

## THEORY AND APPLICATIONS OF TAR MODEL WITH TWO THRESHOLD VARIABLES

Haiqiang Chen<sup>1</sup>, Terence Tai-Leung Chong<sup>2</sup>, and Jushan Bai<sup>3</sup>

<sup>1</sup>*Wang Yanan Institute for Studies in Economics, Xiamen University, Fujian, China*

<sup>2</sup>*Department of Economics and Institute of Global Economics and Finance,  
The Chinese University of Hong Kong, Hong Kong, China*

<sup>3</sup>*Department of Economics, Columbia University, New York, New York, USA*

□ *A growing body of threshold models has been developed over the past two decades to capture the nonlinear movement of financial time series. Most of these models, however, contain a single threshold variable only. In many empirical applications, models with two or more threshold variables are needed. This article develops a new threshold autoregressive model which contains two threshold variables. A likelihood ratio test is proposed to determine the number of regimes in the model. The finite-sample performance of the estimators is evaluated and an empirical application is provided.*

**Keywords** Bootstrapping; Likelihood ratio test; Misspecification; Threshold autoregressive model.

**JEL Classification** C22.

### 1. INTRODUCTION

A growing body of threshold models has been developed over the past two decades to capture the nonlinear movement of financial time series. Tong (1983) develops a threshold autoregressive (TAR) model and uses it to predict stock price movements. A number of new models have been proposed since the seminal work of Tong (1983), including the smooth transition threshold autoregressive model (STAR) of Chan and Tong (1986) and the functional-coefficient autoregressive (FAR) model of Chen and Tsay (1993). Tsay (1998) develops a multivariate TAR model for the arbitrage activities in the security market. Dueker et al. (2007) develop a contemporaneous TAR model for the bond market.

Address correspondence to Terence Tai-Leung Chong, Department of Economics and Institute of Global Economics and Finance, The Chinese University of Hong Kong, Hong Kong, China; E-mail: chong2064@cuhk.edu.hk

Most of the aforementioned models, however, contain a single threshold variable only. In many empirical applications, a model with two or more threshold variables is more appropriate. For example, Leeper (1991) divides the policy parameter space into four disjoint regions according to whether monetary and fiscal policies are active or passive. Given these policy combinations, macroeconomic variables such as real output, inflation, and unemployment have different dynamics. Tiao and Tsay (1994) divide the U.S. quarterly real Gross National Product (GNP) growth rate into four regimes according to the level and sign of the past growth rate. Durlauf and Johnson (1995) split that cross-country GDP growth rate into different regimes according to the level of per capita real GDP and literacy rate. In modelling currency crises, Sachs et al. (1996), Frankel and Rose (1996), Kaminsky (1998), and Edison (2000) argue that the occurrence of currency crises hints at the values of fiscal reserves, foreign reserves, and interest rate differential between home countries and the United States. In these examples, TAR models with multiple threshold variables can be used to describe the dynamics of different regimes.<sup>1</sup>

As the distributional theory is rather involved, no asymptotic result has been developed for TAR models with multiple threshold variables.<sup>2</sup> This article contributes to the literature by developing estimation and inference procedures for TAR models with two threshold variables. Our model is applied to identify the regimes of the Hong Kong stock market. The case of Hong Kong is of interest because of its rising role as a global financial center. In 2006, Hong Kong became the world's second most popular place for Initial Public Offering (IPO) after London. In 2007, the Hong Kong stock market ranked fifth in the world, while its warrant market ranks top worldwide in terms of turnover. Using the historical prices of the Hang Seng index and the market turnover as threshold variables, our estimation shows that the stock market of Hong Kong can be classified into a high-return stable regime, a low-return volatile regime, and a neutral regime. This is different from the conventional bull–bear classification.

The remainder of the article is organized as follows: Section 2 presents the model and discusses the estimation procedure. Section 3 derives the limiting distribution of the threshold estimators. Section 4 proposes a likelihood ratio test to determine the number of regimes. Monte Carlo simulations are conducted, and the performance of the estimation procedure is evaluated in Section 5. An empirical application is provided in Section 6. Section 7 concludes the article.

<sup>1</sup>Threshold model with two threshold variables can also be applied to the cross-section of financial data. For example, in the Fama and French (1992) model, one may use firm size and book-to-market ratio as threshold variables to explain abnormal returns of a stock. Avramov et al. (2006) also sort stocks into different categories according to historical returns and liquidity level.

<sup>2</sup>A related empirical study is the nested threshold autoregressive (NeTAR) models of Astatkie et al. (1997).

## 2. TAR MODEL WITH TWO THRESHOLD VARIABLES

Consider the following TAR model with two threshold variables which classifies the observations  $y_t$  into four regimes:

$$y_t = \begin{cases} \beta_0^{(1)} + \beta_1^{(1)} y_{t-1} + \beta_2^{(1)} y_{t-2}, \dots, + \beta_{p_1}^{(1)} y_{t-p_1} + u_t, & \text{when } z_{1t} \leq \gamma_1^0, z_{2t} \leq \gamma_2^0 \\ \beta_0^{(2)} + \beta_1^{(2)} y_{t-1} + \beta_2^{(2)} y_{t-2}, \dots, + \beta_{p_2}^{(2)} y_{t-p_2} + u_t, & \text{when } z_{1t} \leq \gamma_1^0, z_{2t} > \gamma_2^0 \\ \beta_0^{(3)} + \beta_1^{(3)} y_{t-1} + \beta_2^{(3)} y_{t-2}, \dots, + \beta_{p_3}^{(3)} y_{t-p_3} + u_t, & \text{when } z_{1t} > \gamma_1^0, z_{2t} \leq \gamma_2^0 \\ \beta_0^{(4)} + \beta_1^{(4)} y_{t-1} + \beta_2^{(4)} y_{t-2}, \dots, + \beta_{p_4}^{(4)} y_{t-p_4} + u_t, & \text{when } z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0 \end{cases}, \quad (1)$$

where

$z_t \stackrel{\text{def}}{=} (z_{1t}, z_{2t})$  are the threshold variables;

$\gamma^0 \stackrel{\text{def}}{=} (\gamma_1^0, \gamma_2^0) \in \Omega$ , where  $\Omega = [\gamma_1, \bar{\gamma}_1] \times [\gamma_2, \bar{\gamma}_2]$  is a strict subset of the support of  $z_t$ .  $\gamma^0$  is the threshold parameter vector pending to be estimated;

$p_j$  ( $j = 1, 2, 3, 4$ ) is the order in each regime;

$\beta^{(j)} \stackrel{\text{def}}{=} (\beta_0^{(j)}, \beta_1^{(j)}, \beta_2^{(j)}, \dots, \beta_{p_j}^{(j)})'$  are the structural parameters and  $\beta^{(i)} \neq \beta^{(j)}$  for some  $i \neq j$ .<sup>3</sup>

The model is a linear AR model within each regime.<sup>4</sup> The threshold variables  $z_1$  and  $z_2$  can be exogenous variables or functions of the lags of  $y_t$ .<sup>5</sup> Given  $\{y_t, z_t\}_{t=1}^T$ , our objective is to estimate the threshold parameters  $\gamma^0$  and the structural parameters  $\beta^{(j)}$ . Without loss of generality, we let  $p = \max\{p_1, p_2, p_3, p_4\}$ , and  $\beta_q^{(j)} = 0$  when  $q > p_j$ ,  $j = 1, 2, 3, 4$ . The model can be rewritten as

$$y_t = \sum_{j=1}^4 \Psi_t^{(j)}(\gamma^0) (\beta_0^{(j)} + \sum_{i=1}^p \beta_i^{(j)} y_{t-i} + u_t), \quad (2)$$

<sup>3</sup>Restrictions on the structural parameters can be imposed so that there are less than four regimes. For example, if  $\beta^{(1)} = \beta^{(2)} = \beta^{(3)}$ , the model will have two regimes only.

<sup>4</sup>An empirical example of the Model (1) is Tiao and Tsay's (1994) four-regime TAR model for quarterly U.S. real GNP growth rates  $y_t$ :

$$y_t = \begin{cases} -0.015 - 1.076y_{t-1} + \varepsilon_{1t}, & y_{t-1} \leq y_{t-2} \leq 0 \\ 0.63y_{t-1} - 0.76y_{t-2} + \varepsilon_{2t}, & y_{t-1} > y_{t-2}, y_{t-2} \leq 0 \\ 0.006 + 0.43y_{t-1} + \varepsilon_{3t}, & y_{t-1} \leq y_{t-2}, y_{t-2} > 0 \\ 0.433y_{t-1} + \varepsilon_{4t}, & y_{t-1} > y_{t-2} > 0 \end{cases}.$$

In their model, the process is divided into four regimes by  $z_{1t} = y_{t-2}$  and  $z_{2t} = y_{t-1} - y_{t-2}$ , and the threshold values are set to zero. In practice, we need to estimate the threshold values.

<sup>5</sup>The model is a Self-Exciting Threshold Autoregressive (SETAR) model if the threshold variable is  $y_{t-d}$ .

where  $\Psi_t^{(j)}(\gamma^0)$  is an indicator function which equals one when the threshold condition is satisfied, and equals zero otherwise. Specifically,

$$\begin{aligned} \Psi_t^{(1)}(\gamma^0) &= I(z_{1t} \leq \gamma_1^0, z_{2t} \leq \gamma_2^0), \\ \Psi_t^{(2)}(\gamma^0) &= I(z_{1t} \leq \gamma_1^0, z_{2t} > \gamma_2^0), \\ \Psi_t^{(3)}(\gamma^0) &= I(z_{1t} > \gamma_1^0, z_{2t} \leq \gamma_2^0), \\ \Psi_t^{(4)}(\gamma^0) &= I(z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0). \end{aligned}$$

For analytical reasoning, it is convenient to rewrite the model (2) in the following matrix form:

$$Y = \sum_{j=1}^4 I_j(\gamma^0) X \beta^{(j)} + U, \tag{3}$$

where

$$X = (x'_T, x'_{T-1}, \dots, x'_{p+1})' = \begin{pmatrix} 1, y_{T-1}, y_{T-2}, \dots, y_{T-p} \\ 1, y_{T-2}, y_{T-3}, \dots, y_{T-p-1} \\ \dots \\ 1, y_p, y_{p-1}, \dots, y_1 \end{pmatrix}_{(T-p) \times (p+1)},$$

$$x_t = (1, y_{t-1}, \dots, y_{t-p})' \text{ for } t = p+1, \dots, T.$$

$$I_j(\gamma^0) = \text{diag} \left\{ \Psi_T^{(j)}(\gamma^0), \Psi_{T-1}^{(j)}(\gamma^0), \dots, \Psi_{p+1}^{(j)}(\gamma^0) \right\},$$

$$Y = (y_T, y_{T-1}, \dots, y_{p+1})',$$

$$U = (u_T, u_{T-1}, \dots, u_{p+1})'.$$

We make the following assumptions:

- (A1)  $y_t$  is stationary ergodic and  $E(y_t^4) < \infty$ .
- (A2)  $\{u_t\}$  is a sequence of independent and identical distributed (IID) normal errors with zero mean and finite variance  $\sigma^2$ .
- (A3) The threshold variables  $z_{1t}$  and  $z_{2t}$  are strictly stationary and have a continuous joint distribution  $F(\gamma)$ , which is differentiable with respect to both variables. Let  $f(\gamma)$  denote the corresponding joint density function and  $f_i(\gamma) = \frac{\partial F(\gamma)}{\partial \gamma_i}$ . We assume that  $0 < f(\gamma) \leq \bar{f} < \infty$ ;  $0 < f_i(\gamma) \leq \bar{f}_i < \infty$  for  $i = 1, 2$ .

(A1) assumes that  $y_t$  is stationary ergodic, which allows us to apply the law of large number. A sufficient condition for (A1) to hold is

$\max_j \sum_i (|\hat{\beta}_i^{(j)}|) < 1$ .<sup>6</sup> (A2) assumes that  $\{u_t\}$  is a sequence of i.i.d. normal errors with finite second moment.<sup>7</sup> (A3) requires the stationarity of the threshold variables. We also assume that the threshold variables are continuous with positive density everywhere, so that it is dense near  $\gamma^0$  as the sample size increases. This assumption is needed for the consistent estimation of threshold values.

Given  $\gamma = (\gamma_1, \gamma_2)$ , the conditional least square (CLS) estimator for  $\beta^{(j)}$  is defined as

$$\hat{\beta}^{(j)}(\gamma) = (X' I_j(\gamma) X)^{-1} X' I_j(\gamma) Y, \quad (j = 1, 2, 3, 4), \quad (4)$$

where

$$I_j(\gamma) = \text{diag} \left\{ \Psi_T^{(j)}(\gamma), \Psi_{T-1}^{(j)}(\gamma), \dots, \Psi_{p+1}^{(j)}(\gamma) \right\}.$$

The residual sum of squares is

$$RSS_T(\gamma) = \left\| \sum_{j=1}^4 I_j(\gamma^0) X \beta^{(j)} + U - \sum_{j=1}^4 I_j(\gamma) X \hat{\beta}^{(j)}(\gamma) \right\|^2,$$

and we define the estimator of  $\gamma^0$  as the value that minimizes  $RSS_T(\gamma)$ :

$$\hat{\gamma} = \arg \min_{\gamma \in \Omega} RSS_T(\gamma). \quad (5)$$

The structural estimators evaluated at the estimated threshold values are defined as

$$\hat{\beta}^{(j)}(\hat{\gamma}) = (X' I_j(\hat{\gamma}) X)^{-1} X' I_j(\hat{\gamma}) Y. \quad (6)$$

Appendix 2 shows the consistency of the estimators  $(\hat{\gamma}, \hat{\beta}^{(j)}(\hat{\gamma}))$ .

### 3. LIMITING DISTRIBUTION OF $(\hat{\gamma}_1, \hat{\gamma}_2)$

In this section, the asymptotic joint distribution of the least-squares estimator  $\hat{\gamma}$  is derived under the assumption that the magnitude of change goes to zero at an appropriate rate. As pointed out by Hansen (2000), the assumption of decaying threshold effect is needed in order to obtain an

<sup>6</sup>See Chan (1993) and Hansen (1997).

<sup>7</sup>In this article, we generalize the TAR model to the one with two threshold variables. The error term  $u_t$  is assumed to be i.i.d. normal in order to derive the asymptotic distribution of the threshold estimators. We can relax this assumption and allow for heteroskedasticity of  $u_t$ . The estimators will still be consistent. See Hansen (1997) for more discussion on the heteroskedastic errors.

asymptotic distribution of  $\hat{\gamma}$  free of nuisance parameters.<sup>8</sup> For notational simplicity, we rewrite Model (2) as

$$Y = X\beta^{(1)} + \sum_{j=2}^4 X_{\gamma_0}^{(j)} \delta^{(j)} + U, \quad (7)$$

where

$$X_{\gamma_0}^{(j)} = I_j(\gamma_0) X, \quad (j = 2, 3, 4)$$

and

$$\delta^{(j)} = \beta^{(j)} - \beta^{(1)}, \quad (j = 2, 3, 4).$$

For any given  $\gamma$ , we define

$$X_{\gamma}^{(j)} = I_j(\gamma) X, \quad (j = 1, 2, 3, 4).$$

Observe that  $X_{\gamma}^{(i)'} X_{\gamma}^{(j)} = 0$  if  $i \neq j$ , and  $X' X_{\gamma}^{(j)} = X_{\gamma}^{(j)'} X_{\gamma}^{(j)}$ .

Let  $X_0^{(j)} = X_{\gamma_0}^{(j)}$ , we have

$$X = \sum_{j=1}^4 X_0^{(j)} = \sum_{j=1}^4 X_{\gamma}^{(j)}.$$

We define the following conditional moment functionals:

$$D(\gamma) = E(x_t x_t' | z_t = \gamma), \quad (8)$$

$$V(\gamma) = E(x_t x_t' u_t^2 | z_t = \gamma). \quad (9)$$

Let  $D = D(\gamma_0)$ ,  $V = V(\gamma_0)$ . Under the assumption (A2),  $V = \sigma^2 D$ . We define block diagonal matrices  $D^* = \text{diag}\{D, D\}$  and  $V^* = \text{diag}\{V, V\}$ .<sup>9</sup> We also need the following assumptions before the limiting distribution of  $\hat{\gamma}$  can be obtained. These assumptions mainly follow Hansen (1997, 2000).

(A4)  $M > M_j(\gamma) > 0$  for all  $\gamma \in \Omega$ , where  $M = E(x_t x_t')$ ,  $M_j(\gamma) = E(x_t x_t' \Psi_t^{(j)}(\gamma))$ ,  $j = 1, 2, 3, 4$ .

(A5)  $\delta_T = (\delta^{(2)'}, \delta^{(3)'}, \delta^{(4)'})' = cT^{-\alpha} = (c_2', c_3', c_4')' T^{-\alpha}$ ,  $0 < \alpha < \frac{1}{2}$ ,  $c$  is a  $3p$ -dimensional constant vector and  $c_i$  is a  $p$ -dimensional constant vector for  $i = 2, 3, 4$ .

<sup>8</sup>This approach is first used in the literature of change points (Bai, 1997) and applied to the threshold model by Hansen (2000).

<sup>9</sup>Note that  $D$  and  $V$  are  $p \times p$  matrices and  $D^*$  and  $V^*$  are  $2p \times 2p$  matrices.

(A6)  $D(\gamma)$  and  $V(\gamma)$  are continuous at  $\gamma = \gamma_0$ .

(A7)  $d_1' D^* d_1 > 0, d_2' D^* d_2 > 0$ , where  $d_1 = (c_2' - c_4', c_3')', d_2 = (c_2', c_3' - c_4)'$ .

(A4) is the conventional full-rank condition which excludes perfect collinearity.  $\Omega$  is restricted to be a proper subset of the support of  $z$ . (A5) assumes that the parameter change is small and converges to zero at a slow rate when the sample size is large. Under this assumption, we are able to make the limiting distribution of  $\hat{\gamma}$  free of nuisance parameters (Chan, 1993). By letting  $\delta_T$  go to zero, we reduce the rate of convergence of  $\hat{\gamma}$  from  $O_p(T^{-1})$  to  $O_p(T^{-1+2\alpha})$  and obtain a simpler limiting distribution of  $\hat{\gamma}$ . (A6) requires the moment functionals to be continuous so that one can obtain the Taylor expansion around  $\gamma_0$ . This condition excludes regime-dependent heteroskedasticity. (A7) excludes the continuous threshold model.<sup>10</sup> Moreover,  $d_1' D^* d_1 > 0$  and  $d_2' D^* d_2 > 0$  impose the identification condition for  $\gamma_1^0$  and  $\gamma_2^0$ , respectively.<sup>11</sup>

**Theorem 1.** *Under assumptions (A1) to (A7), we have*

$$T^{1-2\alpha} \lambda_T \left( (\hat{\gamma}_1 - \gamma_1^0), (\hat{\gamma}_2 - \gamma_2^0) \right) = (r_1, r_2) \xrightarrow{d} \arg \max_{-\infty < r_1 < \infty, -\infty < r_2 < \infty} \left[ -\frac{1}{2} |r_1| + W_1(r_1) - \frac{1}{2} |r_2| + W_2(r_2) \right],$$

where

$$\lambda_T = \left( \frac{(d_1' D^* d_1) f_1^0}{\sigma^2}, \frac{(d_2' D^* d_2) f_2^0}{\sigma^2} \right),$$

and  $W_i(r_i)$  is a two-sided Brownian motion on the real line defined as

$$W_i(r_i) = \begin{cases} \Lambda_{i1}(-r_i) & \text{if } r_i < 0 \\ 0 & \text{if } r_i = 0 \\ \Lambda_{i2}(r_i) & \text{if } r_i > 0 \end{cases},$$

$\Lambda_i(r_i)$ ,  $i = 1, 2$ , and  $j = 1, 2$  are four independent standard Brownian motions on  $[0, \infty)$ .

**Proof.** See Appendix 3. □

<sup>10</sup>This article focuses on the discontinuous threshold effect. For continuous threshold models, one is referred to Chan and Tsay (1998).

<sup>11</sup>Note that  $d_1 = (c_2' - c_4', c_3')'$  measures the size of the threshold effect for the first threshold variable  $z_1$ , while  $d_2 = (c_2', c_3' - c_4)'$  measures the size of the threshold effect for the second threshold variable  $z_2$ . When  $c_2 = c_4 \neq 0$  and  $c_3 = 0$ , we obtain a single threshold model with only two regimes separated by  $z_2 = \gamma_2^0$ . In this case,  $\gamma_1^0$  is not identified. When  $c_2 = 0$  and  $c_3 = c_4 \neq 0$ , we have a single threshold model with only two regimes separated by  $z_1 = \gamma_1^0$  and  $\gamma_2^0$  is not identified.

The result of Hansen (1997) is a special case of Theorem 1 with  $d_1 = 0$  or  $d_2 = 0$ . One can also use Theorem 1 to simulate the confidence interval of  $(\hat{\gamma}_1, \hat{\gamma}_2)$ . The parameter ratio  $\lambda_T$  can be estimated by a polynomial regression or kernel regression. See Hansen (1997, 2000).

#### 4. TESTING FOR AND ESTIMATION OF THE THRESHOLD

To determine the number of regimes, we first consider the null hypothesis of no threshold effect:

$$H_0 : \beta^{(1)} = \beta^{(2)} = \beta^{(3)} = \beta^{(4)}.$$

Under the null hypothesis, there is only one regime. We define a likelihood ratio test statistic as

$$J_T = \max_{\gamma \in \Omega} (T - p) \frac{\tilde{\sigma}^2 - \hat{\sigma}^2(\gamma)}{\hat{\sigma}^2(\gamma)}. \quad (10)$$

$(T - p)\tilde{\sigma}^2$  is the residual sum of squares under the null hypothesis, while  $(T - p)\hat{\sigma}^2(\gamma)$  is the residual sum of squares under the alternatives. If  $H_0$  cannot be rejected, then the model is a simple AR model. Rejection of the null hypothesis suggests the existence of more than one regimes. The threshold estimator is defined as  $\hat{\gamma} = \arg \min \hat{\sigma}^2(\gamma) = \arg \max J_T(\gamma)$ . Since  $\gamma$  is not identified under the null hypothesis, the asymptotic distribution of  $J_T(\hat{\gamma})$  is not a standard  $\chi^2$ . Hansen (1996) shows that the asymptotic distribution can be approximated by the following bootstrap procedure.

Let  $u_t^*$  ( $t = 1, \dots, T$ ) be i.i.d.  $N(0, 1)$ , and set  $y_t^* = u_t^*$ . Next, we regress  $y_t^*$  on  $x_t = (1, y_{t-1}^*, y_{t-2}^*, \dots, y_{t-p}^*)$  to obtain the  $J_T^*(\gamma) = (T - p) \frac{\tilde{\sigma}^{*2} - \hat{\sigma}^{*2}(\gamma)}{\hat{\sigma}^{*2}(\gamma)}$  and  $J_T^* = \max_{\gamma \in \Omega} J_T^*(\gamma)$ .

The distribution of  $J_T^*$  converges weakly in probability to the distribution of  $J_T$  under the null hypothesis. Therefore, one can use the bootstrap value of  $J_T^*$  to approximate the asymptotic null distribution of  $J_T$ . The percentage of draws where the simulated statistic under  $H_0$  exceeds the one obtained from the original sample is our bootstrapping  $p$ -value. The null hypothesis will be rejected if the  $p$ -value is small.

Rejection of the null hypothesis implies the presence of threshold effects. To determine the number of regimes, a general-to-specific approach is adopted. First, a three-regime model is tested against a four-regime model. Each of the following hypotheses

- (I)  $H_0 : \beta^{(1)} = \beta^{(2)}$ ;
- (II)  $H_0 : \beta^{(1)} = \beta^{(3)}$ ;
- (III)  $H_0 : \beta^{(1)} = \beta^{(4)}$ ;
- (IV)  $H_0 : \beta^{(2)} = \beta^{(3)}$ ;



- (V)  $H_0 : \beta^{(2)} = \beta^{(4)}$ ;  
 (VI)  $H_0 : \beta^{(3)} = \beta^{(4)}$ .

is tested against the alternative hypothesis

$$H_1 : \text{there are four regimes.}$$

A likelihood ratio test

$$J_T(\hat{\gamma}) = (T - p) \frac{\hat{\sigma}_0^2(\hat{\gamma}) - \hat{\sigma}_1^2(\hat{\gamma})}{\hat{\sigma}_1^2(\hat{\gamma})} \quad (11)$$

is used to test these pairs of hypotheses, where  $(T - p)\hat{\sigma}_0^2(\hat{\gamma})$  is the residual sum of squares under  $H_0$ , and  $(T - p)\hat{\sigma}_1^2(\hat{\gamma})$  is the residual sum of squares under  $H_1$ . A parametric bootstrap method is applied to obtain the critical value.  $\hat{\gamma}$  is the estimated value from the unrestricted model. Let  $y_t^* = \sum_{j=1}^4 (\hat{\beta}_0^{(j)} + \sum_{i=1}^p \hat{\beta}_i^{(j)} y_{t-i}^*) \Psi_t^{(j)}(\hat{\gamma}) + u_t^*$ , where  $u_t^*$  are i.i.d.  $N(0, 1)$  and  $\hat{\beta}_i^{(j)}$ 's are estimated under the restricted model. We regress  $y_t^*$  on  $x_t = (1, y_{t-1}^*, y_{t-2}^*, \dots, y_{t-p}^*)$  to obtain  $J_T^*(\hat{\gamma}) = (T - p) \frac{\hat{\sigma}_0^{*2}(\hat{\gamma}) - \hat{\sigma}_1^{*2}(\hat{\gamma})}{\hat{\sigma}_1^{*2}(\hat{\gamma})}$ , and repeat this procedure a large number of times to calculate the percentage of draws for which the simulated statistic exceeds the actual value. The null is rejected if this  $p$ -value is too small.

Rejection of all the null hypotheses (I)–(VI) implies the existence of four regimes. If any one of them is accepted, then there are less than four regimes, and we proceed to test a two-regime model against a three-regime model. For instance, if (I)  $H_0 : \beta^{(1)} = \beta^{(2)}$  is accepted, then there are at most three regimes, and we proceed to test the two-regime model against the three-regime model. The following three hypotheses are tested using  $J_T(\hat{\gamma})$ :

$$\begin{aligned} H_0 : \beta^{(1)} &= \beta^{(2)} = \beta^{(3)}; \\ H_0 : \beta^{(1)} &= \beta^{(2)} = \beta^{(4)}; \\ H_0 : \beta^{(1)} &= \beta^{(2)}, \quad \beta^{(3)} = \beta^{(4)}. \end{aligned}$$

The alternative hypothesis is:

$$H_1 : \text{There are three regimes with } \beta^{(1)} = \beta^{(2)}.$$

If all the above null hypotheses are rejected, we conclude that there are three regimes. Otherwise, we conclude that the model has two regimes. In empirical studies, one can estimate the autoregressive order, the threshold value, and the coefficients of the TAR model via the following procedure.

**Step 1:** First, a first-order TAR model is estimated:

$$y_t = \sum_{j=1}^4 (\hat{\beta}_0^{(j)} + \hat{\beta}_1^{(j)} y_{t-1}) \Psi_t^{(j)}(\gamma) + \hat{u}_t,$$

and the initial threshold estimate  $\hat{\gamma}_T$  is obtained.

The first-order model is estimated for simplicity purposes (Chong, 2001). The initial threshold estimate will still be consistent even if the true model is not of the first-order (Bai et al., 2008; Chong, 2003).<sup>12</sup>

**Step 2:** Given the threshold values obtained from step 1, we use the Akaike Information Criterion (AIC) (Tsay, 1998) to select the autoregressive order in each regime. In our case,

$$AIC_j(p_j) = n_j \ln[RSS_j(\hat{\gamma}_T)/n_j] + 2(p_j + 1), \quad (12)$$

where

$n_j$  is the number of observations in the  $j$ th regime;

$p_j$  is the order of autoregression in the  $j$ th regime;

$RSS_j(\hat{\gamma}_T)$  is the residual sum of squares for the  $j$ th regime.

Define

$$\hat{p}_j = \arg \min_{p_j \in \{1, 2, \dots, P_{\max}\}} AIC_j(p_j), \quad (13)$$

where  $P_{\max}$  is the maximum order considered in the model. The AIC for the whole model can be written as

$$NAIC = \sum_{j=1}^s AIC_j(\hat{p}_j), \quad (14)$$

where  $s$  is the number of regimes.

**Step 3:** Perform the sequential likelihood ratio test to determine the number of regimes.

**Step 4:** Use the result obtained from step 3 to refine the threshold values, and repeat steps 2 and 3 until all the estimates converge.

<sup>12</sup>The proof is available upon request.

## 5. SIMULATIONS

In the previous section, it is argued the threshold value can be consistently estimated even if we start with a misspecified model in step 1. This result is obtained by Chong (2003) and Bai et al. (2008). The following experiments examine the consistency of the threshold estimator under model misspecifications.

The experiment is set up as follows:

Sample size:  $T = 200$ ;

Number of replications:  $N = 500$ ;

$u_t \sim N(0, 1)$ ,  $\varepsilon_t \sim N(0, 1)$ ,  $z_{1t} \sim N(0, 1)$ ;

$P_{\max} = 10$ .

We consider two cases for  $z_{2t}$ : (i)  $z_{2t} \sim N(0, 1)$ , and (ii)  $z_{2t} = z_{1t} + \varepsilon_t$ .

The following data generating processes are examined:

DGP 1:  $y_t = (0.3y_{t-1} + 0.3y_{t-2})I(z_{1t} \leq 0 \text{ or } z_{2t} \leq 0) + (-0.3y_{t-1} - 0.3y_{t-2})I(z_{1t} > 0 \text{ and } z_{2t} > 0) + u_t$ ;

DGP 2:  $y_t = 0.3y_{t-1}I(z_{1t} \leq 0 \text{ or } z_{2t} \leq 0) - 0.3y_{t-1}I(z_{1t} > 0 \text{ and } z_{2t} > 0) + u_t$ .

Three misspecified models are estimated:

Model A:  $y_t = \sum_{j=1}^4 \hat{\beta}_1^{(j)} y_{t-1} \Psi_t^{(j)}(\gamma) + \hat{u}_t$ ;

Model B:  $y_t = \sum_{j=1}^4 (\hat{\beta}_1^{(j)} y_{t-1} + \hat{\beta}_2^{(j)} y_{t-2}) \Psi_t^{(j)}(\gamma) + \hat{u}_t$ ;

Model C:  $y_t = \hat{\beta}_1^{(1)} y_{t-1} \mathbf{1}(z_{1t} \leq \gamma_1) + \hat{\beta}_1^{(2)} y_{t-1} \mathbf{1}(z_{1t} > \gamma_1) + \hat{u}_t$ .

Model A underestimates the autoregressive order  $p$ , while Model B overestimates the autoregressive order  $p$ . Both of them overestimate the number of regimes. The estimation results are reported in Table 1. For all misspecified estimated models,  $\hat{\gamma}_1$  and  $\hat{\gamma}_2$  converge to the true threshold value 0. The results for models A and B suggest that the consistency of the threshold estimators is unaffected by the misspecification of regressors. Therefore, if the number of threshold variables is known, one can obtain a preliminary and consistent threshold estimate using the simplest model possible. The preliminary estimate of the threshold value can be used to obtain the estimates of other parameters of interest. In the context of our model, such a preliminary threshold value allows us to determine the number of regimes, as well as the order and parameters of the autoregressive model within each regime.

The results of Chong (2003) and Bai et al. (2008) apply to cases where the threshold variables are correctly specified. Model C underspecifies the number of threshold variables. The results for Model C show that the

**TABLE 1** The simulation results

DGP	Estimated model	$z_{2t}$	$\hat{\gamma}_1$	$Var(\hat{\gamma}_1)$	$\hat{\gamma}_2$	$Var(\hat{\gamma}_2)$
1	A	$N(0, 1)$	0.007	0.030	-0.002	0.025
1	A	$z_{1t} + \varepsilon_t$	-0.002	0.036	-0.003	0.029
2	B	$N(0, 1)$	-0.001	0.025	-0.005	0.034
2	B	$z_{1t} + \varepsilon_t$	-0.006	0.025	0.002	0.027
2	C	$N(0, 1)$	-0.015	0.68		
2	C	$z_{1t} + \varepsilon_t$	-0.19	0.28		

estimators of the single threshold-variable model may not be consistent in the presence of two dependent threshold variables.

## 6. EMPIRICAL APPLICATION

Our model is applied to the daily return series of the Hang Seng Index. The Hong Kong stock market is studied because of its rising role as a global financial center. In 2006, Hong Kong becomes the world's second most popular place for IPO after London. In 2007, the Hong Kong stock market ranks fifth in the world, and its warrant market ranks first globally in terms of turnover. Most of the previous studies in the literature use the first lagged return as the threshold variable to identify the market regimes. Such a classification method does not take investors' sentiment into account and does not consider the information of market turnover. In this article, we use the past information of price and market turnover to construct our threshold variables. Our sample period runs from January 3 1995 to January 13, 2005. The return series is defined as the log-difference of the Hang Seng Index (HSI). There are over 2500 observations in our sample. Figures 1 and 2 show the time series data for daily return and market turnover.

The two threshold variables are analogous to those of Granville (1963) and Lee and Swaminathan (2000). We define the first threshold variable as

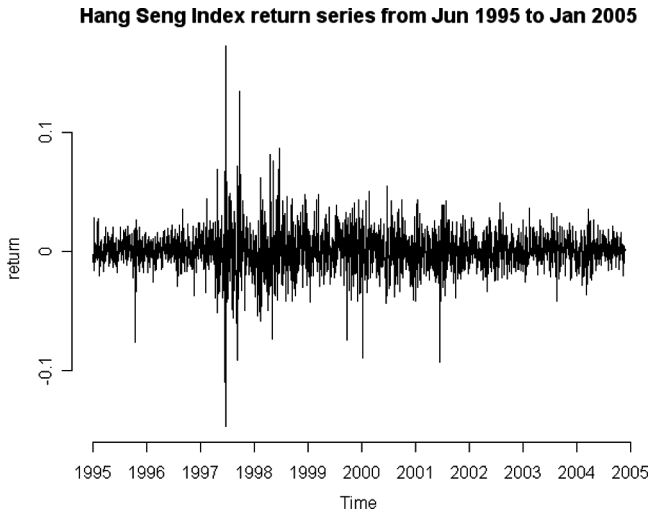
$$Rap_t = \frac{PMA20_t}{PMA250_t}, \quad (15)$$

where

$$PMA250_t = \frac{\sum_{j=1}^{250} p_{t-j}}{250}, \quad PMA20_t = \frac{\sum_{j=1}^{20} p_{t-j}}{20}.$$

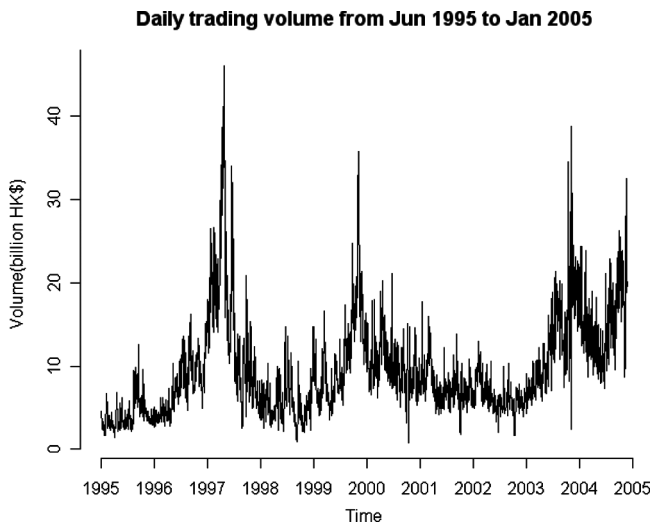
$PMA250_t$  is the average price for the past 250 trading days.

$PMA20_t$  is the average price for the past 20 trading days.



**FIGURE 1** Hand Seng index return series.

The variable is a ratio of two moving averages, which is similar to that of Hong and Lee (2003). In particular, the 250-day moving average, which is widely used by investors to define the market state, is employed. If the price rises above (falls below) the 250-day moving average, an average investor who has taken a long position in the previous year (about 250 trading days) has made a profit (loss), suggesting that the market sentiment should be good (bad). To reduce noise, we use the crossing



**FIGURE 2** Hand Seng index daily trading volume.

of the 20-day and 250-day moving averages to help identify the market regimes.

The second threshold variable contains the information of the market turnover, which has been widely used to measure the liquidity of the market, see Amihud and Mendelson (1986), Brennan et al. (1998), and Amihud (2002) among others. Several studies have shown that the autocorrelation in stock returns is related to turnover or trading volume. For example, Campbell et al. (1993) find that the first-order daily return autocorrelation tends to decline with turnover, and the returns accompanied by high volume tend to be reversed more strongly. Llorente et al. (2002) point out that intensive trading volume can help to identify the periods in which shocks occur. Therefore, we define the second threshold variable as

$$Rav_t = \log(\text{turnover}_{t-1}) - MAV_{t-1}, \quad (16)$$

where

$$MAV_t = \frac{\sum_{j=1}^{250} \log(\text{turnover}_{t-j})}{250}.$$

Figure 3 shows the two threshold variables  $Rap_t$  and  $Rav_t$ .

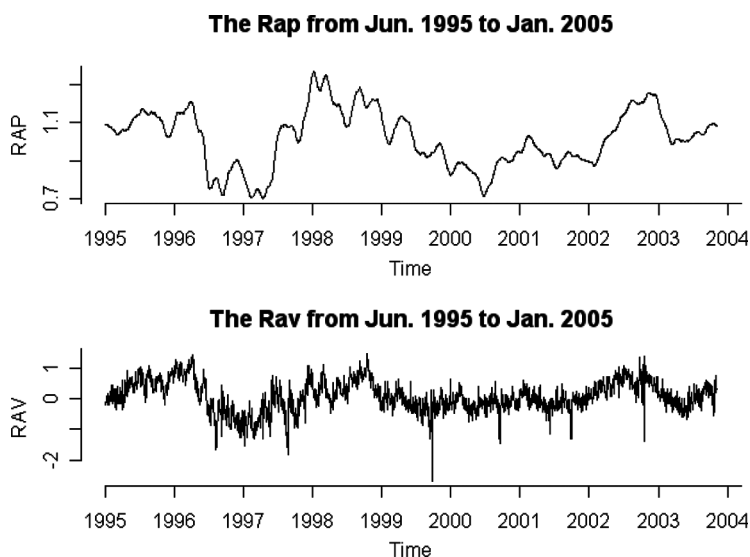


FIGURE 3 The two threshold variables.

**TABLE 2a** Results of the LR test for 4 regimes vs 3 regimes

$H_1 : s = 4$	$\beta^{(i)} \neq \beta^{(j)}$ when $i \neq j$					
$H_0 : s = 3$	$\beta^{(2)} = \beta^{(1)}$	$\beta^{(3)} = \beta^{(2)}$	$\beta^{(4)} = \beta^{(1)}$	$\beta^{(3)} = \beta^{(1)}$	$\beta^{(2)} = \beta^{(4)}$	$\beta^{(3)} = \beta^{(4)}$
$J_T(\hat{\gamma})$	126.7	136.3	5.45	16.59	93.05	18.79
$p$ -value	< 0.01	< 0.01	> 0.05	< 0.05	< 0.01	< 0.05

**TABLE 2b** Results of the LR test for 3 regimes vs. 2 regimes

$H_1 : s = 3$	$\beta^{(4)} = \beta^{(1)}$		
$H_0 : s = 2$	$\beta^{(2)} = \beta^{(3)}, \beta^{(4)} = \beta^{(1)}$	$\beta^{(1)} = \beta^{(2)} = \beta^{(4)}$	$\beta^{(3)} = \beta^{(4)} = \beta^{(1)}$
$J_T(\hat{\gamma})$	107.9	159.5	18.77
$p$ -value	< 0.01	< 0.01	< 0.05

Our four-regime threshold model on the return series is

$$y_t = \sum_{j=1}^4 \Psi_t^{(j)}(\gamma^0) (\beta_0^{(j)} + \beta_1^{(j)} y_{t-1} + \beta_2^{(j)} y_{t-2} + \dots + \beta_p^{(j)} y_{t-p}) + u_t, \quad (17)$$

where:

$y_t$  is the return series defined as the log-difference of the HSI;

$$\begin{aligned} \Psi_t^{(1)}(\gamma^0) &= I(Rap_t \leq \gamma_1^0, Rav_t \leq \gamma_2^0); \\ \Psi_t^{(2)}(\gamma^0) &= I(Rap_t \leq \gamma_1^0, Rav_t > \gamma_2^0); \\ \Psi_t^{(3)}(\gamma^0) &= I(Rap_t > \gamma_1^0, Rav_t \leq \gamma_2^0); \\ \Psi_t^{(4)}(\gamma^0) &= I(Rap_t > \gamma_1^0, Rav_t > \gamma_2^0). \end{aligned}$$

The estimated threshold values from step 1 in Section 3 are  $\hat{\gamma}_{rap} = 1.02$  and  $\hat{\gamma}_{rav} = 0.57$ . The results of the sequential likelihood ratio test are shown in Tables 2a and 2b.

Note from Table 2a that the null hypothesis  $\beta^{(4)} = \beta^{(1)}$  cannot be rejected since  $J_T(\hat{\gamma})$  has a  $p$ -value larger than 0.05.<sup>13</sup> Next, we proceed to test the 3-regime model against the 2-regime model. The results from Table 2b suggest that the movement of the return series can be approximated by a three-regime model.

Table 3 shows the final estimation results. The threshold estimates are revised to  $\hat{\gamma} = (1.02, 0.53)$ .

<sup>13</sup>In some cases, if two or more hypotheses cannot be rejected, we choose the one with the largest  $p$ -value as the candidate model in the subsequent step.

**TABLE 3** The estimated TAR model

Regime	Estimation results
<b>I</b>	$y_t = 0.0003 + 0.065y_{t-1}$ , if $Rap_t > 1.02$ and $Rav_t \leq 0.53$
<b>II</b>	$y_t = 0.0067 - 0.3y_{t-1} - 0.4y_{t-2} + 0.18y_{t-3} + 0.09y_{t-4} - 0.12y_{t-5} + 0.54y_{t-6}$ $- 0.5y_{t-7} - 0.18y_{t-8}$ , if $Rap_t \leq 1.02$ and $Rav_t > 0.53$
<b>III</b>	$y_t = 0.00014 + 0.096y_{t-1}$ otherwise

Figure 4 plots the estimated residuals of the model.<sup>14</sup>

Using the Markov-switching model, Maheu and McCurdy (2000) divide the stock market into a high-return stable regime and a low-return volatile regime. From Table 3, we are able to classify the stock market of Hong Kong into three regimes. Since high turnover is usually associated with volatile returns Karpoff (1987); Foster and Viswanathan (1995), Regime I generated by our model corresponds to the high-return stable regime, while Regime II is the low-return volatile regime.<sup>15</sup> Regime III is a neutral regime. Table 4 shows a chronology of major events affecting the Hong Kong stock market between 1996 and 2005.<sup>16</sup>

## 7. CONCLUSION

Conventional threshold models only allow for a single threshold variable. In many applications, the use of multiple threshold variables is needed. In this article, a new TAR model with two threshold variables is developed. In addition, the consistency and limiting distribution of the estimators are established. A likelihood ratio test is also constructed to detect the threshold effect. Our model is applied to identify the regimes of the Hong Kong stock market. The two threshold variables used in

<sup>14</sup>A Ljung–Box test has been conducted and the results suggest that the residuals are white noise. The details of the test can be obtained from the authors upon request.

<sup>15</sup>Note that the first-order coefficient for Regime I is 0.065, which is positive as compared to that of  $-0.3$  for Regime II. This agrees with Campbell et al. (1993) that the first-order daily return autocorrelation tends to decline when turnover increases.

<sup>16</sup>We associate the estimated regimes with these major events. For example, the establishment of the Hong Kong Special Administration Region in July 1997 falls into Regime I. During the Asian Financial Crisis, the crash of the stock market of Hong Kong and the burst of the property market fall into Regime II. The market experiences a volatile year in the millennium. Driven by the technology bubble and the accession of China to the World Trade Organization, the Hang Seng Index reaches a record high of 18301 in March 2000. However, the burst of the bubble in 2000 brings the stock market back into Regime II again. The market enters Regime I at the end of 2001. Note that the Hong Kong stock market is not seriously affected by the 911 incident. The accession of China to the World Trade Organization in 2001 is a good news for Hong Kong. In the beginning of 2003, the outbreak of the Severe Acute Respiratory Syndrome (SARS) threatens the economy. The market switches from Regime I to Regime II during the SARS period, but it rebounds sharply in the second half of the year. In June, China and Hong Kong sign the Closer Economic Partnership Arrangement (CEPA), a free trade agreement between Hong Kong and China which gives Hong Kong a preferential access to the Chinese market.



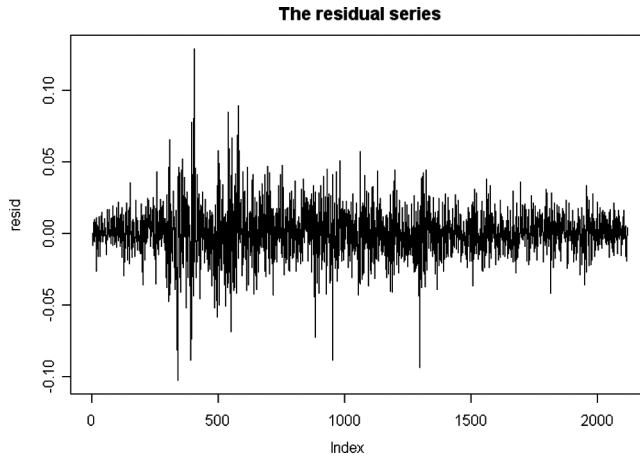


FIGURE 4 The residual series.

this article are analogous to those of Granville (1963) and Lee and Swaminathan (2000). Unlike the conventional bull–bear classification, it is shown that the Hong Kong stock market can be classified into three regimes, namely, a high-return stable regime, a low-return volatile regime, and a neutral regime. It should be mentioned that our model assumes a single threshold for each threshold variable. It can be extended to allow for the existence of multiple thresholds (Gonzalo and Pitarakis, 2002). For example, if there are two threshold variables and each threshold variable has two threshold values, then the model can have at most nine regimes. One may also define the threshold condition as a nonlinear function of the two threshold variables. Finally, one may relax the i.i.d. assumption of the error term to allow for serial dependence and regime-dependent heteroskedasticity. Such extensions, however, are beyond the scope of this article and are left for future research.

TABLE 4 A Chronology of the Hong Kong stock market and the corresponding regimes

Date	Event	Regime
1997.7	The establishment of the Hong Kong Special Administration Region	I
1997.10.23	Asian currency turmoil triggered by the floating of Thai Baht	II
1998.1–1999.3	The burst of the property market	III, II
1999.9–2000.3	Global technology stock boom and the admission of China into the World Trade Organization (WTO)	I, III
2000.4–2000.6	The burst of the high-tech bubble	II, III
2001.9.11	The 911 incident	I
2001.11.13	The accession of China to the WTO	I
2003.2–2003.6	The outbreak of SARS	I, II
2003.6.29	The launch of Closer Economic Partnership Arrangement with China	I

## APPENDIX 1: LEMMAS

Throughout the appendix, let  $\|A\| = (tr(A'A))^{1/2}$  denote the Euclidean norm of a matrix  $A$ . Let  $\|A\|_r = (E|A|^r)^{1/r}$  denote the  $L^r$ -norm of a random matrix and  $\Rightarrow$  denote weak convergence with respect to the uniform metric.

Let

$$\begin{aligned} x_t &= (1, y_{t-1}, y_{t-2}, \dots, y_{t-p})' \quad \text{for } t = p+1, p+2, \dots, T, \\ X &= (x'_T, x'_{T-1}, \dots, x'_{p+1})_{(T-p) \times (p+1)}, \\ Y &= (y_T, \dots, y_{p+1})', \\ U &= (u_T, u_{T-1}, \dots, u_{p+1})', \\ I_j(\gamma) &= \text{diag} \left\{ \Psi_T^{(j)}(\gamma), \Psi_{T-1}^{(j)}(\gamma), \dots, \Psi_{p+1}^{(j)}(\gamma) \right\}, \end{aligned}$$

where  $\Psi_t^{(j)}(\gamma)$  is defined in Section 2.

Let  $M$  and  $M_j(\gamma)$  be moment functionals defined as

$$M = E(x_t x'_t), \quad M_j(\gamma) = E\left(x_t x'_t \Psi_t^{(j)}(\gamma)\right), \quad j = 1, 2, 3, 4.$$

**Lemma 1.** Under assumptions (A1)–(A2), it can be shown that:

- (a)  $\frac{1}{T} X' X \xrightarrow{p} M$ ;
- (b)  $\frac{1}{T} X' U \xrightarrow{p} 0$ .

*Proof.* The proof is straightforward by applying the law of large number for stationary ergodic processes.  $\square$

**Lemma 2.** For any  $\gamma \in \Omega$ , under assumptions (A1)–(A3), we have for  $j = 1, 2, 3, 4$ :

- (a)  $\frac{1}{T} X' I_j(\gamma) X \xrightarrow{p} M_j(\gamma)$ ;
- (b)  $\frac{1}{T} X' I_j(\gamma) U \xrightarrow{p} 0$ ;
- (c)  $\frac{1}{T} (X' I_j(\gamma) U)' (X' I_j(\gamma) U) \xrightarrow{p} E(x_t x'_t u_t^2 \Psi_t^{(j)}(\gamma)) = \sigma^2 M_j(\gamma)$ .

*Proof.* The proof of part (a) for  $j = 1$  is similar to the proof of Lemma A1 in Hansen (1996) by replacing  $\{w_t \leq \gamma\}$  with  $\{z_{1t} \leq \gamma_1, z_{2t} \leq \gamma_2\}$ . For  $j = 2$ , we have  $\frac{1}{T} X' I_2(\gamma) X = \frac{1}{T} \sum x_t x'_t \{z_{1t} \leq \gamma_1\} - \frac{1}{T} \sum x_t x'_t \{z_{1t} \leq \gamma_1, z_{2t} \leq \gamma_2\} \xrightarrow{p} E(x_t x'_t \{z_{1t} \leq \gamma_1\}) - E(x_t x'_t \{z_{1t} \leq \gamma_1, z_{2t} \leq \gamma_2\}) = M_2(\gamma)$ . A similar proof can be applied to the cases where  $j = 3$  and 4. The proofs for (b) and (c) are analogous and are therefore skipped.  $\square$

## APPENDIX 2: CONSISTENCY OF ESTIMATORS

To prove the consistency of the estimator  $\hat{\gamma} = \arg \min_{\gamma \in \Omega} RSS_T(\gamma)$ , it suffices to show that  $RSS_T(\gamma)$  converges uniformly to a function  $b(\gamma)$  which is minimized at  $\gamma_0$ . For simplicity, denote  $\hat{\beta}^{(j)} = \hat{\beta}^{(j)}(\gamma)$  for  $j = 1, 2, 3, 4$ . Let  $\hat{Y}(\gamma) = \sum_{j=1}^4 I_j(\gamma) X \hat{\beta}^{(j)}$ . The residual sum of squares can be written as

$$\begin{aligned} RSS_T(\gamma) &= \|Y - \hat{Y}(\gamma)\|^2 = Y'Y - \hat{Y}(\gamma)' \hat{Y}(\gamma) \\ &= \sum_{j=1}^4 \left( \beta^{(j)'} X' I_j(\gamma^0) X \beta^{(j)} - \hat{\beta}^{(j)'} X' I_j(\gamma) X \hat{\beta}^{(j)} \right) \\ &\quad + 2 \sum_{j=1}^4 U' I_j(\gamma^0) X \beta^{(j)} + U' U. \end{aligned}$$

Next, we prove that  $RSS_T(\gamma)$  has a unique minimum at  $\gamma = \gamma^0$ . We partition the threshold space into four regions.

*Case 1:*  $\gamma_1 \leq \gamma_1^0$  and  $\gamma_2 \leq \gamma_2^0$

Using Lemmas 1 and 2, and the facts that

$$\begin{aligned} I_1(\gamma) I_1(\gamma^0) &= I_1(\gamma), & I_1(\gamma) I_j(\gamma^0) &= 0 \quad \text{for } j = 2, 3, 4, \\ I_2(\gamma) I_1(\gamma^0) &= I_2(\gamma) - I_2(\gamma_1, \gamma_2^0), & I_2(\gamma) I_2(\gamma^0) &= I_2(\gamma_1, \gamma_2^0), \\ I_2(\gamma) I_j(\gamma^0) &= 0, & \text{for } j &= 3, 4, \\ I_3(\gamma) I_1(\gamma^0) &= I_3(\gamma) - I_3(\gamma_1^0, \gamma_2), & I_3(\gamma) I_2(\gamma^0) &= 0, \\ I_3(\gamma) I_3(\gamma^0) &= I_3(\gamma_1^0, \gamma_2), & I_3(\gamma) I_4(\gamma^0) &= 0, \\ I_4(\gamma) I_1(\gamma^0) &= I_1(\gamma^0) + I_1(\gamma) - I_1(\gamma_1^0, \gamma_2) - I_1(\gamma_1, \gamma_2^0), \\ I_4(\gamma) I_2(\gamma^0) &= 0, & I_4(\gamma) I_3(\gamma^0) &= I_4(\gamma_1^0, \gamma_2) - I_4(\gamma^0), \\ I_4(\gamma) I_4(\gamma^0) &= I_4(\gamma^0), \end{aligned}$$

it can be shown that

$$\begin{aligned} \hat{\beta}^{(1)} &= (X' I_1(\gamma) X)^{-1} X' I_1(\gamma) Y \\ &= (X' I_1(\gamma) X)^{-1} X' I_1(\gamma) \left[ \sum_{j=1}^4 I_j(\gamma^0) X \beta^{(j)} + U \right] \\ &= \beta^{(1)} + \frac{1}{\sqrt{T}} \left( \frac{X' I_1(\gamma) X}{T} \right)^{-1} \left( \frac{X' I_1(\gamma) U}{\sqrt{T}} \right) \xrightarrow{p} \beta^{(1)}, \\ \hat{\beta}^{(2)} &= (X' I_2(\gamma) X)^{-1} X' I_2(\gamma) Y \\ &\xrightarrow{p} M_2^{-1}(\gamma) (M_2(\gamma) - M_2(\gamma_1, \gamma_2^0)) (\beta^{(1)} - \beta^{(2)}) + \beta^{(2)}, \end{aligned}$$

$$\begin{aligned}\hat{\beta}^{(3)} &= (X' I_3(\gamma) X)^{-1} X' I_3(\gamma) Y \\ &\xrightarrow{p} M_3^{-1}(\gamma) (M_3(\gamma) - M_3^0(\gamma_1^0, \gamma_2)) (\beta^{(1)} - \beta^{(2)}) \\ &\quad + M_3^{-1}(\gamma) M_3(\gamma_1^0, \gamma_2) (\beta^{(3)} - \beta^{(2)}) + \beta^{(2)},\end{aligned}$$

$$\begin{aligned}\hat{\beta}^{(4)} &= (X' I_4(\gamma) X)^{-1} X' I_4(\gamma) Y \\ &\xrightarrow{p} M_4^{-1}(\gamma) [M_4(\gamma^0) - M_4(\gamma_1^0, \gamma_2) - M_4(\gamma_1, \gamma_2^0) + M_4(\gamma)] (\beta^{(1)} - \beta^{(2)}) \\ &\quad + M_4^{-1}(\gamma) M_4(\gamma^0) (\beta^{(4)} - \beta^{(2)}) \\ &\quad + M_4^{-1}(\gamma) (M_4(\gamma_1^0, \gamma_2) - M_4(\gamma^0)) (\beta^{(3)} - \beta^{(2)}) + \beta^{(2)}.\end{aligned}$$

Therefore,

$$\begin{aligned}&\frac{1}{T} (RSS_T(\gamma) - U' U) \\ &= \frac{1}{T} \sum_{j=1}^4 \left( \beta^{(j)'} X' I_j(\gamma^0) X \beta^{(j)} - \hat{\beta}^{(j)'} X' I_j(\gamma) X \hat{\beta}^{(j)} \right) + \frac{2}{T} \sum_{j=1}^4 U' I_j(\gamma^0) X \beta^{(j)} \\ &= \sum_{j=1}^4 \beta^{(j)'} M_j(\gamma^0) (\beta^{(j)} - \beta^{(2)}) \\ &\quad - \left[ \hat{\beta}^{(1)'} M_1(\gamma) + \hat{\beta}^{(2)'} (M_2(\gamma) - M_2(\gamma_1, \gamma_2^0)) \right] (\beta^{(1)} - \beta^{(2)}) \\ &\quad + \hat{\beta}^{(3)'} [M_3(\gamma) - M_3(\gamma_1^0, \gamma_2)] (\beta^{(1)} - \beta^{(2)}) \\ &\quad + \hat{\beta}^{(4)'} [M_4(\gamma^0) - M_4(\gamma_1^0, \gamma_2) - M_4(\gamma_1, \gamma_2^0) + M_4(\gamma)] (\beta^{(1)} - \beta^{(2)}) \\ &\quad - \left[ \hat{\beta}^{(3)'} M_3(\gamma_1^0, \gamma_2) + \hat{\beta}^{(4)'} (M_4(\gamma_1^0, \gamma_2) - M_4(\gamma^0)) \right] (\beta^{(3)} - \beta^{(2)}) \\ &\quad - \hat{\beta}^{(4)'} M_4(\gamma^0) (\beta^{(4)} - \beta^{(2)}) + o_p(1) \\ &= (\beta^{(1)} - \beta^{(2)})' [M_1(\gamma^0) - M_1(\gamma) - M_2^{-1}(\gamma) (M_2(\gamma) - M_2(\gamma_1, \gamma_2^0))^2 \\ &\quad - M_3^{-1}(\gamma) (M_3(\gamma) - M_3(\gamma_1^0, \gamma_2))^2 \\ &\quad - M_4^{-1}(\gamma) (M_4(\gamma^0) - M_4(\gamma_1^0, \gamma_2) - M_4(\gamma_1, \gamma_2^0) + M_4(\gamma))^2] \times (\beta^{(1)} - \beta^{(2)}) \\ &\quad + (\beta^{(3)} - \beta^{(2)})' [M_3(\gamma^0) - M_4^{-1}(\gamma) (M_4(\gamma_1^0, \gamma_2) - M_4(\gamma^0))^2 \\ &\quad - M_3^{-1}(\gamma) (M_3(\gamma_1^0, \gamma_2))^2] (\beta^{(3)} - \beta^{(2)}) \\ &\quad + (\beta^{(4)} - \beta^{(2)})' [M_4(\gamma^0) - M_4^{-1}(\gamma) (M_4(\gamma^0))^2] (\beta^{(4)} - \beta^{(2)}) + o_p(1) \\ &= (\beta^{(1)} - \beta^{(2)})' Q_1 (\beta^{(1)} - \beta^{(2)}) + (\beta^{(3)} - \beta^{(2)})' Q_2 (\beta^{(3)} - \beta^{(2)}) \\ &\quad + (\beta^{(4)} - \beta^{(2)})' Q_3 (\beta^{(4)} - \beta^{(2)}) + o_p(1) \\ &= b_1(\gamma) + o_p(1).\end{aligned}$$

For any  $\gamma_1 \leq \gamma_1^0$  and  $\gamma_2 \leq \gamma_2^0$ , it is obvious that  $Q_3$  is positive semidefinite since  $M_4(\gamma) > M_4(\gamma^0)$ . Meanwhile, using the following results:

$$\begin{aligned} M_1(\gamma^0) - M_1(\gamma) &= E(x_t x_t' \Psi_t^{(1)}(\gamma)) \\ &= E(x_t x_t' [\Psi_t^{(2)}(\gamma) - \Psi_t^{(2)}(\gamma_1, \gamma_2^0) + \Psi_t^{(3)}(\gamma) \\ &\quad - \Psi_t^{(3)}(\gamma_1^0, \gamma_2) + \Psi_t^{(4)}(\gamma^0) - \Psi_t^{(4)}(\gamma_1^0, \gamma_2) \\ &\quad - \Psi_t^{(4)}(\gamma_1, \gamma_2^0) + \Psi_t^{(4)}(\gamma)]) \\ &= M_2(\gamma) - M_2(\gamma_1, \gamma_2^0) + M_3(\gamma) - M_3(\gamma_1^0, \gamma_2) \\ &\quad + M_4(\gamma^0) - M_4(\gamma_1^0, \gamma_2) - M_4(\gamma_1, \gamma_2^0) + M_4(\gamma), \\ M_3(\gamma^0) &= M_4(\gamma_1^0, \gamma_2) - M_4(\gamma^0) + M_3(\gamma_1^0, \gamma_2), \end{aligned}$$

it can be shown that  $Q_1$  and  $Q_2$  are positive semidefinite. Thus,  $b_1(\gamma) \geq b_1(\gamma^0) = 0$ , and the equation holds if and only if  $\gamma = \gamma^0$ .

By analogy, for the remaining three cases,

$$\begin{aligned} \frac{1}{T} (RSS_T(\gamma) - U'U) &= b_j(\gamma) + o_p(1) \quad \text{and} \\ b_j(\gamma) &\geq b_j(\gamma^0) = 0 \quad \text{for } j = 2, 3, 4. \end{aligned}$$

Define a non-stochastic function  $b(\gamma)$  as  $b_j(\gamma)$  for the  $j$ th case, we have

$$\sup_{\gamma \in \Omega} \left| \frac{1}{T} (RSS_T(\gamma) - U'U) - b(\gamma) \right| = o_p(1). \quad (18)$$

Thus,  $b(\gamma)$  is minimized if and only if  $\gamma = \gamma^0$ . This implies that the limit of  $\frac{1}{T} RSS_T(\gamma)$  is minimized at  $\gamma^0$ . By the superconsistency of  $\hat{\gamma}$ ,  $\hat{\beta}^{(j)}$  will also be consistent.

### APPENDIX 3: THE LIMITING DISTRIBUTION OF $\hat{\gamma}$

To derive the limiting distribution of  $\hat{\gamma}$  for shrinking break, we let  $\delta = (\delta^{(2)'}, \delta^{(3)'}, \delta^{(4)'})' = cT^{-\alpha}$ ,  $0 < \alpha < \frac{1}{2}$ ,  $c = (c'_2, c'_3, c'_4)'$  is a  $3p$ -dimensional vector of constants. Define

$$\hat{\gamma} = \arg \min_{\gamma \in \Omega} RSS_T(\gamma) = \arg \min_{\gamma \in \Omega} [RSS_T(\gamma) - RSS_T(\gamma^0)].$$

To obtain the limiting distribution of  $\hat{\gamma}$ , we first examine the asymptotic behavior of  $RSS_T(\gamma) - RSS_T(\gamma^0)$  in the neighborhood of the true thresholds.

Recall from Eq. (7) that the true model can be written as

$$Y = X\beta^{(1)} + \sum_{j=2}^4 X_{\gamma_0}^{(j)} \delta^{(j)} + U = X\beta^{(1)} + X_0\delta + U,$$

where  $X_0 = (X_{\gamma_0}^{(2)}, X_{\gamma_0}^{(3)}, X_{\gamma_0}^{(4)})$ . Let

$$\begin{aligned} \hat{\delta}^{(j)} &= \hat{\delta}^{(j)}(\gamma), & \hat{\delta}_0^{(j)} &= \hat{\delta}^{(j)}(\gamma^0), & \text{for } j = 2, 3, 4, \\ \hat{\delta}(\gamma) &= (\hat{\delta}^{(2)'}, \hat{\delta}^{(3)'}, \hat{\delta}^{(4)'})' & \text{and} & & \hat{\delta}(\gamma^0) = (\hat{\delta}_0^{(2)'}, \hat{\delta}_0^{(3)'}, \hat{\delta}_0^{(4)'})'. \end{aligned}$$

For any  $\gamma$ , define  $X_\gamma = (X_\gamma^{(2)}, X_\gamma^{(3)}, X_\gamma^{(4)})$ . We have

$$\begin{aligned} \hat{\beta}^{(1)}(\gamma) &= (X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} X_\gamma^{(1)'} Y \\ &= \beta^{(1)} + (X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} X_\gamma^{(1)'} X_0 \delta + (X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} X_\gamma^{(1)'} U, \\ \hat{\beta}^{(1)}(\gamma^0) &= (X_{\gamma_0}^{(1)'} X_{\gamma_0}^{(1)})^{-1} X_{\gamma_0}^{(1)'} Y \\ &= \beta^{(1)} + (X_{\gamma_0}^{(1)'} X_{\gamma_0}^{(1)})^{-1} X_{\gamma_0}^{(1)'} U. \end{aligned}$$

Since  $\hat{\gamma}$  is a consistent estimator, we study its asymptotic behavior in the neighborhood of the true thresholds. Let  $\gamma_1 = \gamma_1^0 + \frac{v}{T^{1-2\alpha}}$ ,  $\gamma_2 = \gamma_2^0 + \frac{\omega}{T^{1-2\alpha}}$ . By Lemmas 1 and 2,

$$\begin{aligned} \hat{\beta}^{(1)}(\gamma) - \hat{\beta}^{(1)}(\gamma^0) &= (X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} X_\gamma^{(1)'} X_0 \delta + (X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} X_\gamma^{(1)'} U \\ &\quad - (X_{\gamma_0}^{(1)'} X_{\gamma_0}^{(1)})^{-1} X_{\gamma_0}^{(1)'} U \\ &= \sum_{j=2}^4 (X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} X_\gamma^{(1)'} (X_0^{(j)} - X_{\gamma_0}^{(j)}) \delta \\ &\quad + (X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} (X_\gamma^{(1)'} U - X_{\gamma_0}^{(1)'} U) \\ &\quad + ((X_\gamma^{(1)'} X_\gamma^{(1)})^{-1} - (X_{\gamma_0}^{(1)'} X_{\gamma_0}^{(1)})^{-1}) X_{\gamma_0}^{(1)'} U \\ &= O_p\left(\frac{1}{T^{1-2\alpha}} \frac{1}{T^\alpha}\right) + O_p\left(\frac{1}{T^{1/2-\alpha}} \frac{1}{T^{1/2}}\right) + O_p\left(\frac{1}{T^{1-2\alpha}} \frac{1}{T^{1/2}}\right) \\ &= O_p\left(\frac{1}{T^{1-\alpha}}\right). \end{aligned}$$

By the  $\sqrt{T}$  consistency of the Ordinary Least Square (OLS) estimator, we have

$$\hat{\beta}^{(1)}(\gamma^0) - \beta^{(1)} = O_p\left(\frac{1}{T^{1/2}}\right)$$

and

$$\begin{aligned}\hat{\beta}^{(1)}(\gamma) - \beta^{(1)} &= \left(\hat{\beta}^{(1)}(\gamma) - \hat{\beta}^{(1)}(\gamma^0)\right) + \left(\hat{\beta}^{(1)}(\gamma^0) - \beta^{(1)}\right) \\ &= O_p\left(\frac{1}{T^{1-\alpha}}\right) + O_p\left(\frac{1}{T^{1/2}}\right) = O_p\left(\frac{1}{T^{1/2}}\right).\end{aligned}\quad (19)$$

Moreover, since

$$\hat{\delta}(\gamma) = (X'_\gamma X_\gamma)^{-1} X'_\gamma (X_0 \delta + U) = (X'_\gamma X_\gamma)^{-1} X'_\gamma X_0 \delta + (X'_\gamma X_\gamma)^{-1} X'_\gamma U$$

and

$$\hat{\delta}(\gamma^0) = (X'_0 X_0)^{-1} X'_0 (X_0 \delta + U) = \delta + (X'_0 X_0)^{-1} X'_0 U,$$

we have

$$\begin{aligned}\hat{\delta}(\gamma) - \hat{\delta}(\gamma^0) &= (X'_\gamma X_\gamma)^{-1} X'_\gamma (X_0 - X_\gamma) \delta + (X'_\gamma X_\gamma)^{-1} X'_\gamma U - (X'_0 X_0)^{-1} X'_0 U \\ &= O_p\left(\frac{1}{T^{1-2\alpha}} T^{-\alpha}\right) + O_p\left(\frac{1}{T^{1/2-2\alpha}} \frac{1}{T^{1/2}}\right) + O_p\left(\frac{1}{T^{1-2\alpha}} T^{-1/2}\right) \\ &= O_p\left(\frac{1}{T^{1-\alpha}}\right).\end{aligned}$$

We also have

$$\hat{\delta}(\gamma^0) - \delta = (X'_0 X_0)^{-1} X'_0 U = O_p\left(\frac{1}{T^{1/2}}\right).$$

Therefore,

$$\begin{aligned}\hat{\delta}(\gamma) - \delta &= \left(\hat{\delta}(\gamma) - \hat{\delta}(\gamma^0)\right) + \left(\hat{\delta}(\gamma^0) - \delta\right) \\ &= O_p\left(\frac{1}{T^{1-\alpha}}\right) + O_p\left(\frac{1}{T^{1/2}}\right) = O_p\left(\frac{1}{T^{1/2}}\right).\end{aligned}\quad (20)$$

By (19) and (20), we have

$$\begin{aligned}RSS_T(\gamma) - RSS_T(\gamma^0) &= \left(Y - X\hat{\beta}^{(1)}(\gamma) - X_\gamma \hat{\delta}(\gamma)\right)' \left(Y - X\hat{\beta}^{(1)}(\gamma) - X_\gamma \hat{\delta}(\gamma)\right) \\ &\quad - \left(Y - X\hat{\beta}^{(1)}(\gamma^0) - X_0 \hat{\delta}(\gamma^0)\right)' \left(Y - X\hat{\beta}^{(1)}(\gamma^0) - X_0 \hat{\delta}(\gamma^0)\right) \\ &= \left(Y - X\hat{\beta}^{(1)}(\gamma) - X_\gamma \hat{\delta}(\gamma)\right)' \left(Y - X\hat{\beta}^{(1)}(\gamma) - X_\gamma \hat{\delta}(\gamma)\right)\end{aligned}$$

$$\begin{aligned}
 & - \left( Y - X\hat{\beta}^{(1)}(\gamma) - X_0\hat{\delta}(\gamma) \right)' \left( Y - X\hat{\beta}^{(1)}(\gamma) - X_0\hat{\delta}(\gamma) \right) + o_p(1) \\
 = & -2\hat{\delta}(\gamma)'(X_\gamma - X_0)U + \hat{\delta}(\gamma)'(X_\gamma - X_0)'(X_\gamma - X_0)\hat{\delta}(\gamma) \\
 & + 2\hat{\delta}(\gamma)'(X_\gamma - X_0)'(X_\gamma - X_0)(\hat{\beta}^{(1)}(\gamma) - \beta^{(1)}) + o_p(1) \\
 = & \delta'(X_\gamma - X_0)'(X_\gamma - X_0)\delta + 2\hat{\delta}(\gamma)'(X_\gamma - X_0)'(X_\gamma - X_0)(\hat{\beta}^{(1)}(\gamma) - \beta^{(1)}) \\
 & - 2\hat{\delta}(\gamma)'(X_\gamma - X_0)U + (\delta + \hat{\delta}(\gamma))'(X_\gamma - X_0)'(X_\gamma - X_0)(\hat{\delta}(\gamma) - \delta) + o_p(1) \\
 = & T^{-2\alpha}c'(X_\gamma - X_0)'(X_\gamma - X_0)c - 2T^{-\alpha}U'(X_\gamma - X_0)c + O_p(T^{-1/2+\alpha}) + o_p(1) \\
 = & R_1 + R_2 + o_p(1),
 \end{aligned}$$

where

$$\begin{aligned}
 R_1 &= T^{1-2\alpha} \frac{[(X_\gamma - X_0) c]' (X_\gamma - X_0) c}{T} \\
 &= T^{1-2\alpha} \frac{1}{T} \sum_{t=p+1}^T \left\| \sum_{j=2}^4 c'_j x_t \left( \Psi_t^{(j)}(\gamma) - \Psi_t^{(j)}(\gamma^0) \right) \right\|^2, \tag{21}
 \end{aligned}$$

and

$$\begin{aligned}
 R_2 &= -2T^{-\alpha}U'(X_\gamma - X_0)c \\
 &= -2T^{-\alpha} \sum_{j=2}^4 \sum_{t=p+1}^T x'_t u_t \left( \Psi_t^{(j)}(\gamma) - \Psi_t^{(j)}(\gamma^0) \right) c_j. \tag{22}
 \end{aligned}$$

To examine the asymptotic behavior of  $RSS_T(\gamma) - RSS_T(\gamma^0)$ , we study the asymptotics of  $R_1$  and  $R_2$ . We consider four different cases and provide the proof for the case where  $v > 0$  and  $\omega > 0$ . The proofs for the other 3 cases are analogous.

Case 1:  $v > 0$  and  $\omega > 0$ :

$$\begin{aligned}
 & c'_2 x_t \left( \Psi_t^{(2)}(\gamma) - \Psi_t^{(2)}(\gamma^0) \right) \\
 &= c'_2 x_t \left( \Psi_t^{(2)}(\gamma_1, \gamma_2) - \Psi_t^{(2)}(\gamma_1^0, \gamma_2) + \Psi_t^{(2)}(\gamma_1^0, \gamma_2) - \Psi_t^{(2)}(\gamma_1^0, \gamma_2^0) \right) \\
 &= c'_2 x_t \left( I(\gamma_1^0 \leq z_{1t} < \gamma_1, z_{2t} > \gamma_2) - I(z_{1t} \leq \gamma_1^0, \gamma_2^0 \leq z_{2t} < \gamma_2) \right), \\
 & c'_3 x_t \left( \Psi_t^{(3)}(\gamma) - \Psi_t^{(3)}(\gamma^0) \right) \\
 &= c'_3 x_t \left( \Psi_t^{(3)}(\gamma_1, \gamma_2) - \Psi_t^{(3)}(\gamma_1^0, \gamma_2) + \Psi_t^{(3)}(\gamma_1^0, \gamma_2) - \Psi_t^{(3)}(\gamma_1^0, \gamma_2^0) \right) \\
 &= c'_3 x_t \left( -I(\gamma_1^0 \leq z_{1t} < \gamma_1, z_{2t} \leq \gamma_2) + I(z_{1t} > \gamma_1^0, \gamma_2^0 \leq z_{2t} < \gamma_2) \right),
 \end{aligned}$$



$$\begin{aligned}
& c'_4 x_t (\Psi_t^{(4)}(\gamma) - \Psi_t^{(4)}(\gamma^0)) \\
&= c'_4 x_t (\Psi_t^{(4)}(\gamma_1, \gamma_2) - \Psi_t^{(4)}(\gamma_1^0, \gamma_2) + \Psi_t^{(4)}(\gamma_1^0, \gamma_2) - \Psi_t^{(4)}(\gamma_1^0, \gamma_2^0)) \\
&= c'_4 x_t (-I(\gamma_1^0 \leq z_1 < \gamma_1, z_2 > \gamma_2) - I(z_1 > \gamma_1^0, \gamma_2^0 \leq z_2 < \gamma_2)).
\end{aligned}$$

Summing up the three terms, we have

$$\begin{aligned}
& \sum_{j=2}^4 c'_j x_t (\Psi_t^{(j)}(\gamma) - \Psi_t^{(j)}(\gamma^0)) \\
&= (c_2 - c_4)' x_t I(\gamma_1^0 \leq z_1 < \gamma_1, z_2 > \gamma_2) - c'_2 x_t I(z_{1t} \leq \gamma_1^0, \gamma_2^0 \leq z_{2t} < \gamma_2) \\
&\quad - c'_3 x_t I(\gamma_1^0 \leq z_{1t} < \gamma_1, z_{2t} \leq \gamma_2) + (c_3 - c_4)' x_t I(z_1 > \gamma_1^0, \gamma_2^0 \leq z_2 < \gamma_2).
\end{aligned}$$

Since the four terms are orthogonal, by Lemma 2, we have

$$\begin{aligned}
& \frac{1}{T} \sum_{t=p+1}^T \left\| \sum_{j=2}^4 c'_j x_t (\Psi_t^{(j)}(\gamma) - \Psi_t^{(j)}(\gamma^0)) \right\|^2 \\
&\xrightarrow{p} (c_2 - c_4)' (M_2(\gamma_1, \gamma_2) - M_2(\gamma_1^0, \gamma_2)) (c_2 - c_4) \\
&\quad + c'_2 ((M_1(\gamma_1^0, \gamma_2) - M_1(\gamma_1^0, \gamma_2^0)) c_2 + c'_3 (M_1(\gamma_1, \gamma_2) - M_1(\gamma_1^0, \gamma_2)) c_3) \\
&\quad + (c_3 - c_4)' (M_3(\gamma_1^0, \gamma_2) - M_3(\gamma_1^0, \gamma_2^0)) (c_3 - c_4).
\end{aligned}$$

By the continuity of  $D(\gamma)$  around  $\gamma^0$ , we can apply the first-order Taylor approximation to the moment functionals and obtain the following results:

$$\begin{aligned}
& (c_2 - c_4)' (M_2(\gamma_1, \gamma_2) - M_2(\gamma_1^0, \gamma_2)) (c_2 - c_4) \\
&\quad = |\gamma_1 - \gamma_1^0| (c_2 - c_4)' Df_1^0 (c_2 - c_4) + o(1), \\
& c'_2 ((M_1(\gamma_1^0, \gamma_2) - M_1(\gamma_1^0, \gamma_2^0)) c_2) = |\gamma_2 - \gamma_2^0| c'_2 Df_2^0 c_2 + o(1), \\
& c'_3 (M_1(\gamma_1, \gamma_2) - M_1(\gamma_1^0, \gamma_2)) c_3 = |\gamma_1 - \gamma_1^0| c'_3 Df_1^0 c_3 + o(1), \\
& (c_3 - c_4)' (M_3(\gamma_1^0, \gamma_2) - M_3(\gamma_1^0, \gamma_2^0)) (c_3 - c_4) \\
&\quad = |\gamma_2 - \gamma_2^0| (c_3 - c_4)' Df_2^0 (c_3 - c_4) + o(1),
\end{aligned}$$

where  $f_i^0 = \frac{\partial F(\gamma)}{\partial \gamma_i} |_{\gamma=\gamma^0}$  for  $i = 1, 2$  and  $D = E(x_t x_t' | z_t = \gamma_0)$ .

Thus,

$$\begin{aligned}
R_1 &= T^{1-2\alpha} \frac{1}{T} \sum_{t=p+1}^T \left\| \sum_{j=2}^4 c'_j x_t (\Psi_t^{(j)}(\gamma) - \Psi_t^{(j)}(\gamma^0)) \right\|^2 \\
&= T^{1-2\alpha} (|\gamma_1 - \gamma_1^0| (c_2 - c_4)' Df_1^0 (c_2 - c_4) + |\gamma_2 - \gamma_2^0| c'_2 Df_2^0 c_2)
\end{aligned}$$

$$\begin{aligned}
 & + |\gamma_1 - \gamma_1^0| c_3' D f_1^0 c_3 + |\gamma_2 - \gamma_2^0| (c_3 - c_4)' D f_2^0 (c_3 - c_4) + o_p(1) \\
 & = |v| d_1' D^* f_1^0 d_1 + |\omega| d_2' D^* f_2^0 d_2 + o_p(1), \tag{23}
 \end{aligned}$$

where  $D^* = \text{diag}\{D, D\}$ ,  $d_1 = ((c_2 - c_4)', c_3')'$ ,  $d_2 = (c_2', (c_3 - c_4)')'$ .

Next, we consider the asymptotic property of  $R_2$  for  $v > 0$  and  $\omega > 0$ :

$$\begin{aligned}
 & T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (\Psi_t^{(2)}(\gamma) - \Psi_t^{(2)}(\gamma^0)) c_2 \\
 & = -2T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (\Psi_t^{(2)}(\gamma_1, \gamma_2) - \Psi_t^{(2)}(\gamma_1^0, \gamma_2) + \Psi_t^{(2)}(\gamma_1^0, \gamma_2) - \Psi_t^{(2)}(\gamma_1^0, \gamma_2^0)) c_2 \\
 & = -2T^{-\alpha} \sum_{t=p+1}^T x_t' u_t ((\gamma_1^0 \leq z_{1t} < \gamma_1, z_{2t} > \gamma_2) - I(z_{1t} \leq \gamma_1^0, \gamma_2^0 \leq z_{2t} < \gamma_2)) c_2 \\
 & \Rightarrow -2(B_1(v) - B_2(\omega)) c_2; \\
 & T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (\Psi_t^{(3)}(\gamma) - \Psi_t^{(3)}(\gamma^0)) c_3 \\
 & = -2T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (\Psi_t^{(3)}(\gamma_1, \gamma_2) - \Psi_t^{(3)}(\gamma_1^0, \gamma_2) + \Psi_t^{(3)}(\gamma_1^0, \gamma_2) - \Psi_t^{(3)}(\gamma_1^0, \gamma_2^0)) c_3 \\
 & = -2T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (-I(\gamma_1^0 \leq z_{1t} < \gamma_1, z_{2t} \leq \gamma_2) + I(z_{1t} > \gamma_1^0, \gamma_2^0 \leq z_{2t} < \gamma_2)) c_3 \\
 & \Rightarrow -2(-B_3(v) + B_4(\omega)) c_3; \\
 & T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (\Psi_t^{(4)}(\gamma) - \Psi_t^{(4)}(\gamma^0)) c_4 \\
 & = -2T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (\Psi_t^{(4)}(\gamma_1, \gamma_2) - \Psi_t^{(4)}(\gamma_1^0, \gamma_2) + \Psi_t^{(4)}(\gamma_1^0, \gamma_2) - \Psi_t^{(4)}(\gamma_1^0, \gamma_2^0)) c_4 \\
 & = -2T^{-\alpha} \sum_{t=p+1}^T x_t' u_t (-I(\gamma_1^0 \leq z_{1t} < \gamma_1, z_{2t} > \gamma_2) - I(z_{1t} > \gamma_1^0, \gamma_2^0 \leq z_{2t} < \gamma_2)) c_4 \\
 & \Rightarrow -2(-B_1(v) - B_4(\omega)) c_4.
 \end{aligned}$$

Summing up the three terms, we have

$$\begin{aligned}
 R_2 & = -2T^{-\alpha} \sum_{j=2}^4 \sum_{t=p+1}^T x_t' u_t (\Psi_t^{(j)}(\gamma) - \Psi_t^{(j)}(\gamma^0)) c_j \\
 & \Rightarrow -2[B_1(v)(c_2 - c_4) - B_2(\omega)c_2 - B_3(v)c_3 + B_4(\omega)(c_3 - c_4)], \tag{24}
 \end{aligned}$$

where  $B_j(\cdot)$ ,  $j = 1, 2, 3, 4$ , are independent Brownian motion vectors corresponding to the four disjointed regions. The covariance matrix of  $B_j(\cdot)$  is given by

$$\begin{aligned} E(B_j(1) B_j(1)') &= Vf_1^0, & \text{for } j = 1, 3, \\ E(B_j(1) B_j(1)') &= Vf_2^0, & \text{for } j = 2, 4, \end{aligned}$$

where  $V = E(x_i x_i' u_i^2 | z_i = \gamma_0) = \sigma^2 D$  and  $f_i^0 = \frac{\partial F(\gamma)}{\partial \gamma_i} |_{\gamma=\gamma_0}$  for  $i = 1, 2$ .

Let  $B_1^*(v) = (B_1(v), -B_3(v))$ ,  $B_2^*(\omega) = (-B_2(\omega), B_4(\omega))$ .  $B_1^*(v)$  and  $B_2^*(\omega)$  are two independent Brownian motion vectors with covariance matrix  $E(B_1^*(1) B_1^*(1)') = V^* f_1^0$ ,  $E(B_2^*(1) B_2^*(1)') = V^* f_2^0$ , respectively, where  $V^* = \text{diag}\{V, V\}$ .

Thus, (24) can be rewritten as

$$\begin{aligned} R_2 &\Rightarrow -2[B_1^*(v)(c_2' - c_4', c_3')' + B_2^*(\omega)(c_2', c_3' - c_4')'] \\ &= -2[B_1^*(v)d_1 + B_2^*(\omega)d_2] \\ &= -2\left(\sqrt{d_1' V^* d_1} f_1^0 W_1(v) + \sqrt{d_2' V^* d_2} f_2^0 W_2(\omega)\right), \end{aligned} \tag{25}$$

where  $W_1(v)$  and  $W_2(\omega)$  are independent standard Brownian motions.

Similarly, for the other three cases, we can show

$$\begin{aligned} R_1 &= |v| d_1' D^* f_1^0 d_1 + |\omega| d_2' D^* f_2^0 d_2 + o_p(1), \\ R_2 &\Rightarrow -2\left(\sqrt{d_1' V^* d_1} f_1^0 W_1(v) + \sqrt{d_2' V^* d_2} f_2^0 W_2(\omega)\right). \end{aligned}$$

Making the change-of-variables

$$\begin{aligned} v &= \frac{d_1' V^* d_1}{(d_1' D^* d_1)^2 f_1^0} r_1, \\ \omega &= \frac{d_2' V^* d_2}{(d_2' D^* d_2)^2 f_2^0} r_2, \end{aligned}$$

and noting  $d_1' V^* d_1 = \sigma^2 d_1' D^* d_1$  and  $d_2' V^* d_2 = \sigma^2 d_2' D^* d_2$ , we have

$$\begin{aligned} RSS_T(\gamma) - RSS_T(\gamma^0) &\xrightarrow{p} R_1 + R_2 \\ &\Rightarrow \sigma^2|r_1| + \sigma^2|r_2| - 2\sigma^2 W_1(r_1) - 2\sigma^2 W_2(r_2) \\ &= \sigma^2(|r_1| + |r_2| - 2W_1(r_1) - 2W_2(r_2)). \end{aligned}$$

Define

$$\lambda_T = \left( \frac{(d_1' D^* d_1) f_1^0}{\sigma^2}, \frac{(d_2' D^* d_2) f_2^0}{\sigma^2} \right).$$

The asymptotic distribution can be expressed as

$$\begin{aligned}
 & T^{1-2\alpha} \lambda_T \left( (\hat{\gamma}_1 - \gamma_1^0), (\hat{\gamma}_2 - \gamma_2^0) \right) \\
 &= (r_1, r_2) \\
 &\Rightarrow \arg \min_{-\infty < r_1 < \infty, -\infty < r_2 < \infty} \left[ \left( \frac{1}{2} |r_1| - W_1(r_1) \right) + \left( \frac{1}{2} |r_2| - W_2(r_2) \right) \right] \\
 &= \arg \max_{-\infty < r_1 < \infty, -\infty < r_2 < \infty} \left[ \left( -\frac{1}{2} |r_1| + W_1(r_1) \right) + \left( -\frac{1}{2} |r_2| + W_2(r_2) \right) \right].
 \end{aligned} \tag{26}$$

## ACKNOWLEDGMENTS

We would like to thank Patrik Guggenberger, Guido Kuersteiner, Yongmiao Hong and the seminar participants of Tsinghua University and Nanyang Technological University for suggestions. All the remaining errors are ours.

## REFERENCES

- Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time series effects. *Journal of Financial Markets* 5:31–56.
- Amihud, Y., Mendelson, H. (1986). Asset pricing and the bid–ask spread. *Journal of Financial Economics* 17:223–249.
- Astatkie, T., Watts, D. G., Watt, W. E. (1997). Nested threshold autoregressive (NeTAR) models. *International Journal of Forecasting* 13:105–116.
- Avramov, D., Chorida, T., Goyal, A. (2006). Liquidity and autocorrelation in individual stock returns. *Journal of Finance* LXI:2365–2390.
- Bai, J. (1997). Estimation of a change point in multiple regressions. *Review of Economics and Statistics* 79:551–563.
- Bai, J., Chen, H., Chong, T. T. L., Wang, X. (2008). Generic consistency of the break-point estimator under specification errors in a multiple-break model. *Econometrics Journal* 11:287–307.
- Brennan, M. J., Cordia, T., Subrahmanyam, A. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics* 49:345–373.
- Campbell, J. Y., Grossman, S. J., Wang, J. (1993). Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics* 108:905–939.
- Caner, M., Hansen, B. E. (2001). Threshold autoregression with a unit root. *Econometrica* 69:1555–1596.
- Chan, K. S. (1993). Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model. *Annals of Statistics* 21:520–533.
- Chan, K. S., Tong, H. (1986). On estimating thresholds in autoregressive models. *Journal of Time Series Analysis* 7:179–190.
- Chan, K. S., Tsay, R. S. (1998). Limiting properties of the Least Square Estimator of a Continuous Threshold Autoregressive Model. *Biometrika* 85:413–426.
- Chen, R., Tsay, R. S. (1993). Functional-coefficient autoregressive models. *Journal of the American Statistical Association* 88:298–308.
- Chong, T. T. L. (2001). Structural change in AR(1) models. *Econometric Theory* 17:87–155.
- Chong, T. T. L. (2003). Generic consistency of the break-point estimator under specification errors. *Econometrics Journal* 6:167–192.

- Dueker, M., Martin, S., Spagnolo, F. (2007). Contemporaneous threshold autoregressive models: estimation, testing and forecasting. *Journal of Econometrics* 141:517–547.
- Durlauf, S. N., Johnson, P. A. (1995). Multiple regimes and cross-country growth behavior. *Journal of Applied Econometrics* 10:365–384.
- Edison, H. J. (2000). Do indicators of financial crises work? An Evaluation of an Early Warning System. Board of Governors of Federal Reserve System, International Finance Discussion Papers Number 675.
- Fama, E. F., French, K. R. (1992). The cross-section of unexpected stock returns. *Journal of Finance* 47:427–465.
- Foster, F. D., Viswanathan, S. (1995). Can speculative trading explain the volume-volatility relation? *Journal of Business and Economic Statistics* 13:379–396.
- Frankel, J. A., Rose, A. (1996). Currency crashes in emerging markets: empirical indicators. *NBER WP5437*.
- Gonzalo, J., Pitarakis, J. (2002). Estimation and model selection based inference in single and multiple threshold models. *Journal of Econometrics* 110:319–352.
- Granville, J. (1963). *Granville's New Key to Stock Market Profits*. Upper Saddle River, NJ: Prentice Hall.
- Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica* 64:413–430.
- Hansen, B. E. (1997). Inference in TAR models. *Studies in Nonlinear Dynamics and Economics* 2:1–14.
- Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica* 68:575–603.
- Hong, Y., Lee, T. H. (2003). Inference on predictability of foreign exchange rate via generalized spectrum and nonlinear time series models. *Review of Economics and Statistics* 85:1048–1062.
- Kaminsky, G. L. (1998). Currency and Banking Crises: The Early Warnings of Distress. Board of Governors of the Federal Reserve System. International Finance Discussion Papers Number 629.
- Karpoff, J. M. (1987). The relation between price changes and trading volume: a survey. *Journal of Financial and Quantitative Analysis* 22:109–125.
- Lee, M. C., Swaminathan, B. (2000). Price momentum and trading volume. *Journal of Finance* IV:2017–2069.
- Leeper, E. M. (1991). Equilibria under “Active” and “Passive” monetary and fiscal policies. *Journal of Monetary Economics* 27:129–147.
- Llorente, G., Michaely, R., Saar, G., Wang, J. (2002). Dynamic volume-return relation of individual stocks. *Review of Financial Studies* 15:1005–1047.
- Maheu, J. M., McCurdy, T. H. (2000). Identifying bull and bear markets in stock returns. *Journal of Business and Economic Statistics* 18:100–112.
- Sachs, J., Tornell, A., Velasco, A. (1996). Financial crises in emerging markets: the lessons from 1995. *Brookings Papers on Economic Activity* 1:147–215.
- Tsay, R. S. (1998). Testing and modeling multivariate threshold models. *Journal of the American Statistical Association* 93:231–240.
- Tiao, G. C., Tsay, R. S. (1994). Some advances in nonlinear and adaptive modeling in time-series. *Journal of Forecasting* 13:109–131.
- Tong, H. (1983). *Threshold Models in Non-Linear Time Series Analysis* (Lecture Notes in Statistics No. 21), New York: Springer-Verlag.