Testing for a Unit Root in Noncausal Autoregressive Models

Pentti Saikkonen† Rickard Sandberg‡

University of Helsinki, Finland†

Stockholm School of Economics, Sweden‡

HECER Time Series Econometrics Seminar

November 4, 2013

Outline

Some definitions and terminology

- Assume that $\{X_t\}$ satisfies the pth-order difference quation

$$X_t - \zeta_1 X_{t-1} - \cdots - \zeta_p X_{t-p} = \epsilon_t$$

where $\epsilon_t \overset{i.i.d.}{\sim} WN(0, \sigma^2)$

- Provided that $\zeta(z) = 1 - \zeta_1 z - \cdots - \zeta_p z^p$ has no roots on the unit circle ($\zeta(z) \neq 0$ for $|z| = 1$), a unique stationary solution to the difference equation exisits

Some defintions and terminology

- This solution is said to be causal if $\zeta(z)$ has no roots inside the unit circle ($\zeta(z) \neq 0$ for $|z| \leq 1$), and we have that $X_t = f(\epsilon_t, \epsilon_{t-1}, ...)$
- On the other hand, the solution is said to be noncausal if $\zeta(z)$ has any root inside the unit circle
- And, the solution is siad to be purely noncausal if $\zeta(z)$ has all the roots inside the unit circle ($\zeta(z) \neq 0$ for $|z| \geq 1$), and we have that $X_t = f(\epsilon_{t+1}, \epsilon_{t+2}, ...)$

Some defintions and terminology

- To handle the mixed case it is convenient to factor the auturegressive polynomial as

$$\zeta(z) = 1 - \zeta_1 z - \cdots - \zeta_p z^p = \zeta^\dagger(z)\zeta^*(z)$$

where

$$\zeta^\dagger(z) = 1 - \phi_1 z - \cdots - \phi_r z^r \neq 0 \quad \text{for } |z| \leq 1$$

$$\zeta^*(z) = 1 - \varphi_1 z - \cdots - \varphi_s z^s \neq 0 \quad \text{for } |z| \geq 1$$

and $r, s \geq 0$ with $r + s = p$

The stationary NCAR model

- This suggets that we may define a stationary NCAR(r, s) model (Lanne and Saikkonen, 2011) as

$$\phi(B)\varphi(B^{-1})X_t = \epsilon_t, \qquad t = 1, 2, ... \tag{1}$$

where

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_r B^r$$

and

$$\varphi(B^{-1}) = 1 - \varphi_1 B^{-1} - \cdots - \varphi_s B^{-s}$$

with $\phi_i \neq 0$ for some $i \in \{1, ..., r\}$ and $\varphi_j \neq 0$ for some $j \in \{1, ..., s\}$, B is the usual backward shift operator, and ϵ_t is an error term (properties discussed later on)

The stationary NCAR model

- Example: the NCAR(1, 1) model

$$\underbrace{X_t}_{\text{current value}} (1+\phi_1\varphi_1) - \phi_1 \underbrace{X_{t-1}}_{\text{past value}} - \varphi_1 \underbrace{X_{t+1}}_{\text{future value}} = \epsilon_t$$

The stationary NCAR model

- The NCAR model in (1) is said to be purely noncausal if $\phi_1 = \cdots = \phi_r = 0$
- On the other hand, the NCAR model in (1) reduces to the conventional casual autoregression when $\varphi_1 = \cdots = \varphi_s = 0$
- Assuming that $\phi(z)$ and $\varphi(z)$ $(z \in \mathbb{C})$ have their roots outside the unit circle the NCAR model in (1) has a stationary solution

$$X_t = f(\ldots, \epsilon_{t-1}, \epsilon_t, \epsilon_{t+1}, \ldots) = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j}$$

i.e. two-sided moving average representation

The stationary NCAR model

- A backward-looking moving average representation: define $u_t \overset{def}{=} \varphi(B^{-1})X_t$ which implies that we can write (1) as

$$u_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$$

- A forward-looking moving average representation: define $v_t \overset{def}{=} \phi(B)X_t$ which yields that we can write (1) as

$$v_t = \sum_{j=0}^{\infty} \beta_j \epsilon_{t+j}$$

Applications of the stationary NCAR model

- The use of the NACR model in practice can be motivated by...
- According to the theory on present value models and rational expectations (say) future values also effect the current value
- For instance, Lanne and Saikkonen (2011) find that the (high) persistence previously found in US inflation is not caused by dependence on past inflation but on the expectations on future inflation
- That is, they find rather strong support for that US inflation series has a purely forward-looking behavior (i.e. New Keynesian modeling)

Applications of the NCAR model

- On may also think of this is that the error term (ϵ_t) in an autoregressive model should not be predictable by the past of the series ($E(\epsilon_t|y_{t-1}, y_{t-2}, ...) = 0$))

- However, if relevant variables are omitted their impact goes (at least partly) to the error term of the model and, as the considered time series may help to predict the omitted variables, the assumed unpredictability condition may break down

- As economic variables are typically interrelated, this point appears particularly pertinent in economic applications

- So, the NCAR model may provide a viable alternative, for it explicitly allows for the predictability of the error term by the past of the considered series

Applications of the NCAR model

- Confined to natural sciences and engineering, NCAR type of models have been studied by Breidt et al. (1991), Lii and Rosenblatt (1996), Huang and Pawitan (2000), Breidt et al. (2001), and Andrews (2006)
- NCAR models have more recently been applied to economic time series Lanne and Saikkonen (2011), and further studied by Lanne et al. (2012a), Lanne et al. (2012b), Lanne et al. (2012c), Lanne and Saikkonen (2013), and Gourieroux and Zakoian (2013)
- In these more recent papers it is demonstrated that noncausal models many times perform better (both in an in sample and out of sample context) then the causal models

A nonstationary NCAR model

- Lanne and Saikkonen (2011) consider specification, estimation and hypothesis testing, using methods of ML, in the stationary NCAR model
- They show that estimators of the NCAR parameters follow standard distributions
- In this work, a ML based unit root test is derived in the NCAR model such that the stationarity assumption in above articles (more or less) can be tested
- It is noticed that we expect traditional unit root tests (DF type of tests) not to be so powerful in the presence of a forward-looking component (they are simply based on a misspecified model

A nonstationary NCAR model

- To derive a unit root test we first define a unit root NCAR model. Hence, assume that $r > 0$ and proceed by writing

$$\phi(B) = \Delta - \phi B - \pi_1 \Delta B - \cdots - \pi_{r-1} \Delta B^{r-1} \qquad (2)$$

  where $\Delta = 1 - B$ is the difference operator

- Using (2) implies that the NCAR model in (1) can be written as

$$\Delta X_t = \phi X_{t-1} + \pi_1 \Delta y_{t-1} + \cdots + \pi_{r-1} \Delta X_{t-r+1} + v_t, \qquad t = 1, 2, ...$$

  where v_t is defined as above

- Remark: this looks like an ADF-type of regression, but the error term is fundamentally different...

A nonstationary NCAR model

- It is also assumed that (no more than a single unit root):

$$\pi(z) \neq 0 \text{ for } |z| \leq 1 \quad \text{and} \quad \varphi(z) \neq 0 \text{ for } |z| \leq 1$$

- A nonstationary NCAR model may now be defined by letting $\phi = 0$ in (2), and entails

$$\Delta X_t = \sum_{j=-\infty}^{\infty} \psi_j \epsilon_{t-j}$$

- This implies that ΔX_t is a stationary and ergodic process with finite second moments

- Therefore, we consider a test-statistic for testing:

$$H_0 : \phi = 0$$

against

$$H_1 : \phi < 0$$

in the NCAR model in (1).

Identification

- A practical complication of NCAR models is that they cannot be identified by second-order properties or a Gaussian likelihood

- That is, even if X_t is noncausal, its spectral density and, hence, ACF cannot be distinguished from those of a causal autoregressive process (Brockwell and Davis, 1987)

- This identification problem is circumvented if non-Gaussian errors are assumed. As such, the following assumption is imposed

- **Assumption 1.** *The zero mean error term* ϵ_t *is a sequence of non-Gaussian IID random variables with a (Lebesgue) density* $\sigma^{-1}f\left(\sigma^{-1}x;\lambda\right)$ *which depends on the error variance* $\sigma^2(>0)$ *and (possibly) on the parameter vector* λ $(d \times 1)$ *taking values in an open set* $\Lambda \subseteq \mathbb{R}^d$.

Large sample properties of unit root MLEs

- Once the error distribution is specified, the parameters of the NCAR model are estimated by ML
- The (approximate) log-likelihood is given by:

$$l_T(\theta) = \sum_{t=r+1}^{T-s} g_t(\theta)$$

where

$$\begin{aligned} & g_t(\theta) \\ = \quad & \log f\left(\sigma^{-1}\left(v_t - \varphi_1 v_{t+1} - \cdots - \varphi_s v_{t+s}\right); \lambda\right) - \log \sigma \\ = \quad & \log f\left(\sigma^{-1}\left(\Delta u_t - \phi u_{t-1} - \pi_1 \Delta u_{t-1} - \cdots - \pi_{r-1} \Delta u_{t-r+1}\right); \lambda\right) \\ & - \log \sigma \end{aligned}$$

where $u_t(\varphi) = \varphi\left(B^{-1}\right) X_t$ and $v_t(\phi, \pi) = \phi(B) X_t$

Specification

- The parameters of the model are given by: $\theta = (\phi, \vartheta) = (\phi, \vartheta_1, \vartheta_2)$ where
  - ϕ is a long-run parameter ; mean equation
  - $\vartheta_1 = (\pi, \varphi)$ is a vector of short-run parameters ; mean equation
  - $\vartheta_2 = (\sigma, \lambda)$ is a vector of parameters related to the error distribution

- The vector of first-order partial derivatives for the log-likelihood (based on a single value) when evaluated at the true parameter value is denoted

$$g_{\theta,t}(\theta_0) = \begin{bmatrix} g_{\phi,t}(\theta_0) \\ g_{\vartheta,t}(\theta_0) \end{bmatrix}$$

  where θ_0 signifies the true value of θ (and we assume that H_0 holds so the true value of ϕ is zero)

- Remark: The score of ϑ (evaluated at θ_0) is clearly a stationary and ergodic process similar to the score in Lanne and Saikkonen (2011)

Large sample properties of unit root MLEs

- Before presenting the large sample results of the score we assume:

- **Assumption 2.** *For all* $(x, \lambda) \in (\mathbb{R}, \Lambda)$, $f(x; \lambda) > 0$ *and* $f(x; \lambda)$ *is twice continuously differentiable with respect to* (x, λ) *and an even function of* x, *that is,* $f(x; \lambda) = f(-x; \lambda)$

- Remark: unlike other authors we assume that $f(.; \lambda)$ is even. This assumption is imposed to simplify the limiting distribution of our test

- It is also convenient to introduce a new error process $e_{x,t}$ which is a normailized derivative of the density of the error distribution

- **Assumption 3.** *(i)* $E[e_{x,t}] = 0$ *and* $E\left[e_{x,t}^2\right] = \mathcal{J}$, *where* $\mathcal{J} = \int (f_x(x;\lambda_0)^2 / f(x;\lambda_0))dx > 1$ *is finite. Moreover,* $Cov[\epsilon_t, e_{x,t}] = -\sigma_0$. *(ii) The score vector* $g_{\vartheta,t}(\theta_0)$ *has zero expectation and finite positive definite covariance matrix* Σ

- These assumptions may be verified by the results of Andrews et al. (2006) and Lanne and Saikkonen (2011)

- Remark: $\mathcal{J}$ will later on enter our limiting distribution as a nuisance parameter

Lemma (large sample results first-order derivatives)

Suppose that Assumptions 1-3 hold. Then,

$$T^{-1}\sum_{t=r+1}^{T-s} g_{\phi,t}(\theta_0) \xrightarrow{d} Z_1 = -\frac{1}{\sigma_0\pi_0(1)}\int_0^1 B_\epsilon(u)\, dB_{e_x}(u) \qquad (3)$$

where $B_\epsilon(u)$ *and* $B_{e_x}(u)$ *are two dependent Brownian motions, and*

$$T^{-1/2}\sum_{t=r+1}^{T-s} g_{\vartheta,t}(\theta_0) \xrightarrow{d} Z_2 \sim N(0,\Sigma) \qquad (4)$$

Moreover, joint weak convergence applies with Z_1 *and* Z_2 *independent.*

- Furthermore, by $g_{\phi\phi,t}(\theta_0)$, $g_{\vartheta\vartheta,t}(\theta_0)$, and $g_{\phi\vartheta,t}(\theta_0)$ we denote second-order partial derivatives evaluated at the true parameter values

- In order to derive the large sample properties of the second-order derivatives we assume:

- **Assumption 4.** *$E[e_{xx,t}] = -E\left[e^2_{x,t}\right]$ and $E[g_{\vartheta\vartheta,t}(\theta_0)] = -\Sigma$ with Σ given in Assumption 3(ii). Moreover, $E[e_{xx,t}\epsilon_t] = 0$ and $E[e_{\lambda x,t}] = 0$.*

Lemma (large sample results second-order derivatives)

Suppose that Assumptions 1-4 hold. Then,

$$-T^{-2}\sum_{t=r+1}^{T-s} g_{\phi\phi,t}\left(\theta_0\right) \xrightarrow{d} \frac{\mathcal{J}}{\sigma_0^2\pi_0\left(1\right)^2}\int_0^1 B_\epsilon^2\left(u\right)d\left(u\right) \qquad (5)$$

$$-T^{-1}\sum_{t=r+1}^{T-s} g_{\vartheta\vartheta,t}\left(\theta_0\right) \xrightarrow{p} \Sigma \qquad (6)$$

and

$$-T^{-3/2}\sum_{t=r+1}^{T-s} g_{\phi\vartheta,t}\left(\theta_0\right) \xrightarrow{p} 0 \qquad (7)$$

Moreover, the weak convergences in (5) and in the previous Lemma hold jointly and, Z_2 in (4) is independent of the limit in (5).

Large sample properties of unit root MLEs

- Denoting $Z = (Z_1, Z_2)$ and $D_T = \text{diag}\left(T, T^{-1/2} I_{r+s+d}\right)$ we obtain from the two lemmas the following Theorem.

Theorem (large sample results the score and the hessian)

Suppose that Assumptions 1-4 hold. Then,

$$S_T(\theta_0) \overset{def}{=} D_T^{-1} \sum_{t=r+1}^{T-s} g_{\theta,t}(\theta_0) \xrightarrow{d} Z \tag{8}$$

and

$$G_T(\theta_0) \overset{def}{=} -D_T^{-1} \sum_{t=r+1}^{T-s} g_{\theta\theta,t}(\theta_0) D_T^{-1} \xrightarrow{d} G(\theta_0) \tag{9}$$

where the (block diagonal) matrix $G(\theta_0)$ contains the weak limits in (5), (6) and (7), and the weak convergences in (8) and (9) hold jointly with $(Z_1, G(\theta_0))$ and Z_2 independent.

Large sample properties of unit root MLEs

- To derive the limiting distribution of the ML estimator of θ under the unit root hypothesis we also need the condition (verified in the paper): for all $c > 0$,

$$\sup_{\theta \in N_{T,c}} \left\| G_T(\theta) - G_T(\theta_0) \right\| \xrightarrow{p} 0 \tag{10}$$

where $N_{T,c} = \{\theta : D_T \|\theta - \theta_0\| \leq c\}$.

Theorem

Suppose that Assumptions 1-4 and condition (10) hold. Then, with probability approaching one, there exists a sequence of local maximizers of the log-likelihood function $\hat{\theta}_T$ such that

$$\left(D_T(\hat{\theta}_T - \theta_0), G_T(\theta_0) \right) \xrightarrow{d} \left(G(\theta_0)^{-1} Z, G(\theta_0) \right). \tag{11}$$

Moreover, $G_T(\hat{\theta}_T) - G_T(\theta_0) \xrightarrow{p} 0.$

Unit root testing

- The following "*t*-ratio" type of test-statistic is used to test the unit root hypothesis:

$$\tau_T \overset{def}{=} \hat{\phi}/\sqrt{G_T^{1,1}(\hat{\theta}_T)}$$

where $G_T^{1,1}(\hat{\theta}_T)$ abbreviates the (1,1)-element of $G_T(\hat{\theta}_T)^{-1}$.

Theorem

Suppose that Assumptions 1-4 and condition (10) hold. Then

$$\tau_T \xrightarrow{d} \frac{\int_0^1 W_\epsilon(u)\, dW_\epsilon(u) - (\mathcal{J}-1)^{1/2} \int_0^1 W_\epsilon(u)\, dW(u)}{\sqrt{\mathcal{J} \int_0^1 W_\epsilon^2(u)\, d(u)}} \overset{def}{=} \tau$$

where $W_\epsilon(u) = \sigma_0^{-1} B_\epsilon(u) \sim BM(1)$, *and* $W(u) \sim BM(1)$ *and independent of* $W_\epsilon(u)$

- Remark: the asymptotic distribution τ is free of nuisance parameters except for the parameter $\mathcal{J}$ which depends on the distribution of the error term

- The nuisance parameter problem will be addressed in our simulation studies

- Tests allowing for trends: by $\tau_T(m)$ we refer to a test based on raw, demeaned, and detrended data, when $m = 0$, $m = 1$, and $m = 2$, respectively. The previously considered large sample results are still valid but demeaned and detrended Brownian motions now apply

- A solution the the nuisance parameter problem. It is noted that the correlation between $B_{\epsilon}(u)$ and $B_{e_x}(u)$ equals $\rho = \mathcal{J}^{-1/2} \in [0, 1]$
- In the following figure the 1st, 5th, and 10th percentiles of τ are shown as a function of ρ (obtained via simulations)

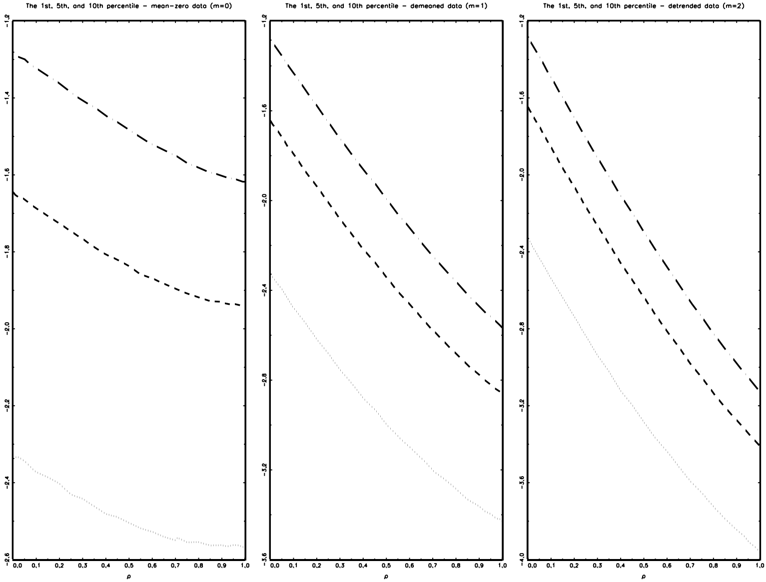

Figure: The role of the nuisance parameter $\mathcal{J}$

Critical values

- In this figure, a monotonically decreasing relationship between the percentiles and ρ is seen
- Due to the percentiles monotonicity in ρ it is obvious that if the value of $\mathcal{J}$ would be known, Figure 1 could be used to determine (conventional) (asymptotic) critical values
- We proceed instead with curve estimation of the percentiles by fitting a second-order polynomial: $cv_{\alpha,m}(\rho) = b_0 + b_1\rho + b_2\rho^2$ for $\alpha \in \{.01, .05.10\}$ and $m \in \{0, 1, 2\}$. The curve estimates, obtained by LS, are presented in the following table:

Critical values

Table 1 Curve estimation results of the percentiles of the $\tau_T(m)$ statistic.

Case	Critical value (α)	b_0	b_1	b_2	R^2
mean-zero data	1%	−2.321	−.492	.251	.998
($m = 0$)	5%	−1.639	−.495	.187	.999
	10%	−1.276	−.480	.131	.999
demeaned data	1%	−2.322	−1.578	.474	1.00
($m = 1$)	5%	−1.639	−1.591	.367	1.00
	10%	−1.276	−1.584	.289	1.00
detrended data	1%	−2.324	−2.201	.575	1.00
($m = 2$)	5%	−1.640	−2.230	.462	1.00
	10%	−1.276	−2.231	.381	1.00

Estimation of $\mathcal{J}$

- In the case of t-distributed errors one can show that:

$$\mathcal{J} = \frac{\lambda(\lambda+1)}{(\lambda-2)(\lambda+3)}$$

and an estimate of $\mathcal{J}$ is readily obtained via the estimate of λ. This estimator will be denoted $\widehat{\mathcal{J}}_1$

- In the case of "other" distributions (where the calculations not are so straightforward) we may use:

$$\widehat{\mathcal{J}} = \frac{1}{T-r-s} \sum_{t=r+1}^{T-s} \left[\frac{f'(\hat{\sigma}^{-1}\hat{\epsilon}_t; \hat{\lambda})}{f(\hat{\sigma}^{-1}\hat{\epsilon}_t; \hat{\lambda})} \right]^2$$

where $\hat{\epsilon}_t = \Delta\hat{u}_t - \hat{\phi}\hat{u}_{t-1} - \hat{\pi}_1\Delta\hat{u}_{t-1} - \cdots - \hat{\pi}_{r-1}\Delta\hat{u}_{t-r+1}$ with $\hat{u}_t = \hat{\varphi}(B^{-1})y_t$. This estimator will be denoted $\widehat{\mathcal{J}}_2$

Empirical size

- In our size studies we use simulate data from a non-stationary NCAR(1, 1) model with $\phi = 0$ and $\varphi_1 \in \{.10, .50, .90\}$
- The error term ϵ_t follows a Student's t-distribution with degrees of freedom λ equal to 3 and standard deviation σ equal to 0.1
- The sample sizes used are $T \in \{100, 250\}$, and the number of replications are set to $10,000$

Empirical size

Table 2 Empirical size of the $\tau_T(m)$ test.

Sample Size		mean-zero data $(m = 0)$ φ_1			demeaned data $(m = 1)$ φ_1			detrended data $(m = 2)$ φ_1		
T		0.1	0.5	0.9	0.1	0.5	0.9	0.1	0.5	0.9
	$\mathcal{J}$	.053	.053	.083	.059	.059	.086	.058	.056	.098
100	$\widehat{\mathcal{J}}_1$	.052	.052	.083	.054	.054	.080	.056	.059	.097
	$\widehat{\mathcal{J}}_2$	.052	.052	.083	.054	.053	.080	.055	.059	.097
	$\mathcal{J}$	.054	.047	.045	.063	.061	.055	.057	.052	.056
250	$\widehat{\mathcal{J}}_1$	.053	.046	.044	.058	.058	.054	.056	.052	.057
	$\widehat{\mathcal{J}}_2$	.053	.046	.044	.058	.058	.054	.056	.052	.057

Empirical power

- In our power studies we use simulate data from a stationary NCAR(1, 1) model with $\phi \in [-.4, 0)$ and $\varphi_1 = .5$
- The error term ϵ_t follows the same distribution as in the size studies
- The sample sizes considered are $T \in \{100, 250\}$.
- For comparison we also choose to report the outcomes of the traditional Dickey-Fuller unit root t-test (based on an AR(2) process) as well as the t-type unit root test of Lucas (1995) (based on M-estimation in an AR(1) model assuming strictly stationary strong-mixing errors)

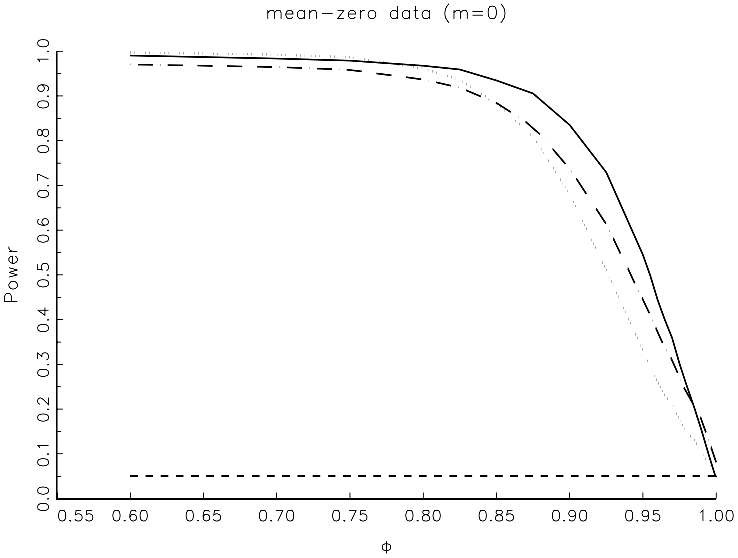

Figure: Power T=100 (solid line $\tau_T(0)$; dotted line $\tau_{DF}(0)$; dashed line $M(0)$)

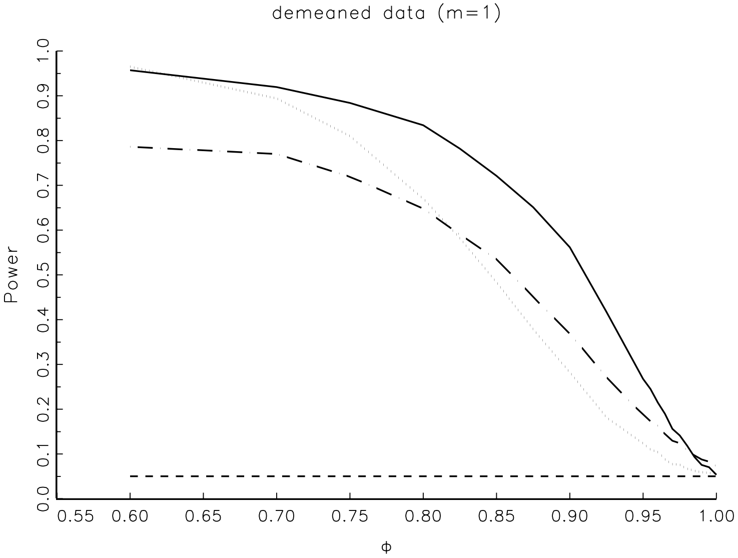

Figure: Power T=100 (solid line $\tau_T(1)$; dotted line $\tau_{DF}(1)$; dashed line $M(1)$)

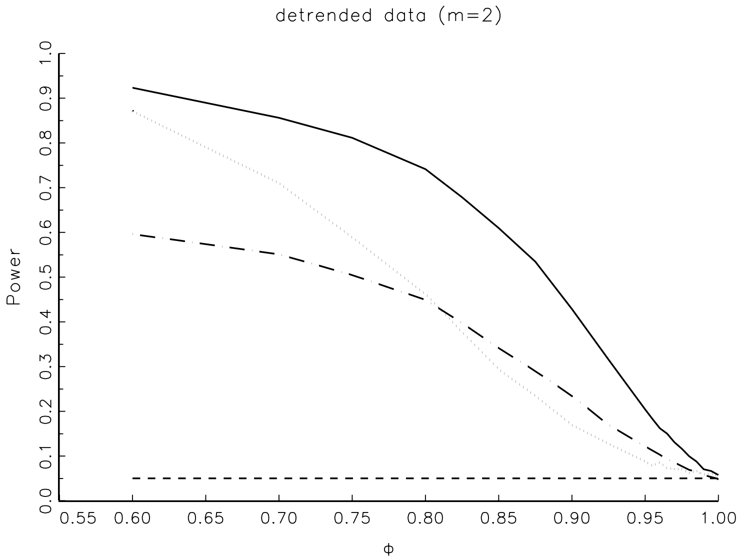

Figure: Power T=100 (solid line $\tau_T(2)$; dotted line $\tau_{DF}(2)$; dashed line $M(2)$)

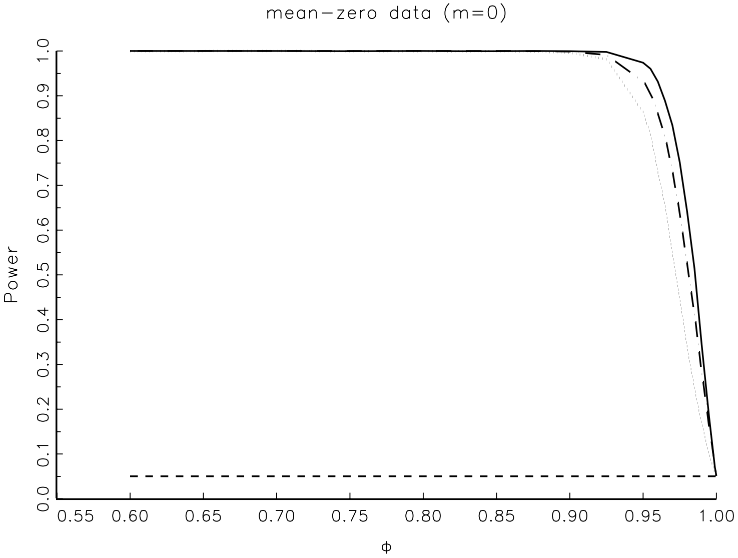

Figure: Power T=250 (solid line $\tau_T(0)$; dotted line $\tau_{DF}(0)$; dashed line $M(0)$)

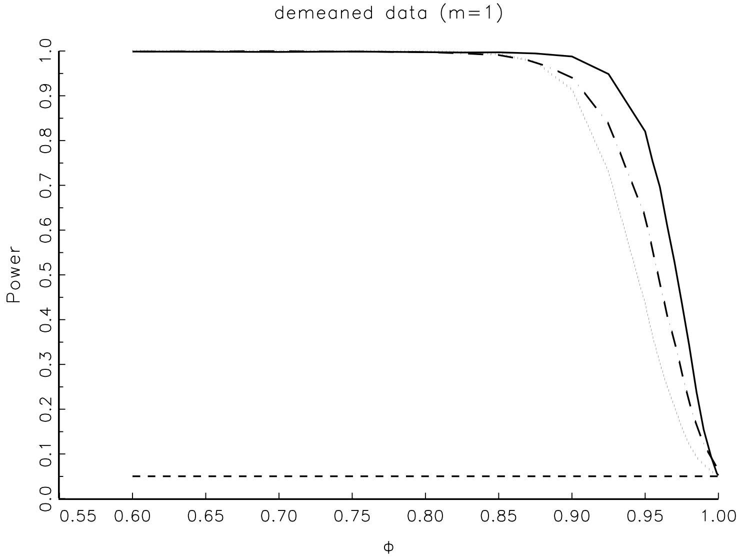

Figure: Power T=250 (solid line $\tau_T(1)$; dotted line $\tau_{DF}(1)$; dashed line $M(1)$)

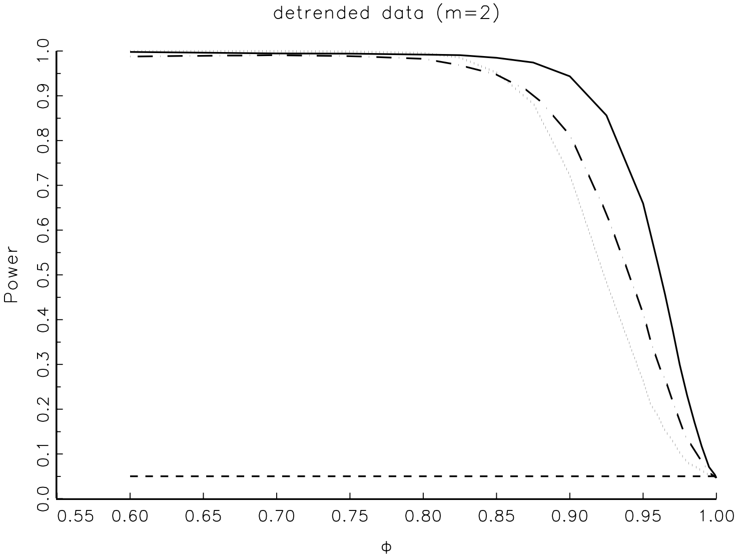

Figure: Power T=250 (solid line $\tau_T(2)$; dotted line $\tau_{DF}(2)$; dashed line $M(2)$)

Application: Finnish interest rates

- In this application the dynamic behavior of Finnish interest rate series (Government bonds), ranging from 1988:Q1 to 2012:Q4 (T=100), is examined.

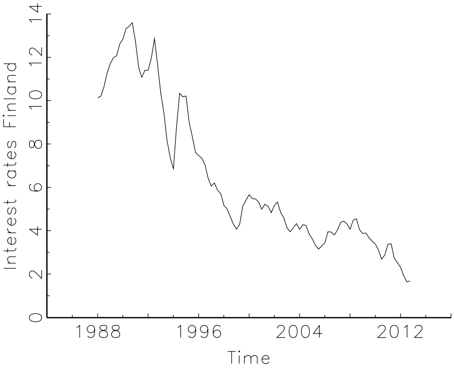

Figure: Government bonds Finland 1988:Q1-2012:Q4

Application: Finnish interest rates

- In this application we shall also take the opportunity and demosntrate/explain the steps of unit root testing in the NCAR model in practiceT. The practioner may proceed as follows...
  - I Fit a causal autoregression (i.e. an AR(p) process) is fitted, and thereafter we check whether the residuals look non-Gaussian
  - II If the residuals appear to be non-Gaussian, r and s must be determined. In our case all combinations of r and s such that $p = r + s$ are considered. That is, several NCAR specifications may be estimated and several unit root tests are conducted
  - III In the case of multiple rejections one may select the model with the highest likelihood
  - IV In the case of no rejection, on e may proceed by estimating a NCAr model using first-differences
- Notice that we must assume that $r > 0$ (otherwise unit root testing makes no sense; purely noncausal models are ruled out).
- Finally, we also abstain from the possibility of unit roots in the noncausal/future part of the process (mainly due to technical complications)

Application: Finnish interest rates

- Following the steps in our procedure we notice that...
  - Using demeaned data ($m = 1$), an AR(3) was selected by both AIC and BIC
  - Using detrended data ($m = 2$) , an AR(2) was selected by both AIC and BIC
  - The Ljung-Box test (4 lags) did not indicate that there are unmodeled serial correlation
  - The McLeod-Li test (4 lags) indicates the presence of conditional heteroscedasticity
  - The Lomnicki, Jarque, and Bera and the Shapiro and Wilk tests reject (strongly) the null hypothesis of normally distributed errors
  - QQ-plots okay

Application: Finnish interest rates

- Inspired by these findings we proceed by unit root testing assuming that the errors have a Student's t-distribution.
- Using demeaned data ($p = 3 = r + s$) implies that the unit root testing is done for two cases, viz. in the NCAR$(1, 2)$ model and in the NCAR$(2, 1)$ model.
- Using detrended data ($p = 2 = r + s$) implies that the unit root testing is done only for one case, viz. in the NCAR$(1, 1)$ model.
- The outcomes of this unit root testing exercise as well as other various estimation results are shown in the following tables.

Application

Table 3 Outcomes unit root test interest rate series Finland

$T^* = 97$			$m = 1$	
		$cv_{.05,1}$	$cv_{.05,1}(\widehat{\mathcal{J}}_1)$	$cv_{0.05,1}(\widehat{\mathcal{J}}_2)$
$ADF(3)$	-1.041	-2.860		
$t_{NCAR(2,1)}$	-5.101^{***}		-2.558	-2.577
$t_{NCAR(1,2)}$	-5.997^{***}		-2.542	-2.531
$T^* = 98$			$m = 2$	
		$cv_{.05,2}$	$cv_{.05,2}(\widehat{\mathcal{J}}_1)$	$cv_{0.05,2}(\widehat{\mathcal{J}}_2)$
$ADF(2)$	-3.213^{*}	-3.410		
$t_{NCAR(1,1)}$	-5.379^{***}		-3.072	-3.048

Application

Table 4 NCAR MLE results for demeaned interest rate series Finland

$m = 1$	NCAR specifications: $p = r + s = 3$				
$T^* = 97$	(3,0)-N	(3,0)-t	(2,1)-t	(1,2)-t	(0,3)-t
ϕ	$-.015$	.003	$-.472$	$-.425$	
π_1	.568	.590	.056		
π_2	$-.180$	$-.084$			
φ_1			.941	.873	1.550
φ_2				.056	$-.695$
φ_3					.130
σ	.447	.457	.477	.498	.461
λ		3.585	3.017	2.831	4.129
LL	-59.556	-49.111	-46.052	**-45.637**	-53.437
$LB(4)$	.492	.324	.674	.348	.522

Table 5 NCAR MLE results for detrended interest series Finland

$m = 2$	NCAR specifications: $p = r + s = 2$			
$T^* = 98$	(2,0)-N	(2,0)-t	(1,1)-t	(0,2)-t
ϕ	$-.104$	$-.052$	-0.418	
π_1	.505	.563		
φ_1			.806	1.445
φ_2				$-.521$
σ	.431	.442	.440	.451
λ		3.861	3.293	3.996
LL	-57.055	-48.998	**-43.480**	-51.528
$LB(4)$	.811	.358	.503	.584

Conclusions

- In this paper we derive unit root tests in the NCAR model by Lanne and Saikkonen (2011).
- Large sample properties of the MLEs in the NCAR model are proved under a unit root assumption
- The finite sample properties of the tests are promising. In particular, the tests are (significantly) more powerful against NCAR alternatives than the DF tests
- In our application to Finnish interest rate series we found evidence in favour of a stationary NCAR model with leptokurtic errors