Working Paper 93-24
Statistics and Econometrics Series 19
October, 1993

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

PREDICTION INTERVALS FOR NEARLY NONSTATIONARY AR(1)-PROCESSES

Nélida Ferretti and Juan Romo*

Abstract

We construct prediction intervals for the observations of first-order autoregressive processes when the model approaches a nonstationary situation with a unit root. The intervals that we propose contain an s-step future value with a given asymptotic probability conditional on the observation. A simulation study has been also carried out to illustrate our results.

Key Words
Prediction intervals, nearly nonstationary time series, coverage probability.

*Ferretti, Universidad de la Plata and CONICET; Romo, Universidad Carlos III de Madrid.
Research partially supported by Cátedra ARGENTARIA (Universidad Carlos III de Madrid, Spain) and by DGICYT PB90-0266 (Spain).

PREDICTION INTERVALS FOR NEARLY NONSTATIONARY AR(1)-PROCESSES

NÉLIDA FERRETTI [1]
Universidad de la Plata, CONICET and Universidad Carlos III de Madrid
Departamento de Matemática, Casilla de Correo Nro. 172, 1900 La Plata, Argentina

and

JUAN ROMO[2]
Universidad Carlos III de Madrid
Departamento de Estadística y Econometría, Madrid 126, 28903 Getafe (Madrid), Spain

Summary

We construct prediction intervals for the observations of first-order autoregressive processes when the model approaches a nonstationary situation with a unit root. The intervals that we propose contain an s-step ahead future value with a given asymptotic probability conditional on the observations. A simulation study has been also carried out to illustrate our results.

Keywords: Prediction intervals; nearly nonstationary time series; coverage probability.
1991 MSC: 62M20.

[1]Research partially supported by Cátedra ARGENTARIA (Universidad Carlos III de Madrid, Spain)
[2]Research partially supported by DGICYT PB90-0266 (Spain).

1 INTRODUCTION

The main goal of this paper is to provide s-step ahead prediction intervals for the observations in nearly nonstationary first-order autoregressive models. We consider the sequence of first-order autoregressive AR(1) models,

$$X_t(n) = \beta_n X_{t-1}(n) + \varepsilon_t, \quad t = 1, \ldots, n, \quad \beta_n = 1 - \frac{\gamma}{n}, \quad \gamma \in I\!R \tag{1.1}$$

where $\{\varepsilon_t\}$ are independent and identically distributed random variables with distribution function F, zero mean and finite variance σ^2. Assume that $X_0(n) = 0$. To simplify notation, $X_t(n)$ will be written as X_t throughout this paper.

This kind of nearly nonstationarity has been previously considered by several authors as Bobkoski (1983), Ahtola and Tiao (1984), Tsay (1985), Chan and Wei (1987), Jeganathan (1987), Phillips (1987), Chan (1988), Cox and Llatas (1991) and Cox (1991). In particular, Bobkoski (1983) has studied the asymptotic behaviour of the least-squares estimate

$$\hat{\beta}_n = \left(\sum_{t=1}^{n} X_{t-1}^2\right)^{-1} \sum_{t=1}^{n} X_{t-1} X_t,$$

showing that

$$n(\hat{\beta}_n - \beta_n) \to_w \left\{\int_0^1 Y^2(t)dt\right\}^{-1} \int_0^1 Y(t)dW(t) \tag{1.2}$$

where $Y(t)$ is the Ornstein-Uhlenbeck process defined by the stochastic differential equation

$$dY(t) = -\gamma Y(t)dt + \sigma dW(t), \quad Y(0) = 0$$

and $\{W(t) : 0 \leq t \leq 1\}$ is a standard Brownian motion. Chan and Wei (1987) obtained that

$$\left(\sum_{t=1}^{n} X_{t-1}^2\right)^{1/2} (\hat{\beta}_n - \beta_n) \to_w \mathcal{L}(\gamma), \tag{1.3}$$

where

$$\mathcal{L}(\gamma) = \left\{\int_0^1 (1 + (e^{2\gamma} - 1)t)^{-2} W^2(t)dt\right\}^{-1/2} \int_0^1 (1 + (e^{2\gamma} - 1)t)^{-1} W(t)dW(t).$$

Finally, Chan (1988) showed that the variables $\mathcal{L}(\gamma)$ and $\left\{\int_0^1 Y^2(t)dt\right\}^{-1/2} \int_0^1 Y(t)dW(t)$ have the same distribution. Cox and Llatas (1991) consider the asymptotic properties of a class of maximum likelihood type estimators of β_n.

Our aim is to construct prediction intervals $I_{n,s}$, $s \geq 1$, such that, for a given coverage level $1-\alpha$, the probability $P\{X_{n+s} \in I_{n,s} \mid \mathcal{F}_n\}$ be asymptotically at least $1-\alpha$, where $\mathcal{F}_n \equiv \sigma(X_1, \ldots, X_n)$ is the σ-algebra generated by the observations $X_1, \ldots, X_n$. Since

$$X_{n+s} = \beta_n^s X_n + \sum_{j=1}^{s} \beta_n^{s-j} \varepsilon_{n+j}, \tag{1.4}$$

let G_s be the distribution function of $\sum_{j=1}^{s} \beta_n^{s-j} \varepsilon_{n+j}$ and consider the interval

$$I_{n,s} = (\beta_n^s X_n + G_s^{-1}(u_1), \beta_n^s X_n + G_s^{-1}(u_2)), \quad 0 < u_1, u_2 < \frac{1}{2}. \tag{1.5}$$

If G_s is continuous then $P\{X_{n+s} \in I_{n,s} \mid \mathcal{F}_n\} = 1 - u_1 - u_2$. Let $\hat{X}_{n+s} = \hat{\beta}_n^s X_n$ be the predictor for X_{n+s}.

Since G_s is unknown, we have to estimate it from the sample; to this end, consider the residuals $\hat{\varepsilon}_t \equiv X_t - \hat{\beta}_n X_{t-1}$, $t = 1, \ldots, n$ and let $\hat{F}_n$ and F_n be the empirical distributions corresponding to $\hat{\varepsilon}_t$ and ε_t, $t = 1, \ldots, n$, respectively. The sample version of (1.5) is

$$\hat{I}_{n,s} = (\hat{\beta}_n^s X_n + \hat{G}_{n,s}^{-1}(u_1), \hat{\beta}_n^s X_n + \hat{G}_{n,s}^{-1}(u_2)), \tag{1.6}$$

where $\hat{G}_{n,s}$ will be an appropriately weighted s-fold convolution of $\hat{F}_n$ and given $0 < \alpha < 1$, we want to get u_1 and u_2 such that

$$P\{P\{X_{n+s} \in \hat{I}_{n,s} \mid \mathcal{F}_n\} \geq 1 - \alpha\} \longrightarrow 1 \quad as \quad n \longrightarrow \infty. \tag{1.7}$$

Analogous prediction intervals in the explosive case have been obtained by Stute and Gründer (1993); the bootstrap version of these intervals can be seen in Stute and Gründer (1990).

The article is organized as follows. In Section 2 we prove that these prediction intervals contain an s-step ahead value with the given asymptotic coverage probability, conditionally on the observations. Section 3 presents the results of a Monte Carlo experiment that gives an estimate of the conditional coverage probability. Finally, the Appendix contains the proofs of some auxiliary results.

2 Results

We will obtain our main result assuming that the distribution function function F of the innovations satisfies

A_1. F is differentiable with $\| F'' \|_\infty < \infty$.

A_2. $F'(x) \leq \frac{c}{|x|}$ for all $| x | > K$ and some finite constant c.

A_3. F^{-1} is continuous.

We start by giving an expression for the conditional probability in (1.7); since $(\varepsilon_{n+1},\ldots,\varepsilon_{n+s})$ and $(X_1,\ldots,X_n)$ are independent, and using (1.4),

$$\begin{aligned} P\{X_{n+s}\le \hat\beta_n^s X_n+\hat G_{n,s}^{-1}(u_2)\mid \mathcal{F}_n\} &= P\{X_{n+s}\le (\hat\beta_n^s-\beta_n^s)X_n+\beta_n^s X_n+\hat G_{n,s}^{-1}(u_2)\mid \mathcal{F}_n\}\\ &= P\{\sum_{j=1}^{s}\beta_n^{s-j}\varepsilon_{n+j}\le (\hat\beta_n^s-\beta_n^s)X_n+\hat G_{n,s}^{-1}(u_2)\mid \mathcal{F}_n\}\\ &= G_s((\hat\beta_n^s-\beta_n^s)X_n+\hat G_{n,s}^{-1}(u_2)),\end{aligned}$$

and the covering probability of the interval in (1.6) is

$$G_s((\hat\beta_n^s-\beta_n^s)X_n+\hat G_{n,s}^{-1}(u_2))-G_s((\hat\beta_n^s-\beta_n^s)X_n+\hat G_{n,s}^{-1}(u_1)).$$

Let us define the function

$$\hat H_{n,s}(x)=G_s(x+\hat G_{n,s}^{-1}(u_2))-G_s(x+\hat G_{n,s}^{-1}(u_1)),$$

we will show that

$$\hat H_{n,s}((\hat\beta_n^s-\beta_n^s)X_n)\to 1 \tag{2.1}$$

in probability as $n\to\infty$ and this implies (1.7).

We will prove (2.1) in our main result; to this end, we need some previous lemmas. The first one gives information on the behaviour of the empirical distribution corresponding to the residuals $\hat\varepsilon_t$, $t=1,\ldots,n$.

LEMMA 2.1. If F satisfies condition A_1 then

$$\|\hat F_n-F_n\|_\infty\to_P 0 \quad as \quad n\to\infty.$$

Proof. See the Appendix.

Boldin (1982) showed that $n^{1/2}\|\hat F_n-F_n\|_\infty\to_P 0$ in the stationary case. The convergence in probability for the explosive case was established by Koul and Levental (1989) and this result was improved to almost sure convergence by Stute and Gründer (1993).

LEMMA 2.2. If F satisfies conditions A_1 and A_2, then

$$\| \hat{G}_{n,s} - G_{n,s} \|_\infty \rightarrow_P 0 \quad as \quad n \rightarrow \infty,$$

where

$$G_{n,s}(x) = F_n(x) * F_n(x/\beta_n) * \ldots * F_n(x/\beta_n^{s-1})$$

and

$$\hat{G}_{n,s}(x) = \hat{F}_n(x) * \hat{F}_n(x/\hat{\beta}_n) * \ldots * \hat{F}_n(x/\hat{\beta}_n^{s-1}).$$

Proof. The Lemma follows in a similar way as Lemma 2.2 in Stute and Gründer (1993). □

LEMMA 2.3. Assume that assumptions A_1 and A_2 hold. Then

$$\| \hat{G}_{n,s} - G_s \|_\infty \rightarrow_P 0 \quad as \quad n \rightarrow \infty$$

and

$$\| G_{n,s} - G_s \|_\infty \rightarrow_P 0 \quad as \quad n \rightarrow \infty.$$

Proof. The proof is similar to the one of Lemma 2.3 in Stute and Gründer (1993). □

LEMMA 2.4. Assume that u_1 and u_2 are continuity points of G_s^{-1}. Then, under assumptions A_1 and A_2,

$$\| \hat{H}_{n,s} - H_s \|_\infty \rightarrow_P 0, \quad as \quad n \rightarrow \infty,$$

where $H_s(x) = G_s(x + G_s^{-1}(u_2)) - G_s(x + G_s^{-1}(u_1))$.

Proof. The proof of this Lemma is an immediate consequence of the Mean Value Theorem, Lemma 2.3 and the fact that G_s has a bounded derivative. □

Now, we establish our main result which gives that the prediction intervals have asymptotically the correct covering probability.

THEOREM 2.1. Assume that assumptions A_1, A_2 and A_3 hold. Then, given $0 < \alpha < 1$ there exist u_1 and u_2 such that

$$P\{P\{X_{n+s} \in \hat{I}_{n,s} \mid \mathcal{F}_n\} \geq 1-\alpha\} \rightarrow 1 \quad as \quad n \rightarrow \infty. \tag{2.2}$$

Proof. We have that

$$P\{X_{n+s} \in \hat{I}_{n,s} \mid \mathcal{F}_n\} = \hat{H}_{n,s}((\hat{\beta}_n^s - \beta_n^s)X_n).$$

First, we will prove that

$$(\hat{\beta}_n^s - \beta_n^s)X_n \rightarrow_P 0 \quad as \quad n \rightarrow \infty. \tag{2.3}$$

By Taylor's formula we get

$$(\hat{\beta}_n^s - \beta_n^s)X_n = s\beta_n^{s-1}(\hat{\beta}_n - \beta_n)X_n + \frac{1}{2}s(s-1)\tilde{\beta}_n^{s-2}(\hat{\beta}_n - \beta_n)^2 X_n \tag{2.4}$$

where $\tilde{\beta}_n$ is an intermediate point beetwen $\hat{\beta}_n$ and β_n.

From (1.2) and Lemma 2.1 of Chan and Wei (1987) we obtain

$$s\beta_n^{s-1}(\hat{\beta}_n - \beta_n)X_n \rightarrow_P 0 \quad as \quad n \rightarrow \infty \tag{2.5}$$

and

$$\frac{1}{2}s(s-1)\tilde{\beta}_n^{s-2}(\hat{\beta}_n - \beta_n)^2 X_n \rightarrow_P 0 \quad as \quad n \rightarrow \infty. \tag{2.6}$$

So, from (2.4), (2.5) and (2.6) one concludes (2.3).

Since by Lemma 2.5 of Stute and Gründer (1993), H_s is continuous, it follows from Lemma 2.4 that

$$P\{X_{n+s} \in \hat{I}_{n,s} \mid \mathcal{F}_n\} \rightarrow_w H_s(0) \quad as \quad n \rightarrow \infty. \tag{2.7}$$

Therefore, if we choose u_1 and u_2 such that

$$H_s(0) = u_2 - u_1 \geq 1 - \alpha,$$

and if $1-\alpha$ is a point of continuity of the limit distribution, we obtain (2.2). This completes our proof. □

3 A MONTE CARLO STUDY

We have performed a Monte Carlo study to investigate the approximation (1.7). The proportion of $\{X_{n+s} \in \hat{I}_{n,s}\}$ -values, based on 20,000 Monte Carlo replications, has been considered and $P(X_{n+s} \in \hat{I}_{n,s} \mid \mathcal{F}_n)$ was estimated by

$$\hat{P}_{n,s,\gamma} = \#\{X_{n+s} \in \hat{I}_{n,s}\}/20,000,$$

where $\#A$ indicates the cardinality of the set A and recall that γ controls the size of β_n. We have combined three different sample sizes n ($n = 25, 50, 100$) with ten values of s ($s = 1, \ldots, 10$) and F is the standard normal distribution function.

We have used some routines from IMSL Library : GGUBS (basic uniform $(0,1)$ pseudo-random number generator) and GGNML (normal random deviate generator). The computer programs were written in FORTRAN and performed in a PC/AT/486 at the Departamento de Matemática, Universidad de La Plata.

For each choice of n and s, Table 1 and Table 2 contain values of $\hat{P}_{n,s,\gamma}$ with $\alpha = 0.05$ and $\gamma = 0, \pm 1, \pm 2, \pm 5, \pm 10$ and with $\alpha = 0.1$ and $\gamma = 0, 1, 2, 5, 10$, respectively.

If n and γ are fixed $\hat{P}_{n,s,\gamma}$ increases as s increases. For large positive (large negative) values of γ (i.e., β_n stays away from 1) and for a fixed n $\hat{P}_{n,s,\gamma}$ is getting closer to the nominal level $(1 - \alpha)$ as s increases (decreases).

For $n = 25$ and $0 \leq \gamma \leq 2$ $(0.92 \leq \beta_n \leq 1)$, $n = 50$ and $0 \leq \gamma \leq 5$ $(0.90 \leq \beta_n \leq 1)$ and for $n = 100$ and $0 \leq \gamma \leq 10$ $(0.90 \leq \beta_n \leq 1)$ the values of $\hat{P}_{n,s,\gamma}$ are close to the nominal level $1 - \alpha = 0.95$ or 0.90 for almost all considered values of s.

Table 1 shows that the values of $\hat{P}_{n,s,\gamma}$ are larger or equal to 0.95 for $\gamma < 0$ in all values of s. These quantities are close to 0.95 for $1 < \beta_n \leq 1.05$ $(\gamma = -1 : n = 25$ and $s \leq 2, n = 50$ and $s \leq 4, n = 100$ and $s \leq 7; \gamma = -2 : n = 50$ and $s \leq 2, n = 100$ and $s \leq 5; \gamma = -5 : n = 100$ and $s = 1)$.

TABLE 1
An estimator of $P(X_{n+s} \in \hat{I}_{n,s}|\mathcal{F}_n)$
when $\alpha = 0.05$ and F is Normal

n	γ	s = 1	2	3	4	5	6	7	8	9	10
25	-10	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	-5	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	-2	0.97	0.97	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	-1	0.96	0.96	0.97	0.98	0.98	0.99	0.99	0.99	0.99	0.99
	0	0.94	0.95	0.96	0.97	0.97	0.97	0.98	0.98	0.98	0.98
	1	0.93	0.94	0.95	0.96	0.96	0.97	0.97	0.97	0.97	0.98
	2	0.92	0.93	0.95	0.95	0.96	0.96	0.97	0.97	0.97	0.97
	5	0.87	0.91	0.93	0.94	0.94	0.95	0.95	0.95	0.96	0.96
	10	0.77	0.86	0.89	0.89	0.90	0.90	0.90	0.91	0.91	0.91
50	-10	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	-5	0.97	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	-2	0.96	0.96	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.99
	-1	0.95	0.96	0.96	0.96	0.97	0.97	0.98	0.98	0.98	0.98
	0	0.94	0.95	0.95	0.96	0.96	0.96	0.97	0.97	0.97	0.98
	1	0.94	0.95	0.95	0.95	0.95	0.96	0.96	0.96	0.96	0.97
	2	0.93	0.94	0.94	0.94	0.95	0.96	0.96	0.96	0.96	0.96
	5	0.91	0.93	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	10	0.87	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95
100	-10	0.98	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
	-5	0.96	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.99	0.99
	-2	0.95	0.96	0.96	0.96	0.96	0.97	0.97	0.97	0.98	0.98
	-1	0.94	0.95	0.95	0.96	0.96	0.96	0.96	0.97	0.97	0.97
	0	0.94	0.95	0.95	0.95	0.95	0.95	0.96	0.96	0.96	0.96
	1	0.94	0.94	0.94	0.95	0.95	0.95	0.95	0.95	0.96	0.96
	2	0.93	0.94	0.94	0.95	0.95	0.95	0.95	0.95	0.95	0.96
	5	0.92	0.93	0.94	0.94	0.94	0.94	0.95	0.95	0.95	0.95
	10	0.91	0.93	0.93	0.94	0.94	0.94	0.94	0.95	0.95	0.96

TABLE 2
An estimator of $P(X_{n+s} \in \hat{I}_{n,s}|\mathcal{F}_n)$
when $\alpha = 0.1$ and F is Normal

		s									
n	γ	1	2	3	4	5	6	7	8	9	10
25	0	0.89	0.90	0.92	0.93	0.94	0.94	0.95	0.96	0.96	0.96
	1	0.87	0.89	0.90	0.92	0.92	0.93	0.93	0.93	0.94	0.95
	2	0.85	0.88	0.89	0.91	0.91	0.91	0.92	0.92	0.93	0.93
	5	0.79	0.84	0.87	0.88	0.89	0.89	0.89	0.89	0.91	0.91
	10	0.66	0.76	0.79	0.80	0.80	0.82	0.82	0.82	0.82	0.82
50	0	0.89	0.90	0.90	0.91	0.92	0.92	0.93	0.93	0.94	0.94
	1	0.88	0.89	0.89	0.89	0.91	0.91	0.91	0.92	0.92	0.93
	2	0.87	0.89	0.89	0.89	0.90	0.90	0.90	0.91	0.91	0.92
	5	0.85	0.87	0.88	0.88	0.89	0.89	0.89	0.90	0.90	0.91
	10	0.80	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.89	0.89
100	0	0.89	0.90	0.90	0.90	0.91	0.91	0.91	0.91	0.91	0.92
	1	0.88	0.89	0.89	0.90	0.90	0.90	0.90	0.91	0.91	0.91
	2	0.88	0.89	0.89	0.90	0.90	0.90	0.90	0.90	0.90	0.91
	5	0.87	0.89	0.89	0.89	0.89	0.89	0.89	0.90	0.90	0.90
	10	0.84	0.87	0.88	0.89	0.89	0.89	0.89	0.89	0.89	0.90

4 Conclusion

We have provided prediction intervals for a nearly nonstationary $AR(1)$ model and we have proved that these intervals contain an s-step ahead future value with a given asymptotic probability conditionally on the observations.

Moreover we have presented results from a Monte Carlo study that confirm the theoretical results. The approximation (1.7) is good even for moderate sample sizes and for $0.90 < \beta_n < 1.05$.

APPENDIX

In the sequel, $o(1)$ ($o_p(1)$) will represent a sequence of numbers (r.v.'s) converging (in probability) to zero; $O(1)$ stands for a bounded sequence of numbers. For a real number x, $[x]$ is the greatest integer smaller than x.

In Lemma A1 we state a preliminary technical result that we need to prove Lemma 2.1.

LEMMA A1. Let $\beta_n = 1 - \gamma/n, \quad \gamma \in I\!R, \gamma \neq 0$. Then

$$\frac{1}{n}\sum_{j=1}^{n} |\beta_n|^{2(n-j)} \to (1 - e^{-2\gamma})/2\gamma, \quad as \quad n \to \infty. \tag{A.1}$$

Proof. The following proof is similar to that of (2.3) in Chan and Wei (1987). For any $\delta > 0$, choose $0 = t_0 \leq t_1 \leq \ldots \leq t_k = 1$ such that

$$\max_{1 \leq i \leq k} |e^{-2\gamma(1-t_i)} - e^{-2\gamma(1-t_{i-1})}| < \delta. \tag{A.2}$$

Hence,

$$\begin{aligned} &\left| \int_0^1 e^{-2\gamma(1-t)} dt - \sum_{i=1}^{k} e^{-2\gamma(1-t_{i-1})}(t_i - t_{i-1}) \right| \\ &\leq \sum_{i=1}^{k} \int_{t_i-1}^{t_i} |e^{-2\gamma(1-t)} - e^{-2\gamma(1-t_{i-1})}| dt < \delta. \end{aligned} \tag{A.3}$$

Let $I_i = \{ l \in I\!N : [nt_{i-1}] < l \leq [nt_i] \}$ and $c_\gamma = (1 - e^{-2\gamma})/2\gamma$. Then

$$\begin{aligned} B_n &= \frac{1}{n}\sum_{j=1}^{n} |\beta_n|^{2(n-j)} - c_\gamma \\ &= \frac{1}{n}\sum_{i=1}^{k}\sum_{l \in I_i} |\beta_n|^{2(n-l)} - c_\gamma \\ &= \frac{1}{n}\sum_{i=1}^{k}\sum_{l \in I_i} [|\beta_n|^{2(n-l)} - |\beta_n|^{2(n-[nt_{i-1}])}] \\ &+ \frac{1}{n}\sum_{i=1}^{k}\sum_{l \in I_i} |\beta_n|^{2(n-[nt_{i-1}])} - \sum_{i=1}^{k} |\beta_n|^{2(n-[nt_{i-1}])}(t_i - t_{i-1}) \\ &+ \sum_{i=1}^{k} |\beta_n|^{2(n-[nt_{i-1}])}(t_i - t_{i-1}) - c_\gamma \\ &= B_{1n} + B_{2n} + B_{3n}, \end{aligned} \tag{A.4}$$

where

$$B_{1n} = \frac{1}{n}\sum_{i=1}^{k}\sum_{l\in I_i}[|\beta_n|^{2(n-l)} - |\beta_n|^{2(n-[nt_{i-1}])}],$$

$$B_{2n} = \frac{1}{n}\sum_{i=1}^{k}\sum_{l\in I_i}|\beta_n|^{2(n-[nt_{i-1}])} - \sum_{i=1}^{k}|\beta_n|^{2(n-[nt_{i-1}])}(t_i - t_{i-1})$$

and

$$B_{3n} = \sum_{i=1}^{k}|\beta_n|^{2(n-[nt_{i-1}])}(t_i - t_{i-1}) - c_\gamma.$$

From

$$|\beta_n|^{2(n-[nt_j])} \longrightarrow e^{-2\gamma(1-t_j)} \quad as \quad n \to \infty \tag{A.5}$$

and (A.2), we have

$$\begin{aligned}
&\max_{1\le i\le k}\max_{l\in I_i}||\beta_n|^{2(n-l)} - |\beta|^{2(n-[nt_{i-1}])}| \\
&\max_{1\le i\le k}||\beta_n|^{2(n-[nt_i])} - |\beta_n|^{2(n-[nt_{i-1}])}| \\
&\le \max_{1\le i\le k}|e^{-2\gamma(1-t_i)} - e^{-2\gamma(1-t_{i-1})}| + o(1) \\
&\le \delta + o(1).
\end{aligned}$$

Thus,

$$|B_{1n}| \le \delta + o(1). \tag{A.6}$$

Since

$$\max_{1\le i\le k}|\beta_n|^{2(n-[nt_{i-1}])} \le e^{|2\gamma|}$$

we obtain

$$|B_{2,n}| \le e^{2|\gamma|}\sum_{i=1}^{k}|\frac{1}{n}([nt_i] - [nt_{i-1}]) - (t_i - t_{i-1})| = o(1) \tag{A.7}$$

Moreover, since $c_\gamma = \int_0^1 e^{-2\gamma(1-t)}dt$ it follows that

$$|B_{3n}| \leq |\sum_{i=1}^{k} |\beta|_n^{2(n-[nt_{i-1}])}(t_i - t_{i-1}) - \sum_{i=1}^{k} e^{-2\gamma(1-t_{i-1})}(t_i - t_{i-1})|$$
$$+|\sum_{i=1}^{k} e^{-2\gamma(1-t_{i-1})}(t_i - t_{i-1}) - \int_0^1 e^{-2\gamma(1-t)}dt|. \tag{A.8}$$

Then from (A.3) and (A.5) there exits N_0 such that

$$|B_{3n}| < 2\delta, \quad \forall n \geq N_0. \tag{A.9}$$

So, from (A.4), (A.6), (A.7) and (A.9) one concludes (A.1) This completes our proof. □

Proof of Lemma 2.1. We have

$$F_n(x) = \frac{1}{n}\sum_{t=1}^{n} \Delta_t(x)$$

where $\Delta_t(x) = 1$ if $\epsilon_t < x$ and 0 otherwise. Since $\hat{\varepsilon}_{n,t} = \varepsilon_t - (\hat{\beta}_n - \beta_n)X_{t-1}$,

$$\hat{F}_n(x) = \frac{1}{n}\sum_{t=1}^{n} \Delta_t(x + (\hat{\beta}_n - \beta_n)X_{t-1}).$$

Then

$$\hat{F}_n(x) - F_n(x) = Z_{1n}(x) + Z_{2n}(x),$$

where

$$Z_{1n} = \frac{1}{n}\sum_{t=1}^{n}[\Delta_t(x + (\hat{\beta}_n - \beta_n)X_{t-1}) - \Delta_t(x) - F(x + (\hat{\beta}_n - \beta_n)X_{t-1}) + F(x)]$$

and

$$Z_{2n} = \frac{1}{n}\sum_{t=1}^{n}[F(x + (\hat{\beta}_n - \beta_n)X_{t-1}) - F(x)].$$

First, we will show that

$$\sup_x \mid Z_{2n}(x) \mid \rightarrow_P 0 \quad as \quad n \rightarrow \infty. \tag{A.10}$$

By Taylor's formula we get

$$Z_{2n}(x) = \frac{1}{n} F'(x)(\hat{\beta}_n - \beta) \sum_{t=1}^{n} X_{t-1} + \frac{1}{2n} F''(\xi)(\hat{\beta}_n - \beta)^2 \sum_{t=1}^{n} X_{t-1}^2 \tag{A.11}$$

where ξ is an intermediate point between x and $x + (\hat{\beta}_n - \beta)X_{t-1}$.

As in Chan and Wei (1987), we have

$$\begin{aligned} \frac{1}{n^{3/2}} \sum_{t=1}^{n} X_{t-1} &= \frac{1}{n} \sum_{i=1}^{n-1} \beta_n^{(i-n)} X_n\left(\frac{i}{n}\right) - \int_0^1 e^{\gamma(1-t)} X_n(t) dt + \int_0^1 e^{\gamma(1-t)} X_n(t) dt \\ &= o_P(1) + \int_0^1 e^{\gamma(1-t)} X_n(t) dt \end{aligned} \tag{A.12}$$

where $X_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} \beta_n^{n-i} \epsilon_i$.

From (A.12), Lemma 2.1 of Chan and Wei (1987) and the continuous mapping theorem,

$$\frac{1}{n^{3/2}} \sum_{t=1}^{n} X_{t-1} \to_w \int_0^1 e^{\gamma(1-t)} \tilde{W}(t) dt \quad as \quad n \to \infty \tag{A.13}$$

where $\{\tilde{W}(t) : 0 \le t \le 1\}$ is a standard Brownian motion.

From Lemma 2.2 of Chan and Wei (1987) we deduce that

$$\begin{aligned} \frac{1}{n^2} \sum_{t=1}^{n} X_{t-1}^2 &= \frac{1}{n} \sum_{i=1}^{n-1} \beta_n^{2(i-n)} X_n^2\left(\frac{i}{n}\right) - \int_0^1 e^{2\gamma(1-t)} X_n^2(t) dt + \int_0^1 e^{2\gamma(1-t)} X_n^2(t) dt \\ &= o_P(1) + \int_0^1 e^{2\gamma(1-t)} X_n^2(t) dt \end{aligned} \tag{A.14}$$

By (A.14), Lemma 2.1 of Chan and Wei (1987) and the continuous mapping theorem we obtain

$$\frac{1}{n^2} \sum_{t=1}^{n} X_{t-1}^2 \to_w \int_0^1 e^{2\gamma(1-t)} \tilde{W}^2(t) dt \quad as \quad n \to \infty. \tag{A.15}$$

Hence from (1.2), (A.11), (A.13), (A.15) and the fact that $\sup_x F'(x) < \infty$, we obtain (A.10). So to prove the lemma, it suffices to show that

$$\sup_x \mid Z_{1n}(x) \mid \to_P 0 \quad as \quad n \to \infty. \tag{A.16}$$

Consider the auxiliary process (Boldin (1982))

$$Z_n(x, \eta_n, \mathbf{X}) = \frac{1}{n} \sum_{t=1}^{n} [\Delta_t(x + \eta_n X_{t-1}) - \Delta_t(x) - F(x + \eta_n X_{t-1}) + F(x)]$$

depending on x, the non-random sequence $\{\eta_n\}_n, n \in I\!N$, and the vector $\mathbf{X} = (X_0, \ldots, X_{n-1})$.

For all $\varepsilon > 0$

$$P(\sup_x \mid Z_{1n}(x) \mid > \varepsilon) \leq P(\sup_x \sup_{|\eta_n| \leq n^{-\gamma_0}} \mid Z_n(x, \eta_n, \mathbf{X}) \mid > \varepsilon) + P(n \mid \hat{\beta}_n - \beta_n \mid > n^{1-\gamma_0}) \quad (A.17)$$

where $3/5 < \gamma_0 < 1$.

From (1.2), $n(\hat{\beta}_n - \beta_n)$ is bounded in probability. Then

$$P(n \mid \hat{\beta}_n - \beta_n \mid > n^{1-\gamma_0}) \to 0 \quad as \quad n \to \infty. \quad (A.18)$$

So, to prove (A.16) it suffices to show that

$$\sup_x \sup_{|\eta_n| \leq n^{-\gamma}} \mid Z_n(x, \eta_n, \mathbf{X}) \mid \to_P 0 \quad as \quad n \to \infty \quad (A.19)$$

for all $\gamma > 3/5$.

Let $\{m_n : n \in zN\}$ and $\{N_n : n \in I\!N\}$, be sequences of integers such that $n^{-\gamma/4} 3^{m_n} \to 1$ and $n^{-1/2-\gamma/4} N_n \to 1$, $as \ n \to \infty$. We divide the interval $[-n^{-\gamma}, n^{-\gamma}]$ into 3^{m_n} parts using the points

$$\eta_{sn} = -n^{-\gamma} + 2n^{-\gamma} 3^{-m_n} s, \quad s = 0, 1, \ldots, 3^{m_n},$$

and the real line into N_n parts using the points

$$-\infty = x_0 < x_1 < \ldots < x_{N_n} = +\infty, \quad F(x_i) = i N_n^{-1}.$$

We can go over from the supremum in (A.19) to the supremum for a finite set of the points x and η_n such that

$$x \in \{x_i : i = 0, 1, \ldots, N_n\} \quad and \quad \eta_n \in \{\eta_{sn} : s = 0, 1, \ldots, 3^{m_n}\}.$$

Consider the sequences of random variables

$$\tilde{V}_{sk} = X_{k-1}(1 - 2n^{-\gamma} 3^{-m_n} \eta_{sn}^{-1} I_{\{X_{k-1} > 0\}}) \quad and \quad \hat{V}_{sk} = X_{k-1}(1 - 2n^{-\gamma} 3^{-m_n} \eta_{sn}^{-1} I_{\{X_{k-1} \leq 0\}})$$

where I_A denotes the indicator of the event A. If $\eta_{sn} = 0$, we set $\tilde{V}_{sk} = \hat{V}_{sk} = X_{k-1}, k = 1, \ldots, n$. Let

$$\tilde{V}_s = (\tilde{V}_{s1}, \ldots, \tilde{V}_{sn}) \quad and \quad \hat{V}_s = (\hat{V}_{s1}, \ldots, \hat{V}_{sn}).$$

If η_{jn} satisfies the condition $0 \leq \eta_{jn} - \eta_n \leq 2n^{-\gamma} 3^{-m_n}$, it follows that

$$\eta_{jn} \tilde{V}_{jt} \leq \eta_n X_{t-1} \leq \eta_{tn} \hat{V}_{jt}, \quad t = 1, \ldots, n.$$

Thus, for $x \in [x_r, x_{r+1}]$, we have

$$x_r + \eta_{jn}\tilde{V}_{jt} \leq x + \eta_n X_{t-1} \leq x_{r+1} + \eta_{jn}\hat{V}_{jt}, \quad t = 1, \ldots, n.$$

We obtain

$$\begin{aligned} Z_n(x, \eta_n, \mathbf{X}) \;\; \geq \;\; & Z_n(x_r, \eta_{jn}, \tilde{V}_j) + \frac{1}{n}\sum_{t=1}^{n}[F(x_r + \eta_{jn}\tilde{V}_{jt}) - F(x_{r+1} + \eta_{jn}\hat{V}_{jt})] \\ & + \frac{1}{n}\sum_{t=1}^{n}[\Delta_t(x_r) - F(x_r) - \Delta_t(x) + F(x)] \end{aligned}$$

and

$$\begin{aligned} Z_n(x, \eta_n, \mathbf{X}) \;\; \leq \;\; & Z_n(x_{r+1}, \eta_{jn}, \hat{V}_j) + \frac{1}{n}\sum_{t=1}^{n}[F(x_{r+1} + \eta_{jn}\hat{V}_{jt}) - F(x_r + \eta_{jn}\tilde{V}_{jt})] \\ & + \frac{1}{n}\sum_{t=1}^{n}[\Delta_t(x_{r+1}) - F(x_{r+1}) - \Delta_t(x) + F(x)]. \end{aligned}$$

Hence,

$$\begin{aligned} \sup_x \;\; \sup\nolimits_{|\eta_n| \leq n^{-\gamma}} \;\; & | \, Z_n(x, \eta_n, \mathbf{X}) \, | \\ & \leq \sup_{i \leq N_n} \sup_{s \leq 3^{m_n}} | \, Z_n(x_i, \eta_{sn}, \tilde{V}_s) \, | & \text{(A.20)} \\ & + \sup_{i \leq N_n - 1} \sup_{s \leq 3^{m_n}} | \, Z_n(x_{i+1}, \eta_{sn}, \hat{V}_s) \, | & \text{(A.21)} \\ & + \sup_{|t_1 - t_2| \leq N_n^{-1}} \frac{1}{n} \, \Big| \sum_{k=1}^{n}[\Delta_t(F^{-1}(t_1)) - t_1 - \Delta_t(F^{-1}(t_2)) + t_2] \, \Big| & \text{(A.22)} \\ & + \sup_{i \leq N_n - 1} \sup_{s \leq 3^{m_n}} \frac{1}{n}\sum_{t=1}^{n}[F(x_{i+1} + \eta_{jn}\hat{V}_{jt}) - F(x_i + \eta_{jn}\tilde{V}_{jt})]. & \text{(A.23)} \end{aligned}$$

Term (A.22) converges to zero in probability by Theorem 13.1 of Billingsley (1968) (p. 105 -108). Now we will show that (A.23) tends to zero in probability as $n \to \infty$. By Taylor's formula, (A.23) is bounded above by

$$\begin{aligned} & \sup\nolimits_{i \leq N_n - 1}[F(x_{i+1}) - F(x_i)] + \sup\nolimits_x | \, F'(x) \, | \sup\nolimits_{s \leq 3^{m_n}} \left(\tfrac{1}{n} \left| \textstyle\sum_{t=1}^{n} \eta_{sn}\tilde{V}_{st} \right| + \tfrac{1}{n} \left| \textstyle\sum_{t=1}^{n} \eta_{sn}\hat{V}_{st} \right| \right) \\ & \quad + \tfrac{1}{2} \sup\nolimits_x | \, F''(x) \, | \sup\nolimits_{s \leq 3^{m_n}} \tfrac{1}{n} \textstyle\sum_{t=1}^{n} \eta_{sn}^2(\tilde{V}_{st}^2 + \hat{V}_{st}^2). & \text{(A.24)} \end{aligned}$$

From definitions of $\tilde{V}_{st}$ and $\hat{V}_{st}$ we have

$$\max\left(\left|\sum_{t=1}^{n}\eta_{sn}\tilde{V}_{st}\right|,\left|\sum_{t=1}^{n}\eta_{sn}\hat{V}_{st}\right|\right) \leq n^{-\gamma}\left|\sum_{t=1}^{n}X_{t-1}\right| + 2n^{-\gamma}3^{-m}\sum_{t=1}^{n} \mid X_{t-1} \mid . \tag{A.25}$$

Moreover, since $\mid 1-2n^{-\gamma}3^{-m_n}\eta_{sn}^{-1} \mid \leq 3$ for any s,

$$\max(\mid \tilde{V}_{st} \mid, \mid \hat{V}_{st} \mid) \leq 3 \mid X_{t-1} \mid .$$

Then

$$\sum_{t=1}^{n}\eta_{sn}^2(\tilde{V}_{st}^2+\hat{V}_{st}^2) \leq 18n^{-2\gamma}\sum_{t=1}^{n}X_{t-1}^2. \tag{A.26}$$

Therefore (A.24) is bounded by

$$N_n^{-1}+c\left(n^{-\gamma}\left|\frac{1}{n}\sum_{t=1}^{n}X_{t-1}\right| + n^{-\gamma}3^{-m_n}\frac{1}{n}\sum_{t=1}^{n} \mid X_{t-1} \mid + n^{-2\gamma}\frac{1}{n}\sum_{t=1}^{n}X_{t-1}^2\right), \tag{A.27}$$

where c is a constant.

From Lemma 2.1 and 2.2 in Chan and Wei (1987) and from the continuous mapping theorem, it follows that $\frac{1}{n^{3/2}}\sum_{t=1}^{n} \mid X_{t-1} \mid$ is bounded in probability. Hence from (A.13), (A.15) and the fact that $n^{-\gamma/4}3^{m_n} \to 1$ and $n^{-1/2-\gamma/4}N_n \to 1$, as $n \to \infty$, one concludes that (A.27) tends to zero in probability as $n \to \infty$. Thus, the convergence of (A.23) to zero in probability is proved. Only remains to show that

$$\sup_{i\leq N_n}\sup_{s\leq 3^{m_n}} \mid Z_n(x_i,\eta_{sn},\tilde{V}_s) \mid \to_P 0 \quad as \quad n\to\infty \tag{A.28}$$

and

$$\sup_{i\leq N_n-1}\sup_{s\leq 3^{m_n}} \mid Z_n(x_{i+1},\eta_{sn},\hat{V}_s) \mid \to_P 0 \quad as \quad n\to\infty. \tag{A.29}$$

We will establish (A.28) and (A.29) can be proved similarly. For each sequence $\{l_n : n \in \mathbb{N}\}$ of non-negative integers we define

$$X_{t-1}^* = \sum_{r=0}^{l_n-1}\beta_n^r\varepsilon_{t-1-r}$$

and

$$\tilde{V}_{st}^* = X_{t-1}^*(1-2n^{-\gamma}3^{-m_n}\eta_{sn}^{-1}I_{\{X_{t-1}^*>0\}}).$$

So, we can write $Z_n(x_i,\eta_{sn},\tilde{V}_s)$ in the form

$$Z_n(x_i, \eta_{sn}, \tilde{V}_s) = \frac{1}{n}\sum_{t=1}^{n} \nu_t(i,s) + \frac{1}{n}\sum_{t=1}^{n} \xi_t(i,s),$$

where

$$\nu_t(i,s) = \Delta_t(x_i + \eta_{sn}\tilde{V}_{st}) - F(x_i + \eta_{sn}\tilde{V}_{st}) - \Delta_t(x_i + \eta_{sn}\tilde{V}^*_{st}) + F(x_i + \eta_{sn}\tilde{V}^*_{st})$$

and

$$\xi_t(i,s) = \Delta_t(x_i + \eta_{sn}\tilde{V}^*_{st}) - F(x_i + \eta_{sn}\tilde{V}^*_{st}) - \Delta_t(x_i) + F(x_i).$$

Hence

$$\sup_{i\le N_n}\sup_{s\le 3^{m_n}} \mid Z_n(x_i, \eta_{sn}, \tilde{V}_s) \mid \le \sup_{i\le N_n}\sup_{s\le 3^{m_n}} \left(\left| \frac{1}{n}\sum_{t=1}^{n} \nu_t(i,s)\right| + \left| \frac{1}{n}\sum_{t=1}^{n} \xi_t(i,s)\right| \right).$$

So, for all $\delta > 0$,

$$\begin{aligned} & P\left(\sup_{i\le N_n}\sup_{s\le 3^{m_n}} \mid Z_n(x_i, \eta_{sn}, \tilde{V}_s) \mid > 2\delta \right) \\ & \le \sum_{i=0}^{N_n}\sum_{s=0}^{3^{m_n}} \left[P\left(\left| \frac{1}{n}\sum_{t=1}^{n} \nu_t(i,s)\right| > \delta \right) + P\left(\left| \frac{1}{n}\sum_{t=1}^{n} \xi_t(i,s)\right| > \delta \right) \right] \\ & \le \sum_{i=0}^{N_n}\sum_{s=0}^{3^{m_n}} \left[\delta^{-2} E\left(\left| \frac{1}{n}\sum_{t=1}^{n} \nu_t(i,s)\right|^2 \right) + \delta^{-4} E\left(\left| \frac{1}{n}\sum_{t=1}^{n} \xi_t(r,s)\right|^4 \right) \right]. \end{aligned} \tag{A.30}$$

As in Lemma 1 of Boldin (1982) we obtain

$$\frac{1}{\lambda_n}\sup_{i\le N_n}\sup_{s\le 3^{m_n}} E\left(\left| \frac{1}{n^{1/2}}\sum_{t=1}^{n} \xi_t(i,s)\right|^4 \right) = O(1) \tag{A.31}$$

where $\lambda_n = n^{-1-\gamma} l_n^3 + n^{-2\gamma}$. Then if we choose $l_n = n^{4/15}$ it follows from (A.31) that

$$\sum_{i=0}^{N_n}\sum_{s=0}^{3^{m_n}} E\left(\left| \frac{1}{n}\sum_{t=1}^{n} \xi_t(i,s)\right|^4 \right) \to 0 \quad as \quad n \to \infty \tag{A.32}$$

for any $\gamma > 3/5$.

From Lemma A1 and Lemma 2 in Boldin (1982) we have

$$n^{1/2+\gamma}\sup_{i\le N_n}\sup_{s\le 3^{m_n}} E\left(\left| \frac{1}{n^{1/2}}\sum_{t=1}^{n} \nu_t(i,s)\right|^2 \right) = O(1).$$

Therefore

$$\sum_{i=0}^{N_n} \sum_{s=0}^{3^{m_n}} E\left(\left|\frac{1}{n}\sum_{t=1}^{n} \nu_t(i,s)\right|^2\right) \to 0 \quad as \quad n \to \infty \tag{A.33}$$

for any $\gamma > 3/5$.

Hence from (A.32) and (A.33) we obtain that (A.30) tends to zero in probability for any $\gamma > 3/5$. Then (A.28) holds and therefore (A.19) follows. Thus, from (A.17), (A.18) and (A.19) one now concludes (A.16). This completes our proof. □

References

AHTOLA, J.A. and TIAO, G.C. (1984). Some aspects of parameter inference for nearly nonstationary and nearly noninvertible ARMA models II, *Qüestiió*, **8**, 155-163.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons.

BOBKOSKI, M. J. (1983). Hypothesis testing in nonstationary time series. Unpublished Ph. D. thesis, University of Wisconsin, Dept. of Statistics.

BOLDIN, M. V. (1982). Estimation of the distribution of noise in an autoregressive scheme. *Theor. Prob. Appl.*, **27**, 866-871.

CHAN, N. H. (1988). The parameter inference for nearly nonstationary time series. *J. Am. Statist. Assoc.*, **83**, 857-862.

CHAN, N. H. and WEI, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes. *Ann. Statist.*, **15**, 1050-1063.

COX, D. D. (1991). Gaussian likelihood estimation for nearly nonstationary AR(1) processes. *Ann. Statist.*, **19**, 1129-1142.

COX, D. D. and LLATAS, I. (1991). Maximum likelihood type estimation for nearly nonstationary autoregressive time series. *Ann. Statist.*, **19**, 1109-1128.

JEGANATHAN, P. (1987). On the asymptotic behavior of least squares estimators in AR time series with roots near the unit circle. Technical Report 150, Dept. Statistics, Univ. Michigan, Ann Arbor.

KOUL, H. L. AND LEVENTHAL, S. (1989). Weak convergence of the residual empirical process in explosive autoregression. *Ann. Statist.*, **17**, 1784-1794.

PHILLIPS, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika*, **74**, 535-547.

STUTE, W. and GRÜNDER, B. (1990). Bootstrap approximation to prediction intervals for explosive AR(1) processes. *Lectures Notes in Economics and Mathematical Systems*, **376**, 121-130, Springer Verlag, New York.

STUTE, W. and GRÜNDER, B. (1993). Nonparametric prediction intervals for explosive AR(1)-processes. *Nonparametric Statistics*, **2**, 155-167.

TSAY, R. S. (1985). Asymptotic inference of least squares estimates in nearly nonstationary time series. Technical Report, Dept. Statistics, Carnegie Mellon Univ.