Econometric Theory, **28**, 2012, 705–717.
doi:10.1017/S0266466611000727

NOTES AND PROBLEMS

TOWARD A UNIFIED INTERVAL ESTIMATION OF AUTOREGRESSIONS

NGAI HANG CHAN
Chinese University of Hong Kong

DEYUAN LI
Fudan University

LIANG PENG
Georgia Institute of Technology

An empirical likelihood–based confidence interval is proposed for interval estimations of the autoregressive coefficient of a first-order autoregressive model via weighted score equations. Although the proposed weighted estimate is less efficient than the usual least squares estimate, its asymptotic limit is always normal without assuming stationarity of the process. Unlike the bootstrap method or the least squares procedure, the proposed empirical likelihood–based confidence interval is applicable regardless of whether the underlying autoregressive process is stationary, unit root, near-integrated, or even explosive, thereby providing a unified approach for interval estimation of an AR(1) model to encompass all situations. Finite-sample simulation studies confirm the effectiveness of the proposed method.

1. INTRODUCTION

Consider the first-order autoregressive process

$$Y_i = \beta Y_{i-1} + \epsilon_i, \qquad i = 1, \ldots, n, \qquad Y_0 = 0, \tag{1}$$

where $\epsilon_1, \ldots, \epsilon_n$ are independent and identically distributed random variables with mean zero and finite variance σ^2. The process is called stationary, unit root, near-integrated, and explosive for the cases $|\beta| < 1$, $|\beta| = 1$, $\beta = \beta(n) = 1 - \gamma/n$ (for some constant $\gamma \in R$), and $|\beta| > 1$, respectively. Although the behavior of the

We thank three anonymous referees, the editor, and the co-editor Giuseppe Cavaliere for helpful references and constructive suggestions, which led to an improved version of this note. This research was supported in part by grants from HKSAR-RGC-GRF nos. 400306, 400308, and 400410, NSA grant no. H98230-10-1-0170, NSF grant no. DMS1005336, and NNSFC grant no. 10801038. Address correspondence to Ngai Hang Chan, Department of Statistics, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong; e-mail: nhchan@sta.cuhk.edu.hk.

© Cambridge University Press 2011

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

process depends on the value of β, the parameter β can be estimated by the usual least squares procedure:

$$\hat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} (Y_i - \beta Y_{i-1})^2 = \sum_{i=1}^{n} Y_i Y_{i-1} \Big/ \sum_{i=1}^{n} Y_{i-1}^2.$$

The asymptotic behavior of $\hat{\beta}$ has been studied extensively in the literature. Specifically,

$$\left(\sum_{i=1}^{n} Y_{i-1}^2 \right)^{1/2} \{\hat{\beta} - \beta\} \xrightarrow{d} \begin{cases} N(0, \sigma^2), & \text{if } |\beta| < 1, \\ \sigma \frac{\int_0^1 X(t)\, dW(t)}{(\int_0^1 X^2(t)\, dt)^{1/2}}, & \text{if } \beta = \beta(n) = 1 - \gamma/n, \\ (1 - \beta^{-2})^{-1/2} \operatorname{sgn}(U) V, & \text{if } |\beta| > 1, \end{cases} \tag{2}$$

where $\gamma \in R$, $\{X(t) : 0 \le t \le 1\}$ is an Ornstein–Uhlenbeck process satisfying the stochastic differential equation $dX(t) = -\gamma X(t)\, dt + \sigma\, dW(t)$ with $X(0) = 0$, $\{W(t) : 0 \le t \le 1\}$ is a standard Brownian motion, and U and V are independent random variables having the same distribution as $\sum_{j=1}^{\infty} \beta^{-(j-1)} \epsilon_j$. For background information and a succinct review of these developments, see Chan and Wei (1987), Phillips (1987), Datta (1996), and Chan (2009). Further studies of near-integrated infinite variance AR(1) process can be found in Chan (1990), Phillips (1990), Chan, Peng, and Qi (2006), and Chan and Zhang (2009), and the references therein.

To construct a confidence interval for the autoregressive coefficient β or to test $H_0 : \beta = \beta_0$, one usually employs (2) either by means of simulating the limit distribution or using a bootstrap procedure. Because the limit distribution depends on the location of β, one needs to distinguish the different cases of β before simulating from the limit distribution. However, differentiating the near-integrated case ($\beta = 1 - \gamma/n$) from the stationary or the integrated case ($|\beta| \le 1$) through testing for β remains a challenging task as the limit distribution depends on the nuisance parameter γ in (2). It is also known that the full-sample bootstrap method is not applicable to construct a confidence interval for β without knowing the true location of β a priori.

The empirical likelihood method is a powerful nonparametric technique for interval estimation and hypothesis testing. For an overview on empirical likelihood methods, see the monograph by Owen (2001). Because the limit distribution in (2) is nonnormal for $|\beta| \ge 1$, applying the empirical likelihood method to the score equation does not lead to a chi-square limit; that is, the Wilks theorem does not hold for the near-integrated case and the explosive case; see Chuang and Chan (2002) for the application of the empirical likelihood method to the near-integrated case. As a result, directly applying the empirical likelihood method would not lead to a unified procedure; one still needs to determine the value of β first.

Recently, the study of uniform convergence of β for nonstationary processes has been receiving considerable attention. For example, Phillips and Han (2008) and Han, Phillips, and Sul (2010) proposed the idea of differencing the sequence to

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

obtain an estimator that has a normal limit when the process is stationary or even has a unit root. A more general approach was proposed by Mikusheva (2007), where it was shown that the grid bootstrap method based on the statistic $\sum_{i=1}^{n} Y_{i-1}(Y_i - \beta Y_{i-1})/\{\sum_{i=1}^{n} Y_{i-1}^2\}^{1/2}$ is valid for $|\beta| \le 1$ and for β slightly bigger than one. On the other hand, a weighted least squares estimate proposed by So and Shin (1999) has a normal limit regardless of whether the process is stationary, unit root, near-integrated, or explosive. The simulation study reported in Section 3 indicates, however, that the pivotal statistic in So and Shin (1999) becomes problematic for explosive processes.

In this note, we propose an empirical likelihood method for constructing confidence intervals for all $\beta \in R$ based on a weighted least squares estimate that is more general than the one proposed in So and Shin (1999). Using this method, a unified approach for interval estimation of β is established for all values of β, thereby providing a much more convenient and effective method for interval estimations of β. The proposed empirical likelihood method is also applicable to certain nonstationary AR(p) processes, whereas it is not clear a priori if So and Shin's estimate is applicable under such circumstances.

The note is organized as follows. Methodologies and main results for AR(1) processes are given in Section 2, and a simulation study is presented in Section 3. Proofs are collected in the Appendix.

2. METHODOLOGIES

Weighted estimation procedures have been studied in the literature, and a normal limit distribution is obtained; see Ling (2005, 2007) and Chan and Peng (2005). In this note, the possibility of using a weighted estimation method to estimate the autoregressive coefficient β is explored so that the resulting limit distribution will still be normal.

Specifically, consider applying a weighted estimation method to the near-integrated case. Write $\hat{\beta} - \beta = \sum_{i=1}^{n} \epsilon_i Y_{i-1} / \sum_{i=1}^{n} Y_{i-1}^2$. When $\beta = 1 - \gamma/n$, instead of converging to a constant in probability as in the usual case, the quantity $n^{-2}\sum_{i=1}^{n} \mathrm{E}(\epsilon_i^2 Y_{i-1}^2 | \mathcal{F}_{i-1})$ ($\mathcal{F}_i = \sigma(\epsilon_1, \ldots, \epsilon_i)$) converges in distribution to a random variable. As a result, the limit of $n^{-1}\sum_{i=1}^{n} \epsilon_i Y_{i-1}$ can no longer be a normal distribution. However, as the proofs of Chan and Wei (1987) indicate, the following equality still holds:

$$n^{-1}\sum_{i=1}^{n} \mathrm{E}\left(\frac{Y_{i-1}^2}{1+Y_{i-1}^2}\epsilon_i^2 | \mathcal{F}_{i-1}\right) = n^{-1}\sum_{i=1}^{n} \mathrm{E}(\epsilon_i^2 | \mathcal{F}_{i-1}) + o_p(1) = \sigma^2 + o_p(1).$$

Motivated by this equation, a weighted least squares estimate is proposed as follows:

$$\hat{\beta}^w(\delta) = \arg\min_{\beta} \sum_{i=1}^{n} \frac{(Y_i - \beta Y_{i-1})^2}{(\delta + Y_{i-1}^2)^{1/2}} = \sum_{i=1}^{n} \frac{Y_{i-1}Y_i}{(\delta + Y_{i-1}^2)^{1/2}} \Big/ \sum_{i=1}^{n} \frac{Y_{i-1}^2}{(\delta + Y_{i-1}^2)^{1/2}},$$

where δ is a nonnegative constant. When $\delta = 0$, it reduces to the Cauchy estimate

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

proposed by So and Shin (1999). For such a weighted least squares estimate, asymptotic normality is achieved.

THEOREM 1. *Assume model (1) holds with* $\mathrm{E}|\epsilon_1|^{2+d} < \infty$ *for some* $d > 0$. *Further assume that* β *is either a constant or* $\beta = \beta(n) = 1 - \gamma/n$ *for some constant* $\gamma \in R$. *Then for any* $\delta \geq 0$

$$\left(\sum_{i=1}^{n} \frac{Y_{i-1}^2}{\delta + Y_{i-1}^2}\right)^{-1/2} \left(\sum_{i=1}^{n} \frac{Y_{i-1}^2}{(\delta + Y_{i-1}^2)^{1/2}}\right) \times (\hat{\beta}^w(\delta) - \beta) \xrightarrow{d} N(0, \sigma^2) \quad \text{as } n \to \infty.$$

Based on this theorem, one can construct a confidence interval for β or test $H_0 : \beta = \beta_0$ via estimating σ^2 by means of $\hat{\sigma}^2 = n^{-1} \sum_{i=1}^{n} (Y_i - \tilde{\beta} Y_{i-1})^2$, where $\tilde{\beta}$ is the least squares estimate as proposed by So and Shin (1999). However, the simulation study reported in Section 3 indicates that $\hat{\sigma}^2$ behaves erratically when the process is explosive. To circumvent such a difficulty, in this note, we propose applying the empirical likelihood method to the weighted score equation of the weighted least squares estimate. Specifically, define

$$Z_i(\beta; \delta) = \frac{Y_{i-1}}{(\delta + Y_{i-1}^2)^{1/2}} (Y_i - \beta Y_{i-1}) \quad \text{for } i = 2, \ldots, n.$$

Because the weighted least squares estimate $\hat{\beta}^w(\delta)$ is the solution to the score equation $\sum_{i=2}^{n} Z_i(\beta; \delta) = 0$, define the empirical likelihood function as

$$L(\beta; \delta) = \sup \left\{ \prod_{i=2}^{n} (n p_i) : p_2 \geq 0, \ldots, p_n \geq 0, \sum_{i=2}^{n} p_i = 1, \sum_{i=2}^{n} p_i Z_i(\beta; \delta) = 0 \right\}.$$

After applying the Lagrange multiplier technique, we obtain $p_i = n^{-1}(1 + \lambda Z_i(\beta; \delta))^{-1}$, and the log empirical likelihood ratio becomes $l(\beta; \delta) = -2 \log L(\beta; \delta) = 2 \sum_{i=2}^{n} \log(1 + \lambda Z_i(\beta; \delta))$, where $\lambda = \lambda(\beta, \delta)$ satisfies

$$n^{-1} \sum_{i=2}^{n} \frac{Z_i(\beta; \delta)}{1 + \lambda Z_i(\beta; \delta)} = 0. \tag{3}$$

The following theorem shows that the Wilks theorem holds for the proposed empirical likelihood method.

THEOREM 2. *Under the conditions of Theorem 1,* $l(\beta_0; \delta) \xrightarrow{d} \chi^2(1)$ *as* $n \to \infty$ *for any given* $\delta \geq 0$, *where* β_0 *denotes the true value of* β.

In particular, using this theorem, one can construct a confidence interval for β_0 with level α as

$$I_\alpha = \{\beta : l(\beta; \delta) \leq \chi_{1,\alpha}^2\},$$

where $\chi_{1,\alpha}^2$ denotes the α-quantile of $\chi^2(1)$.

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

Remark 1. When the sample size n is not large, the rate of convergence of $Z_i(\beta;\delta)$ to infinity is not very fast in the near-integrated case or the explosive case. Hence a small δ would be preferred. In the simulation study, $\delta = 1$ works well.

Remark 2. Theorem 2 remains true when $\{\epsilon_i\}$ is a martingale difference sequence with respect to an increasing sequence of σ-fields $\{\mathcal{F}_t\}$ such that

$$\begin{cases} \frac{1}{n}\sum_{t=1}^{n} \mathrm{E}(\epsilon_t^2|\mathcal{F}_{t-1}) \xrightarrow{p} \sigma^2, \\ \frac{1}{n}\sum_{t=1}^{n} \mathrm{E}(|\epsilon_t|^{2+d}|\mathcal{F}_{t-1}) \xrightarrow{p} c < \infty, \end{cases}$$

for some $d > 0$ as $n \to \infty$.

Remark 3. By using a similar argument as given in So and Shin (1999), it can be shown that the proposed empirical likelihood method is also applicable to a seasonal AR(1) model, thereby extending the application to seasonal autoregressive models.

Next we generalize the proposed empirical likelihood method to an AR(p) model in the so-called augmented Dickey–Fuller form as follows:

$$Y_t = \beta Y_{t-1} + \sum_{j=1}^{p-1} \theta_j (Y_{t-j} - Y_{t-j-1}) + \epsilon_t, \tag{4}$$

where $|\beta| \le 1$ and β is the only possible unit root. For such a model, Mikusheva (2007) proposed a unified method to construct confidence intervals for $|\beta| \le 1$. As pointed out in Section 1, this method requires the construction of a t-statistic for each β and then uses a grid bootstrap method to find the critical values and the confidence intervals. As a result, the method is computationally intensive. Although So and Shin (1999) made use of the Cauchy estimate for model (4) when $\beta = 1$, it still remains unclear a priori if it is possible to construct a valid pivotal statistic for all $|\beta| \le 1$ in this circumstances. Alternatively, we propose applying the empirical likelihood method to the weighted score equations as follows.

Define $Z_i(\beta,\theta) = (Z_{i,1}(\beta,\theta),\ldots,Z_{i,p}(\beta,\theta))^T$, $\theta = (\theta_1,\ldots,\theta_{p-1})^T$,

$$Z_{i,1}(\beta,\theta) = \frac{Y_{i-1}}{\left(\delta + Y_{i-1}^2\right)^{1/2}} \left\{ Y_i - \beta Y_{i-1} - \sum_{j=1}^{p-1} \theta_j (Y_{i-j} - Y_{i-j-1}) \right\},$$

$$Z_{i,l}(\beta,\theta) = (Y_{i-l+1} - Y_{i-l}) \left\{ Y_i - \beta Y_{i-1} - \sum_{j=1}^{p-1} \theta_j (Y_{i-j} - Y_{i-j-1}) \right\}$$

for $l = 2,\ldots,p$,

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

and the empirical likelihood function as

$$\tilde{L}(\beta,\theta) = \sup\left\{ \prod_{i=p+3}^{n} ((n-p-2)p_i) : p_i \geq 0, \right.$$
$$\left. \times \sum_{i=p+3}^{n} p_i = 1, \sum_{i=p+3}^{n} p_i Z_i(\beta,\theta) = 0 \right\},$$

where $\theta = (\theta_1, \ldots, \theta_{p-1})^T$. It follows from the Lagrange multiplier techniques that

$$\tilde{l}(\beta,\theta) = -2\log \tilde{L}(\beta,\theta) = 2 \sum_{i=p+3}^{n} \log\{1 + \lambda^T Z_i(\beta,\theta)\},$$

where $\lambda = \lambda(\beta,\theta)$ satisfies

$$\sum_{i=p+3}^{n} \frac{Z_i(\beta,\theta)}{1+\lambda^T Z_i(\beta,\theta)} = 0. \tag{5}$$

Because we are only interested in β, consider the profile empirical likelihood ratio statistic $\tilde{l}^p(\beta) = \min_\theta \tilde{l}(\beta,\theta)$.

Throughout we use $\beta_0, \theta_0 = (\theta_{1,0}, \ldots, \theta_{p-1,0})^T$ to denote the true values of β and θ, respectively.

THEOREM 3. *Assume that the ϵ_i's in (4) are independent and identically distributed random variables with mean zero and $\mathrm{E}|\epsilon_i|^3 < \infty$. Further assume that (4) defines a strictly stationary AR(p) model for $\{Y_i\}$ when $|\beta_0| < 1$ and that (4) defines a strictly stationary AR($p-1$) model for $\{Y_i - Y_{i-1}\}$ when $|\beta_0| = 1$. Then, as $n \to \infty$, $\tilde{l}^p(\beta_0) \xrightarrow{d} \chi^2(1)$ for any given $\delta \geq 0$.*

Based on this theorem, a unified confidence interval for $|\beta| \leq 1$ can be constructed.

3. SIMULATIONS

To investigate the finite-sample behavior of the proposed empirical likelihood method, a total of 10,000 random samples each with sizes $n = 200$ and 500 are drawn from model (1), where the distribution function of ϵ_i is $N(0,1)$. Consider the cases $\beta = 0.5, 0.9, 0.99, 1, 1.01, 1.1$, and 1.5.

For comparison, we implement the partially aggregating estimator with $L = n^{1/3}$ and $n^{1/2}$ in Han, Phillips, and Sul (2010) because the fully aggregating estimator has a nonstandard limit for unit root or near unit root processes. Denote this method as PAM(L). We also calculate the interval based on the grid bootstrap method and statistic $\sum_{i=1}^{n} Y_{i-1}(Y_i - \beta Y_{i-1}) / \left\{ \sum_{i=1}^{n} Y_{i-1}^2 \right\}^{1/2}$ used in Mikusheva (2007) by drawing 5,000 bootstrap samples from the estimated errors. Denote this method as GBM. The pivotal statistic in So and Shin (1999) is calculated also and

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

denoted by CEM. In calculating the coverage probability for the proposed empirical likelihood method, $\delta = 1, 0.1, 2$ are used in the "emplik" package in the R software.

In Tables 1 and 2, the empirical coverage probabilities for these confidence intervals are reported. Several observations can be deduced from these tables. First, the proposed empirical likelihood method works reasonably well under all circumstances of β, whereas the other three methods do not work for the explosive cases. Second, the proposed empirical likelihood method has similar performance as the grid bootstrap method of Mikusheva (2007) and the Cauchy estimator of So and Shin (1999) for $\beta \leq 1.01$. Third, the choice of $\delta \leq 1$ in the proposed empirical likelihood method is preferred. Finally, these three methods perform better than the partially aggregating method when $\beta = 0.99$ and 1.

TABLE 1. Empirical coverage probabilities calculated for the intervals based on the proposed empirical likelihood method (ELM1, ELM2, ELM3 denoted for $\delta = 1, 0.1, 2$, respectively), the pivotal Cauchy estimate (CEM), the grid bootstrap method (GBM), partially aggregating methods with $L = n^{1/3}$ (PAM1) and $L = n^{1/2}$ (PAM2) under model (1) with levels $\alpha = 0.55, 0.9, 0.95$, sample size $n = 200$, and ϵ_i being standard normal

β	0.5	0.9	0.99	1	1.01	1.1	1.5
ELM1, $\alpha = 0.55$	0.5487	0.5485	0.5469	0.5379	0.5412	0.5512	0.5413
ELM1, $\alpha = 0.9$	0.8967	0.8971	0.8959	0.8951	0.8912	0.8967	0.8949
ELM1, $\alpha = 0.95$	0.9477	0.9504	0.9498	0.9481	0.9450	0.9457	0.9466
ELM2, $\alpha = 0.55$	0.5535	0.5491	0.5467	0.5422	0.5515	0.5532	0.5418
ELM2, $\alpha = 0.9$	0.9058	0.8984	0.8949	0.8955	0.9007	0.8982	0.8991
ELM2, $\alpha = 0.95$	0.9512	0.9493	0.9462	0.9483	0.9468	0.9502	0.9504
ELM3, $\alpha = 0.55$	0.5497	0.5479	0.5383	0.5239	0.5350	0.5448	0.5451
ELM3, $\alpha = 0.9$	0.8959	0.8977	0.8908	0.8865	0.8847	0.8970	0.8952
ELM3, $\alpha = 0.95$	0.9486	0.9509	0.9468	0.9415	0.9405	0.9490	0.9458
CEM, $\alpha = 0.55$	0.5514	0.5532	0.5537	0.5536	0.5551	0.5539	0.7791
CEM, $\alpha = 0.9$	0.8985	0.9016	0.9002	0.8994	0.9007	0.8983	0.7843
CEM, $\alpha = 0.95$	0.9489	0.9495	0.9507	0.9509	0.9507	0.9493	0.7896
GBM, $\alpha = 0.55$	0.5484	0.5470	0.5457	0.5503	0.5493	0.5474	0.6761
GBM, $\alpha = 0.9$	0.8980	0.8960	0.8950	0.8958	0.8971	0.8968	0.9427
GBM, $\alpha = 0.95$	0.9469	0.9467	0.9480	0.9468	0.9477	0.9480	0.9731
PAM1, $\alpha = 0.55$	0.5439	0.5400	0.5233	0.5177	0.2693	0	0
PAM1, $\alpha = 0.9$	0.8979	0.8908	0.8785	0.8790	0.5657	0	0
PAM1, $\alpha = 0.95$	0.9498	0.9424	0.9380	0.9344	0.6600	0	0
PAM2, $\alpha = 0.55$	0.5440	0.5326	0.5204	0.5021	0.2524	0	0
PAM2, $\alpha = 0.9$	0.8965	0.8896	0.8735	0.8695	0.5980	0	0
PAM2, $\alpha = 0.95$	0.9483	0.9432	0.9332	0.9264	0.7192	0	0

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

TABLE 2. Empirical coverage probabilities calculated for the intervals based on the proposed empirical likelihood method (ELM1, ELM2, ELM3 denoted for $\delta = 1, 0.1, 2$, respectively), the pivotal Cauchy estimate (CEM), the grid bootstrap method (GBM), partially aggregating methods with $L = n^{1/3}$ (PAM1) and $L = n^{1/2}$ (PAM2) under model (1) with levels $\alpha = 0.55, 0.9, 0.95$, sample size $n = 500$, and ϵ_i being standard normal

β	0.5	0.9	0.99	1	1.01	1.1	1.5
ELM1, $\alpha = 0.55$	0.5451	0.5481	0.5442	0.5427	0.5428	0.5497	0.5352
ELM1, $\alpha = 0.9$	0.8973	0.8995	0.8983	0.8938	0.8974	0.8953	0.8946
ELM1, $\alpha = 0.95$	0.9478	0.9471	0.9501	0.9458	0.9480	0.9487	0.9473
ELM2, $\alpha = 0.55$	0.5511	0.5479	0.5576	0.5491	0.5503	0.5410	0.5440
ELM2, $\alpha = 0.9$	0.9017	0.8970	0.9019	0.8985	0.9002	0.8995	0.8946
ELM2, $\alpha = 0.95$	0.9488	0.9484	0.9521	0.9467	0.9497	0.9506	0.9462
ELM3, $\alpha = 0.55$	0.5579	0.5506	0.5438	0.5413	0.5536	0.5518	0.5444
ELM3, $\alpha = 0.9$	0.9037	0.8947	0.8961	0.8970	0.9001	0.8965	0.8989
ELM3, $\alpha = 0.95$	0.9530	0.9495	0.9481	0.9491	0.9485	0.9489	0.9491
CEM, $\alpha = 0.55$	0.5441	0.5454	0.5492	0.5520	0.5501	0.8568	0.7785
CEM, $\alpha = 0.9$	0.8939	0.8968	0.9015	0.8998	0.9010	0.8568	0.7830
CEM, $\alpha = 0.95$	0.9458	0.9472	0.9509	0.9505	0.9481	0.8568	0.7872
GBM, $\alpha = 0.55$	0.5461	0.5494	0.5466	0.5555	0.5523	0.6180	0.2419
GBM, $\alpha = 0.9$	0.8987	0.8985	0.8972	0.9013	0.8991	0.9205	0.5981
GBM, $\alpha = 0.95$	0.9497	0.9503	0.9494	0.9499	0.9498	0.9586	0.6230
PAM1, $\alpha = 0.55$	0.5444	0.5386	0.5204	0.5160	0.0216	0	0
PAM1, $\alpha = 0.9$	0.8969	0.8912	0.8780	0.8719	0.0523	0	0
PAM1, $\alpha = 0.95$	0.9460	0.9453	0.9372	0.9341	0.0674	0	0
PAM2, $\alpha = 0.55$	0.5429	0.5402	0.5174	0.4987	0.0241	0	0
PAM2, $\alpha = 0.9$	0.8969	0.8915	0.8793	0.8668	0.0642	0	0
PAM2, $\alpha = 0.95$	0.9467	0.9452	0.9352	0.9286	0.0893	0	0

We also examine the length of the intervals by drawing 1,000 random samples from model (1) with a sample size $n = 200$ for $\beta = 0.9, 0.99, 1$, and 1.01. The average interval lengths are reported in Table 3, which shows that the grid bootstrap method produces the shortest intervals. The partially aggregating method and the Cauchy estimate have longer intervals than the proposed empirical likelihood method for $\delta = 1$.

In addition, we also study the finite-sample behavior of the proposed profile empirical likelihood method with $\delta = 1$ for general AR(p) models following (4). We draw 10,000 random samples from model (4) with $p = 2$, $\theta_1 = 0.5$. Here, $\beta = 0.5, 0.9, 0.99$, and 1 and the distribution of ϵ_i is $N(0, 1)$. The coverage probabilities in Table 4 show that the proposed method performs reasonably well.

In summary, the proposed empirical likelihood method seems to be the only method that is valid for all cases, including explosive processes. It is also applicable to seasonable AR(1) processes or AR(p) processes such as (4).

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

TABLE 3. Average interval lengths calculated for the intervals based on the proposed empirical likelihood method (ELM with $\delta = 1$), the pivotal Cauchy estimate (CEM), the grid bootstrap method (GBM), partially aggregating methods with $L = n^{1/3}$ (PAM1) and $L = n^{1/2}$ (PAM2) under model (1) with levels $\alpha = 0.9, 0.95$, and sample size $n = 200$

β	0.9	0.99	1	1.01
ELM, $\alpha = 0.9$	0.1054	0.0515	0.0381	0.0217
ELM, $\alpha = 0.95$	0.1279	0.0617	0.0456	0.0261
CEM, $\alpha = 0.9$	0.1323	0.0570	0.0420	0.0241
CEM, $\alpha = 0.95$	0.1577	0.0680	0.0500	0.0287
GBM, $\alpha = 0.9$	0.0966	0.0484	0.0358	0.0197
GBM, $\alpha = 0.95$	0.1167	0.0577	0.0424	0.0234
PAM1, $\alpha = 0.9$	0.1213	0.0801	0.0722	0.0552
PAM1, $\alpha = 0.95$	0.1446	0.9540	0.0860	0.0658
PAM2, $\alpha = 0.9$	0.1162	0.0707	0.0614	0.0438
PAM2, $\alpha = 0.95$	0.1385	0.0843	0.0732	0.0522

TABLE 4. Empirical coverage probabilities calculated for the intervals based on the proposed empirical likelihood method (ELM with $\delta = 1$) under model (4) with $p = 2$ and $\theta_1 = 0.5$

β	0.5	0.9	0.99	1
$(n, \alpha) = (200, 0.55)$	0.5496	0.5494	0.5501	0.5388
$(n, \alpha) = (200, 0.9)$	0.8964	0.8988	0.8940	0.8944
$(n, \alpha) = (200, 0.95)$	0.9437	0.9502	0.9496	0.9470
$(n, \alpha) = (500, 0.55)$	0.5432	0.5551	0.5547	0.5471
$(n, \alpha) = (500, 0.9)$	0.8967	0.9032	0.9012	0.8979
$(n, \alpha) = (500, 0.95)$	0.9487	0.9525	0.9507	0.9484

REFERENCES

Chan, N.H. (1990) Inference for near-integrated time series with infinite variance. *Journal of the American Statistical Association* 85, 1069–1074.

Chan, N.H. (2009) Time series with roots on or near the unit circle. In T.G. Andersen, R.A. Davis, J. Kreiss, & T. Mikosch (eds.), *Springer Handbook of Financial Time Series*, pp. 695–707. Springer-Verlag.

Chan, N.H. & L. Peng (2005) Weighted least absolute deviations estimation for an AR(1) process with ARCH(1) errors. *Biometrika* 92, 477–484.

Chan, N.H., L. Peng, & Y. Qi (2006) Quantile inference for near-integrated autoregressive time series with infinite variance. *Statistica Sinica* 16, 15–28.

Chan, N.H. & C.Z. Wei (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.

Chan, N.H. & R.M. Zhang (2009) Inference for nearly nonstationary processes under strong dependence and infinite variance. *Statistica Sinica* 19, 925–947.

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

Chuang, C. & N.H. Chan (2002) Empirical likelihood for autoregressive models, with applications to unstable time series. *Statistica Sinica* 12, 387–407.
Datta, S. (1996) On asymptotic properties of bootstrap for AR(1) processes. *Journal of Statistical Planning and Inference* 53, 361–374.
Hall, P. & C. Heyde (1980) *Martingale Limit Theory and Its Applications*. Academic Press.
Han, C., P.C.B. Phillips, & D. Sul (2010) Uniform Asymptotic Normality in Stationary and Unit Root Autoregression. Working paper, Yale University.
Ling, S. (2005) Self-weighted least absolute deviation estimation for infinite variance autoregressive models. *Journal of the Royal Statistical Society, Series B* 67, 381–393.
Ling, S. (2007) Self-weighted and local quasi-maximum likelihood estimators for ARMA-GARCH/IGARCH models. *Journal of Econometrics* 140, 849–873.
Mikusheva, A. (2007) Uniform inference in autoregressive models. *Econometrica* 75, 1411–1452.
Owen, A. (2001) *Empirical Likelihood*. Chapman and Hall.
Phillips, P.C.B. (1987) Toward a unified asymptotic theory of autoregression. *Biometrika* 74, 535–574.
Phillips, P.C.B. (1990) Time series regression with a unit-root and infinite-variance errors. *Econometric Theory* 6, 44–62.
Phillips, P.C.B. & C. Han (2008) Gaussian inference in AR(1) time series with or without a unit root. *Econometric Theory* 63, 1023–1078.
Qin, J. & J. Lawless (1994) Empirical likelihood and general estimating equations. *Annals of Statistics* 22, 300–325.
So, B.S. & D.W. Shin (1999) Cauchy estimators for autoregressive processes with applications to unit root tests and confidence intervals. *Econometric Theory* 15, 165–176.

APPENDIX: Proofs

Proof of Theorem 1. Define

$$X_i = \frac{Y_{i-1}}{(\delta + Y_{i-1}^2)^{1/2}}\epsilon_i \quad \text{and} \quad \mathcal{F}_i = \sigma(\epsilon_1, \ldots, \epsilon_i)$$

for $i = 1, \ldots, n$. Then

$$\frac{1}{\sqrt{n}}\left(\sum_{i=1}^{n} \frac{Y_{i-1}^2}{(\delta + Y_{i-1}^2)^{1/2}}\right)(\hat{\beta}^{w}(\delta) - \beta) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i. \tag{A.1}$$

When $|\beta| < 1$, it is easy to check that

$$\frac{1}{n}\sum_{i=1}^{n} \frac{Y_{i-1}^2}{\delta + Y_{i-1}^2} \xrightarrow{p} \mathrm{E}\frac{Y_1^2}{\delta + Y_1^2}. \tag{A.2}$$

When $\beta = 1 - \gamma/n$ for some constant γ, Lemma 2.1 in Chan and Wei (1987) shows that there is a standard Brownian motion $W_1(t)$ such that as $n \to \infty$,

$$n^{-1/2}\beta^{n-[nt]}Y_{[nt]} \xrightarrow{D} W_1\left(e^{-2\gamma}\left(e^{2t\gamma} - 1\right)/(2\gamma)\right)$$

in $D[0, 1]$, which implies that

$$\frac{1}{n}\sum_{i=1}^{n} \frac{Y_{i-1}^2}{\delta + Y_{i-1}^2} \xrightarrow{p} 1 \tag{A.3}$$

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

as $n \to \infty$. When $|\beta| > 1$, it follows that as $T \to \infty$,

$$\beta^{2-T} Y_{T-1} \xrightarrow{d} \sum_{i=1}^{\infty} \beta^{1-i} \epsilon_i,$$

which in turn implies that as $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} \frac{Y_{i-1}^2}{\delta + Y_{i-1}^2} \xrightarrow{p} 1. \tag{A.4}$$

Hence, by (A.2)–(A.4), as $n \to \infty$,

$$\frac{1}{n} \sum_{i=1}^{n} \mathrm{E}(X_i^2 | \mathcal{F}_{i-1}) = \sigma^2 \frac{1}{n} \sum_{i=1}^{n} \frac{Y_{i-1}^2}{\delta + Y_{i-1}^2} \xrightarrow{p} \begin{cases} \sigma^2 \mathrm{E} \frac{Y_1^2}{\delta + Y_1^2}, & \text{if } |\beta| < 1, \\ 1, & \text{if } |\beta| > 1, \\ 1, & \text{if } \beta = 1 - \gamma/n. \end{cases} \tag{A.5}$$

Similar to the proofs of (A.2)–(A.4), it can be shown that for any $c > 0$,

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}(X_i^2 I(|X_i| > c\sqrt{n}) | \mathcal{F}_{i-1}) &\le (c\sqrt{n})^{-d} n^{-1} \sum_{i=1}^{n} \mathrm{E}(|X_i|^{2+d} | \mathcal{F}_{i-1}) \\ &= (c\sqrt{n})^{-d} \mathrm{E}|\epsilon_1|^{2+d} \frac{1}{n} \sum_{i=1}^{n} \frac{|Y_{i-1}|^{2+d}}{(\delta + Y_{i-1}^2)^{(2+d)/2}} \\ &\xrightarrow{p} 0 \end{aligned} \tag{A.6}$$

as $n \to \infty$. Hence, Theorem 1 follows from (A.1)–(A.6) and Corollary 3.1 of Hall and Heyde (1980). ■

Proof of Theorem 2. It follows from (A.2)–(A.4) that

$$\frac{1}{n} \sum_{i=1}^{n} Z_i^2(\beta_0; \delta) \xrightarrow{p} \begin{cases} \sigma^2 \mathrm{E} \frac{Y_1^2}{\delta + Y_1^2}, & \text{if } |\beta| < 1, \\ 1, & \text{if } \beta = 1 - \gamma/n, \\ 1, & \text{if } |\beta| > 1. \end{cases} \tag{A.7}$$

By Corollary 3.1 of Hall and Heyde (1980),

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i(\beta_0; \delta) \xrightarrow{d} \begin{cases} N\left(0, \sigma^2 \mathrm{E} \frac{Y_1^2}{\delta + Y_1^2}\right), & \text{if } |\beta| < 1, \\ N(0, \sigma^2), & \text{if } \beta = 1 - \gamma/n, \\ N(0, \sigma^2), & \text{if } |\beta| > 1. \end{cases} \tag{A.8}$$

It is easy to check that

$$\max_{1 \le i \le n} |Z_i(\beta_0; \delta)| = o(n^{-1/2}) \quad \text{a.s.} \tag{A.9}$$

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

Hence, Theorem 2 can be proved by using (A.7)–(A.9) and standard arguments in Chapter 11 of Owen (2001). ■

Proof of Theorem 3. Define

$$Q_{1n}(\theta,\lambda)=\frac{1}{n-p-2}\sum_{i=p+3}^{n}\frac{Z_i(\beta_0,\theta)}{1+\lambda^T Z_i(\beta_0,\theta)}$$

and

$$Q_{2n}(\theta,\lambda)=\frac{1}{n-p-2}\sum_{i=p+3}^{n}\frac{1}{1+\lambda^T Z_i(\beta_0,\theta)}\left\{\frac{\partial}{\partial\theta}Z_i(\beta_0,\theta)\right\}^T\lambda.$$

Note that when $\beta_0=1$,

$$\frac{1}{n-p-2}\sum_{i=p+3}^{n}\frac{Y_{i-1}^2}{\delta+Y_{i-1}^2}\xrightarrow{p}1.$$

Similar to the proof of Theorem 1, it can be shown that

$$\begin{cases}\dfrac{1}{\sqrt{n-p-2}}\displaystyle\sum_{i=p+3}^{n}Z_i(\beta_0,\theta_0)\xrightarrow{d}N(0,\mathrm{E}\{Z_1(\beta_0,\theta_0)Z_1^T(\beta_0,\theta_0)\}),\\ \dfrac{1}{n-p-2}\displaystyle\sum_{i=p+3}^{n}Z_i(\beta_0,\theta_0)Z_i^T(\beta_0,\theta_0)\xrightarrow{p}\mathrm{E}\{Z_1(\beta_0,\theta_0)Z_1^T(\beta_0,\theta_0)\}.\end{cases}\tag{A.10}$$

Define $\tilde{\theta}=\arg\min_\theta \tilde{l}(\beta_0,\theta)$ and $\tilde{\lambda}=\lambda(\beta_0,\tilde{\theta})$. Then, using the same arguments in the proof of Theorem 1 of Qin and Lawless (1994), it follows from (A.10) that

$$\begin{pmatrix}\tilde{\lambda}\\ \tilde{\theta}-\theta_0\end{pmatrix}=S_n^{-1}\begin{pmatrix}-Q_{1n}(\theta_0,0)+o_p\left(n^{-1/2}\right)\\ o_p\left(n^{-1/2}\right)\end{pmatrix}\tag{A.11}$$

and

$$S_n\xrightarrow{p}\begin{pmatrix}S_{11}&S_{12}\\ S_{21}&0\end{pmatrix}=\begin{pmatrix}-\mathrm{E}\{Z_1(\beta_0,\theta_0)Z_1^T(\beta_0,\theta_0)\}&\mathrm{E}\left\{\frac{\partial Z_1(\beta_0,\theta_0)}{\partial\theta}\right\}\\ \mathrm{E}\left\{\frac{\partial Z_1(\beta_0,\theta_0)}{\partial\theta}\right\}^T&0\end{pmatrix},\tag{A.12}$$

where

$$S_n=\begin{pmatrix}\frac{\partial Q_{1n}(\theta_0,0)}{\partial\lambda^T}&\frac{\partial Q_{1n}(\theta_0,0)}{\partial\theta}\\ \frac{\partial Q_{2n}(\theta_0,0)}{\partial\lambda^T}&0\end{pmatrix}.$$

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press

By the standard arguments in the proof of the empirical likelihood method (see Owen, 2001), it follows from (A.10)–(A.12) that

$$\begin{aligned}\tilde{l}^p(\beta_0) &= 2n(\tilde{\lambda}^T, \tilde{\theta}^T - \theta_0^T)(Q_{1n}^T(\theta_0, 0), 0)^T \\ &\quad + n(\tilde{\lambda}^T, \tilde{\theta}^T - \theta_0^T)S_n(\tilde{\lambda}^T, \tilde{\theta}^T - \theta_0^T)^T + o_p(1) \\ &= -n(Q_{1n}^T(\theta_0, 0), 0)S_n^{-1}(Q_{1n}^T(\theta_0, 0), 0)^T + o_p(1) \\ &= -(W^T, 0)\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & 0 \end{pmatrix}^{-1}(W^T, 0)^T + o_p(1),\end{aligned} \tag{A.13}$$

where W is a multivariate normal random variable with mean zero and covariance $-S_{11}$. Because

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & 0 \end{pmatrix}^{-1} = \begin{pmatrix} S_{11}^{-1} - S_{11}^{-1}S_{12}\Delta^{-1}S_{21}S_{11}^{-1} & S_{11}^{-1}S_{12}\Delta^{-1} \\ \Delta^{-1}S_{21}S_{11}^{-1} & -\Delta^{-1} \end{pmatrix},$$

where $\Delta = S_{21}S_{11}^{-1}S_{12}$, we have

$$\begin{aligned}&-(W^T, 0)\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & 0 \end{pmatrix}^{-1}(W^T, 0)^T \\ &\quad = -W^T\{S_{11}^{-1} - S_{11}^{-1}S_{12}\Delta^{-1}S_{21}S_{11}^{-1}\}W \\ &\quad = \{(-S_{11})^{-1/2}W\}^T S_{11}^{1/2}\{S_{11}^{-1} - S_{11}^{-1}S_{12}\Delta^{-1}S_{21}S_{11}^{-1}\}S_{11}^{1/2}\{(-S_{11})^{-1/2}W\} \\ &\quad = \{(-S_{11})^{-1/2}W\}^T\{I - S_{11}^{-1/2}S_{12}\Delta^{-1}S_{21}S_{11}^{-1/2}\}\{(-S_{11})^{-1/2}W\}.\end{aligned} \tag{A.14}$$

Because

$$\mathrm{tr}(S_{11}^{-1/2}S_{12}\Delta^{-1}S_{21}S_{11}^{-1/2}) = \mathrm{tr}(\Delta^{-1}S_{21}S_{11}^{-1}S_{12}) = p - 1,$$

we have $\mathrm{tr}\left(I - S_{11}^{-1/2}S_{12}\Delta^{-1}S_{21}S_{11}^{-1/2}\right) = 1$. Hence it follows from (A.13) and (A.14) that $\tilde{l}^p(\beta_0) \xrightarrow{p} \chi^2(1)$ as $n \to \infty$. ■

https://doi.org/10.1017/S0266466611000727 Published online by Cambridge University Press