Threshold Autoregression with a Near Unit Root[1]

Mehmet Caner
Bilkent University[2]

and

Bruce E. Hansen
University of Wisconsin[3]

January 2000

Technical Appendices

[1]Caner thanks TUBITAK and Hansen thanks the National Science Foundation and the Alfred P. Sloan Foundation for research support. We thank Frank Diebold, Peter Pedroni, Pierre Perron, Simon Potter, four referees and the co-editor Alain Monfort for stimulating comments on an earlier draft.

[2]Department of Economics, Bilkent University, Bilkent, Ankara, Turkey, 06533. mcaner@bilkent.edu.tr

[3]Department of Economics, University of Wisconsin, Madison, WI 53706. bhansen@ssc.wisc.edu.

Appendix A

This appendix establishes the rate of convergence of the threshold estimate in the TAR model of Caner and Hansen (1998) (henceforth, CH).

To simplify the notation, we use the TAR(1) model, which is their equation (1) with $x_{t-1} = (y_{t-1}\ t\ \ 1)$. It is assumed that the threshold variable $U_{t-1} = g(\Delta y_{t-1}, ..., \Delta y_{t-m})$ has been transformed so that it has a marginal uniform distribution. The tests are invariant if we replace x_{t-1} with

$$x_{Tt-1} = \begin{pmatrix} \frac{1}{\sqrt{T}} y_{t-1} \\ \frac{t}{T} \\ 1 \end{pmatrix},$$

and it will be convenient to make this switch. The model is hence

$$\Delta y_t = \theta_1' x_{Tt-1} 1_{\{U_{t-1}<u\}} + \theta_2' x_{Tt-1} 1_{\{U_{t-1}\geq u\}} + e_t. \tag{1}$$

Partitioning $\theta_1 = (\rho_1\ \beta_1\ \mu_1)$, and similarly θ_2, we assume $\rho_1 = 0$ and $\rho_2 = 0$ to impose the unit root, and impose the maintained assumption $\beta_1 = \beta_2$. We assume $\delta = \mu_2 - \mu_1 \neq 0$, so the threshold u is identified. Let $u_0 \neq 0$ denote the true value of the threshold. The test statistic is invariant to $\sigma^2 = Ee_t^2$, so we normalize $\sigma^2 = 1$.

The model is estimated by least squares. Let $\hat{\theta}_1(u)$ and $\hat{\theta}_2(u)$ denote the conditional OLS estimates given a value of u, let $S_T(u)$ be the residual sum of squared errors given u, let $Q_T(u) = S_T(u_0) - S_T(u)$, and let

$$\hat{u} = \underset{\pi_1 \leq u \leq \pi_2}{\operatorname{argmax}} Q_T(u)$$

be the LS estimate of u. Let $\hat{\theta}_1 = \hat{\theta}_1(\hat{u})$ and $\hat{\theta}_2 = \hat{\theta}_2(\hat{u})$ be the LS estimate of θ_1 and θ_2.

To establish the rate of convergence of $\hat{u}$, we first show that $\hat{u}$ is consistent.

Lemma 1 $\hat{u} \to_p u_0$.

Proof:

Let $x_{Tt-1}(u) = x_{Tt-1} 1_{\{U_{t-1}<u\}}$. Let e be the $T \times 1$ vector of the stacked e_t, X be the $T \times 2$ matrix of the stacked x_{Tt-1}, and similarly define X_u. Let $X_0 = X_{u_0}$ and let Δ_u be stacked $1_{\{u_0<U_{t-1}<u\}}$. Standard linear algebra can show that

$$Q_T(u) = e' (P_u - P_0) e + 2\delta\iota' (I - P_u) e - \delta^2 \left[\Delta_u' \Delta_u - \Delta_u' X_u^* (X_u^{*\prime} X_u^*)^{-1} X_u^{*\prime} \Delta_u\right]$$

where $P_u = X_u^* \left(X_u^{*\prime} X_u^*\right)^{-1} X_u^{*\prime}$ and $X_u^* = (X \quad X_u)$. Since $x_{T[Tr]} \Rightarrow (W(r) \quad r \quad 1)' \equiv X(r)$, say, CH, Theorem 3 yields

$$T^{-1} X_u^{*\prime} X_u^* = \begin{bmatrix} T^{-1}X'X & T^{-1}X'X_u \\ T^{-1}X_u'X & T^{-1}X_u'X_u \end{bmatrix} \Rightarrow \begin{bmatrix} \int_0^1 XX' & u\int_0^1 XX' \\ u\int_0^1 XX' & u\int_0^1 XX' \end{bmatrix},$$

and CH, Theorem 2, yields

$$T^{-1/2} X_u^{*\prime} e = \begin{bmatrix} T^{-1/2}X'e \\ T^{-1/2}X_u'e \end{bmatrix} \Rightarrow \begin{bmatrix} \int_0^1 X(s)dW(s,1) \\ \int_0^1 X(s)dW(s,u) \end{bmatrix}.$$

Using these results, some tedious calculations show that for $u_0 \le u \le \bar{u}$,

$$T^{-1} Q_T(u) \Rightarrow -\delta^2 \left|u - u_0\right| \left[1 - \frac{|u - u_0|}{u}\right].$$

Note that the right-hand-side is monotonically decreasing in u. Similarly, for $\underline{u} \le u \le u_0$, the probability limit is monotonically increasing in u. Thus the maximizer of the limit function is u_0. Since the convergence is uniform it follows that $\hat{u} \rightarrow_p u_0$ as desired. □

We next show that the regression estimates from model (1) are consistent.

Lemma 2 $\left(\hat{\theta}_1 - \theta_1\right) = o_p(1)$ *and* $\left(\hat{\theta}_2 - \theta_2\right) = o_p(1)$.

Proof: We show the result for $\hat{\theta}_1$. For $u < u_0$,

$$\left(\hat{\theta}_1(u) - \theta_1\right) = \left(T^{-1}X_u'X_u\right)^{-1}\left(T^{-1}X_u'e\right) \Rightarrow 0.$$

For $u > u_0$,

$$\begin{aligned} \left(\hat{\theta}_1(u) - \theta_1\right) &= \left(T^{-1}X_u'X_u\right)^{-1}\left(T^{-1}X_u'e - T^{-1}X_u'\Delta_u\delta\right) \\ &\Rightarrow -\left(u\int_0^1 XX'\right)^{-1}\int_0^1 X\left|u - u_0\right|\delta. \end{aligned}$$

Since the right-hand-side is continuous in u, equals zero at u_0, and $\hat{u} \rightarrow_p u_0$, it follows that $\left(\hat{\theta}_1(u) - \theta_1\right) \Rightarrow 0$ as desired. □

We now can establish the rate of convergence of the estimator $\hat{u}$.

Theorem 1 $T(\hat{u} - u_0) = O_p(1)$.

Proof: We appeal to the proof of Proposition 1 of Chan (1993). Chan showed this result for a stationary TAR. The primary difference in our case is the presence of the non-stationary variables x_{Tt}. Chan showed that under the consistency results of Lemmas 1 and 2 above, $T(\hat{u} - u_0) = O_p(1)$ follows from his three inequalities (4.4a)-(4.4c). The first two are identical in our setting, implying that the Theorem holds if we verify an analog of Chan's (4.4c). Chan shows that this holds if for some $\Delta > 0$, there is some $H < \infty$ such that for all $u_0 \leq u_1 < u_2 \leq u_0 + \Delta$,

$$E\left|x_{Tt-1}e_t 1_{\{u_1 < U_{t-1} \leq u_2\}}\right| \leq H(u_2 - u_1) \tag{2}$$

and

$$E\left(\left|x_{Tt-1}e_t\right|^2 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) \leq H(u_2 - u_1). \tag{3}$$

The arguments for (2) and (3) are similar. We verify (2). Since e_t is iid and $|x_{Tt-1}| \leq 2 + \left|T^{-1/2}y_{t-1}\right|$,

$$\begin{aligned} E\left|x_{Tt-1}e_t 1_{\{u_1 < U_{t-1} \leq u_2\}}\right| &= E\left|e_t\right| E\left(\left|x_{Tt-1}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) \\ &\leq E\left|e_t\right| E\left(\left(2 + \left|T^{-1/2}y_{t-1}\right|\right) 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) \\ &= 2E\left|e_t\right|(u_2 - u_1) + E\left|e_t\right| T^{-1/2} E\left(\left|y_{t-1}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right), \end{aligned}$$

we simply need to to show that

$$T^{-1/2}E\left(\left|y_{t-1}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) \leq H(u_2 - u_1). \tag{4}$$

Since y_t is a random walk and $E(\Delta y_t) = \sigma^2 < \infty$, it follows that $T^{-1/2}E\left|y_{t-m}\right| \leq B$ for some $B < \infty$. Since y_{t-m} is independent of $U_{t-1} = g(\Delta y_{t-1}, ..., \Delta y_{t-m})$, it follows that

$$T^{-1/2}E\left(\left|y_{t-m}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) = T^{-1/2}E\left|y_{t-m}\right| E1_{\{u_1 < U_{t-1} \leq u_2\}} \leq B(u_2 - u_1).$$

Next, note that $E\left(\left|e_{t-k}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) \leq G_k(u_2 - u_1)$, where

$$G_k = \sup_{u_0 \leq u \leq u_0 + \Delta} \frac{d}{du} E\left(\left|e_{t-k}\right| 1_{\{U_{t-1} \leq u\}}\right) < \infty$$

(the finiteness follows from the assumption that e_t and U_t have continuous distributions). Hence

$$\begin{aligned} T^{-1/2}E\left(\left|y_{t-1}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) &\leq T^{-1/2}E\left(\left|y_{t-m}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) + T^{-1/2}E\left|y_{t-1} - y_{t-m}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}} \\ &\leq B(u_2 - u_1) + \sum_{k=1}^{m} E\left(\left|e_{t-k}\right| 1_{\{u_1 < U_{t-1} \leq u_2\}}\right) \\ &\leq \left[B + \sum_{k=1}^{m} G_k\right](u_2 - u_1), \end{aligned}$$

as needed. We have established (4) and thus (2). □

Appendix B

In this Appendix, we derive equations (16) and (A.24) from the paper.

Let

$$W(s) = \begin{pmatrix} W_1(s) \\ W_2(s) \\ W_3(s) \end{pmatrix}$$

be a vector Brownian motion with covariance matrix

$$E\left[W(1)W(1)'\right] = \begin{pmatrix} 1 & 0 & \delta_1 \\ 0 & 1 & \delta_2 \\ \delta_1 & \delta_2 & 1 \end{pmatrix}$$

(equation (A.21) in the paper). Define

$$\begin{aligned} a_1 &= \left(1-\delta_1^2\right)^{-1/2} \\ a_2 &= \left(1-\delta_2^2\right)^{-1/2} \end{aligned}$$

and let

$$W_{1\cdot 3}(s) = a_1 W_1(s) - a_1\delta_1 W_3(s) \tag{5}$$

and

$$W_{2\cdot 3}(s) = a_2 W_2(s) - a_2\delta_2 W_3(s). \tag{6}$$

Then

$$\begin{pmatrix} W_{1\cdot 3}(s) \\ W_{2\cdot 3}(s) \\ W_3(s) \end{pmatrix} = \begin{pmatrix} a_1 & 0 & -a_1\delta_1 \\ 0 & a_2 & -a_2\delta_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} W_1(s) \\ W_2(s) \\ W_3(s) \end{pmatrix}$$

is a vector Brownian motion with covariance matrix

$$\begin{aligned} & \begin{pmatrix} a_1 & 0 & -a_1\delta_1 \\ 0 & a_2 & -a_2\delta_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \delta_1 \\ 0 & 1 & \delta_2 \\ \delta_1 & \delta_2 & 1 \end{pmatrix} \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ -a_1\delta_1 & -a_2\delta_2 & 1 \end{pmatrix} \\ = & \begin{pmatrix} 1 & -a_1a_2\delta_1\delta_2 & 0 \\ -a_1a_2\delta_1\delta_2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ = & \begin{pmatrix} 1 & \sigma_{21} & 0 \\ \sigma_{21} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned}$$

where

$$\sigma_{21} = -a_1 a_2 \delta_1 \delta_2 = -\frac{\delta_1 \delta_2}{\sqrt{(1-\delta_1^2)(1-\delta_2^2)}},$$

which is the definition given in equation (16) of the paper.

Thus the pair

$$\begin{pmatrix} W_{1\cdot 3}(s) \\ W_{2\cdot 3}(s) \end{pmatrix}$$

is a vector Brownian motion with covariance matrix $\begin{pmatrix} 1 & \sigma_{21} \\ \sigma_{21} & 1 \end{pmatrix}$, and is independent of $W_3(s)$. Rewriting (5) and (6), we have the decompositions

$$\begin{aligned} W_1(s) &= a_1^{-1} W_{1\cdot 3}(s) + \delta_1 W_3(s) \\ &= \left(1-\delta_1^2\right)^{1/2} W_{1\cdot 3}(s) + \delta_1 W_3(s) \end{aligned}$$

and

$$\begin{aligned} W_2(s) &= a_2^{-1} W_{2\cdot 3}(s) + \delta_2 W_3(s) \\ &= \left(1-\delta_2^2\right)^{1/2} W_{2\cdot 3}(s) + \delta_2 W_3(s), \end{aligned}$$

which together are equation (A.24) in the paper.