Threshold Models

Mei-Yuan Chen
Department of Finance
National Chung Hsing University

February 25, 2013

Threshold Models

The threshold autoregressive model was proposed by Tong (1978) and discussed in detail by Tong and Lim (1980) and Tong (1983). A useful package for threshold models in **R** is **tsDyn**. In **tsDyn**, the functions, **setar()**, **lstar()**, **nneTs()**, and **aar()**, are useful. There are some useful **R** codes for simulating TAR time series (**tar.sim()**), estimating TAR models (**tar()**), doing Likelihood ratio test for threshold nonlinearity (**tlrt()**), doing Tsay's test for nonlinearity (**Tsay.test()**), doing model diagnostics for a Fitted ARIMA Model (**tsdiag.Arima**), doing model diagnostics for a fitted TAR model (**tsdiag.TAR**), and doing prediction based on a fitted TAR model (**predict.TAR**).

Distinguishing SETAR Model Classes

Let y_t be a univariate time series and $\boldsymbol{Y}_{t-1} = (1\, y_{t-1}\, y_{t-2}\, \cdots\, y_{t-p})$, a $k \times 1$ vector with $k = p + 1$. A SETAR(m) model with m regimes takes the form

$$y_t = \boldsymbol{\alpha}_1' \boldsymbol{Y}_{t-1} I_{1t}(\boldsymbol{\gamma}, d) + \cdots + \boldsymbol{\alpha}_m' \boldsymbol{Y}_{t-1} I_{mt}(\boldsymbol{\gamma}, d) + e_t, \tag{1}$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_{m-1})$ with $\gamma_1 < \gamma_2 < \cdots < \gamma_{m-1}$ is a vector of threshold parameters, and $I_{jt}(\boldsymbol{\gamma}, d) = I(\gamma_{j-1} < q_{t-1} \leq \gamma_j)$, where $I(\cdot)$ is the indicator function and $q_{t-1} = q(y_{t-1}, \ldots, y_{t-p})$ is a known function of the data. Three functions are considered in Koop and Potter (1997):

1. $q_{t-1} = y_{t-d}$;
2. $q_{t-1} = y_{t-1} - y_{t-d}$;
3. $q_{t-1} = (y_{t-1} - y_{t-d})/d$.

Therefore, d is called the delay parameter. Conventionally, $\gamma_0 = -\infty$ and $\gamma_m = \infty$. The error e_t is a uniformly square integrable martingale difference sequence, hence

$$E(e_t|\mathcal{F}_{t-1}) = 0, \tag{2}$$

where $\mathcal{F}_t$ denotes the natural filtration, and $\text{var}(e_t) = \sigma^2 < \infty$. The method provided by Hansen (1999) is to determine the number of regimes m. This method is outlined as follows.

SETAR(1) Models

The class SETAR(1) is the class of linear autoregressions, which can be written as

$$y_t = \boldsymbol{\alpha}_1' \boldsymbol{Y}_{t-1} + e_t. \tag{3}$$

Thus testing for linearity (within the SETAR class of models) is a test of the null hypothesis of SETAR(1) against the alternative of SETAR(m) for some $m > 1$. Similarly, the null of the SETAR(2) model

$$y_t = \boldsymbol{\alpha}_1' \boldsymbol{Y}_{t-1} I_{1t}(\boldsymbol{\gamma}, d) + \boldsymbol{\alpha}_2' \boldsymbol{Y}_{t-1} I_{2t}(\boldsymbol{\gamma}, d) + e_t, \tag{4}$$

can be tested against the alternative of a SETAR(m) for some $m > 2$.

Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m, \boldsymbol{\gamma}, d)$ denote the collection of parameters in (1). Under assumption (2) the appropriate estimation method is the least square estimator $\hat{\boldsymbol{\theta}}_T$ obtained by

$$\hat{\boldsymbol{\theta}}_T = \arg\min_{\boldsymbol{\theta}} \sum_{t=1}^{T} (y_t - \boldsymbol{\alpha}_1' \boldsymbol{Y}_{t-1} I_{1t}(\boldsymbol{\gamma}, d) - \cdots - \boldsymbol{\alpha}_m' \boldsymbol{Y}_{t-1} I_{mt}(\boldsymbol{\gamma}, d))^2. \quad (5)$$

Denote $S_m = \hat{e}_m' \hat{e}_m$ as the sum of squared residuals where $\hat{e}_m$ is the LS residuals from model SETAR(m). The natural LS test of the hypothesis of SETAR(j) against SETAR(k) ($k > j$) is to reject for large values of

$$F_{jk} = T \left(\frac{S_j - S_k}{S_k} \right). \quad (6)$$

This is the likelihood ratio test when the errors e_t are independent $N(0, \sigma^2)$. It is also the conventional F (or Wald) test, and is equivalent to the conventional Lagrange multiplier (or score) test.

Estimation for a SETAR(1) Model

The solution to the LS problem of (5) for the SETAR(1) model is:

$$\hat{\boldsymbol{\alpha}}_1 = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y},$$

where $\boldsymbol{X}$ is the $T \times k$ matrix whose ith row is $\boldsymbol{Y}'_{t-1}$ and $\boldsymbol{y}$ is the $T \times 1$ vector. The residual vector is $\hat{\boldsymbol{e}}_1 = \boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\alpha}}_1$ and the sum of squared residuals is $S_1 = \hat{\boldsymbol{e}}'_1\hat{\boldsymbol{e}}_1$.

Estimation for a SETAR(2) Model

In the SETAR(2) model, $\boldsymbol{\gamma} = \gamma_1$, so we let $I_{1t}(\boldsymbol{\gamma}, d) = I(q_{t-1} \leq \gamma)$ and $I_{2t}(\boldsymbol{\gamma}, d) = I(q_{t-1} > \gamma)$. Let $\alpha = (\alpha_1', \alpha_2')'$ and

$$\boldsymbol{X}_{t-1}(\boldsymbol{\gamma}, d) = \left(\begin{array}{c} \boldsymbol{X}_{t-1} I_{1t}(\boldsymbol{\gamma}, d) \\ \boldsymbol{X}_{t-1} I_{2t}(\boldsymbol{\gamma}, d) \end{array} \right).$$

Let $\boldsymbol{X}(\boldsymbol{\gamma}, d)$ be the $T \times 2k$ matrix whose ith row is $\boldsymbol{X}_{t-1}(\gamma, d)'$. Then the minimization problem (5)

$$S_2 = \min_{d,\boldsymbol{\gamma},\boldsymbol{\alpha}} (\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{\gamma}, d)\boldsymbol{\alpha})'(\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{\gamma}, d)\boldsymbol{\alpha})$$

can be solved sequentially through concentration. That is, for given $(d, \boldsymbol{\gamma})$, minimization over $\boldsymbol{\alpha}$ is an OLS regression of $\boldsymbol{y}$ on $\boldsymbol{X}(\boldsymbol{\gamma}, d)$. The solution can be written as

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}, d) = [\boldsymbol{X}(\boldsymbol{\gamma}, d)'\boldsymbol{X}(\boldsymbol{\gamma}, d)]^{-1}[\boldsymbol{X}(\boldsymbol{\gamma}, d)'\boldsymbol{y}]. \tag{7}$$

Let

$$S_2(\boldsymbol{\gamma}, d) = [\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{\gamma}, d)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}, d)]'[\boldsymbol{y} - \boldsymbol{X}(\boldsymbol{\gamma}, d)\hat{\boldsymbol{\alpha}}(\boldsymbol{\gamma}, d)]$$

be the sum of squared residuals given $(d, \boldsymbol{\gamma})$. Then

$$(\hat{\boldsymbol{\gamma}}_1, \hat{d}) = \arg\min_{\boldsymbol{\gamma}, d} S_2(\boldsymbol{\gamma}, d). \tag{8}$$

Once the solution to (8) is found, we find $\hat{\boldsymbol{\alpha}}$ through (7), i.e., the LS estimation of $\boldsymbol{\alpha}$ are then found as $\hat{\boldsymbol{\alpha}} = \hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\gamma}}_1, \hat{d})$. Since the parameter space for d is discrete, the LS estimator $\hat{d}$ is super-consistent, and for the purpose of inference on the other parameters we can act as if d is known.

Testing TAR(1) against TAR(2): i.i.d. Case

The null for testing TAR(1) (a linear AR(p)) against TAR(2) is $H_0 : \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2$. The testing problem is tainted by the difficulty that the threshold γ is not identified under H_0. Hansen (1996) suggested a testing methodology for doing the inference. When the error e_t is i.i.d., consider the F-statistic

$$F_{12} = \sup_{\gamma \in \Gamma} F_{12}(\gamma, d)$$

where

$$F_{12}(\gamma, d) = \left(\frac{S_1 - S_2(\gamma, d)}{S_2(\gamma, d)} \right)$$

which has the asymptotic distribution as $\chi^2(k)$ when γ is identified.

Hansen (1996)'s Bootstrap Procedure

1. Let $u_t^*, t = 1, \ldots, T$ be i.i.d. $N(0, 1)$ random draws, and set $y_t^* = u_t^*$.
2. Using the observations $\boldsymbol{x}_t, t = 1, \ldots, T$, regress y_t^* on $\boldsymbol{x}_t$ to obtain the error sum of squares S_1^*;
3. Using the observations $\boldsymbol{x}_t(\gamma, d), t = 1, \ldots, T$, regress y_t^* on $\boldsymbol{x}_t(\gamma, d)$ to obtain the error sum of squares $S_2^*(\gamma, d)$;
4. Calculate $F_{12}^*(\gamma, d) = (S_1^* - S_2^*(\gamma, d))/S_2^*(\gamma, d)$ and $F_{12}^* = \sup_\gamma F_{12}^*(\gamma, d)$;
5. Repeat above steps B times and then the empirical distribution of F_{12}^* is established.

Hansen (1996) shows the empirical distribution of F_{12}^* converges weakly in probability to the null distribution of F_{12}.

Testing TAR(1) against TAR(2): Non i.i.d. Case

If the error e_t is conditionally heteroskedastic, consider the Wald statistic

$$W_{12} = \sup_{\gamma \in \Gamma} W_{12}(\gamma, d)$$

where

$$\begin{aligned} W_{12}(\gamma, d) &= (\boldsymbol{R}\hat{\boldsymbol{\alpha}}(\gamma, d))' \left[\boldsymbol{R}\left(\boldsymbol{M}_T(\gamma, d)^{-1}\boldsymbol{V}_T(\gamma, d)\boldsymbol{M}_T(\gamma, d)^{-1}\right)\boldsymbol{R}'\right] \\ & \quad (\boldsymbol{R}\hat{\boldsymbol{\alpha}}(\gamma, d)), \end{aligned}$$

$\boldsymbol{R} = [(\boldsymbol{I} \quad -\boldsymbol{I}]$, $M_T(\gamma, d) = \sum_{t=1}^{T} \boldsymbol{x}_t(\gamma, d)\boldsymbol{x}(\gamma, d)'$, and $\boldsymbol{V}_T(\gamma, d) = \sum_{t=1}^{T} \boldsymbol{x}_t(\gamma, d)\boldsymbol{x}_t(\gamma, d)'\hat{e}_t(\gamma, d)$, where $\hat{e}_t(\gamma, d)$ is residual from estimated TAR(2). To obtain critical values, bootstrap the data as previously, but instead set $y_t^* = \hat{e}(\gamma, d)u_t^*$. Hansen (1996) shows that this procedure produces the asymptotically correct null distribution.

SETAR(3) Model

The SETAR(3) model is

$$y_t = \boldsymbol{\alpha}_1' \boldsymbol{Y}_{t-1} I_{1t}(\boldsymbol{\gamma}, d) + \boldsymbol{\alpha}_2' \boldsymbol{Y}_{t-1} I_{2t}(\boldsymbol{\gamma}, d) + \boldsymbol{\alpha}_3' \boldsymbol{Y}_{t-1} I_{3t}(\boldsymbol{\gamma}, d) + e_t, \quad (9)$$

where $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$. In principle, this model can be estimated using the same techniques described in the previous section, namely conditional on $(\boldsymbol{\gamma}, d)$, the parameters $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$ nay be estimated by OLS, and then a grid search over $(\boldsymbol{\gamma}, d)$ yields the LS estimates. To reduce works on grid search over $(\boldsymbol{\gamma}, d)$, a computational short-cut was proposed by Bai (1997) and Bai and Perron (1997) in the change-point literature.

However, if true model is (9) but the (misspecified) SETAR(2) model (4) is actually estimated, least-squares estimate $\hat{d}$ will be consistent for d and $\hat{\gamma}_1$ will be consistent for one of the pair $(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$. Bai and Perron show that if $\boldsymbol{\gamma}(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$ is estimated by least-squares on (9), enforcing that $d = \hat{d}$ and that one element of $\boldsymbol{\gamma}$ equals $\hat{\gamma}_1$, then the second-stage estimate $\hat{\gamma}_2$ will be consistent for the remaining element of the pair $(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$. Thus this two-step method yields consistent estimation of $\hat{d}$ and $\hat{\boldsymbol{\gamma}} = (\hat{\boldsymbol{\gamma}}_1, \hat{\boldsymbol{\gamma}}_2)$.

Furthermore, Bai (1997) shows that these estimates can be made asymptotically efficient, in the sense of having the same asymptotic distribution as estimates obtained from joint estimation of (9), if this model is iterated at least once. That is, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$ is estimated by least-squares on (9), enforcing the constraint that $d = \hat{d}$ and that one element of $\boldsymbol{\gamma}$ equals $\hat{\gamma}_2$, yielding a refined estimate $\hat{\gamma}_1$. Further iteration does not affect the asymptotic distribution, but may yield finite-sample improvements. This "one-step-at-a-time" approach yields enormous computational savings. Rather than pN^2 function evaluations, it involves approximately $pN + 2N$ function evaluations, which is only a minor increase over the requirements for estimation of the SETAR(2) model.

Testing SETAR(j) Against SETAR(k)

The test statistic of the SETAR(1) model against the SETAR(2) is

$$F_{12} = \max_{\boldsymbol{\gamma}, d} F_{12}(\boldsymbol{\gamma}, d), \tag{10}$$

where

$$F_{12}(\boldsymbol{\gamma}, d) = T \left(\frac{S_1 - S_2(\boldsymbol{\gamma}, d)}{S_2(\boldsymbol{\gamma}, d)} \right).$$

Hansen (1997) has shown that the asymptotic distribution of the empirical process $F_{12}(\boldsymbol{\gamma}, d)$ and described an algorithm to calculate the asymptotic distribution of F_{12} under homoskedasticity and heteroskedasticity of the error terms. However, the asymptotic distribution is derived under the assumption of the process y_t being stationary, excluding unit roots or near unit roots. Thus, in addition to the calculation of asymptotic distribution for the empirical process $F_{12}(\boldsymbol{\gamma}, d)$, Hansen (1999) also explores the distribution by bootstrap. Furthermore, Caner and Hansen (2000) show that the bootstrap achieves a good approximation even if there is a unit root or near unit root.

Similarly, the test statistic of the SETAR(1) against SETAR(3) is

$$F_{13} = \max_{\boldsymbol{\gamma}, d} F_{13}(\boldsymbol{\gamma}, d),$$

where

$$F_{13}(\boldsymbol{\gamma}, d) = T\left(\frac{S_1 - S_3(\boldsymbol{\gamma}, d)}{S_3(\boldsymbol{\gamma}, d)}\right).$$

And then, the test statistic of the SETAR(2) against SETAR(3) is

$$F_{23} = \max_{\boldsymbol{\gamma}, d} F_{23}(\boldsymbol{\gamma}, d),$$

where

$$F_{23}(\boldsymbol{\gamma}, d) = T\left(\frac{S_2(\boldsymbol{\gamma}, d) - S_3(\boldsymbol{\gamma}, d)}{S_3(\boldsymbol{\gamma}, d)}\right).$$

Empirical Application: Threshold CAPM

Akdeniz, L., A. Altay-Salih, and M. Caner (2003) introduces the threshold CAPM model to capture time-varying betas. The Capital Asset Pricing Model (CAPM) of Sharpe (1964), Lintner (1965), and Black (1972) has constituted one of the cornerstones of modern finance theory for the last four decades. The CAPM posits a simple and stable linear relationship between an asset's systematic risk and its expected return. However, recent studies, notably Banz (1981), Basu (1983), Bhandari (1988), and Fama and French (1992), have found weak or no statistical evidence in support of this simple relationship. Stimulated by these findings, a number of researchers have sought to find alternative

Alternative explanations of CAPM

1. Fama and French (1993, 1995): concluded that fundamental variables, namely Book-to-Market equity ratio and Market Equity, found to explain the variation in returns must be proxies for some unidentified risk factors.
2. Ferson (1989), Ferson and Harvey (1991, 1993), Ferson and Korajczyk (1995), and Jaganathan and Wang (1996): argue that beta and market risk premium vary over time, therefore, static CAPM should be improved by incorporating time variation in beta in the model.

Why Thershold CAPM

1. Jagannathan and Wang (1996) argue, time variation in beta may come from two different sources. During high interest rate periods highly leveraged firms are more likely to face financial problems so their betas are more likely to rise.
2. At the same time, the decrease in the uncertainty about the growth prospects of firms can cause their betas to decrease.

It is not immediately clear which cause will dominate. Hence betas might be higher or lower during high interest rate periods. In addition, some industries may be more capital intensive than others, which makes them more sensitive to

R Package: tsDyn

Nonlinear autoregressive time series models:

1. "nnetTs"
2. "setar"
3. "star"
4. "star"
5. "aar"