A Unit Root Test for Nonlinear ESTAR(2) Models

Heni Puspaningrum

School of Mathematics and Applied Statistics
Centre for Statistical and Survey Methodology
University of Wollongong
Australia
Outline

1. Introduction
2. Model, null hypothesis and auxiliary regression
3. Limit results and asymptotic test
4. Small sample properties
5. Empirical example
6. Conclusion
Dickey and Fuller (DF) (1979, 1981): unit root test based on linear ARMA models

Consider AR(1) model: $y_t = \theta y_{t-1} + \varepsilon_t$

Unit root test: $H_0 : \theta = 1$ vs $H_1 : \theta < 1$

In other form, $\Delta y_t = \psi y_{t-1} + \varepsilon_t$

and the unit root test: $H_0 : \psi = 0$ vs $H_1 : \psi < 0$

For autocorrelated $\varepsilon_t$, use an augmented Dickey Fuller (ADF) test:

$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + \varepsilon_t$
Disadvantage of the ADF test: lack power when the model specification under the alternative hypothesis is nonlinear (see Nelson and Plosser, 1982; Taylor et al., 2001; and Rose, 1988)

As a result, the literature in nonlinear unit root tests has been growing rapidly. For examples:


2. LSTAR(1) and LSTAR(2) models: Eklund (2003a, 2003b)

3. ESTAR(1) models: Kapetanios et al. (2003)
Purpose:

1. Extend the work in Kapetanios et al. (2003)

2. Consider an exponential smooth transition autoregressive of order 2, ESTAR(2) model

3. May be used to develop a unit root test for ESTAR(p) models, $p > 2$
Why ESTAR models?

1. Real exchange rates and purchasing power parity (PPP) deviations (Michael et al., 1997; Taylor et al., 2001; Paya et al., 2003; and Kapetanios et al., 2003)

2. Empirical analysis of deviations from optimal money holdings (Terasvirta and Eliasson, 2001; Sarno et al., 2003)

3. Monetary policy rules (Bec et al., 2000)

4. Symmetric deviations from arbitrage processes such as stock index futures (Monoyios and Sarno, 2002)
Consider ESTAR(2) model as follows,

\[ y_t = \theta_{10} + \theta_{11} y_{t-1} + \theta_{12} y_{t-2} + (\theta_{20} + \theta_{21} y_{t-1} + \theta_{22} y_{t-2}) F(y_{t-1}, \theta) + \varepsilon_t \]

where

\[ \varepsilon_t \sim \text{iid } N(0, \sigma^2_\varepsilon) \]

and

\[ F(y_{t-1}, \theta) = 1 - \exp(-\theta y_{t-1}^2), \quad \theta > 0 \]

0 \leq F(y_{t-1}, \theta) \leq 1 and it has symmetrical U–shaped around zero
We assume

1. $y_t$ is a mean zero stochastic process

2. $\theta_{10} = \theta_{20} = 0$ as we consider a unit root test without a drift

Rearrange $y_t$ as follow,

$$y_t = (\theta_{11} + \theta_{12})y_{t-1} - \theta_{12}(y_{t-1} - y_{t-2}) +$$

$$[(\theta_{21} + \theta_{22})y_{t-1} - \theta_{22}(y_{t-1} - y_{t-2})] F(y_{t-1}, \theta) + \varepsilon_t$$  \quad (2)

To test the null hypothesis $H_0 : \theta = 0$, we use a first-order Taylor approximation of $F(y_{t-1}, \theta)$ around $\theta = 0$ (Luukkonen et al., 1988)

$$F(y_{t-1}, \theta) \approx \theta y_{t-1}^2 + R_1(y_{t-1}, \theta)$$  \quad (3)
Substitute (3) into (2) and then subtracting the both sides by $y_{t-1}$ will become

$$\Delta y_t = (\theta_{11} + \theta_{12} - 1)y_{t-1} - \theta_{12}\Delta y_{t-1} +$$

$$(\theta_{21} + \theta_{22})\theta y_{t-1}^3 - \theta_{22}\theta y_{t-1}^2 \Delta y_{t-1} + \varepsilon^*_t$$

(4)

where

$$\varepsilon^*_t = \varepsilon_t + R_1(y_{t-1}, \theta)[(\theta_{21} + \theta_{22})y_{t-1} - \theta_{22}\Delta y_{t-1}]$$

and $R_1$ is the remainder
Imposing $\theta_{11} + \theta_{12} - 1 = 0$ in (4), $y_t$ will follow a unit root process in the middle regime (see e.g. Balke and Fomby, 1997, Michael et al., 1997 and Kapetanios et al., 2003). Thus,

$$\Delta y_t = \delta_1 \Delta y_{t-1} + \delta_2 y^3_{t-1} + \delta_3 y^2_{t-1} \Delta y_{t-1} + \epsilon^*_t$$  \hspace{1cm} (5)

where $\delta_1 = -\theta_{12}$, $\delta_2 = (\theta_{21} + \theta_{22})\theta$, $\delta_3 = -\theta_{22}\theta$

$H_0 : \delta_2 = \delta_3 = 0$

$H_1 : \delta_2 \neq 0$ or/and $\delta_3 \neq 0$

Under $H_1$, $y_t$ follows a nonlinear but globally stationary process provided that the combination of parameters $\delta_1, \delta_2$ and $\delta_3$ fulfil conditions for stationarity
Sufficient stationarity conditions for ESTAR(2) models:

\((\theta_{21} + \theta_{22}) < 0\)  \\
\(\theta_{11} \geq 0\)  \\
\(0 \leq (\theta_{22} - \theta_{21}) < 2\theta_{11}\) or \((\theta_{22} - \theta_{21}) \leq 0\)  \\
\(-1 < \theta_{12} < 1\)  \\
\(-1 - \theta_{12} < \theta_{22} \leq 0\) or \(0 \leq \theta_{22} < 1 - \theta_{12}\)  \\
(6)
Under $H_0$,

$$\Delta y_t = \delta_1 \Delta y_{t-1} + \delta_2 y_{t-1}^3 + \delta_3 y_{t-1}^2 \Delta y_{t-1} + \varepsilon^*_t$$

becomes

$$\Delta y_t = \delta_1 \Delta y_{t-1} + \varepsilon_t$$

$$= (1 - \delta_1 L)^{-1} \varepsilon_t = v_t$$ (7)

where $L$ is the lag operator, i.e. $Ly_t = y_{t-1}$. We assume that the sequence $\{v_t\}$ satisfies the following assumption.

**Assumption 1** For some $p > \beta > 2$, $\{v_t\}$ is a zero mean $m \times 1$ vector, strong mixing sequence with mixing coefficient $\alpha_k$ of size $-p\beta/(p - \beta)$ and $\sup_{t \geq 1} \|v_t\|_p = C < \infty$. In addition, $(1/T)E(V_T V'_T) \rightarrow \Pi < \infty$ where $V_T = \sum_{t=1}^{T} v_t$. 
Limit results and asymptotic test

Rearrange (5) in matrices as follow,

\[
\begin{pmatrix}
\Delta y_1 \\
\Delta y_2 \\
\vdots \\
\Delta y_T
\end{pmatrix}
= 
\begin{pmatrix}
\Delta y_0 & y_0^3 & y_0^2 \Delta y_0 \\
\Delta y_1 & y_1^3 & y_1^2 \Delta y_1 \\
\vdots & \vdots & \vdots \\
\Delta y_{T-1} & y_{T-1}^3 & y_{T-1}^2 \Delta y_{T-1}
\end{pmatrix}
\begin{pmatrix}
\delta_1 \\
\delta_2 \\
\delta_3 \\
\vdots
\end{pmatrix}
+ 
\begin{pmatrix}
\varepsilon_1^* \\
\varepsilon_2^* \\
\vdots \\
\varepsilon_T^*
\end{pmatrix}
\] (8)

Using ordinary least square (OLS), the estimator of \( \delta \) is

\[
\hat{\delta} = (X'X)^{-1} X'Y
\] (9)
The null hypothesis, $H_0 : \delta_2 = \delta_3 = 0$, has the alternative representation $H_0 : R\delta = r$ where

$$R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \delta = (\delta_1, \delta_2, \delta_3)', \quad r = \begin{bmatrix} 0 \\ 0 \end{bmatrix}'$$

A $F_{nl}$ test statistics $H_0 : \delta_2 = \delta_3 = 0$ vs $H_1 : (\delta_2, \delta_3) \neq (0, 0)$ is then defined in the usual way as

$$F_{nl} = \frac{1}{k} (\hat{\delta} - \delta)' (RY)' \left[ \hat{\sigma}^2 RY (X'X)^{-1} YR' \right]^{-1} RY (\hat{\delta} - \delta)$$

where $k = 2$ and $Y = \text{diag} \left( T^{1/2}, T^2, T^{3/2} \right)$
For large sample, under $H_0$,

$$F_{nl} \Rightarrow \frac{1}{2} \left[ \frac{(B_2(1) \int_0^1 B_{1,2}^2(s)ds - \int_0^1 B_{1,2}^2(s)dB_2(s))^2}{\int_0^1 B_{1,2}^4(s)ds - (\int_0^1 B_{1,2}^2(s)ds)^2} + \frac{(\int_0^1 B_{1,1}^3(s)dB_1(s))^2}{\int_0^1 B_{1,1}^6(s)ds} \right]$$ (11)

as $T \to \infty$ where $B_1$ and $B_2$ are two independent standard Brownian motions.
Table 1: Critical values for the $F_{nl}$ test statistics with $\delta_1 = 0$ based on 100,000 replications

<table>
<thead>
<tr>
<th>T</th>
<th>0.1</th>
<th>0.05</th>
<th>0.025</th>
<th>0.01</th>
<th>0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>2.684264</td>
<td>3.433036</td>
<td>4.236435</td>
<td>5.295368</td>
<td>8.02343</td>
</tr>
<tr>
<td>100</td>
<td>2.680106</td>
<td>3.403369</td>
<td>4.117247</td>
<td>5.139271</td>
<td>7.510085</td>
</tr>
<tr>
<td>250</td>
<td>2.708653</td>
<td>3.424705</td>
<td>4.139673</td>
<td>6.096821</td>
<td>7.375901</td>
</tr>
<tr>
<td>500</td>
<td>2.731896</td>
<td>3.439074</td>
<td>4.156963</td>
<td>5.108723</td>
<td>7.435359</td>
</tr>
<tr>
<td>10000</td>
<td>2.765274</td>
<td>3.494263</td>
<td>4.188189</td>
<td>5.138637</td>
<td>7.739216</td>
</tr>
</tbody>
</table>
Power simulation

ADF unit root test

$H_0$: unit root $\quad H_1$: no unit root

linear stationary

non-linear unit root test

$H_0$: linear unit root $\quad H_1$: nonlinear stationary
Table 2: The percentage of accepting globally nonlinear stationary by $F_{nl}$ test when it is concluded that the data (T=50) has a unit root by the ADF test using 5% significance level. The number in the brackets denote the number of series from 10000 replications accepting a unit root by the ADF test.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>(1.2,-0.2)</th>
<th>(1.2,-0.2)</th>
<th>(1.2,-0.2)</th>
<th>(1.0)</th>
<th>(1.2,-0.2)</th>
<th>(0.8,0.2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-1.4,-0.4)</td>
<td>(-0.4,-0.4)</td>
<td>(-1.4,0.4)</td>
<td>(-0.4,-0.4)</td>
<td>(-0.5,0.4)</td>
<td>(-0.5,0.4)</td>
</tr>
<tr>
<td>0.001</td>
<td>8.01(9043)</td>
<td>5.19(9328)</td>
<td>5.81(9270)</td>
<td>4.30(9431)</td>
<td>3.39(9480)</td>
<td>3.49(9604)</td>
</tr>
<tr>
<td>0.005</td>
<td>26.46(6180)</td>
<td>13.63(8229)</td>
<td>17.19(7901)</td>
<td>9.06(9006)</td>
<td>4.96(9381)</td>
<td>4.07(9469)</td>
</tr>
<tr>
<td>0.01</td>
<td>40.81(3373)</td>
<td>23.38(6625)</td>
<td>28.27(6105)</td>
<td>14.98(8232)</td>
<td>6.34(9269)</td>
<td>4.82(9421)</td>
</tr>
<tr>
<td>0.05</td>
<td>63.72(113)</td>
<td>48.64(1137)</td>
<td>54.62(866)</td>
<td>41.45(3089)</td>
<td>13.54(8410)</td>
<td>9.99(9017)</td>
</tr>
<tr>
<td>0.1</td>
<td>62.5(8)</td>
<td>54.42(226)</td>
<td>54.86(144)</td>
<td>48.72(938)</td>
<td>16.99(7753)</td>
<td>13.62(8510)</td>
</tr>
<tr>
<td>0.15</td>
<td>NA(0)</td>
<td>61.11(72)</td>
<td>65.71(35)</td>
<td>50.53(378)</td>
<td>17.53(7297)</td>
<td>15.38(8090)</td>
</tr>
<tr>
<td>0.5</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>35.71(14)</td>
<td>14.17(6430)</td>
<td>15.23(6890)</td>
</tr>
<tr>
<td>1</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>10.99(6390)</td>
<td>11.97(6642)</td>
</tr>
<tr>
<td>2</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>9.34(6539)</td>
<td>9.30(6730)</td>
</tr>
<tr>
<td>50</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>7.33(7303)</td>
<td>6.61(7963)</td>
</tr>
<tr>
<td>100</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>NA(0)</td>
<td>7.35(7385)</td>
<td>6.58(7963)</td>
</tr>
</tbody>
</table>
Empirical example

Figure 1: AUS$ 1 to HK$
Figure 2: Correlogram
Table 3: The ADF unit root test

Unit-root tests 12 to 304
Critical values: 5%=-1.941, 1%=-2.573

<table>
<thead>
<tr>
<th></th>
<th>t-adf</th>
<th>beta Y_1</th>
<th>sigma</th>
<th>lag</th>
<th>t-DY_lag</th>
<th>t-prob</th>
<th>F-prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>HK</td>
<td>-0.75989</td>
<td>0.99854</td>
<td>0.18457</td>
<td>10</td>
<td>0.68092</td>
<td>0.4965</td>
<td>0.4965</td>
</tr>
<tr>
<td>HK</td>
<td>-0.74367</td>
<td>0.99857</td>
<td>0.18439</td>
<td>9</td>
<td>0.057017</td>
<td>0.9546</td>
<td>0.7920</td>
</tr>
<tr>
<td>HK</td>
<td>-0.74387</td>
<td>0.99857</td>
<td>0.18407</td>
<td>8</td>
<td>0.67293</td>
<td>0.5015</td>
<td>0.8212</td>
</tr>
<tr>
<td>HK</td>
<td>-0.73364</td>
<td>0.99859</td>
<td>0.18389</td>
<td>7</td>
<td>0.92285</td>
<td>0.3569</td>
<td>0.8795</td>
</tr>
<tr>
<td>HK</td>
<td>-0.72958</td>
<td>0.99860</td>
<td>0.18385</td>
<td>6</td>
<td>-0.086123</td>
<td>0.9314</td>
<td>0.7792</td>
</tr>
<tr>
<td>HK</td>
<td>-0.73062</td>
<td>0.99860</td>
<td>0.18353</td>
<td>5</td>
<td>-1.1650</td>
<td>0.2450</td>
<td>0.8795</td>
</tr>
<tr>
<td>HK</td>
<td>-0.71296</td>
<td>0.99863</td>
<td>0.18364</td>
<td>4</td>
<td>-1.5933</td>
<td>0.1122</td>
<td>0.7940</td>
</tr>
<tr>
<td>HK</td>
<td>-0.70028</td>
<td>0.99866</td>
<td>0.18413</td>
<td>3</td>
<td>1.2102</td>
<td>0.2272</td>
<td>0.5849</td>
</tr>
<tr>
<td>HK</td>
<td>-0.69851</td>
<td>0.99866</td>
<td>0.18428</td>
<td>2</td>
<td>0.49163</td>
<td>0.6234</td>
<td>0.5292</td>
</tr>
<tr>
<td>HK</td>
<td>-0.70468</td>
<td>0.99865</td>
<td>0.18404</td>
<td>1</td>
<td>1.9336</td>
<td>0.0541</td>
<td>0.6039</td>
</tr>
<tr>
<td>HK</td>
<td>-0.72486</td>
<td>0.99860</td>
<td>0.18490</td>
<td>0</td>
<td></td>
<td></td>
<td>0.3588</td>
</tr>
</tbody>
</table>

Accept a unit root
\[ \Delta y_t = \delta_1 \Delta y_{t-1} + \delta_2 y_{t-1}^3 + \delta_3 y_{t-1}^2 \Delta y_{t-1} + \varepsilon_t^* \]

\begin{align*}
0.14623 & -0.019242 & -0.022724 \\
(2.016) & (-2.646) & (0.072506)
\end{align*}

\[ R^2 = 0.0398075 \quad \text{sigma} = 0.181605 \]

\[ \text{DW} = 2.01 \quad \text{RSS} = 9.828111 \]

\[ y_t = HK_t - \text{mean} (HK_t) \]

\[ F_{nl} \text{ test statistics} = 4.096 > 3.4247 \text{ critical value from Table 1} \]

Conclude in fact \( HK \) is a nonlinear ESTAR(2) model
Conclusion

1. We have extended ESTAR(1) models to ESTAR(2) models
2. Using our $F_{nl}$ test, we have more chance to correctly identify the alternative nonlinear stationary process when the ADF test concludes that the data has a unit root
3. The empirical example of the exchange rate of Australian dollar to Hongkong dollar shows that the method presented in this paper can be applied to empirical data