# Threshold Stochastic Unit Root Models (TARSUR)

*The Case of Stock Prices*

**Jesús Gonzalo and Raquel Montesinos**
**Dept of Economics**
**Universidad Carlos III de Madrid**

**http://halweb.uc3m.es/jgonzalo**

***UC Riverside 2004***

# 0. Outline

- Introduction
  - Introduction to TAR models
  - Review of Unit Root Models
  - OUR contribution
- Definition and Properties of TARSUR models
- Estimating and Testing TARSUR models
- Finite Sample Performance
- Applications:
  - US Stock Prices
  - International Interest Rates
- Conclusions

# I. Introduction to TAR Models

$$Y_t = \begin{cases} \phi_{10} + \phi_{11} Y_{t-1} + \sigma_1 \epsilon_t & \text{if } Y_{t-1} \leq r \\ \phi_{20} + \phi_{21} Y_{t-1} + \sigma_2 \epsilon_t & \text{if } Y_{t-1} > r \end{cases}$$

- References:
  Tong (1990), Chan (1990, 1993), Hansen (1996, 2000)

- Maintained Assumption: $Y_t$ stationary ergodic, finite second moments and its density is positive everywhere.

# Review of Unit Root Models

- *Unit Roots:*

  Nelson and Plosser (1982); Phillips and Xiao (1998) (survey)

- *Extensions:*

  – *Fractional unit roots:* $(1-L)^d Y_t = \varepsilon_t$

  Granger & Joujeux(1980), Robinson (1994), Beran (1994), Dolado, Gonzalo & Mayoral (2002).

  – *Stochastic unit roots (STUR):*

  $$Y_t = \rho_t Y_{t-1} + \varepsilon_t$$
  $$\text{with } E(\rho_t) = 1$$

  Leybourne, McCabe & Tremayne (1996), and Granger & Swanson (1997).

  – *Threshold unit roots (TARUR):*

  $$Y_t = \rho_1 I(Z_{t-d} < r) Y_{t-1} + \rho_2 I(Z_{t-d} > r) Y_{t-1} + \varepsilon_t$$

  with $\rho_1 < 1, \rho_2 = 1$

  González & Gonzalo (1998); Caner & Hansen (2001)

## Our Contribution

$\cdot$*Threshold stochastic unit roots (TARSUR):*

$$Y_t = \rho_1 I(Z_{t-d} < r)Y_{t-1} + \rho_2 I(Z_{t-d} > r)Y_{t-1} + \varepsilon_t$$

$$Y_t = \delta_t Y_{t-1} + \varepsilon_t,$$

with $\delta_t = \rho_1 I(Z_{t-d} < r) + \rho_2 I(Z_{t-d} > r)$,
and $E(\delta_t) = 1$, $V(\delta_t) > 0$.

## Advantages

1. Threshold variable is suggested by Economic Theory: we can find an explanation or cause for the existence of unit roots.

2. Its computational simplicity: parameter estimation is done by least squares.

3. A simple t-statistic is used to test the hypothesis of exact unit root versus stochastic unit root.

4. We are able to introduce deterministic terms with threshold effects. Therefore we could potentially forecast the changes in those deterministic elements.

5. Threshold models are easier to use for forecasting than random coefficient models: Assuming we can forecast the different regimes.

# II. Definition of the TARSUR model

$$Y_t = [\rho_1 I(Z_{t-d} \leq r_1) + \cdots + \rho_n I(Z_{t-d} > r_{n-1})] Y_{t-1} + \varepsilon_t$$

$$= \delta_{\mathbf{t}} \mathbf{Y}_{\mathbf{t-1}} + \varepsilon_{\mathbf{t}}, \quad (1)$$

where $t = 1, \cdots, T$, $I(\cdot)$ is an indicator function, and $\varepsilon_t$ is an innovation term. $Z_{t-d}$ is the threshold variable, $d$ the delay parameter, and $r_1 < r_2 < \cdots < r_{n-1}$ are the threshold values.

**Definition 1:** A *TARSUR* process is defined by equation (1) with $E(\delta_t) = \sum_{i=1}^{n} \rho_i p_i = 1$, where $p_i$ is the probability of $Z_{t-d}$ being in regime $i$, and $V(\delta_t) > 0$.

## Assumptions

**(A.1)** $\{\varepsilon_t, Z_t\}$ is strictly stationary, ergodic, adapted to the *sigma-field* $\Im_t \stackrel{def}{=} \{(\varepsilon_j, Z_j), j \leq t\}$.

**(A.2)** $\{\varepsilon_t, Z_t\}$ is strong mixing with mixing coefficients $\alpha_m$ satisfying $\sum_{m=1}^{\infty} \alpha_m^{1/2-1/r} < \infty$ for some $r > 2$.

**(A.3)** $\varepsilon_t$ is independent of $\Im_{t-1}$, $E(\varepsilon_t) = 0$ and $E\left|\varepsilon_t\right|^4 = k < \infty$.

**(A.4)** $Z_t$ has a continuous marginal distribution.

**(A.5)** $E(\max(0, \log \varepsilon_1)) < \infty$.

**(A.6)** $ess.\sup|\varepsilon_1| < \infty$[a].

[a] The essential supremum of X is $ess\ sup\ X = \inf\{x : P(|X| > x) = 0\} = ||x||_{\infty}$.

# Properties of the TARSUR model

$$\mathbf{Y_t} = \delta_\mathbf{t} \mathbf{Y_{t-1}} + \varepsilon_\mathbf{t},$$

$$\begin{aligned} Y_t(Y_0) &= \varepsilon_t + \sum_{j=1}^{n-1} \left( \prod_{i=0}^{j-1} \delta_{t-i} \right) \varepsilon_{t-j} + \left( \prod_{i=0}^{n-1} \delta_{t-i} \right) Y_0 \\ &= C_{1,t}(n) + C_{2,t}(n) \end{aligned}$$

**(a)** If $C_{1,t}(n)$ converges in $L^p$ for $p \in [0, \infty]$, then $C_{1,t} = \varepsilon_t + \sum_{j=0}^{\infty} \left( \prod_{i=0}^{j-1} \rho_{t-i} \right) \varepsilon_{t-j}$ is a strictly stationary solution.

**(b)** if $C_{2,t}(n)$ converges in probability to zero, then this solution is unique.

**(c)** if $p > 0$ in result (a) then $\{Y_t\}$ has a finite $p$th order moment.

From (a) and (b):

$$Y_t = \sum_{j=0}^{\infty} \psi_{t,j} \varepsilon_{t-j}$$

where $\psi_{t,j} = \prod_{i=0}^{j-1} \delta_{t-i}$.

## Stationarity of the TARSUR model

**Theorem 1** *If the sequence* $\{\varepsilon_t, Z_t\}$ *satisfies assumptions (A.1), (A.5) and*

$$-\infty \leq E \log |\delta_1| < 0 \tag{2}$$

*holds, then process (1) is strictly stationary.*

*Moreover, if (A.6) is also satisfied*

$$\sum_{j=0}^{\infty} \left( E \left| \psi_{t,j} \right|^2 \right)^{\frac{1}{2}} < \infty, \tag{3}$$

*where* $\psi_{t,j} = \prod_{i=0}^{j-1} \delta_{t-i}$, *then process (1) is covariance stationary.*

**Corollary 1** *A TARSUR process with* $\rho_i \geq 0$, $\forall i$, *is strictly stationary.*

## Weak Stationarity of TARSUR model

**Case I**: $Z_t$ is an *i.i.d.* process

**Case II**: $Z_t$ is a *1$^{st}$-order stationary Markov Chain,* and $\{Y_t\}$ has only two regimes.

**Corollary 2** *Consider Case I, then process (1) is covariance stationary if and only if* $E(\delta_t^2) < 1$.

**Corollary 3** *Consider Case II and without loss of generality that* $\rho_1 < \rho_2$. *Denote by* $p_{ji}$ *the probability of being in regime* $i$ *given that* $Z_t$ *has been in regime* $j$ *the previous period. Define the following* $2 \times 2$ *matrix*

$$F_2 = \left\{ \rho_i^2 p_{ji}, (i, j = 1, 2) \right\}.$$

*Then, process (1) is covariance stationary, if the spectral radius of* $F_2$ *is less than one.*

# IRF of a TARSUR model

From the MA($\infty$) representation:

$$\xi_h = E\left(\frac{\delta Y_{t+h}}{\delta \varepsilon_t}\right) = E\left(\psi_{t,h}\right), \quad h = 1, 2, \cdots.$$

**Proposition 1** *Under the conditions of Corollary 3, the IRF is given by*

$$\xi_h = \begin{pmatrix} 1 & 1 \end{pmatrix} F_1^h \begin{pmatrix} \rho_1 p_1 \\ \rho_2 p_2 \end{pmatrix}, \quad h = 1, 2, \cdots.$$

*where* $F_1 = \{\rho_i p_{ji}, (i, j = 1, 2)\}$. *Then Shocks have transitory effects* ($lim_{h\to\infty}\xi_h = 0$) *if and only if the spectral radius of* $F_1$ *is less than one.*

## IRF of a TARSUR model (cont)

For a *TARSUR* process, the following implications hold:

1. If $p_{22} > p_{12} : lim_{h\rightarrow\infty}\xi_h = \infty$, as it happens in an explosive model. $Y_t$ is not covariance stationary.

2. If $p_{22} = p_{12} : \xi_h = 1, \quad \forall h$, as it happens in a random walk model. $Y_t$ is not covariance stationary.

3. If $p_{22} < p_{12} : lim_{h\rightarrow\infty}\xi_h = 0$ as it happens in a stationary model. $Y_t$ could be covariance stationary.

## Differencing a TARSUR model

$$\Delta Y_t = (\delta_t - 1)\, y_{t-1} + \varepsilon_t \qquad (4)$$

Iterating backwards (4)

$$\Delta Y_t = \sum_{j=0}^{\infty} \Psi_{t,j} \varepsilon_{t-j}$$

where $\Psi_{t,0} = 1$ and $\Psi_{t,j} = (\delta_t - 1)\psi_{t-1,j-1},\ j \geq 1$.

**Proposition 2** *Assume that $Y_t$ follows model (1). If $\delta_t$ has a strictly positive variance, $\Delta Y_t$ is strictly (covariance) stationary if and only if $Y_t$ is strictly (covariance) stationary.*

# IRF of a Difference TARSUR model

IRF:

$$\Upsilon_h = E\left(\frac{\delta \Delta Y_{t+h}}{\delta \varepsilon_t}\right) = E\left(\Psi_{t,h}\right)$$

1. If $p_{22} > p_{12}$ : $lim_{h\to\infty}\Upsilon_h = \infty$ . $\Delta Y_t$ is not covariance stationary.

2. If $p_{22} = p_{12}$ : $\Upsilon_h = 0, \forall h \geq 1$. $\Delta Y_t$ is not covariance stationary. Note that in this case $Z_t$ is an *i.i.d.* process.

3. If $p_{22} < p_{12}$ : $lim_{h\to\infty}\Upsilon_h = 0$. $\Delta Y_t$ could be covariance stationary.

# III. Estimation of TAR Models

Parameters $\vartheta = (\Phi_1', \Phi_2', r)'$.

Let $W_t = (1, Y_{t-1})'$ and $I(+, r) = 1(Y_{t-1} > r)$.

$$Y_t = I(-, r)W_t'\Phi_1 + I(+, r)W_t'\Phi_2 + error$$

- First step, minimize $S_n(\Phi_1', \Phi_2', r) =$

$$\sum_{t=2}^{n}(Y_t - I(-, r)W_t'\Phi_1 - I(+, r)W_t'\Phi_2)^2$$

  For a given $r$, the OLS solution is $\hat{\Phi}_{1,n}(r)$ and $\hat{\Phi}_{2,n}(r)$.

- Second step, minimize $S_n(\hat{\Phi}_{1,n}(r), \hat{\Phi}_{2,n}(r), r)$. The solution is

$$\hat{r}_n$$

- The CLS estimator of $\vartheta = (\Phi_1', \Phi_2', r)'$ is

$$\hat{\vartheta}_n = (\hat{\Phi}_{1,n}', \hat{\Phi}_{2,n}', \hat{r}_n)' \equiv (\hat{\Phi}_{1,n}(\hat{r}_n)', \hat{\Phi}_{2,n}(\hat{r}_n)', \hat{r}_n)'$$

# Testing TARSUR Models

**Goal:** To test **exact unit root** versus **threshold stochastic unit root**

**DGP**:

$$\begin{aligned} Y_t &= (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} \leq r)) + \\ &\quad + (\rho_1 I(Z_{t-d} \leq r) + \rho_2 I(Z_{t-d} > r))\, Y_{t-1} + \varepsilon_t, \end{aligned}$$

$$\begin{aligned} \Delta Y_t &= (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \\ &\quad + ((\rho_1 - \rho_2) I(Z_{t-d} \leq r) + (\rho_2 - 1))\, Y_{t-1} + \varepsilon_t. \end{aligned} \tag{5}$$

Substituting the maintained hypothesis into (5)

$$E(\rho_t) = \rho_1 p(r) + \rho_2 (1 - p(r)) = 1,$$

$$(\boldsymbol{\rho_2 - 1 = -(\rho_1 - \rho_2) p(r)}).$$

$$\begin{aligned} \Delta Y_t &= (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \\ &\quad + \gamma \mathbf{U_t(r) Y_{t-1}} + \varepsilon_\mathbf{t}, \end{aligned} \tag{6}$$

where $\gamma = (\boldsymbol{\rho_1} - \boldsymbol{\rho_2})$, $U_t(r) = I(Z_{t-d} \leq r) - p(r)$.

# Testing TARSUR Models (cont)

**Exact unit root** versus **threshold stochastic unit root**

$$\mathbf{H_0}: \begin{array}{l} E(\delta_t) = 1 \\ V(\delta_t) = 0 \end{array} \Longleftrightarrow \rho_1 = \rho_2 = 1 \Leftrightarrow \gamma = \mathbf{0}$$

$$\mathbf{H_1}: \begin{array}{l} E(\delta_t) = 1 \\ V(\delta_t) > 0 \end{array} \Longleftrightarrow \rho_1 \neq \rho_2 \Leftrightarrow \gamma \neq \mathbf{0}$$

$$E(\rho_t) = \gamma p(r) + \rho_2 = 1$$

$$V(\rho_t) = \gamma^2 p(r)(1 - p(r))$$

- Regression model:

$$\begin{aligned} \Delta Y_t \quad = \quad & (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \quad (7) \\ & + (\beta_1 t I(Z_{t-d} \leq r) + \beta_2 t I(Z_{t-d} > r)) \\ & + \gamma \mathbf{U_t(r) Y_{t-1}} + \varepsilon_{\mathbf{t}}. \end{aligned}$$

## Asymptotic Distribution

It depends on whether

1. threshold value is known, or

2. threshold value is unknown:
   (a) threshold value is identified under the null: DGP has $\mu_1 \neq \mu_2$,

   (b) threshold value is unidentified under the null: DGP has $\mu_1 = \mu_2$.

- **Threshold value is known**:

  Model (7) is estimated by least squares.

**Proposition 2** *Suppose that the threshold value is known and that assumptions (A.1), (A.2) (A.3) and (A.4) hold. Under the null of no threshold the* $t_{\gamma=0}$ *statistic has the following asymptotic distribution*

$$t_{\gamma=0}(r) \Rightarrow N(0, 1).$$

## Asymptotic Distribution (cont)

**Threshold value is unknown**:

Model (7) is estimated by **sequential least squares:** we will assume that this parameter lies in a bounded interval $R^*$

$$\widehat{r} = \arg_{r \in R*} \min \widehat{\sigma}^2(r)$$

**(1) Threshold value is unknown but identified**: under the null $\gamma = 0$ ($\rho_1 = \rho_2 = 1$) but $\mu_1 \neq \mu_2$.

**Proposition 3** *Suppose that assumptions (A.1), (A.2) (A.3) and (A.4) hold. Under* $H_0 : \gamma = 0$, $\mu_1 \neq \mu_2$. *Then the* $t_{\gamma=0}$ *statistic in regression (7) has the following asymptotic distribution*

$$t_{\gamma=0}(\widehat{r}) \Rightarrow N(0, 1).$$

## III. Asymptotic Distribution (cont)

**(2) Threshold value is unknown and unidentified**: under the null $\gamma = 0$ $(\rho_1 = \rho_2 = 1)$ and $\mu_1 = \mu_2$.

$$\widehat{r} = \arg_{r \in R*} \min \widehat{\sigma}^2(r) = \arg_{r \in R^*} \sup W_T(r),$$

where $W_T(r) = t^2_{\gamma=0}(r)$ is the test statistic of the null of no threshold.

The appropiate test statistic is

$$W_T = \sup_{r \in R*} t^2_{\gamma=0}(r)$$

# Asymptotic Distribution (cont)

**Proposition 4** *Suppose that assumptions (A.1), (A.2) (A.3) and (A.4) hold.*

1. *DGP (6) with* $\mu_1 = \mu_2 = 0$, *and regression model (7) with no deterministic terms. Then under* $H_0 : \gamma = 0$

$$W_T \Rightarrow \sup_{r \in R^*} \frac{\left(\int W(s) dV(s, p(r))\right)^2}{p(r)(1 - p(r)) \int W(s)^2 ds},$$

*where* $W(\cdot)$ *is a standard Brownian motion and*$V(s, p(r))$ *is a standard Kiefer-Müller process on*$[0, 1]$.[a]

---

[a] A *standard Kieffer-Müller* process $V$ on $[0, 1]^2$ is given by $V(t_1, t_2) = W(t_1, t_2) - t_2 W(t_1, 1)$ where $W(t_1, t_2)$ is a *standard Brownian sheet.* A *standard Brownian sheet* $W(t_1, t_2)$ is a zero-mean Gaussian process with continuous sample paths and covariance function $Cov\,[W(s, t), W(u, v)] = (s \wedge t)(u \wedge v)$.

# Asymptotic Distribution (cont)

2. *DGP (6) with* $\mu_1 = \mu_2 = \mu$, *and regression model (7) with a threshold constant term and a threshold deterministic trend. Then under* $H_0 : \gamma = 0$

$$W_T \Rightarrow \sup_{r \in R^*} \frac{\left(\int W^{**}(s) dV(s, p(r))\right)^2}{p(r)(1 - p(r)) \int W^{**}(s)^2 ds},$$

*where*

$W^*(s) =$
$W(s) - \int_0^1 W(a) g(a)' da (\int_0^1 g(a) g(a)' da^{-1}) g(s)$ *and*
$g(s) = (1\ s)'$.

# IV. Finite Sample Performance

Simulations = 10,000; sample sizes = $T = 100$, $T = 250$ and 500.

- **Empirical size** $\sim$ nominal one.
- **Power test**:

  $|\gamma| = 0.02,\ 0.06,\ 0.2,\ 0.6$

  $|\Delta\mu| = 0, 0.3, 0.6, 1, 2$

  (a) Power increases with the size of $|\gamma|$

  (b) Better power in the presence of threshold effects.

- **Alternatives models with different values of:**

  $E(\delta_t) = 0.3, 0.5, 0.7, 0.9$

  $V(\delta_t)$ varies from 0 to 0.3

  $|\Delta\mu| = 0, 0.3, 0.6, 1.0, 2.0$

  Power increases with $E(\delta_t)$,$V(\delta_t)$ and $\Delta\mu$. Very low when the parameter is not stochastic ($V(\delta_t) = 0$), but increases considerably when $V(\delta_t)$ is high.

- **Power of the DF test:** DF test hardly distinguishes between an exact unit root and a threshold stochastic unit root.

# V. A Simple Empirical Strategy

**Three Steps:**

1. DF test
2. Selection of the threshold variables
3. TARSUR test

# US Stock Prices

Quarterly series of Standard and Poor Composite (1947:1-1999:4) (www.econ.yale.edu/˜shiller), deflated by the Producer Price Index for January, base 1996=100.

The real GDP (S.A.) series from the U.S. Department of Commerce, Bureau of Economic Analysis (www.bea.doc.gov).

$Y_t$ = real stock price index

$$Z_{t-d} = \begin{cases} \text{real dividend changes} \quad (\Delta div_{t-d}) \\ \text{real earnings changes} \quad (\Delta ear_{t-d}) \\ \text{real GDP changes} \quad (\Delta gdp_{t-d}) \end{cases}$$

**Estimated model:**

$$\begin{aligned} \Delta Y_t &= (\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)) \\ &+ (\beta_1 t I(Z_{t-d} \leq r) + \beta_2 t I(Z_{t-d} > r)) \\ &+ \gamma \left( I(Z_{t-d} \leq r) - p(r) \right) Y_{t-1} + \varepsilon_t . \end{aligned}$$

# US Stock Prices (cont)

## Results

**Graphs of the Time Series and Tables with the results will be shown during the presentation (see Paper)**

# US Stock Prices (comments)

**Real GNP changes is a candidate to explain the existence of a stocahstic unit root in stock prices**: when increments are negative (”down state”) the stock price index is in the stationary regime, and when increments are positive (”up state”), prices follow a mildly explosive model.

**The stochastic root of the autorregresive representation is on average the unity.**

Looking at the transition probabilities, it seems that $\delta_t$ follows a $1^{sr}$ order Markov process ($\Delta gnp_t$ is an *AR(1)* process with positive parameter).Therefore **stock prices will not follow a martingale process** and there is a chance for **some predictability of the future return**:

$$E_{t-1}\left(\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right) = E_{t-1}\left(\delta_t - 1\right) \tag{8}$$

There exist a **positive relation between the expected returns and the real activity of the economy** (N. Chen and E. F. Fama (1990), or R. Roll and S.A. Ross (1986))

# International Bond Yield Data

·*Countries* : USA (BUS), UK (BUK), Japan (BJ) and West Germany (BWG).

·*Frecuency*: daily close of trade observations from 1 april of 1986 to 29 of December of 1989, $T = 980$.

·*Maturity* : redemption yields for goverment bonds with less than 5 years to maturity.

·*Source* : Mills (1993).

- Results from Leybourne, McCabe and Mills (1996):
  (i) BUS does not reject the null hypothesis of a fixed unit root in favour of a stochastic unit root
  (ii) BUK and BWG weakly reject the null of a fixed unit root
  (iii) BJ strongly rejects the null of a fixed unit root.

- There is evidence of BUS Granger causing the other yields (Mills (1993).

$$Y_t = \begin{cases} BUK \\ BJ \\ BWG \end{cases} \quad \text{and } Z_{t-d} = \Delta BUS$$

# International Bond Yield Results

**Estimated model:**

$$\begin{aligned}\Delta Y_t &= \left(\mu_1 I(Z_{t-d} \leq r) + \mu_2 I(Z_{t-d} > r)\right) \\ &\quad + \left(\beta_1 t I(Z_{t-d} \leq r) + \beta_2 t I(Z_{t-d} > r)\right) \\ &\quad + \gamma \left(I(Z_{t-d} \leq r) - p(r)\right) Y_{t-1} + \varepsilon_t.\end{aligned}$$

**Graphs of the Time Series and Tables with the results will be shown in the presentation**

# International Bond Yield (Comments)

**BUS changes is a candidate to explain the existence of a stochastic unit root in BJ**: when BUS increments are negative BJ is in the stationary regimen, and when BUS increments are positive, BJ follows a mildly explosive model.

**The stochastic root of the autorregresive representation is on average the unity.**

Looking at the transition probabilities, it seems that $\rho_t$ follows an *i.i.d* process (the correlogram of $\Delta BUS_t$ also suggests these results). Nevertheless, the delay parameter, $d$, is equal to 1, therefore **bond yields will not follow a martingale process and there is a chance for some predictability.**

$$E_{t-1}\left(\frac{Y_t - Y_{t-1}}{Y_{t-1}}\right) = E_{t-1}\left(\delta_t - 1\right)$$

$$= \begin{cases} \rho_1 - 1 < 0, if \Delta BUS_{t-1} \leq r \\ \rho_2 - 1 > 0, if \Delta BUS_{t-1} > r \end{cases}$$

# VI. Conclusions

**Steps:**

1. DF test
2. Selection of the threshold variables
3. TARSUR test

Given that many economic variables are well approximated by models with AR unit roots, this paper introduces a methodology designed to find possible causes of the existence of those unit roots.

This methodology has:

- Explanatory power
- Forecast power