore  $\frac{V_n^2}{s_n^2} \xrightarrow{P} 1$  and this proves (68).

Remains to show (69), ie  $\frac{1}{s_n^2} \sum_{t=1}^{n-1} E \left[ \tau_{t+1}^2 X_t^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n) \right] \longrightarrow 0$

$$\begin{aligned}
|X_t| &\leq \sum_{k=0}^t \tau_k |\theta_0|^{t-k} |\epsilon_k| \\
X_t^2 &\leq \sum_{k=0}^t \tau_k^2 |\theta_0|^{2(t-k)} \epsilon_k^2 + 2 \sum_{0 \leq i < j \leq t} \tau_i \tau_j |\theta_0|^{2t-i-j} \epsilon_i \epsilon_j
\end{aligned}$$

Hence for  $1 \leq t \leq n$ ,

$$\begin{aligned}
& E [\tau_{t+1}^2 X_t^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \\
& \leq \sum_{k=0}^t \tau_k^2 \tau_{t+1}^2 |\theta_0|^{2(t-k)} E [\epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] + 2 \sum_{0 \leq i < j \leq t} \tau_i \tau_j \tau_{t+1}^2 |\theta_0|^{2t-i-j} E [\epsilon_i \epsilon_j \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \\
& \leq A_1 \max_{0 \leq k \leq t} E [\epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] + A_2 \max_{0 \leq i < j \leq t} E [\epsilon_i \epsilon_j \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \\
& \leq A \max_{0 \leq k \leq t} E [\epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)]
\end{aligned}$$

$\frac{s_n}{\sqrt{n}}$  is bounded below by say,  $M > 0$ . Hence for a fixed  $k_0$ ,  $0 \leq k_0 \leq t$ ,

$$\begin{aligned}
& E [\epsilon_{k_0}^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \\
& \leq E [\epsilon_{k_0}^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon M \sqrt{n})] \\
& \leq E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( |\epsilon_{t+1}| \sum_{k=0}^t \tau_k \tau_{t+1} |\epsilon_k| \geq \epsilon M \sqrt{n} \right) \right] \\
& \leq E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( |\epsilon_{t+1}| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 \epsilon_{t+1}^2 I \left( \sum_{k=0}^t \tau_k \tau_{t+1} |\theta_0|^{t-k} |\epsilon_k| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] \\
& = E \left[ \epsilon_{t+1}^2 I \left( |\epsilon_{t+1}| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( \sum_{k=0}^t \tau_k \tau_{t+1} |\theta_0|^{t-k} |\epsilon_k| \geq \sqrt{\epsilon M} n^{1/4} \right) \right] \\
& \leq E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_1 n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( \sum_{k=0}^t |\theta_0|^{t-k} |\epsilon_k| \geq c_2 n^{1/4} \right) \right] \\
& \leq E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_1 n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( |\epsilon_{k_0}| \geq \frac{c_2}{2} n^{1/4} \right) \right] + E \left[ \epsilon_{k_0}^2 I \left( \sum_{k \neq k_0} |\theta_0|^{t-k} |\epsilon_k| \geq \frac{c_2}{2} n^{1/4} \right) \right] \\
& \leq E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_1 n^{1/4} \right) \right] + E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq \frac{c_2}{2} n^{1/4} \right) \right] + P \left[ \sum_{k \neq k_0} |\theta_0|^{t-k} |\epsilon_k| \geq \frac{c_2}{2} n^{1/4} \right] \\
& \leq E \left[ \epsilon_1^2 I \left( |\epsilon_1| \geq c_3 n^{1/4} \right) \right] + \frac{c_4}{n^{1/4}}
\end{aligned}$$

Hence  $\max_{0 \leq k \leq t} E [\epsilon_k^2 \epsilon_{t+1}^2 I(\tau_{t+1} |X_t \epsilon_{t+1}| \geq \epsilon s_n)] \rightarrow 0$ . Using the fact that  $\frac{s_n}{\sqrt{n}}$  is bounded below, this proves (69).

Using Result4, from (68) and (69) we deduce that  $\frac{S_n}{s_n} \xrightarrow{d} N(0, 1)$ , ie

$$\frac{1}{s_n} \sum_{t=2}^n X_{t-1} Z_t \xrightarrow{d} N(0, 1) \tag{74}$$

**Step2.**

$$\frac{1}{n} \sum_{t=2}^n X_{t-1}^2 \xrightarrow{P} \frac{\tau^2}{(1 - \theta_0^2)} \tag{75}$$

This follows once we show

$$\frac{1}{n} \sum_{t=1}^{n-1} (X_t^2 - EX_t^2) \xrightarrow{P} 0 \tag{76}$$

$$\text{and } \frac{1}{n} \sum_{t=1}^{n-1} EX_t^2 \rightarrow \frac{\tau^2}{(1 - \theta_0^2)} \tag{77}$$

Using the expressions for  $X_t$  and  $EX_t^2$  from equations(58) and (59), we have

$$\begin{aligned} & \frac{1}{n} \sum_{t=1}^{n-1} (X_t^2 - EX_t^2) \\ = & \frac{1}{n} \sum_{k=0}^{n-1} (\epsilon_k^2 - 1) \left( \sum_{t=k}^{n-1} \tau_k^2 \theta_0^{2(t-k)} \right) + \frac{2}{n} \sum_{0 \leq i < j \leq n-1} \epsilon_i \epsilon_j \left( \sum_{t=j}^{n-1} \tau_i \tau_j \theta_0^{2t-i-j} \right) \\ \xrightarrow{P} & 0 \end{aligned}$$

The above steps can be justified by proceeding as in the proof of (68). This complete (76). To see (77), note that

$$\frac{1}{n} \sum_{t=1}^{n-1} EX_t^2 = \frac{EX_0^2}{n} \sum_{t=1}^{n-1} \theta_0^{2t} + \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=1}^t \tau_k^2 \theta_0^{2t-2k} \quad (78)$$

$$\approx \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=1}^t \tau_k^2 \theta_0^{2t-2k} \quad (79)$$

$$= \frac{1}{n} \sum_{k=1}^{n-1} \tau_k^2 \sum_{t=0}^{n-k-1} \theta_0^{2t} \quad (80)$$

$$= \frac{1}{n} \sum_{k=1}^{n-1} \tau_k^2 \frac{(1 - \theta_0^{2(n-k)})}{(1 - \theta_0^2)} \quad (81)$$

$$\approx \frac{1}{(1 - \theta_0^2)} \frac{1}{n} \sum \tau_k^2 \quad (82)$$

$$\rightarrow \frac{\tau^2}{(1 - \theta_0^2)} \quad (83)$$

This proves (77). (76) and (77) together prove (75) and this completes Step2.

Dividing (74) by (75) we deduce,

$$\left( \frac{s_n}{\sqrt{n}} \right)^{-1} \sqrt{n} (\hat{\theta}_2 - \theta_0) \xrightarrow{d} N \left( 0, \frac{(1 - \theta_0^2)^2}{\tau^4} \right)$$

$$\begin{aligned} \frac{s_n^2}{n} &= \frac{1}{n} \sum_{t=2}^n \tau_t^2 EX_{t-1}^2 \\ &= \frac{1}{n} \sum_{t=1}^{n-1} \sum_{k=1}^t \tau_k^2 \tau_{t+1}^2 \theta_0^{2(t-k)} \end{aligned}$$

Hence if we assume (A3), we have  $\frac{s_n^2}{n} \rightarrow \frac{\rho^2}{\theta_0^2}$ , and then,

$$\sqrt{n} (\hat{\theta}_2 - \theta_0) \xrightarrow{d} \mathbf{N} \left( \mathbf{0}, \frac{\rho^2 (1 - \theta_0^2)^2}{\tau^4 \theta_0^2} \right)$$

This completes the proof.

### 2.3 Consistency of the Weighted Bootstrap technique

Now we prove that the Weighted Bootstrap estimator is asymptotically normal with the same asymptotic distribution as of the least squares estimate. In particular WB provides a consistent resampling scheme in the AR model with introduced heteroscedasticity.

**Theorem 6** Let  $\hat{\theta}_n^*$  be the weighted bootstrap estimator of  $\hat{\theta}_n$  as defined in (4). Then under the conditions (A1)-(A4) on the weights,

$$P \left[ \sqrt{n}\sigma_n^{-1}(\hat{\theta}_n^* - \hat{\theta}_n) \leq x \mid X_1, \dots, X_n \right] - P[Y \leq x] = o_P(1) \quad \forall x \in \mathbb{R} \quad (84)$$

where  $Y \sim N(0, \sigma^2)$ ,  $\sigma^2$  being defined in Theorem (??).

**Proof** As in (13),

$$\begin{aligned} & \sqrt{n}\sigma_n^{-1}(\hat{\theta}_n^* - \hat{\theta}_n) \\ &= \frac{\sum W_t X_{t-1} Z_t / \sqrt{n}}{\sum w_t X_{t-1}^2 / n} - \sqrt{n}(\hat{\theta}_n - \theta) \sigma_n^{-1} \frac{\sum (w_t - 1) X_{t-1}^2 / n}{\sum w_t X_{t-1}^2 / n} \\ &= T_1 - T_2 \end{aligned} \quad (85)$$

**Claim 1.** There exists  $\tau > 0$  such that

$$P \left[ \frac{1}{\sqrt{n}} \sum_{t=2}^n W_t X_{t-1} Z_t \leq x \mid X_1, \dots, X_n \right] \xrightarrow{P} \Phi\left(\frac{x}{\tau}\right) \quad \forall x \in \mathbb{R}$$

To see this let us verify the first three conditions of Result(1) with  $c_{nj} = X_j Z_{j+1}$  and  $U_{nj} = W_j$  for  $j = 1, \dots, n-1$ .

**Condition 1**  $\frac{1}{n} \sum_{t=2}^n X_{t-1} Z_t \xrightarrow{P} 0$

Follows from Theorem ??.

**Condition 2**  $\frac{1}{n} \sum_{t=2}^n (X_{t-1} Z_t)^2 \xrightarrow{P} \tau^2$

$$\begin{aligned} \text{Let } S_n &= \frac{1}{n} \sum_{t=2}^n X_{t-1}^2 Z_t^2 \\ U_n &= \frac{1}{n} \sum_{t=2}^n X_{t-1}^2 (Z_t^2 - \sigma_t^2) \\ V_n &= \frac{1}{n} \sum_{t=2}^n X_{t-1}^2 \sigma_t^2 \end{aligned}$$

Then  $S_n = U_n + V_n$ .

$$\begin{aligned} E(U_n) &= 0 \\ V(nU_n) &= \sum_{t=2}^n V(X_{t-1}^2 (Z_t^2 - \sigma_t^2)) + 2 \sum_{2 \leq s < t \leq n} \sum_{t=2}^n \text{Cov}(X_{t-1}^2 (Z_t^2 - \sigma_t^2), X_{s-1}^2 (Z_s^2 - \sigma_s^2)) \\ &= \sum_{t=2}^n V(X_{t-1}^2 (Z_t^2 - \sigma_t^2)) \\ &\leq M \sum_{t=2}^n E X_{t-1}^4 \quad \text{where } E Z_t^4 \leq M \end{aligned}$$

Therefore  $V(U_n) \leq \frac{M}{n^2} \sum EX_{t-1}^4 \rightarrow 0$  since  $\frac{1}{n} \sum EX_{t-1}^4$  is bounded.

Hence  $U_n \xrightarrow{P} 0$ .

$$\begin{aligned} V_n &= \frac{1}{n} \left[ \sigma_1^2 \sum_{t \text{ odd}} X_{t-1}^2 + \sigma_2^2 \sum_{t \text{ even}} X_{t-1}^2 \right] \\ &= \frac{1}{2} \left[ \sigma_1^2 \frac{\sum_{t \text{ odd}} X_{t-1}^2}{n/2} + \sigma_2^2 \frac{\sum_{t \text{ even}} X_{t-1}^2}{n/2} \right] \\ &\xrightarrow{P} \frac{1}{2} \left[ \sigma_1^2 \sigma_e^2 (1 - \theta^4)^{-1} + \sigma_2^2 \sigma_o^2 (1 - \theta^4)^{-1} \right] \\ &= \frac{1}{2(1 - \theta^4)} \left[ \sigma_1^2 \sigma_e^2 + \sigma_2^2 \sigma_o^2 \right] \end{aligned}$$

Therefore

$$\begin{aligned} S_n &\xrightarrow{P} \frac{1}{2} \frac{[\sigma_1^2 \sigma_e^2 + \sigma_2^2 \sigma_o^2]}{(1 - \theta^4)} \\ &= \frac{\theta^2(\sigma_1^4 + \sigma_2^4) + 2\sigma_1^2 \sigma_2^2}{2(1 - \theta^4)} = \tau^2 \end{aligned}$$

**Condition 3**  $n^{-1} \max(X_{t-1}^2 Z_t^2) \xrightarrow{P} 0$

Given  $\epsilon$  positive,

$$\begin{aligned} P(n^{-1} \max(X_{t-1}^2 Z_t^2) > \epsilon) &\leq \sum_2^n P(Y_t > n\epsilon) \text{ where } Y_t = X_{t-1}^2 Z_t^2 \\ &\leq \frac{1}{n^2 \epsilon^2} \sum_{t=2}^n E(X_{t-1}^4 Z_t^4) \\ &= \frac{M}{n^2 \epsilon^2} \sum_{t=2}^n EX_{t-1}^4 \\ &\rightarrow 0 \text{ since } \frac{1}{n} \sum EX_t^4 \text{ is bounded.} \end{aligned}$$

This proves Claim 1.

$$\text{Hence for } \tau^2 = \frac{\theta^2(\sigma_1^4 + \sigma_2^4) + 2\sigma_1^2 \sigma_2^2}{2(1 - \theta^4)},$$

$$\mathbf{P}_B \left[ \frac{1}{\sqrt{n}} \sum_{t=2}^n \mathbf{W}_t \mathbf{X}_{t-1} \mathbf{Z}_t \leq \mathbf{x} \right] \xrightarrow{P} \Phi \left( \frac{\mathbf{x}}{\tau} \right) \quad \forall \mathbf{x} \in \mathbb{R} \quad (86)$$

**Claim 2.**

$$\text{With } c = \frac{1}{2} \frac{(\sigma_1^2 + \sigma_2^2)}{(1 - \theta^2)}$$

$$P_B \left[ \left| \frac{1}{n} \sum_{t=2}^n w_t X_{t-1}^2 - c \right| > \epsilon \right] \xrightarrow{P} 0 \quad \forall \epsilon > 0 \quad (87)$$

Note that

$$\frac{1}{n} \sum w_t X_{t-1}^2 = \frac{1}{2} \left[ \frac{\sum_{t \text{ odd}} w_t X_{t-1}^2}{n/2} + \frac{\sum_{t \text{ even}} w_t X_{t-1}^2}{n/2} \right]$$

Using the fact that  $\{X_t\}_{t \text{ even}}$  and  $\{X_t\}_{t \text{ odd}}$  form two homoscedastic AR(1) processes, from Claim 2(Theorem 1) and Remark 1, we get,

$$\frac{\sum_{t \text{ odd}} w_t X_{t-1}^2}{n/2} \xrightarrow{P_B} \frac{\theta^2 \sigma_1^2 + \sigma_2^2}{1 - \theta^4} \text{ a.s.}$$

and

$$\frac{\sum_{t \text{ even}} w_t X_{t-1}^2}{n/2} \xrightarrow{P_B} \frac{\sigma_1^2 + \theta^2 \sigma_2^2}{1 - \theta^4} \text{ a.s.}$$

Hence

$$\frac{1}{n} \sum w_t X_{t-1}^2 \xrightarrow{P_B} \frac{1}{2} \frac{(\sigma_1^2 + \sigma_2^2)}{(1 - \theta^2)} \text{ a.s.}$$

This proves Claim 2.

**Claim 3.**

$$P_B \left[ \left| \frac{1}{n} \sum_{t=2}^n (w_t - 1) X_{t-1}^2 \right| > \epsilon \right] \xrightarrow{P} 0 \quad \forall \epsilon > 0 \quad (88)$$

This follows from equations (??) and (87).

Note that as defined in (85),

$$\sqrt{n} \sigma_n^{-1} (\hat{\theta}_n^* - \hat{\theta}_n) = T_1 - T_2$$

Then from (86) and (87),

$$P_B(T_1 \leq x) - P(T \leq x) = o_P(1)$$

where  $T \sim \left[ \frac{1}{2} \frac{(\sigma_1^2 + \sigma_2^2)}{(1 - \theta^2)} \right]^{-1} N \left( 0, \frac{\sigma_1^2 \sigma_2^2 + \theta^2 (\sigma_1^4 + \sigma_2^4)/2}{(1 - \theta^4)} \right)$

Moreover using equations (87) and (88), from Claim 3(Theorem 1), we get,

$$P_B(|T_2| > \epsilon) = o_P(1) \quad \forall \epsilon > 0$$

Combining

$$P_B[\sqrt{n} \sigma_n^{-1} (\hat{\theta}_n^* - \hat{\theta}_n) \leq x] - P[Y \leq x] = o_P(1) \quad \forall x \in \mathbb{R} \quad (89)$$

where

$$Y \sim N \left( 0, 4 \frac{(1 - \theta^2) \sigma_1^2 \sigma_2^2 + \theta^2 (\sigma_1^4 + \sigma_2^4)/2}{(\sigma_1^2 + \sigma_2^2)^2} \right)$$

and this completes the proof.

**Remark 2.** In Theorems 1 and 3, we have established the consistency of the Weighted Bootstrap estimator in probability, ie we have proved,  $\forall x \in \mathbb{R}$ ,

$$P_B(\sqrt{n} \sigma_n^{-1} (\hat{\theta}_n^* - \hat{\theta}_n) \leq x) - P(\sqrt{n}(\hat{\theta}_n - \theta) \leq x) = o_P(1)$$

The same results can be achieved almost surely. One can prove that,  $\forall x \in \mathbb{R}$ ,

$$P_B(\sqrt{n} \sigma_n^{-1} (\hat{\theta}_n^* - \hat{\theta}_n) \leq x) - P(\sqrt{n}(\hat{\theta}_n - \theta) \leq x) \longrightarrow 0 \text{ a.s.}$$

To prove this, one needs to verify the conditions of Result(1) almost surely, and replace all convergence of sample moments of  $\{X_t\}$  in probability, by almost sure convergence in the proofs.

### 3 Numerical Calculations

In this section, we compare numerically the performance of the Weighted Bootstrap and Residual Bootstrap techniques for an heteroscedastic AR(1) model, and exhibit numerically, the consistency of the Weighted Bootstrap estimator. We simulated 50 observations from the AR process,

$$X_t = \theta X_{t-1} + Z_t, \quad t = 1, 2, \dots, n.$$

where  $Z_t$  is a sequence of independent Normal mean-zero random variables with  $EZ_t^2 = \sigma_1^2$  if  $t$  is odd and  $EZ_t^2 = \sigma_2^2$  if  $t$  is even. For simulation purpose, we used  $\theta = 0.5$ ,  $\sigma_1^2 = 1$ , and  $\sigma_2^2 = 2$ .

The unknown  $\theta$  is estimated by its LSE  $\hat{\theta}_n$  which came to be 0.4418.

Let  $V_n = \sqrt{n}(\hat{\theta}_n - \theta)$  be the quantity of interest which is to be estimated using resampling techniques. Let  $V_n^* = \sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n)$  denote its bootstrap estimate for two different bootstrap techniques: the Residual Bootstrap (which tacitly assumes that all the  $Z_t$ 's have same variance) and the Weighted Bootstrap. In case of WB, we used i.i.d Normal(1,1) weights. We used 200 simulations to estimate the distribution of  $V_n^*$  in both the cases.

We performed the KS test to compare the distributions of  $V_n$  and  $V_n^*$ . To estimate the distribution of  $V_n$ , we used 200 simulations from the above process. The results of the test are as follows.

#### Two-Sample Kolmogorov-Smirnov Test

Data:  $V_n$  and  $V_n^*$

Alternative hypothesis:

cdf of  $V_n$  does not equal the cdf of  $V_n^*$  for at least one sample point

BS Technique	KS value	p-value
RB	0.12	0.0945
WB	0.1	0.234

Figure 1a) presents the estimated densities of  $V_n$  and  $V_n^*$ , with  $\hat{\theta}_n^*$  being the residual bootstrap estimator, while Figure 1b) presents the estimated densities with  $\hat{\theta}_n^*$  being the weighted bootstrap estimator. From the table it can be seen that both the estimators pass the test, but WB does reasonably better. This is also obvious from the density plots.

Next we introduced more heteroscedasticity in the model. This time we took  $\sigma_1^2$  to be 1, and  $\sigma_2^2$  as 10.  $\hat{\theta}_n$  came to be 0.47083. Again we estimate  $V_n$  by  $V_n^*$  and performed a KS test to determine the goodness of the fit. Now the results are as follows:

#### Two-Sample Kolmogorov-Smirnov Test

Data:  $V_n$  and  $V_n^*$

Alternative hypothesis:

cdf of  $V_n$  does not equal the cdf of  $V_n^*$  for at least one sample point

BS Technique	KS value	p-value
RB	0.135	0.0431
WB	0.125	0.0734

Figure 2a) presents the estimated densities of  $V_n$  and  $V_n^*$  for RB, while Figure 2b) presents the estimated densities for WB. From the table, it can be seen that RB fails. This is expected since it is not adapted for heteroscedasticity. It fails to capture the true model in such a situation. WB still performs well, but its performance also falls. This is also reflected from the density plots. Perhaps a larger sample size is required in case of substantial heteroscedasticity.

This illustrates the point that for small sample sizes, at small levels of heteroscedasticity, many Bootstrap techniques perform well, but at substantial levels a careful choice is needed. The success of WB for both levels of heteroscedasticity lends further support to our theoretical results.

## 4 ARCH models

In this section, we first present the basic probabilistic properties of ARCH models. Then we introduce various estimation procedures for the parameters involved, and study their properties. The asymptotic properties of the listed estimators under different error distributions are also introduced. To approximate the distribution of the estimators and draw inference based on an observed sample, various resampling techniques are also listed along with their properties. Finally we supplement our theoretical results with numerical calculations based on a simulated ARCH data set.

### 4.1 Basic Properties of ARCH Processes

**Definition 1** An autoregressive conditional heteroscedastic (ARCH) model with order  $p$  ( $\geq 1$ ) is defined as

$$X_t = \sigma_t \epsilon_t \quad \text{and} \quad \sigma_t^2 = c_0 + b_1 X_{t-1}^2 + \dots + b_p X_{t-p}^2 \quad (90)$$

where  $c_0 \geq 0$ ,  $b_j \geq 0$  are constants,  $\epsilon_t \sim iid(0, 1)$ , and  $\epsilon_t$  is independent of  $\{X_{t-k}, k \geq 1\}$  for all  $t$ .

The necessary and sufficient condition for (90) to define a unique stationary process  $\{X_t\}$  with  $EX_t^2 < \infty$  is

$$\sum_{i=1}^p b_i < 1 \quad (91)$$

Furthermore, for such a stationary solution,  $EX_t = 0$  and  $V(X_t) = c_0 / (1 - \sum_{i=1}^p b_i)$ .

### 4.2 Estimation

We always assume that  $\{X_t\}$  is a strictly stationary solution of the ARCH model (90). Based on observations  $X_1, X_2, \dots, X_n$ , we discuss various methods for estimating parameters in the model. Listed below are four types of estimators for parameters  $c_0$  and  $b_i$ . They are the *Conditional Maximum Likelihood Estimator*, and three *Least Absolute Deviations Estimators*.

**(a) Conditional Maximum Likelihood Estimator** If  $\epsilon_t$  is normal in model (90), the negative logarithm of the (conditional) likelihood function based on observations  $X_1, X_2, \dots, X_n$ , ignoring constants, is

$$\sum_{t=p+1}^n (\log \sigma_t^2 + X_t^2 / \sigma_t^2) \quad (92)$$

The (Gaussian) maximum likelihood estimators are defined as the minimizers of the function above. Note that this likelihood function is based on the conditional probability density function of  $X_{p+1}, \dots, X_n$ , given  $X_1, \dots, X_p$ , since the unconditional probability density function, which involves the joint density of  $X_1, \dots, X_p$  is unattainable.

**(b) Least Absolute Deviations Estimators** The estimator discussed in (a) is derived from maximizing an approximate Gaussian likelihood. In this sense, it is an  $L_2$ -estimator. It is well known that  $L_1$ -estimators are more robust with respect to heavy-tailed distributions than  $L_2$ -estimators. This motivates the study of various *least absolute deviations estimators* for  $c_0$  and  $b_i$  in model (90).

Now we reparametrize the model (90) in such a way that the median of  $\epsilon_t^2$ , instead of the variance of  $\epsilon_t$ , is equal to 1 while  $E\epsilon_t = 0$  remains unchanged. Under this new reparametrization, the parameters  $c_0$  and  $b_i$  differ from those in the old setting by a common positive constant factor. Write

$$\frac{X_t^2}{\sigma_t(\theta)^2} = 1 + e_{t1} \quad (93)$$

where  $e_{t1} = (\epsilon_t^2 - 1)$  has median 0. This leads to the first absolute deviations estimator

$$\hat{\theta}_1 = \arg \min_{\theta} \sum_{t=p+1}^n |X_t^2 / \sigma_t(\theta)^2 - 1| \quad (94)$$

which is an  $L_1$  estimator based on the regression relationship (93).

Alternatively, we can define another form of least absolute estimator as

$$\hat{\theta}_2 = \arg \min_{\theta} \sum_{t=p+1}^n |\log(X_t^2 - \log(\sigma_t(\theta)^2))| \quad (95)$$

which is motivated by the regression relationship

$$\log(X_t^2) = \log(\sigma_t(\theta)^2) + e_{t2} \quad (96)$$

where  $e_{t2} = \log(\epsilon_t^2)$ . Hence median of  $e_{t2}$  is equal to  $\log\{\text{median}(\epsilon_t^2)\}$ , which is 0 under the reparameterisation.

The third L-1 estimator is motivated by the regression equation

$$X_t^2 = \sigma_t^2 + e_{t3} \quad (97)$$

where  $e_{t3} = \sigma_t^2(\epsilon_t^2 - 1)$ . Again under the new parameterisation, the median of  $e_{t3}$  is 0. This leads to the estimator

$$\hat{\theta}_3 = \arg \min_{\theta} \sum_{t=p+1}^n |X_t^2 - \sigma_t(\theta)^2| \quad (98)$$

Intuitively we prefer the estimator  $\hat{\theta}_2$  to  $\hat{\theta}_3$  since the error terms  $e_{t2}$  in regression model (96) are independent and identically distributed while the errors  $e_{t3}$  in model (97) are not independent. Another intuitive justification for using  $\hat{\theta}_2$  is that, the distribution of  $X_t^2$  is confined to the nonnegative half axis and is typically skewed. Hence the log-transformation will make the distribution less skewed. The minimization in (94), (95) and (98) is taken over all  $c_0 > 0$  and all nonnegative  $b_i$ 's.

### 4.3 Asymptotic Properties

In this section we discuss the asymptotic properties of the estimators listed above.

The conditional maximum likelihood estimation remains as one of the most frequently-used methods in fitting ARCH models. To establish the asymptotic normality of the likelihood estimator some regularity conditions are required. Let  $\{X_t\}$  be the unique strictly stationary solution from ARCH(p) model (90) in which  $\epsilon_t$  may not be normal. We assume that  $p \geq 1$ ,  $c_0 > 0$  and  $b_i > 0$  for  $i = 1, 2, \dots, p$ . Let  $(\hat{c}_0, \hat{\mathbf{a}}^T)^T$  be the estimator derived from minimizing (92), which should be viewed as a (conditional) quasimaximum likelihood estimator.

Let  $\theta = (c_0, \mathbf{a}^T)^T$ ,  $\hat{\theta} = (\hat{c}_0, \hat{\mathbf{a}}^T)^T$ , and  $\mathbf{U}_t = \frac{d\sigma_t^2}{d\theta}$ . It may be shown that  $\mathbf{U}_t / \sigma_t^4$  has all its moments finite. We assume that the matrix

$$\mathbf{M} \equiv E(\mathbf{U}_t \mathbf{U}_t^T / \sigma_t^4)$$

is positive definite. Further we assume that the errors are not very heavy tailed, ie  $E(\epsilon_t^4) < \infty$ . Then under the above regularity conditions, it can be established that (see Hall and Yao 2003)

$$\frac{\sqrt{n}}{(E(\epsilon_t^4) - 1)^{1/2}} (\hat{\theta} - \theta) \xrightarrow{d} N(0, M^{-1})$$

If  $E(\epsilon_t^4) = \infty$  the convergence rate of  $\sqrt{n}$  is no longer observable. Then the convergence rate of the likelihood estimator is dictated by the distribution tails of  $\epsilon_t^2$ ; the heavier the tails, the slower the

convergence. Moreover, the asymptotic normality of the estimator is only possible if  $E(|\epsilon_t|^{4-\delta}) < \infty$  for any  $\delta > 0$ .

The asymptotic normality of the least absolute deviations estimator  $\hat{\theta}_2$  in (95) can be established under milder conditions. To do so we will use the reparameterized model. Let  $\theta = (c_0, \mathbf{a}^T)^T$  be the true value under which the median of  $\epsilon_t^2$  equals 1, or equivalently the median of  $\log(\epsilon_t^2)$  equals 0. Define  $\mathbf{U}_t$  and  $\mathbf{M}$  as before. Again we assume there exists a unique strictly stationary solution  $\{X_t\}$  of model (90) with  $E_\theta(X_t^2) < \infty$ . The parameters  $c_0$  and  $b_i$ ,  $i = 1, 2, \dots, p$  are positive.  $\mathbf{M}$  is positive definite.  $\log(\epsilon_t^2)$  has median zero, and its density function  $f$  is continuous at zero.

Under the above conditions, there exists a sequence of local minimizers  $\hat{\theta}_2$  of (95) for which

$$\sqrt{n}(\hat{\theta}_2 - \theta) \xrightarrow{d} N(0, M^{-1}/\{4f(0)^2\})$$

(see Peng and Yao 2003). Thus the least absolute deviations estimator  $\hat{\theta}_2$  is asymptotically normal with convergence rate  $\sqrt{n}$  under very mild conditions. In particular, the tail-weight of the distribution of  $\epsilon_t$  is irrelevant as no condition is imposed on the moments of  $\epsilon_t$  beyond  $E(\epsilon_t^2) < \infty$

Similar to the above result,  $\sqrt{n}(\hat{\theta}_1 - \theta)$  is also asymptotically normal with mean

$$E[\epsilon_t^2 I(\epsilon_t^2 > 1) - \epsilon_t^2 I(\epsilon_t^2 < 1)] [E|m_{11}|, \dots, E|m_{(p+1)(p+1)}|]^T \text{ where } M = (m_{ij})_{i,j}$$

(see Peng and Yao 2003) which is unlikely to be 0. This shows that  $\hat{\theta}_1$  is often a biased estimator.

It can also be shown that  $\sqrt{n}(\hat{\theta}_3 - \theta)$  is also asymptotically normal under the additional condition  $EX_t^4 < \infty$ .

#### 4.4 Bootstrap in ARCH models

As indicated in the earlier section, the range of possible limit distributions for a (conditional) Gaussian maximum likelihood estimator is extraordinarily vast. In particular the limit laws depend intimately on the error distribution. This makes it impossible in heavy tailed cases to perform statistical tests or estimation based on asymptotic distributions in any conventional sense. Bootstrap methods seem the best option for tackling these problems.

**Residual Bootstrap(m-out-of-n) for likelihood estimator:** Let  $\tilde{\epsilon}_t = X_t/\sigma_t(\hat{\theta})$  for  $t = p+1, \dots, n$  and let  $\{\hat{\epsilon}_t\}$  be the standardized version of  $\{\tilde{\epsilon}_t\}$  such that the sample mean is zero and the sample variance is 1. We define

$$\hat{\tau}^2 = \frac{1}{n} \sum_{t=1}^n \tilde{\epsilon}_t^4 - \left( \frac{1}{n} \sum_{t=1}^n \tilde{\epsilon}_t^2 \right)^2$$

Now we draw  $\{\epsilon_t^*\}$  with replacement from  $\{\hat{\epsilon}_t\}$  and define  $X_t^* = \sigma_t^* \epsilon_t^*$  for  $t = p+1, \dots, m$  with

$$(\sigma_t^*)^2 = \hat{c}_0 + \sum_{i=1}^p \hat{b}_i (X_{t-i}^*)^2$$

and form the statistic  $(\hat{\theta}^*, \hat{\tau}^*)$  based on  $\{X_{p+1}^*, \dots, X_m^*\}$  in the same way as  $(\hat{\theta}, \hat{\tau})$  based on  $\{X_{p+1}, \dots, X_n\}$ . It has been proved that (Hall and Yao (2003)) as  $n \rightarrow \infty$ ,  $m \rightarrow \infty$ , and  $m/n \rightarrow 0$ , it holds for any convex set  $C$  that

$$\left| P \left\{ \sqrt{m} \frac{(\hat{\theta}^* - \hat{\theta})}{\hat{\tau}^*} \in C | X_1, \dots, X_n \right\} - P \left\{ \sqrt{n} \frac{(\hat{\theta} - \theta)}{\hat{\tau}} \in C \right\} \right| \rightarrow 0$$

**Weighted Bootstrap for likelihood estimator** For every  $n \geq 1$ , let  $\{w_{nt}\}$ ,  $t = 1, \dots, n$ , be real valued row-wise exchangeable random variables independent of  $\{X_t\}$ . Then we define the weighted bootstrap estimators,  $\hat{\theta}^*$  of  $\hat{\theta}$  as the minimizers of

$$\sum_{t=p+1}^n w_{nt} [\log \sigma_t^2(\hat{\theta}) + X_t^2 / \sigma_t^2(\hat{\theta})] \quad (99)$$

Under suitable regularity conditions on the weights, we can expect the consistency of  $\hat{\theta}^*$ .

It is well known that in the settings where the limiting distribution of a statistic is not normal, standard bootstrap methods are generally not consistent when used to approximate the distribution of the statistic. In particular when the the distribution of  $\epsilon_t$  is very heavy-tailed in the sense that  $E(|\epsilon_t|^d) = \infty$  for some  $2 < d < 4$ , the Gaussian likelihood estimator is no longer asymptotically normal. However the least absolute deviations estimator  $\hat{\theta}_2$  is asymptotically normal under very mild conditions. Hence we expect the Bootstrap methods to work under larger range of possible distributions for  $\hat{\theta}_2$ .

**Weighted Bootstrap for  $\hat{\theta}_2$**  As in (99) we define the weighted bootstrap estimators,  $\hat{\theta}_2^*$  of  $\hat{\theta}_2$  as the minimizers of

$$\sum_{t=p+1}^n w_{nt} |\log(X_t^2 - \log(\sigma_t(\theta)^2))| \quad (100)$$

Let  $\sigma_n^2 = V_B w_{ni}$ ,  $W_{ni} = \sigma_n^{-1}(w_{ni} - 1)$ , where  $P_B$ ,  $E_B$  and  $V_B$ , respectively, denote probabilities, expectations and variances with respect to the distribution of the weights, conditional on the given data  $\{X_1, \dots, X_n\}$ . The following conditions on the weights are assumed:

$$E_B(w_1) = 1 \quad (101)$$

$$0 < k < \sigma_n^2 = o(n) \quad (102)$$

$$c_{1n} = Cov(w_i, w_j) = O(n^{-1}) \quad (103)$$

Also assume that  $\sigma_n^2/n$  decreases to 0 as  $n \rightarrow \infty$ . Further assume that the conditions of Result 1 hold with  $U_{nj} = W_{nj}$ . Then it is plausible that

$$\left| P\{\sqrt{n}\sigma_n^{-1}(\hat{\theta}_2^* - \hat{\theta}_2) \leq x | X_1, \dots, X_n\} - P\{\sqrt{n}(\hat{\theta}_2 - \theta) \leq x\} \right| \xrightarrow{P} 0 \quad \forall x \in \mathbb{R}$$

## 4.5 Numerical Properties

In this section, we compare numerically the three least absolute deviation estimators with the conditional Gaussian maximum likelihood estimator for ARCH(1) model. Then we check the consistency of their Bootstrap analogues.

We took the errors  $\epsilon_t$  to have either a standard normal distribution or a standardised Student's  $t$ -distribution with  $d = 3$  or  $d = 4$  degrees of freedom. We standardized the  $t$ -distributions to ensure that their first two moments are, respectively, 0 and 1. We took  $c_0 = 1$  and  $c_1 = 0.5$  in the models. Setting the sample size  $n = 100$ , we drew 200 samples for each setting. We used different algorithms to find estimates for different estimation procedures. Since the values of the parameters  $c_0$  and  $c_1$  estimated by the least absolute deviations methods differ from the numerical values specified above by a common factor (namely the median of the square of the distribution of  $\epsilon_t$ ), for a given sample, we define the absolute error as  $\left| \frac{\hat{c}_0}{\hat{c}_1} - \frac{c_0}{c_1} \right|$  where  $\hat{c}_0$  and  $\hat{c}_1$  are the respective sample estimates. We average the error over all our samples to obtain the sample average absolute error for an estimation procedure.

The table below displays the average absolute error for the different estimation procedures. The first column indicates distribution of  $\epsilon_t$ , the second column are the estimation procedures, and in the third column are the corresponding average error values.

Distn.	Estimate	Average error
Normal	$\hat{\theta}_{ml}$	2.548
Normal	$\hat{\theta}_1$	6.936
Normal	$\hat{\theta}_2$	5.274
Normal	$\hat{\theta}_3$	16.559
t-3	$\hat{\theta}_{ml}$	11.097
t-3	$\hat{\theta}_1$	5.750
t-3	$\hat{\theta}_2$	2.307
t-3	$\hat{\theta}_3$	56.259
t-4	$\hat{\theta}_{ml}$	13.107
t-4	$\hat{\theta}_1$	7.054
t-4	$\hat{\theta}_2$	4.528
t-4	$\hat{\theta}_3$	24.253

Figures 3a), 3b) and 3c) present the boxplots for the absolute errors with error distributions being normal,  $t_3$  and  $t_4$  respectively. For models with heavy-tailed errors, eg  $\epsilon_t \sim t_d$  with  $d = 3, 4$  the least absolute deviation estimator  $\hat{\theta}_2$  performed best. Furthermore, the gain was more pronounced when the tails were very heavy, eg  $\epsilon_t \sim t_3$ . From the boxplot, it can be seen that, when  $\epsilon_t \sim t_4$ , except for a few outliers, the Gaussian maximum likelihood estimator  $\hat{\theta}_{ml}$  was almost as good as  $\hat{\theta}_1$  and  $\hat{\theta}_2$ . However, when  $\epsilon_t \sim t_3$ ,  $\hat{\theta}_{ml}$  was no longer desirable. On the other hand, when the error  $\epsilon_t$  was normal,  $\hat{\theta}_{ml}$  was of course the best. In fact the absolute error of  $\hat{\theta}_{ml}$  was larger when the tail of the error distribution was heavier, which reflects the fact that, heavier the tails are, slower is the convergence rate; see Hall and Yao (2003). However this is not the case for the least absolute deviations estimators as they are more robust against heavy tails.

Overall the numerical results suggest that we should use the least absolute deviations estimator  $\hat{\theta}_2$  when  $\epsilon_t$  has heavy and especially very heavy tails, eg  $E(|\epsilon_t|^3) = \infty$ , while in general the Gaussian maximum likelihood estimator  $\hat{\theta}_{ml}$  is desirable as long as  $\epsilon_t$  is not very heavy-tailed.

Next we check the consistency of the bootstrap estimators,  $\hat{\theta}_{mle}^*$  and  $\hat{\theta}_2^*$  of  $\hat{\theta}_{mle}$  and  $\hat{\theta}_2$  respectively. We fixed a sample of size 100 from the ARCH(1) process with standard normal errors, and used 200 simulations for four different resampling techniques: the RB, the m-out-of-n RB and the WB. For the m-out-of-n RB, we took  $m$  to be 50. Comparing the values of  $V_n$  and  $V_n^*$ , the results of the KS test are:

### Two-Sample Kolmogorov-Smirnov Test

Data:  $V_n$  and  $V_n^*$

Alternative hypothesis:

cdf of  $V_n$  does not equal the cdf of  $V_n^*$  for at least one sample point

Estimate	BS Technique	KS value	p-value
$\hat{c}_{0ml}$	WB	0.095	0.286
$\hat{c}_{1ml}$	WB	0.110	0.152
$\hat{c}_{0ml}$	RB	0.170	0.005
$\hat{c}_{1ml}$	RB	0.125	0.073
$\hat{c}_{0ml}$	RB(m/n)	0.1	0.234
$\hat{c}_{1ml}$	RB(m/n)	0.095	0.286
$\hat{c}_{02}$	WB	0.095	0.286
$\hat{c}_{12}$	WB	0.130	0.057

In the table above,  $\hat{c}_{0ml}$  and  $\hat{c}_{1ml}$  denote the estimates of  $c_0$  and  $c_1$  respectively using the maximum likelihood estimation procedure, while  $\hat{c}_{02}$  and  $\hat{c}_{12}$  denote the corresponding estimates using

the least absolute deviations estimator. From the table, it can be seen that the full sample (i.e. n-out-of-n)bootstrap fails, while m-out-of-n RB fares better. The reason that the full-sample RB fails to be consistent is that it does not accurately model relationships among extreme order statistics in the sample; see Fan and Yao 2003. WB does reasonably well for both maximum likelihood and least absolute deviations estimation procedures.

## References

- [1] Bose, A. and Chatterjee, S. (2003). Generalized Bootstrap for estimators of minimizers of convex functions. *Journal of Stat. Planning and Inference* 117, 225-239.
- [2] Brockwell, J.P. and Davis, A.R. (1990) *Time Series: Theory and Methods*. New York: Springer.
- [3] Chatterjee, S. and Bose, A. (2004). Generalized Bootstrap for Estimating Equations. To appear in *Ann. Statist.*
- [4] Davis, R.A., Knight, K. and Liu, J. (1992). M-estimation for autoregressions with infinite variances. *Stoch. Proces. Applic.* 40, 145-180.
- [5] Fan, J. and Yao, Q. (2003) *Nonlinear Time Series: Nonparametric and Parametric Methods* . New York: Springer, 125-168.
- [6] Hall, P. and Heyde, C.C. (1980) *Martingale Limit Theory and its Applications*. New York: Academic Press, 6-53.
- [7] Hall, P and Yao, Q. (2003). Inference in ARCH and GARCH models with heavy-tailed errors. *Econometrica* 71, 285-317.
- [8] Peng, L and Yao, Q (2003). Least absolute deviations estimations for ARCH and GARCH models. Preprint.
- [9] Praestgaard, J. and Wellner, J.A. (1993). Exchangeably Weighted Bootstrap of the General Empirical Process. *Ann. Probab.* 21, 2053-2086.

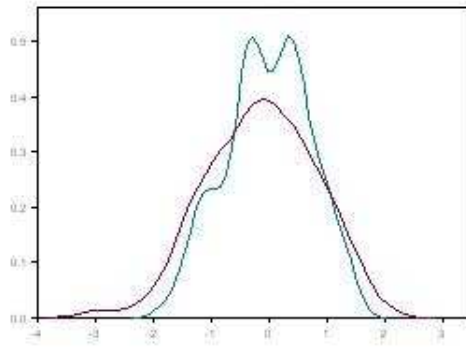


Figure 1(a)

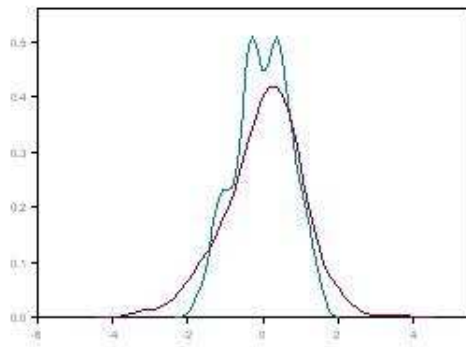


Figure 1(b)

**Figure1:** Sample density plots of  $V_n$  and  $V_n^*$  with  $\sigma_1^2 = 1$  and  $\sigma_2^2 = 2$ . The green line denotes density of  $V_n$ , the red line for density of  $V_n^*$ . (a)  $\hat{\theta}_n^*$  is the residual bootstrap estimator, (b)  $\hat{\theta}_n^*$  is the weighted bootstrap estimator.

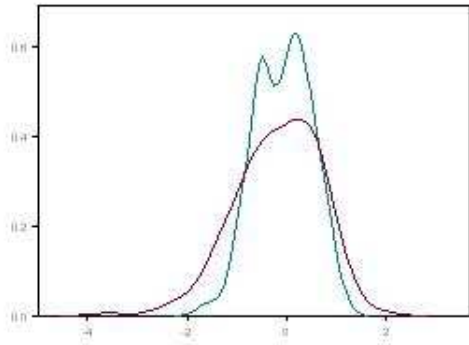


Figure 2(a)

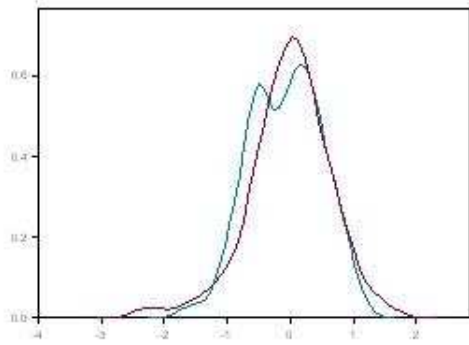


Figure 2(b)

**Figure2:** Sample density plots of  $V_n$  and  $V_n^*$  with  $\sigma_1^2 = 1$  and  $\sigma_2^2 = 10$ . The green line denotes density of  $V_n$ , the red line for density of  $V_n^*$ . (a)  $\hat{\theta}_n^*$  is the residual bootstrap estimator, (b)  $\hat{\theta}_n^*$  is the weighted bootstrap estimator.

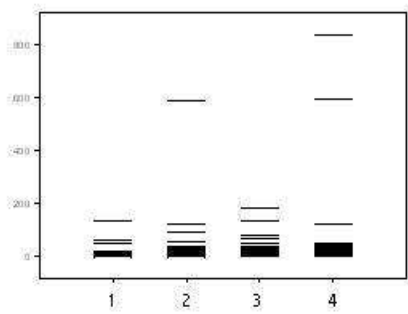


Figure 3(a)

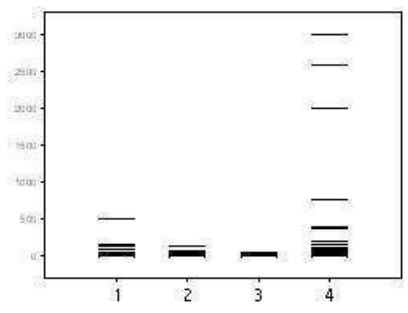


Figure 3(b)

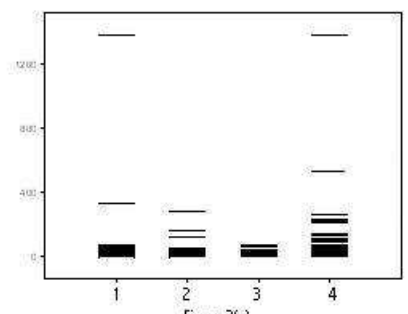


Figure 3(c)

**Figure 3:** Box plots of the absolute errors of the maximum likelihood estimates (MLE), and the three least absolute deviations estimates (LADE). Labels 1, 2, 3 and 4 denote respectively the MLE, LADE1 -  $\hat{\theta}_1$ , LADE2 -  $\hat{\theta}_2$  and LADE3 -  $\hat{\theta}_3$ . (a) Error  $\epsilon_t$  has normal distribution, (b) Error  $\epsilon_t$  has  $t_3$  distribution, (c) Error  $\epsilon_t$  has  $t_4$  distribution.