Estimating and Testing Threshold Regression Models with Multiple Threshold Variables

Chong, Terence Tai Leung and Yan, Isabel K.

The Chinese University of Hong Kong, City University of Hong Kong

24 March 2014

Estimating and Testing Threshold Regression Models with Multiple Threshold Variables

Terence Tai-Leung Chong[*] and **Isabel Kit-Ming Yan**[†‡]

March 22, 2014

Abstract

Conventional threshold models contain only one threshold variable. This paper provides the theoretical foundation for threshold models with multiple threshold variables. The new model is very different from a model with a single threshold variable as several novel problems arisefrom having an additional threshold variable. First, the model is not analogous to a change-point model. Second, the asymptotic joint distribution of the threshold estimators is difficult to obtain. Third, the estimation time increases exponentially with the number of threshold variables. This paper derives the consistency and the asymptotic joint distribution of the threshold estimators. A fast estimation algorithm to estimate the threshold values is proposed. We also develop tests for the number of threshold variables. The theoretical results are supported by simulation experiments. Our model is applied to the study of currency crises.

Keywords: Threshold Model; Multiple Threshold Variables; Currency Crisis; Panel Data

JEL Classification Number: C33; C12; C13

[*]Corresponding author. Department of Economics, Esther Lee Building, Room 930, the Chinese University of Hong Kong, Shatin, New Territories, Hong Kong. Phone: +(852)3943-8193; Fax: +(852)2603-5805; E-mail: chong2064@cuhk.edu.hk.

[†]Department of Economics and Finance, City University of Hong Kong, Kowloon, Hong Kong. Phone: +(852)2788-7315; Fax: +(852)2788-8806; E-mail: efyan@cityu.edu.hk.

[‡]The authors are thankful to Gregory Chow, W.K. Li, and the seminar participants at the Chinese University of Hong Kong and the University of Hong Kong for helpful comments. We would also like to thank Gilbert Lui and Haiqiang Chen for able research assistance. All remaining errors are our own.

1 Introduction

Threshold regression models have developed rapidly over the three decades since the seminal work of Tong (1983). Some later extensions of the model include the smooth transition threshold model (Chan and Tong, 1986), the functional-coefficient autoregressive model (Chen and Tsay, 1993) and the nested threshold autoregressive model (Astatkie, Watts and Watt, 1997). Hansen (1999) develops a threshold model for non-dynamic panels with individual fixed-effects. Tsay (1998) and Gonzalo and Pitarakis (2002) study models with multiple threshold values.

Most of these studies, however, focus on models with only one threshold variable and have limited applications when two or more threshold variables are needed. For instance, it has long been observed that the foreign debt level and interest rate cross certain threshold values before a currency crisis occurs. Studies in the literature of currency crises, such as those of Eichengreen, Rose and Wyplosz (1995), Frankel and Rose (1996), Kaminsky (1998) and Edison (2000) also suggest that the occurrence of a currency crisis depends critically on the values of multiple factors. However, none of these papers has estimated and tested those important threshold values, due to the lack of proper modelling techniques in the literature. To our knowledge, very few studies have been devoted to models with multiple threshold variables[1], and no theoretical results on the consistency and the asymptotic joint distribution of the threshold estimators in the general threshold regression model are available in the related literature.[2]

This paper explores the estimation and inference of a threshold model with multiple threshold variables. This model is not a simple extension of the model with a single threshold variable. The inclusion of an additional threshold variable increases the complexity of the model. Tsay (1998) suggests that a threshold model can be transformed into a change-point model by re-indexing the threshold variable. This analogy, however, cannot generally be carried over to cases involving more than one threshold variable, which this paper explores. Second, the asymptotic joint distribution of the threshold estimators is difficult to obtain. Third, since the threshold estimates are obtained by grid search, the estimation time increases exponentially with the number of threshold variables. The contributions of this paper are two-fold. First, the paper provides the estimation and distributional theories for threshold models with multiple threshold variables. Second, it develops a test for the

[1] Some related studies in this regard include Astatkie, Watts and Watt (1997), Xia and Li (1999) and Xia, Li and Tong (2004) and Chen *et al.* (2012).

[2] Chen *et al.* (2012) provide theoretical results on the consistency and the asymptotic joint distribution of the threshold estimators for TAR models.

number of threshold variables.

We apply our model to the study of currency crises in 16 countries. We take the threshold variables implied by the three generations of currency crisis models, and test for the existence of threshold effects. If there is evidence of a threshold effect, we estimate the threshold values using panel data from the 16 countries. We find overwhelming evidence of threshold effects in the ratio of short-term external liabilities to reserves and the lending rate differential, which is consistent with the implications of the currency crisis models. Our empirical study provides useful estimates of the joint threshold values that can be adopted as policy guidelines in the regulation of short-term external borrowing and interest rate differentials.

The rest of the paper is organized as follows. Section 2 presents the model and the major assumptions. The consistency and the asymptotic distribution of the threshold estimators are established. Section 3 proposes a fast estimation algorithm. The model is extended in Section 4 to allow for panel data. Section 5 develops a sequential test for the number of threshold variables and an LR test for the threshold values. Asymptotic distributions of these tests are obtained. Section 6 provides experimental evidence to support our theory. Section 7 provides an empirical application of our new model. The last section concludes the paper and discusses directions for future research. All proofs are relegated to the Appendices.

Before proceeding to the next section, we present the mathematical notation that is frequently used in this paper. $[x]$ denotes the greatest integer $\leq x$. The symbol '$\xrightarrow{p}$' represents convergence in probability, '$\xrightarrow{d}$' represents convergence in distribution, and '$\Rightarrow$' signifies weak convergence in $D[0,1]$: see Billingsley (1968) and Pollard (1984). All limits are as the sample size $T \rightarrow \infty$ unless otherwise stated.

2 The Model

To begin with, consider the following model:

$$y_t = \beta_1' x_t + (\beta_2' - \beta_1') x_t \Psi\left(\gamma^0, Z_t\right) + \varepsilon_t, \tag{1}$$

where β_1 and β_2 are the pre-shift and post-shift regression slope parameters respectively, with $\beta_i = (\beta_{1i}\ \beta_{2i}\ ...\ \beta_{Ki})'$ being a K by 1 vector of true parameters, $i = 1, 2$;

y_t is the dependent variable.

x_t is a K by 1 vector of covariates.

$(\varepsilon_1\ \varepsilon_2\ ...\ \varepsilon_T)'$ is a T by 1 vector of error term ε_t with $E|\varepsilon_t|^{4r} < \infty$ for some $r > 1$. The errors are assumed to be independent of both the regressors and the threshold variables.

$Z_t = (z_{1t}, ..., z_{mt})$ is a vector of m threshold variables, where $0 < m < \infty$.

$\gamma_0 = \left(\gamma^0_{1,...,}\gamma^0_m\right) \in \Pi^m_{j=1}\left[\underline{\gamma_j}, \overline{\gamma_j}\right]$ is a vector of m true threshold parameters to be estimated.

The observations $\{y_t, x_t, Z_t\}^T_{t=1}$ are real-valued.

$\Psi(\gamma^0, Z_t)$ is an indicator function, which equals one when the threshold variables satisfy some required conditions, and equals zero otherwise. For example, if the parameters change when all of the threshold variables exceed some critical values, then we have:

$$\Psi\left(\gamma^0, Z_t\right) = I\left(z_{1t} > \gamma^0_1, ..., z_{mt} > \gamma^0_m\right). \tag{2}$$

In the scenario of currency crises, imposing such a threshold condition implies that the crisis will not be triggered until all the threshold variables exceed the critical thresholds.[3] For illustration purposes, we will study the case where $m = 2$. The methods extend in a straightforward manner to models with more than two threshold variables. For notational simplicity, we let

$$\Psi_t\left(\gamma^0\right) = \Psi\left(\gamma^0, Z_t\right) = I\left(z_{1t} > \gamma^0_1, z_{2t} > \gamma^0_2\right). \tag{3}$$

Define

$$F(\gamma_1, \gamma_2) = \Pr(z_1 \leq \gamma_1, z_2 \leq \gamma_2) \tag{4}$$

and

$$F_i(\gamma_i) = \Pr(z_i \leq \gamma_i), \quad (i = 1, 2). \tag{5}$$

We assume that the joint distribution of z_1 and z_2 is continuous and differentiable with respect to both variables, and that**:**

(a) $\frac{1}{T}\sum^T_{t=1} I(z_{1t} > \gamma_1, z_{2t} > \gamma_2) \xrightarrow{p} \Pr(z_1 > \gamma_1, z_2 > \gamma_2) \overset{def}{=} \overline{F}(\gamma_1, \gamma_2)$;

(b) $\frac{1}{T}\sum^T_{t=1} I(z_{1t} > \gamma_1) \xrightarrow{p} \Pr(z_1 > \gamma_1) \overset{def}{=} \overline{F}_1(\gamma_1)$;

(c) $\frac{1}{T}\sum^T_{t=1} I(z_{2t} > \gamma_2) \xrightarrow{p} \Pr(z_2 > \gamma_2) \overset{def}{=} \overline{F}_2(\gamma_2)$.

Define

$$\overline{F}_{i\gamma} = \frac{\partial}{\partial\gamma_i}\overline{F}(\gamma_1, \gamma_2), \quad (i = 1, 2), \tag{6}$$

$$\overline{F}^0_i = \overline{F}_i\left(\gamma^0_1, \gamma^0_2\right), \quad (i = 1, 2). \tag{7}$$

[3] If the condition states that at least one threshold variable exceeds the critical value, then $\Psi\left(\gamma^0, Z_t\right) = 1 - I\left(z_{1t} \leq \gamma^0_1, ..., z_{mt} \leq \gamma^0_m\right)$.

Define the moment functionals:

$$\overline{M}_\gamma = \overline{M}(\gamma_1, \gamma_2) = E\left(x_t x_t' I\left(z_{1t} > \gamma_1, z_{2t} > \gamma_2\right)\right), \tag{8}$$

$$\overline{M}_0 = \overline{M}\left(\gamma_1^0, \gamma_2^0\right), \tag{9}$$

$$M = E\left(x_t x_t'\right), \tag{10}$$

$$D_\gamma = D(\gamma_1, \gamma_2) = E\left(x_t x_t' | z_{1t} = \gamma_1, z_{2t} = \gamma_2\right), \tag{11}$$

$$D = D\left(\gamma_1^0, \gamma_2^0\right), \tag{12}$$

$$V(\gamma_1, \gamma_2) = E\left(x_t x_t' \varepsilon_t^2 | z_{1t} = \gamma_1, z_{2t} = \gamma_2\right), \tag{13}$$

$$V = V\left(\gamma_1^0, \gamma_2^0\right), \tag{14}$$

$$\overline{G}(\gamma_1, \gamma_2) = M^{-1}\overline{M}(\gamma_1, \gamma_2). \tag{15}$$

We impose the following assumptions:

$(A1)$ $(x_t, Z_t, \varepsilon_t)$ is strictly stationary, ergodic and ρ-mixing, with ρ-mixing coefficients satisfying $\sum_{j=1}^{\infty} \rho_j^{\frac{1}{2}} < \infty$;
$(A2)$ $E\left(\varepsilon_t | \Im_{t-1}\right) = 0$;
$(A3)$ $E\left|x_t\right|^4 < \infty$ and $E\left|x_t \varepsilon_t\right|^4 < \infty$;
$(A4)$ For all $\gamma \in \Gamma$ and $i = 1, 2$, $E\left(\left|x_t\right|^4 |Z_t = \gamma\right)$, $\overline{F}_i(\gamma_1, \gamma_2)$ are bounded;
$(A5)$ At $\gamma = \gamma_0$ and $i = 1, 2$, $\overline{F}_{i\gamma}$, D_γ and $V(\gamma_1, \gamma_2)$ are continuous;
$(A6)$ $\delta = \beta_2 - \beta_1 = cT^{-\alpha}$ where $c \neq 0$ and $0 < \alpha < \dfrac{1}{2}$;
$(A7)$ $c'Dc$, $c'Vc$, $\overline{F}_1^0$ and $\overline{F}_2^0$ are positive;
$(A8)$ $M > \overline{M}(\gamma_1, \gamma_2) > 0$ for all $\gamma \in \Gamma$.

$(A1)$ implies that all of the regressors are stationary and ergodic. It is used to establish the uniform convergence result, and will be automatically satisfied for i.i.d observations. $(A2)$ requires that model (1) is correctly specified. Assumptions $(A3)$ and $(A4)$ are conditional and unconditional moment bounds. $(A5)$ requires the threshold variable to have a continuous distribution and excludes regime-dependent heteroskedasticity. $(A6)$ assumes that the parameter change is small and converges to zero at a slow rate when the sample size is large. This assumption and $(A7)$ are needed in order for the threshold estimators to have a non-degenerating distribution. $(A8)$ is the conventional full-rank condition which excludes multicollinearity.

We will derive the least squares estimators of β_1, β_2 and γ^0. Given $\gamma = (\gamma_1, \gamma_2)$, the estimators for β are

$$\widehat{\beta}_1'(\gamma) = \sum_{t=1}^{T} y_t x_t' (1 - \Psi_t(\gamma)) \left(\sum_{t=1}^{T} x_t x_t' (1 - \Psi_t(\gamma)) \right)^{-1} \tag{16}$$

and

$$\widehat{\beta}_2'(\gamma) = \sum_{t=1}^{T} y_t x_t' \Psi_t(\gamma) \left(\sum_{t=1}^{T} x_t x_t' \Psi_t(\gamma) \right)^{-1}. \tag{17}$$

Define

$$S_T(\gamma) = \sum_{t=1}^{T} \left(y_t - \widehat{\beta}_1'(\gamma) x_t - \left(\widehat{\beta}_2'(\gamma) - \widehat{\beta}_1'(\gamma) \right) x_t \Psi_t(\gamma) \right)^2, \tag{18}$$

$$\widehat{\gamma} = (\widehat{\gamma}_1, \widehat{\gamma}_2) = \arg \min_{(\gamma_1, \gamma_2) \in \Gamma_T} S_T(\gamma_1, \gamma_2), \tag{19}$$

where

$$\Gamma_T = \Pi_{j=1}^{2} \left(\left[\underline{\gamma_j}, \overline{\gamma_j} \right] \cap \{ z_{j1}, ..., z_{jT} \} \right). \tag{20}$$

The final structural estimators are then defined as

$$\widehat{\beta}_1(\widehat{\gamma}_1, \widehat{\gamma}_2)$$

and

$$\widehat{\beta}_2(\widehat{\gamma}_1, \widehat{\gamma}_2).$$

The behavior of the $\widehat{\beta}_1(\gamma)$ and $\widehat{\beta}_2(\gamma)$ will be affected by the pair-wise relationship between x_t, z_{1t} and z_{2t}. To give a simple illustration, consider the case of a single regressor where

$$M = E\left(x_t^2\right), \tag{21}$$

$$\overline{G}(\gamma_1, \gamma_2) = \frac{\overline{M}_\gamma}{M}. \tag{22}$$

Note from Appendix A1 that

$$\widehat{\beta}_1(\gamma_1,\gamma_2) = \beta_1 + \delta \times \frac{\sum_{t=1}^T x_t^2 \left[I\left(z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0\right) - I\left(z_{1t} > \max\left\{\gamma_1^0, \gamma_1\right\}, z_{2t} > \max\left\{\gamma_2^0, \gamma_2\right\}\right)\right]}{\sum_{t=1}^T x_t^2 \left(1 - I\left(z_{1t} > \gamma_1, z_{2t} > \gamma_2\right)\right)} + o_p(1). \tag{23}$$

Similarly, we have

$$\widehat{\beta}_2(\gamma_1,\gamma_2) = \beta_2 - \delta \times \frac{\sum_{t=1}^T x_t^2 \left[I\left(z_{1t} > \gamma_1, z_{2t} > \gamma_2\right) - I\left(z_{1t} > \max\left\{\gamma_1^0, \gamma_1\right\}, z_{2t} > \max\left\{\gamma_2^0, \gamma_2\right\}\right)\right]}{\sum_{t=1}^T x_t^2 I\left(z_{1t} > \gamma_1, z_{2t} > \gamma_2\right)} + o_p(1). \tag{24}$$

We can partition the space of γ into four regions and discuss four separate cases:

Case 1: $\gamma_1 \leq \gamma_1^0, \gamma_2 \leq \gamma_2^0$

$$\widehat{\beta}_1(\gamma_1,\gamma_2) \xrightarrow{p} \beta_1, \tag{25}$$

$$\widehat{\beta}_2(\gamma_1,\gamma_2) \xrightarrow{p} \beta_2 - \delta\left(1 - \frac{\overline{G}\left(\gamma_1^0, \gamma_2^0\right)}{\overline{G}\left(\gamma_1, \gamma_2\right)}\right). \tag{26}$$

Case 2: $\gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0$

$$\widehat{\beta}_1(\gamma_1,\gamma_2) \xrightarrow{p} \beta_1 + \delta\frac{\overline{G}\left(\gamma_1^0, \gamma_2^0\right) - \overline{G}\left(\gamma_1, \gamma_2^0\right)}{1 - \overline{G}\left(\gamma_1, \gamma_2\right)}, \tag{27}$$

$$\widehat{\beta}_2(\gamma_1,\gamma_2) \xrightarrow{p} \beta_2 - \delta\left(1 - \frac{\overline{G}\left(\gamma_1, \gamma_2^0\right)}{\overline{G}\left(\gamma_1, \gamma_2\right)}\right). \tag{28}$$

Case 3: $\gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0$

$$\widehat{\beta}_1(\gamma_1,\gamma_2) \xrightarrow{p} \beta_1 + \delta\frac{\overline{G}\left(\gamma_1^0, \gamma_2^0\right) - \overline{G}\left(\gamma_1^0, \gamma_2\right)}{1 - \overline{G}\left(\gamma_1, \gamma_2\right)}, \tag{29}$$

$$\widehat{\beta}_2(\gamma_1,\gamma_2) \xrightarrow{p} \beta_2 - \delta\left(1 - \frac{\overline{G}\left(\gamma_1^0, \gamma_2\right)}{\overline{G}\left(\gamma_1, \gamma_2\right)}\right). \tag{30}$$

Case 4: $\gamma_1 > \gamma_1^0, \gamma_2 > \gamma_2^0$

$$\widehat{\beta}_1 \left(\gamma_1,\gamma_2\right) \xrightarrow{p} \beta_1 + \delta \frac{\overline{G}\left(\gamma_1^0, \gamma_2^0\right) - \overline{G}\left(\gamma_1, \gamma_2\right)}{1 - \overline{G}\left(\gamma_1, \gamma_2\right)}, \tag{31}$$

$$\widehat{\beta}_2 \left(\gamma_1,\gamma_2\right) \xrightarrow{p} \beta_2. \tag{32}$$

Note that $\widehat{\beta}_i \left(\gamma_1^0,\gamma_2^0\right) \xrightarrow{p} \beta_i$, $i = 1, 2$. This implies that the structural estimators can be consistently estimated if the threshold estimators are super-consistent.

2.1 Asymptotic behavior of $\frac{1}{T^{1-2\alpha}} \left(S_T\left(\gamma\right) - \varepsilon'\varepsilon\right)$

To study the behavior of the residual sum of squares, let

$$x_t\left(\gamma\right) = x_t \Psi_t\left(\gamma\right) \tag{33}$$

and let X and X_γ be T by K matrices formed by stacking the vectors x_t' and $x_t\left(\gamma\right)'$.

Thus, our model can be rewritten as

$$Y = X\beta_1 + X_\gamma \boldsymbol{\delta} + \varepsilon. \tag{34}$$

The residual sum of squares can also be written as

$$\begin{aligned} S_T\left(\gamma\right) &= \left(Y - X\widehat{\beta}_1\left(\gamma\right) - X_\gamma \widehat{\boldsymbol{\delta}}\left(\gamma\right)\right)' \left(Y - X\widehat{\beta}_1\left(\gamma\right) - X_\gamma \widehat{\boldsymbol{\delta}}\left(\gamma\right)\right) \\ &= Y'\left(I - P_\gamma\right) Y, \end{aligned} \tag{35}$$

where

$$P_\gamma = \widetilde{X}_\gamma \left(\widetilde{X}_\gamma' \widetilde{X}_\gamma\right)^{-1} \widetilde{X}_\gamma',$$

$$\widetilde{X}_\gamma = \left[X \; X_\gamma\right].$$

As $Y - X\beta_1 - X_\gamma \boldsymbol{\delta}$ and X lies in the space spanned by P_γ,

$$S_T\left(\gamma\right) - \varepsilon'\varepsilon = -\varepsilon' P_\gamma \varepsilon + 2\boldsymbol{\delta}' X_0' \left(I - P_\gamma\right)\varepsilon + \boldsymbol{\delta}' X_0' \left(I - P_\gamma\right) X_0 \boldsymbol{\delta}$$

and

$$\frac{1}{T^{1-2\alpha}} \left(S_T\left(\gamma\right) - \varepsilon'\varepsilon\right) = \frac{1}{T} c' X_0' \left(I - P_\gamma\right) X_0 c + o_p\left(1\right),$$

where $X_0 = X_{\gamma_0}$.

From Appendix B, in each of the following four cases, $\frac{1}{T^{1-2\alpha}} \left(S_T\left(\gamma\right) - \varepsilon'\varepsilon\right) \xrightarrow{p} b_i\left(\gamma\right)$, with $b_i\left(\gamma_0\right) = 0$, $i = 1, 2, 3, 4$.

Case 1: $\gamma_1 \leq \gamma_1^0, \gamma_2 \leq \gamma_2^0 \quad b_1(\gamma) = c'\left(\overline{M}_0 - \overline{M}_0 \overline{M}_\gamma^{-1} \overline{M}_0\right) c \geq 0.$

$$\frac{\partial}{\partial \gamma_1} b_1(\gamma) = c' \overline{M}_0 \overline{M}_\gamma^{-1} D_\gamma \overline{F}_{1\gamma} \overline{M}_\gamma^{-1} \overline{M}_0 c \leq 0,$$
$$\frac{\partial}{\partial \gamma_2} b_1(\gamma) = c' \overline{M}_0 \overline{M}_\gamma^{-1} D_\gamma \overline{F}_{2\gamma} \overline{M}_\gamma^{-1} \overline{M}_0 c \leq 0.$$

Case 2: $\gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0$

$$\begin{aligned} & b_2(\gamma) \\ &= c' \begin{pmatrix} \overline{M}_0 - \left(\overline{M}_0 - \overline{M}\left(\gamma_1, \gamma_2^0\right)\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - \overline{M}\left(\gamma_1, \gamma_2^0\right)\right) \\ -\overline{M}\left(\gamma_1, \gamma_2^0\right) \overline{M}_\gamma^{-1} \overline{M}\left(\gamma_1, \gamma_2^0\right) \end{pmatrix} c \\ &> 0. \end{aligned}$$

Case 3: $\gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0$

$$\begin{aligned} & b_3(\gamma) \\ &= c' \begin{pmatrix} \overline{M}_0 - \left(\overline{M}_0 - \overline{M}\left(\gamma_1^0, \gamma_2\right)\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - \overline{M}\left(\gamma_1^0, \gamma_2\right)\right) \\ -\overline{M}\left(\gamma_1^0, \gamma_2\right) \overline{M}_\gamma^{-1} \overline{M}\left(\gamma_1^0, \gamma_2\right) \end{pmatrix} c \\ &> 0. \end{aligned}$$

Case 4: $\gamma_1 > \gamma_1^0, \gamma_2 > \gamma_2^0$

$$b_4(\gamma) = c'\left(M - \overline{M}_0 - \left(M - \overline{M}_0\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(M - \overline{M}_0\right)\right) c.$$

$$\begin{aligned} \frac{\partial}{\partial \gamma_1} b_4(\gamma) &= -c'\left(M - \overline{M}_0\right)\left(M - \overline{M}_\gamma\right)^{-1} D_\gamma \overline{F}_{1\gamma}\left(M - \overline{M}_\gamma\right)^{-1}\left(M - \overline{M}_0\right) c \\ &> 0, \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial \gamma_2} b_4(\gamma) &= -c'\left(M - \overline{M}_0\right)\left(M - \overline{M}_\gamma\right)^{-1} D_\gamma \overline{F}_{2\gamma}\left(M - \overline{M}_\gamma\right)^{-1}\left(M - \overline{M}_0\right) c \\ &> 0. \end{aligned}$$

The threshold estimators are consistent because all of the four functions are minimized at the true thresholds, and it can be shown that $b_i(\gamma) \neq b_i(\gamma_0)$ iff $\gamma \neq \gamma_0$ for $i = 1, 2, 3, 4$.

If x_t are independent of z_{1t} and z_{2t}, we can express $b_1(\gamma)$ to $b_4(\gamma)$ by the joint distribution of the threshold variables. Consider the case in which there is only one regressor. We have

$$\overline{M}_\gamma = E\left(x_t^2\right) \overline{F}\left(\gamma_1, \gamma_2\right)$$

and

$$\overline{G}(\gamma_1,\gamma_2)=\overline{F}(\gamma_1,\gamma_2),$$

$$b_1(\gamma_1,\gamma_2)=c^2\overline{F}\left(\gamma_1^0,\gamma_2^0\right)\left(1-\frac{\overline{F}(\gamma_1^0,\gamma_2^0)}{\overline{F}(\gamma_1,\gamma_2)}\right),$$

$$b_2(\gamma_1,\gamma_2)=c^2\left[\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\frac{\left(\overline{F}(\gamma_1^0,\gamma_2^0)-\overline{F}(\gamma_1,\gamma_2^0)\right)^2}{1-\overline{F}(\gamma_1,\gamma_2)}-\frac{\overline{F}(\gamma_1,\gamma_2^0)^2}{\overline{F}(\gamma_1,\gamma_2)}\right],$$

$$b_3(\gamma_1,\gamma_2)=c^2\left[\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\frac{\left(\overline{F}(\gamma_1^0,\gamma_2^0)-\overline{F}(\gamma_1^0,\gamma_2)\right)^2}{1-\overline{F}(\gamma_1,\gamma_2)}-\frac{\overline{F}(\gamma_1^0,\gamma_2)^2}{\overline{F}(\gamma_1,\gamma_2)}\right],$$

$$b_4(\gamma_1,\gamma_2)=c^2\left(\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}(\gamma_1,\gamma_2)\right)\frac{1-\overline{F}(\gamma_1^0,\gamma_2^0)}{1-\overline{F}(\gamma_1,\gamma_2)}.$$

2.2 Asymptotic joint distribution of $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ when z_{1t} and z_{2t} are independent

The threshold estimator is analogous to the change-point estimator in the structural-change model. The distribution of the change-point estimator will degenerate to the true change point for any fixed magnitude of change because of the superconsistency of the change-point estimator (Chong, 2001). Thus, to obtain a non-degenerate distribution, one needs to let the magnitude of change go to zero at an appropriate rate. For the threshold model, in order to obtain the distribution of the threshold estimators, we also let the threshold effect go to zero at a certain rate. The distribution of the threshold estimator for small threshold effect has been obtained by Hansen (1999, 2000) for the case of a single threshold variable. In our case of two threshold variables, the following theorem states the joint distribution of the threshold estimators. The details can be found in Appendix C.

Theorem 1 *Under assumptions (A1)-(A8),*

$$\begin{aligned}&T^{1-2\alpha}\frac{(c'Dc)^2}{c'Vc}\left(\left(\widehat{\gamma}_1-\gamma_1^0\right)\overline{F}_1^0,\left(\widehat{\gamma}_2-\gamma_2^0\right)\overline{F}_2^0\right)\\&=(\widehat{r}_1,\widehat{r}_2)\\&\xrightarrow{d}\underset{(r_1,r_2)\in R^2}{\arg\max}\sum_{j=1}^{2}\left(-\frac{1}{2}|r_j|+W_j(r_j)\right),\end{aligned}\tag{36}$$

where $W_j(r_j)$, $j = 1, 2$, are double-sided independent standard Brownian motion on $(-\infty, \infty)$.

For $a_1 > 0$ and $a_2 > 0$, the above joint distribution equals

$$F_{(\widehat{r}_1, \widehat{r}_2)}(a_1, a_2)$$
$$= \Pi_{j=1}^{2}\left(1 + \sqrt{\frac{a_j}{2\pi}}\exp\left(-\frac{a_j}{8}\right) + \frac{3}{2}\exp(a_j)\Phi\left(-\frac{3\sqrt{a_j}}{2}\right) - \frac{a_j+5}{2}\Phi\left(-\frac{\sqrt{a_j}}{2}\right)\right), \tag{37}$$

where $\Phi(\cdot)$ is the cdf of a standard normal distribution.

The joint density function, which is depicted in Figure 4b, can be shown as

$$f_{(\widehat{r}_1, \widehat{r}_2)}(a_1, a_2) = \Pi_{j=1}^{2}\left(\frac{3}{2}\exp(a_j)\Phi\left(-\frac{3\sqrt{a_j}}{2}\right) - \frac{1}{2}\Phi\left(-\frac{\sqrt{a_j}}{2}\right)\right). \tag{38}$$

For cases where some of the $a_j < 0$, we can replace those items in the above expression by $F_{\widehat{r}_j}(a_j) = 1 - F_{\widehat{r}_j}(-a_j)$ and $f_{\widehat{r}_j}(a_j) = f_{\widehat{r}_j}(-a_j)$.

Corollary 2 *In general, if we have m threshold variables,*

$$-T^{1-2\alpha}\frac{(c'Dc)^2}{c'Vc}(\widehat{\gamma} - \gamma_0)\circ\frac{\partial F(\gamma_1^0, ..., \gamma_m^0)}{\partial \gamma} \xrightarrow{d} \underset{(r_1,...,r_m)\in R^m}{\arg\max}\sum_{j=1}^{m}\left(-\frac{1}{2}|r_j| + W_j(r_j)\right). \tag{39}$$

where $\circ$ is the Hadamard product operator that multiplies on an element by element basis, and

$$F_{(\widehat{r}_1,...,\widehat{r}_m)}(a_1, ..., a_m)$$
$$= \Pi_{j=1}^{m}\left(1 + \sqrt{\frac{a_j}{2\pi}}\exp\left(\frac{-a_j}{8}\right) + \frac{3\exp(a_j)}{2}\Phi\left(\frac{-3\sqrt{a_j}}{2}\right) - \frac{a_j+5}{2}\Phi\left(\frac{-\sqrt{a_j}}{2}\right)\right), \tag{40}$$

$$f_{(\widehat{r}_1,...,\widehat{r}_m)}(a_1, ..., a_m) = \Pi_{j=1}^{m}\left(\frac{3}{2}\exp(a_j)\Phi\left(-\frac{3\sqrt{a_j}}{2}\right) - \frac{1}{2}\Phi\left(-\frac{\sqrt{a_j}}{2}\right)\right). \tag{41}$$

It should be noted that if the threshold variables are dependent, the consistency and the joint distribution of the threshold estimators might be difficult to obtain. For example, if $z_2 = -z_1$, we might not be able to partition the data into four groups according to the values of the two threshold variables, and the above distributional result will not hold.

3 A Fast Estimation Algorithm for γ when $\mathbf{x}_t$, $\mathbf{z}_{1t}$ and $\mathbf{z}_{2t}$ are Independent

The threshold values are obtained by grid search, which implies that the estimation time increases exponentially with the number of threshold variables. We propose a fast estimation method when the regressors and the threshold variables are all mutually independent. Consider the following simple model with a single regressor and two threshold variables:

$$y_t = \beta_1 x_t + \delta x_t \Psi_t\left(\gamma^0\right) + \varepsilon_t. \tag{42}$$

Note that the asymptotic results are the same as the case where all $x_t = 1$. From Appendix A2, it can be shown that

$$\sup_{(\gamma_1,\gamma_2)\in R^2}\left|\frac{1}{T}S_T\left(\gamma_{1,}\gamma_2\right) - g\left(\gamma_1,\gamma_2\right)\right| = o_p\left(1\right). \tag{43}$$

Let

$$b_i\left(\gamma_1,\gamma_2\right) = T^{2\alpha}\left(g\left(\gamma_1,\gamma_2\right) - \sigma^2\right) \quad \text{for } i = 1,2,3,4. \tag{44}$$

When x_t, z_{1t} and z_{2t} are independent, the following can be shown.

Case 1: $\gamma_1 \le \gamma_1^0, \gamma_2 \le \gamma_2^0$

$$b_1\left(\gamma_1,\gamma_2\right) = c^2\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\left(1 - \frac{\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right), \tag{45}$$

$$\frac{\partial b_1\left(\gamma_1,\gamma_2\right)}{\partial\gamma_1} = -c^2\frac{\overline{F}_1\left(\gamma_1^0\right)^2\overline{F}_2\left(\gamma_2^0\right)^2}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}H_1\left(\gamma_1\right) \le 0, \tag{46}$$

$$\frac{\partial b_1\left(\gamma_1,\gamma_2\right)}{\partial\gamma_2} = -c^2\frac{\overline{F}_1\left(\gamma_1^0\right)^2\overline{F}_2\left(\gamma_2^0\right)^2}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}H_2\left(\gamma_2\right) \le 0, \tag{47}$$

where $H_1\left(\gamma_1\right)$ and $H_2\left(\gamma_2\right)$ are the hazard functions of z_1 and z_2 respectively.

Case 2: $\gamma_1 > \gamma_1^0, \gamma_2 \le \gamma_2^0$

$$b_2\left(\gamma_1,\gamma_2\right) = c^2\overline{F}_2\left(\gamma_2^0\right)^2\left(\frac{\overline{F}_1\left(\gamma_1^0\right)}{\overline{F}_2\left(\gamma_2^0\right)} - \frac{\overline{F}_1\left(\gamma_1\right)}{\overline{F}_2\left(\gamma_2\right)} - \frac{\left(\overline{F}_1\left(\gamma_1^0\right) - \overline{F}_1\left(\gamma_1\right)\right)^2}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right), \tag{48}$$

$$\frac{\partial b_2\left(\gamma_1,\gamma_2\right)}{\partial\gamma_1} = c^2\frac{\overline{F}_1\left(\gamma_1^0\right)^2}{\overline{F}_2\left(\gamma_2\right)}\left(\frac{1 - \overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2\right)}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right)^2 f_1\left(\gamma_1\right) > 0, \tag{49}$$

$$
\begin{aligned}
&\frac{\partial b_2(\gamma_1,\gamma_2)}{\partial \gamma_2} \\
&= -c^2\overline{F}_2\left(\gamma_2^0\right)^2\left(\frac{\overline{F}_1\left(\gamma_1^0\right)-\overline{F}_1\left(\gamma_1\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}+\frac{1}{\overline{F}_2\left(\gamma_2\right)}\right)\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\overline{F}_1\left(\gamma_1\right)H_2\left(\gamma_2\right) \\
&< 0.
\end{aligned} \tag{50}
$$

Case 3: $\gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0$

$$
b_3(\gamma_1,\gamma_2) = c^2\overline{F}_1\left(\gamma_1^0\right)^2\left(\frac{\overline{F}_2\left(\gamma_2\right)}{\overline{F}_1\left(\gamma_1\right)}-\frac{\overline{F}_2\left(\gamma_2^0\right)}{\overline{F}_1\left(\gamma_1^0\right)}-\frac{\left(\overline{F}_2\left(\gamma_2^0\right)-\overline{F}_2\left(\gamma_2\right)\right)^2}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right), \tag{51}
$$

$$
\begin{aligned}
&\frac{\partial b_3(\gamma_1,\gamma_2)}{\partial \gamma_1} \\
&= -c^2\overline{F}_1\left(\gamma_1^0\right)^2\left(\frac{\overline{F}_2\left(\gamma_2^0\right)-\overline{F}_2\left(\gamma_2\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}+\frac{1}{\overline{F}_1\left(\gamma_1\right)}\right)\frac{1-\overline{F}_2\left(\gamma_2^0\right)\overline{F}_1\left(\gamma_1\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\overline{F}_2\left(\gamma_2\right)H_1\left(\gamma_1\right) \\
&< 0,
\end{aligned} \tag{52}
$$

$$
\frac{\partial b_3(\gamma_1,\gamma_2)}{\partial \gamma_2} = c^2\overline{F}_1\left(\gamma_1^0\right)^2\left(\frac{\overline{F}_2\left(\gamma_2^0\right)-\overline{F}_2\left(\gamma_2\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}-\frac{1}{\overline{F}_1\left(\gamma_1\right)}\right)^2\overline{F}_1\left(\gamma_1\right)f_2\left(\gamma_2\right) > 0. \tag{53}
$$

Case 4: $\gamma_1 > \gamma_1^0, \gamma_2 > \gamma_2^0$

$$
b_4(\gamma_1,\gamma_2) = c^2\left(1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\right)\left(1-\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right), \tag{54}
$$

$$
\frac{\partial b_4(\gamma_1,\gamma_2)}{\partial \gamma_1} = c^2\left(\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right)^2\overline{F}_2\left(\gamma_2\right)f_1\left(\gamma_1\right) > 0, \tag{55}
$$

$$
\frac{\partial b_4(\gamma_1,\gamma_2)}{\partial \gamma_2} = c^2\left(\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right)^2\overline{F}_1\left(\gamma_1\right)f_2\left(\gamma_2\right) > 0. \tag{56}
$$

Note that given γ_2, the value of $b(\gamma_1,\gamma_2)$ reduces whenever γ_1 approaches γ_1^0 from both directions. Similarly, given γ_1, the value of $b(\gamma_1,\gamma_2)$ reduces whenever γ_2 approaches γ_2^0. This implies that

$$
\underset{\gamma_1\in R}{Arg\min} b(\gamma_1,\gamma_2) = \gamma_1^0 \quad \forall\gamma_2 \tag{57}
$$

and

$$\underset{\gamma_2 \in R}{Arg\min} b\left(\gamma_1, \gamma_2\right) = \gamma_2^0 \quad \forall \gamma_1. \tag{58}$$

Thus, if x_t, z_{1t} and z_{2t} are independent, we can search for the critical threshold value of one threshold variable by assigning an arbitrary value to another threshold estimate. This will dramatically shorten the time of estimation.

4 Model with Panel Data

In this section, we consider a model for a balanced panel with n individuals over T periods. We assume that the threshold values are the same across individuals for each of the threshold variables. In the panel model here, n is the cross-sectional sample size. The analysis is asymptotic with fixed T and as $n \to \infty$.

We let

$$\Psi_{it}\left(\gamma\right) = I\left(z_{1it} > \gamma_1, z_{2it} > \gamma_2\right).$$

The observations are divided into two regimes depending on whether the threshold variable vector satisfies the threshold conditions. We assume that x_{it} and Z_{it} are not time invariant. The model is

$$y_{it} = \mu_i + \beta_1' x_{it} + \varepsilon_{it}, \quad \Psi_{it}\left(\gamma\right) = 0, \tag{59}$$

$$y_{it} = \mu_i + \beta_2' x_{it} + \varepsilon_{it}, \quad \Psi_{it}\left(\gamma\right) = 1. \tag{60}$$

The following assumptions are imposed:

$(B1)$ For each t, $\left(x_{it}, Z_{it}, \varepsilon_{it}\right)$ are i.i.d. across i.

$(B2)$ For each i, ε_{it} is i.i.d. over t, is independent of $\left\{x_{ij}, Z_{ij}\right\}_{j=1}^{T}$, and $E\left(\varepsilon_{it}\right) = 0$;

$(B3)$ For each $j = 1, ..., k$, $\Pr\left(x_{i1}^j = x_{i2}^j = ... = x_{iT}^j\right) < 1$, where x_{it}^j is the jth element of x_{it}.

$(B4)$ $E\left|x_{it}\right|^4 < \infty$ and $E\left|\varepsilon_{it}\right|^4 < \infty$;

$(B5)$ $\delta = cn^{-\alpha}$ where $c \neq 0$ and $0 < \alpha < \dfrac{1}{2}$;

$(B6)$ At $\gamma = \gamma_0$ and $i = 1, 2$, $\overline{F}_i\left(\gamma_1, \gamma_2\right)$, D_γ and $V\left(\gamma_1, \gamma_2\right)$ are continuous;

$(B7)$ $0 < D < \infty$;

$(B8)$ For $k > t$, $f_{k|t}\left(\gamma_1^0, \gamma_2^0 | \gamma_1^0, \gamma_2^0\right) < \infty$, where $f_{k|t}\left(\gamma_1^0, \gamma_2^0 | \gamma_1^0, \gamma_2^0\right)$ is the value of the conditional joint density of Z_{ik} evaluated at the true thresholds given that Z_{it} equals the true thresholds.

Assumptions $(B1)$-$(B4)$ are standard for fixed effect panel models with exogenous regressors. Assumption $(B5)$ implies that the threshold effect tends to zero at a specified rate, which gives a well-defined distribution of the threshold estimators. Assumption $(B6)$ excludes threshold effects that occur simultaneously in the marginal distribution of the regressors and in the regression function. Assumption $(B7)$ excludes continuous threshold models (Chan and Tsay, 1998). $(B8)$ rules out the possibility that all observations of the threshold variables equal the true threshold values.

Let

$$x_{it}(\gamma) = x_{it}\Psi_{it}(\gamma), \tag{61}$$

$$y_{it} = \mu_i + \beta_1' x_{it} + \delta' x_{it}\Psi_{it}(\gamma) + \varepsilon_{it}. \tag{62}$$

Averaging the above panel equation over t, we have

$$\overline{y}_i = \mu_i + \beta_1'\overline{x}_i + \delta'\overline{x}_i(\gamma) + \overline{\varepsilon}_t, \tag{63}$$

where

$$\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}, \tag{64}$$

$$\overline{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}, \tag{65}$$

$$\overline{x}_i(\gamma) = \frac{1}{T}\sum_{t=1}^{T} x_{it}\Psi_{it}(\gamma), \tag{66}$$

$$\overline{\varepsilon}_t = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it}. \tag{67}$$

Taking the difference, we have

$$y_{it}^* = \beta_1' x_{it}^* + \delta' x_{it}^*(\gamma) + \varepsilon_{it}^*, \tag{68}$$

where

$$y_{it}^* = y_{it} - \overline{y}_i, \tag{69}$$

$$x_{it}^* = x_{it} - \overline{x}_i, \tag{70}$$

$$x_{it}^{*}(\gamma) = x_{it}(\gamma) - \overline{x}_i(\gamma), \tag{71}$$

$$\varepsilon_{it}^{*} = \varepsilon_{it} - \overline{\varepsilon}_i. \tag{72}$$

Let

$$y_i^{*} = \begin{bmatrix} y_{i2}^{*} \\ \vdots \\ y_{iT}^{*} \end{bmatrix}, x_i^{*} = \begin{bmatrix} x_{i2}^{*} \\ \vdots \\ x_{iT}^{*} \end{bmatrix}, x_i^{*}(\gamma) = \begin{bmatrix} x_{i2}^{*}(\gamma) \\ \vdots \\ x_{iT}^{*}(\gamma) \end{bmatrix}, \varepsilon_i^{*} = \begin{bmatrix} \varepsilon_{i2}^{*} \\ \vdots \\ \varepsilon_{iT}^{*} \end{bmatrix}$$

denote the stacked data and errors for an individual, with one time period deleted. Let Y^{*}, $X^{*}(\gamma)$ and ε^{*} denote the data that is stacked over all individuals, i.e.,

$$Y^{*} = \begin{bmatrix} y_1^{*} \\ \vdots \\ y_i^{*} \\ \vdots \\ y_n^{*} \end{bmatrix}, X^{*} = \begin{bmatrix} x_1^{*} \\ \vdots \\ x_i^{*} \\ \vdots \\ x_n^{*} \end{bmatrix}, X^{*}(\gamma) = \begin{bmatrix} x_1^{*}(\gamma) \\ \vdots \\ x_i^{*}(\gamma) \\ \vdots \\ x_n^{*}(\gamma) \end{bmatrix}, \varepsilon^{*} = \begin{bmatrix} \varepsilon_1^{*} \\ \vdots \\ \varepsilon_i^{*} \\ \vdots \\ \varepsilon_n^{*} \end{bmatrix}.$$

Thus, our model becomes

$$Y^{*} = X^{*}\beta_1 + X^{*}(\gamma)\boldsymbol{\delta} + \varepsilon^{*}. \tag{73}$$

As the panel model can be rewritten in the form given in Section (2.1), with n corresponding to T, and as assumption set B is weaker than assumption set A, the estimation method and the asymptotic results in the previous section apply in the panel model. Thus, we have

$$S_{nT}(\gamma) = (Y - X^{*}\beta_1 - X^{*}(\gamma)\boldsymbol{\delta})'(Y - X^{*}\beta_1 - X^{*}(\gamma)\boldsymbol{\delta})$$

$$\widehat{\gamma} = (\widehat{\gamma}_1, \widehat{\gamma}_2) = \arg\min_{\gamma\in\Gamma_n} S_T(\gamma_1, \gamma_2). \tag{74}$$

$$\Gamma_n = \Pi_{j=1}^{2}\left(\left[\underline{\gamma_j}, \overline{\gamma_j}\right] \cap \left(\cup_{i=1}^{n}\{z_{ji1}, ..., z_{jiT}\}\right)\right) \tag{75}$$

The final structural estimators are then defined as

$$\widehat{\beta}_1(\widehat{\gamma}_1, \widehat{\gamma}_2)$$

and

$$\widehat{\beta}_2\left(\widehat{\gamma}_1, \widehat{\gamma}_2\right).$$

and the residual variance is

$$\widehat{\sigma}^2 = \frac{1}{n\left(T-1\right)} S_{nT}\left(\widehat{\gamma}\right). \tag{76}$$

5 Inference

5.1 Testing the number of threshold variables

We start with a threshold model without thresholds, and sequentially test whether this model can be rejected in favor of a threshold model with one additional threshold variable. For the test of no threshold against one threshold variable,

$$H_0 : m = 0$$
$$H_1 : m = 1$$

Define

$$F\left(0,1,1\right) = T\frac{S_T\left(-\infty,-\infty\right) - S_T\left(\widehat{\gamma}_1,-\infty\right)}{S_T\left(\widehat{\gamma}_1,-\infty\right)}, \tag{77}$$

$$F\left(0,1,2\right) = T\frac{S_T\left(-\infty,-\infty\right) - S_T\left(-\infty,\widehat{\gamma}_2\right)}{S_T\left(-\infty,\widehat{\gamma}_2\right)}, \tag{78}$$

where

$S_T\left(-\infty,-\infty\right)$ is the residual sum of squares from the regression without any threshold variable;

$S_T\left(\widehat{\gamma}_1,-\infty\right)$ is the residual sum of squares from the regression without the second threshold variable; and

$S_T\left(-\infty,\widehat{\gamma}_2\right)$ is the residual sum of squares from the regression without the first threshold variable.

For the notation $F\left(\cdot,\cdot,\cdot\right)$, the first entry in the parenthesis stands for the value of m under the null hypothesis, the second represents the value of m under the alternative hypothesis, and the last indicates that the test is on the i^{th} threshold variable.

If the null cannot be rejected for both threshold variables, then we conclude that there is no threshold effect. If the null is rejected for at least one of the threshold variables, then we proceed to the second step of testing one threshold variable against two threshold variables:

$$H_0 : m = 1$$
$$H_1 : m = 2$$

Define

$$F(1,2,1) = T\frac{S_T(\widehat{\gamma}_1, -\infty) - S_T(\widetilde{\gamma}_1, \widetilde{\gamma}_2)}{S_T(\widetilde{\gamma}_1, \widetilde{\gamma}_2)}, \tag{79}$$

$$F(1,2,2) = T\frac{S_T(-\infty, \widehat{\gamma}_2) - S_T(\widetilde{\gamma}_1, \widetilde{\gamma}_2)}{S_T(\widetilde{\gamma}_1, \widetilde{\gamma}_2)}, \tag{80}$$

where $S_T(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ is the residual sum of squares from the regression by imposing both threshold variables. If the null is rejected in both cases, then we conclude that there are two threshold variables. If we reject the null in the first step for the first threshold variable and cannot reject it in the second test, then the first variable is the only threshold variable. A similar argument applies to the second threshold variable. The problem arises when we reject the null in the first step but accept it in the second step for both variables, which should not occur in large samples. In a finite sample where such a situation occurs, we choose the threshold variable that best fits the model.

The asymptotic distributions of the above tests are non-standard. For the case of H_0: $m = 0$ against H_1: $m = 1$, the bootstrapping method in Hansen (1999) is conducted for each potential candidates of threshold variables. For the tests in the following steps, first we treat the regressors and the threshold variables as given, holding their values fixed in repeated bootstrap samples. We then use the regression residuals under H_1 as the empirical distribution. A sample of size T with replacement is drawn from this empirical distribution and the errors are used to create a bootstrap sample under H_0. The values of structural and threshold parameters are fixed at their estimated values under H_0. We repeat this procedure for a large number of times and calculate the percentage of draws for which the simulated statistic exceeds the actual. This is the bootstrap estimate of the asymptotic p-value under H_0. The null is rejected if the p-value is too small.

For illustration, consider a panel model, we test

$$H_0 : m = 1$$
$$H_1 : m = 2$$

We estimate the threshold model with two threshold variables, take its residuals and draw the bootstrap $\widehat{\varepsilon}_{it}^{b*}$ residuals from them ($i = 1, 2, ..., n$;

$t = 1, 2, ..., T$). We then use the bootstrap residuals along with the estimated threshold model with one threshold variable to generate the bootstrap dependent variable:

$$y_{it}^{b*} = \widehat{\beta}_1' x_{it}^* + \left(\widehat{\beta}_2' - \widehat{\beta}_1'\right) x_{it}^* \Psi\left(z_{it}, \widehat{\gamma}_1\right) + \widehat{\varepsilon}_{it}^{*b}. \tag{81}$$

Using the set of dependent and independent variables $\left\{\mathbf{x}_{it}^*, y_{it}^{b*}\right\}$, we can estimate the model under the alternative hypothesis (in this case, a threshold model with two threshold variables) and compute its sum of squared residuals $S_{nT}\left(\widetilde{\gamma}_1^b, \widetilde{\gamma}_2^b\right)$. The sum of squared residuals under the null is $S_{nT}\left(\widehat{\gamma}_1, -\infty\right) = \sum_{t=1}^{T}\sum_{i=1}^{n}\left(\widehat{\varepsilon}_{it}^{*b}\right)^2$. The test statistic for testing two threshold variables under the alternative hypothesis against the null hypothesis that only the first threshold variable should appear in the model is

$$F(1,2,1) = T\frac{S_{nT}\left(\widehat{\gamma}_1, -\infty\right) - S_{nT}\left(\widetilde{\gamma}_1^b, \widetilde{\gamma}_2^b\right)}{S_{nT}\left(\widetilde{\gamma}_1^b, \widetilde{\gamma}_2^b\right)}. \tag{82}$$

For testing whether only the second threshold variable should appear in the model, the test statistic is

$$F(1,2,2) = T\frac{S_{nT}\left(-\infty, \widehat{\gamma}_2\right) - S_{nT}\left(\widetilde{\gamma}_1^b, \widetilde{\gamma}_2^b\right)}{S_{nT}\left(\widetilde{\gamma}_1^b, \widetilde{\gamma}_2^b\right)}. \tag{83}$$

5.2 Testing the threshold values

After obtaining the number of threshold variables, we proceed to test the hypothesis that

$$H_0 : \gamma = \gamma^0.$$

Under the assumption that ε_t is i.i.d. $N\left(0, \sigma^2\right)$, we have

$$LR_T\left(\gamma_1, \gamma_2\right) = T\frac{S_T\left(\gamma_1, \gamma_2\right) - S_T\left(\widehat{\gamma}_1, \widehat{\gamma}_2\right)}{S_T\left(\widehat{\gamma}_1, \widehat{\gamma}_2\right)}. \tag{84}$$

H_0 is rejected for a large $LR_T\left(\gamma_1^0, \gamma_2^0\right)$.

If the threshold variables are independent, one can show that

$$LR_T\left(\gamma_1^0, \gamma_2^0\right) \xrightarrow{d} \eta^2 \xi, \tag{85}$$

where

$$\xi = \xi_1 + \xi_2, \tag{86}$$

$$\xi_1 = \max_{-\infty<r_1<\infty} \left(-|r_1| + 2W_1(r_1) \right), \tag{87}$$

$$\xi_2 = \max_{-\infty<r_2<\infty} \left(-|r_2| + 2W_2(r_2) \right) \tag{88}$$

and

$$\eta^2 = \frac{c'Vc}{\sigma^2 c'Dc}. \tag{89}$$

The distribution of ξ_i $(i = 1, 2)$ is

$$\Pr(\xi_i \leq x) = \left(1 - e^{-\frac{1}{2}x}\right)^2, \tag{90}$$

$$f_{\xi_i}(x) = \left(1 - e^{-\frac{1}{2}x}\right) e^{-\frac{1}{2}x}. \tag{91}$$

Thus,

$$\begin{aligned} \Pr(\xi \leq x) &= \Pr(\xi_1 + \xi_2 \leq x) \\ &= \int_0^x \Pr(\xi_1 \leq x - y) f_{\xi_2}(y)\, dy \\ &= 1 - (x+5)e^{-x} - 2(x-2)e^{-\frac{1}{2}x}, \end{aligned} \tag{92}$$

The density function is given by

$$f_\xi(x) = (x+4)e^{-x} + (x-4)e^{-\frac{1}{2}x}. \tag{93}$$

For $m = 3$, we have

$$\begin{aligned} \Pr(\xi \leq x) &= \Pr(\xi_1 + \xi_2 + \xi_3 \leq x) \\ &= \int_0^x \Pr(\xi_1 + \xi_2 \leq x - y) f_{\xi_3}(y)\, dy \\ &= \frac{1}{2}e^{-2x}\left(62e^x + 14xe^x + 2e^{2x} + x^2e^x - 64e^{\frac{3}{2}x} + 16xe^{\frac{3}{2}x} - 2x^2e^{\frac{3}{2}x}\right) \end{aligned}$$

and

$$f(x) = \frac{1}{2}e^{-2x}\left(48e^{\frac{3}{2}x} - 12xe^x - x^2e^x - 48e^x - 12xe^{\frac{3}{2}x} + x^2e^{\frac{3}{2}x}\right).$$

In our case, we do not have a closed form solution to compute the critical values analogous to those in Table 1 of Hansen (2000). We solve the critical values by simulations. The values are tabulated in Table A:

$\Pr(\xi \leq x)$	0.800	0.85	0.90	0.925	0.95	0.975	0.99
$m=2$	8.33	9.13	10.21	10.96	11.98	13.68	15.85
$m=3$	11.95	12.90	14.17	15.03	16.20	18.12	20.55
$m=4$	15.47	16.54	17.96	18.92	20.21	22.32	24.96
$m=5$	18.93	20.10	21.65	22.69	24.10	26.38	29.20
$m=6$	22.34	23.61	25.28	26.39	27.90	30.32	33.33
$m=7$	25.71	27.07	28.85	30.04	31.63	34.20	37.35
$m=8$	29.06	30.50	32.38	33.63	35.31	38.00	41.31
$m=9$	32.39	33.90	35.88	37.19	38.95	41.76	45.21
$m=10$	35.70	37.28	39.35	40.72	42.55	45.48	49.06

Table A: Asymptotic Critical Values x

In general, if there are m threshold variables, we can derive the distribution function of ξ uniquely from the moment generating function

$$MGF_{\xi}(t) = \left(\frac{1}{(1-t)(1-2t)}\right)^{m} \quad \text{for } t < 0.5. \tag{94}$$

For the estimation of the nuisance η^2, we can extend the results of Hansen (2000). In our case for $m = 2$, it can be estimated via a polynomial regression with $(z_1, z_1^2, z_2, z_2^2, z_1 z_2)$ as the set of regressors, or via the the Nadaraya-Watson kernel estimator with a bivariate Epanechnikov kernel.

6 Simulations

In all of the experiments below, we set $x_t = 1$ for all t, so that the model becomes

$$y_t = \beta_1 + \delta\Psi_t(\gamma) + \varepsilon_t.$$

We simulate the case where $\Psi_t(\gamma) = \Pi_{j=1}^{2} I\left(z_{jt} > \gamma_j\right)$.

z_{jt} is set to be $i.i.d.$ $N(0,1)$, $\gamma_1^0 = 0$, $\gamma_2^0 = 0, \beta_1 = 1, T = 1000$ (sample size); $N = 10000$ (number of replications); $\varepsilon_t \sim i.i.d. N(0,1)$, $c = 1$, $\alpha = \frac{1}{8}$. All of the simulations are done in GAUSS. The codes are available from the authors upon request.

Experiment A. This experiment simulate the behavior of the residual sum of squares and the distribution of $\widehat{\beta}_1(\widehat{\gamma})$ and $\widehat{\beta}_2(\widehat{\gamma})$ for a fixed break with $\beta_1 = 1$, $\beta_2 = 2$. We estimate the following model:

$$y_t = \beta_1 + \delta\Psi_t(\gamma) + \varepsilon_t.$$

Figure 1a plots the 3D graph of $\frac{1}{T}S_T\left(\gamma_{1,}\gamma_2\right).$

Figure 1b plots the 3D graph of $g\left(\gamma_{1,}\gamma_2\right).$

FIGURES 1a and 1b HERE

Figure 2a plots the distribution of $T^{1/2}\left(\widehat{\beta}_1\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_1\right).$

Figure 2b plots the distribution of $T^{1/2}\left(\widehat{\beta}_2\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_2\right).$

Figure 2c plots the 3D distribution of $T^{1/2}\left(\widehat{\beta}_1\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_1,\widehat{\beta}_2\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_2\right).$

FIGURES 2a – 2c HERE

Experiment B. This experiment studies the distribution of $\widehat{\gamma}$ for a shrinking break. Let $\delta=\beta_2-\beta_1=T^{-\frac{1}{8}}$.

In this case, we have

$$(\widehat{\gamma}_1,\widehat{\gamma}_2)=\arg\min_{(\gamma_1,\gamma_2)\in\Gamma_T}S_T\left(\gamma_{1,}\gamma_2\right)=\arg\min_{(\gamma_1,\gamma_2)\in\Gamma_T}\left[S_T\left(\gamma_{1,}\gamma_2\right)-S_T\left(\gamma_{1,}^0\gamma_2^0\right)\right].$$

From Appendix A3, for $\gamma_1=\gamma_1^0+\frac{\upsilon_1}{T^{1-2\alpha}}$, $\gamma_2=\gamma_2^0+\frac{\upsilon_2}{T^{1-2\alpha}}$, $[\upsilon_1,\upsilon_2]\in R^2$, we have

$$\begin{aligned}&S_T\left(\gamma_{1,}\gamma_2\right)-S_T\left(\gamma_{1,}^0\gamma_2^0\right)\\&\overset{d}{=}-\overline{F}_1^0\left(\upsilon_1+2T^{-\alpha}\sum_{t=1}^{|\upsilon_1|T^{2\alpha}}\varepsilon_t^a\right)-\overline{F}_2^0\left(\upsilon_2+2T^{-\alpha}\sum_{t=1}^{|\upsilon_2|T^{2\alpha}}\varepsilon_t^b\right),\end{aligned}$$

where ε_t^a and ε_t^b are independent. Let

$$r_1=-\overline{F}_1^0\upsilon_1,$$

$$r_2=-\overline{F}_2^0\upsilon_2.$$

When z_{1t} and z_{2t} are independent, we have:

$$\begin{aligned}&T^{1-2\alpha}\left(f_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\left(\widehat{\gamma}_1-\gamma_1^0\right),f_2\left(\gamma_2^0\right)\overline{F}_1\left(\gamma_1^0\right)\left(\widehat{\gamma}_2-\gamma_2^0\right)\right)\\&\overset{d}{\to}\arg\max_{(r_1,r_2)\in R^2}\sum_{j=1}^2\left(-\frac{1}{2}\left|r_j\right|+W_j\left(r_j\right)\right).\end{aligned}$$

Under this setting, we have

$$f_1\left(\gamma_1^0\right) = f_2\left(\gamma_2^0\right) = f_1(0) = \frac{1}{\sqrt{2\pi}},$$

$$\overline{F}_1\left(\gamma_1^0\right) = \overline{F}_2\left(\gamma_2^0\right) = \overline{F}_1(0) = \frac{1}{2}.$$

Figure 3a plots the distribution of $T^{3/4} f_1\left(\gamma_1^0\right) \overline{F}_2\left(\gamma_2^0\right)\left(\widehat{\gamma}_1 - \gamma_1^0\right)$ when $\beta_2 - \beta_1 = T^{-1/8}$.

Figure 3b plots the distribution of $T^{3/4} f_2\left(\gamma_2^0\right) \overline{F}_1\left(\gamma_1^0\right)\left(\widehat{\gamma}_2 - \gamma_2^0\right)$ when $\beta_2 - \beta_1 = T^{-1/8}$.

FIGURES 3a and 3b HERE

Figure 4a plots the 3D distribution of

$$T^{1-2\alpha}\left(f_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\left(\widehat{\gamma}_1 - \gamma_1^0\right), f_2\left(\gamma_2^0\right)\overline{F}_1\left(\gamma_1^0\right)\left(\widehat{\gamma}_2 - \gamma_2^0\right)\right)$$

when $\beta_2 - \beta_1 = T^{-1/8}$.

Figure 4b plots the joint density $f_{(\widehat{r}_1, \widehat{r}_2)}(a_1, a_2)$, the definition of which is provided in eqt.(38).

FIGURES 4a and 4b HERE

Experiment C. This experiment studies the distribution of the LR test and plot the confidence interval around the estimated thresholds for shrinking break with $\delta = \beta_2 - \beta_1 = T^{-\frac{1}{8}}$.

Figure 5a plots the finite sample distribution of the LR statistics

$$LR_T\left(\gamma_1^0, \gamma_2^0\right) = T\frac{S_T\left(\gamma_1^0, \gamma_2^0\right) - S_T\left(\widehat{\gamma}_1, \widehat{\gamma}_2\right)}{S_T\left(\widehat{\gamma}_1, \widehat{\gamma}_2\right)}. \tag{95}$$

FIGURE 5a HERE

With $N = 1$, $T = 1000$, and using the 95% critical value obtained from Table 1 for $m = 2$, Figure 5b plots the simulated 95% confidence interval for (γ_1, γ_2) around the estimated thresholds such that $LR_T(\gamma_1, \gamma_2) = 11.98$.

FIGURE 5b HERE

7 Empirical Application

The empirical relevance of our findings is studied through an application to the currency crisis models. The threshold model is particularly appropriate for the currency crisis issue as all the relevant models in the literature suggest that there are significant threshold effects in the crisis indicators. Identifying critical thresholds in the crisis indicators has important policy implications as the thresholds provide guidelines for policy makers, allowing them to formulate regulatory policies to minimize the stampede of currency crises.

Much of the empirical work on currency crises is concerned with finding relevant crisis indicators (Kaminsky, Lizondo and Reinhart, 1997; Kaminsky, 1998; Hali, 2000). While it is helpful to understand the relevance of different crisis indicators, it is equally important for policy makers to know the critical thresholds of the indicators, above which the economy is unable to sustain a stable exchange rate amidst high pressure in the foreign exchange market. Based on the theory developed in this paper, we can estimate the joint threshold values of several crisis indicators simultaneously. We apply our model to the study of currency crises in 16 countries. It is specified as follows:

$$y_{it} = \mu_i + \beta_1' x_{it} + \left(\beta_2' - \beta_1'\right) x_{it} \Psi\left(\mathbf{z}_{it}, \boldsymbol{\gamma}\right) + \varepsilon_{it}, \tag{96}$$

for $i = 1, 2..., 16$, where

$$\Psi\left(\mathbf{z}_t, \gamma^0\right) = \Pi_{j=1}^{m} I\left(\mathbf{z}_{jt} > \gamma_j^0\right), \tag{97}$$

for $j = 1, 2, ..., m$.

A currency crisis will not be triggered until all of the threshold variables exceed the critical thresholds. The number of threshold variables (m) to be included in the model is determined by the tests that have been discussed in Section 5.1. The fixed effect transformation described in Section 4 is used to remove the individual-specific means from the panel data.

7.1 Exchange market pressure index as the dependent variable y_{it}

In the threshold model, the dependent variable y_{it} is taken to be the exchange market pressure index (EMP_{it}), which is measured as the weighted average of the percentage change in the nominal exchange rate, the change in the differential between the domestic and foreign discount rate (the "policy rate"), and the percentage change in the foreign exchange reserves of a country. This index has been employed in a number of studies, including those of Eichengreen, Rose and Wyplosz (1996), Frankel and Rose (1996), Sachs, Tornell and Velasco (1996), and Goldstein, Kaminsky and Reinhart (2000). The exchange market pressure index is defined as:

$$EMP_{it} \equiv [(\alpha_1 \ \%\Delta e_{it}) + (\alpha_2 \ \Delta(i_{it} - i_{US,t})) - (\alpha_3 \ \%\Delta r_{it})] \tag{98}$$

where

$\%\Delta e_{it}$ denotes the percentage change in the exchange rate of country i with respect to the U.S. dollar at time t;

$\Delta(i_{it} - i_{US,t})$ denotes the change in the differential between the short-term discount rate in country i and the US at time t;

$\%\Delta r_{it}$ denotes the percentage change in the foreign exchange reserves of country i at time t; and

α_1, α_2 and α_3 are the weights that are defined as the inverse of the standard deviations of the respective components over the past ten years. The weights are assigned in order to equalize the volatilities of these three components.

7.2 Choice of the threshold variables z_{it} and regressors x_{it}

The threshold variables z_{it} should be exogenous indicators of currency crisis, and they are selected based on the insights from the three generations of currency crisis model. The first generation model (Krugman, 1979; Flood and Garber, 1984) suggests that exogenous government budget deficits lie at the root of balance of payment crises. The empirical implication of this is that the pressure on the foreign exchange market becomes significantly higher once the fiscal deficit exceeds a certain threshold. Second generation models (Obstfeld, 1986) formulate the possibility of a self-fulfilling currency crisis. In this model, there can be multiple equilibria in the foreign exchange market, and the change from the good equilibrium to the bad equilibrium is self-fulfilling. Threats of attacks generate expectation-driven increases in interest rates. The second generation model implies that one should see a drastic increase in the domestic interest rate before the attack. As a result, the relevant threshold variable is the differential between the domestic interest rate and the foreign interest rate. The third generation model suggests that international illiquidity in a country's financial system precipitates the collapse of the exchange rate. A financial system is internationally illiquid if its short-term foreign currency obligations exceed the amount of foreign currency to which it can have access within a short period of time. The empirical implication of the third generation model is that external illiquidity is a crucial threshold variable in financial and currency crises (McKinnon and Huw, 1996; Chang and Velasco, 1998a and 1998b). Given the implications of these theories, an important empirical question should be whether there are threshold effects in the threshold variables and, if so, what the threshold values are.

Table 1 summarizes the threshold variables z_{it} for the three generations of currency crisis models. The ratio of fiscal deficit to GDP is measured as the total government expenditure minus the total government revenue, normalized by GDP. The interest rate differential is constructed as the difference between the 3-month domestic and the US lending rates. Short-term external liabilities are measured as the sum of the short-term external debt, the cumulative portfolio liabilities and six-month imports. When the values of these threshold variables are too high, the economy enters into an unstable state and the risk of currency depreciation is imminent.

We include two fundamentals as the explanatory variables (x_{it}) in the threshold regression model. These variables include the real exchange rate and the ratio of domestic credit to GDP (Edwards, 1989; Dornbusch, Goldgajn and Valdĕs, 1995; Frankel and Rose, 1996; Sachs, Tornell and Velasco, 1996), both in natural log. The real exchange rate index measures the change in the real exchange rate relative to the base period (1986 Q1) and is employed to capture the degree of exchange rate misalignment over the sample period. In the literature, it is presumed that a large cumulative appreciation in the real exchange rate index signifies a high possibility that the real exchange rate is overvalued; hence, there is a stronger pressure for the real exchange rate to revert to the mean. This measure of misalignment is only an indirect measure and does not control for long-run productivity changes; nevertheless, it remains common in the literature because it helps to identify countries that have experienced extreme overvaluations.

The domestic credit variable is defined as the claims on the private sector by deposit money banks and monetary authorities. It reflects the vulnerability of the banking sector to non-performing loans. As there are no internationally comparable ratios of non-performing loans to total assets, the ratio of domestic credit to GDP is utilized because it is presumed that a sharp bank lending boom over a short period reduces the banks' ability to screen out marginal projects. This makes the banks vulnerable to the vagaries of economic fluctuations. To avoid the endogeneity problem in the estimation, we use the average of the lags in the previous four quarters for all of the regressors and threshold variables. Appendix D provides a detailed description of the sources of the variables.

Crisis models	$\mathbf{z}_{it}$
First generation crisis model	ratio of fiscal deficits to GDP
Second generation crisis model	differentials between the domestic interest rate and foreign interest rate
Third generation crisis model	ratio of short-term external liabilities to foreign exchange reserves

Table 1: Threshold variables implied by the three generations of currency crisis models

7.3 Testing the number of threshold variables

In this section, we apply the tests described in Section 5.1 to test for the presence of threshold effects. The asymptotic distribution of the test statistic is non-standard and generally depends on the moments of the sample. We conduct a bootstrap procedure as follows: First, we estimate the model under the alternative hypothesis. Then, we group the regression residuals (after fixed-effect transformation) $\widehat{\varepsilon}_{it}^{*}$ by individual $\widehat{\varepsilon}_{i}^{*}$ =($\widehat{\varepsilon}_{it}^{*}$, $\widehat{\varepsilon}_{i2}^{*}$,....., $\widehat{\varepsilon}_{iT}^{*}$) and draw, with replacement, error sample of individual i $\widehat{\varepsilon}_{it}^{b*}$ $(t = 1, 2, ..., T)$ from this empirical distribution $\widehat{\varepsilon}_{i}^{*}$. This gives the bootstrap errors. The bootstrap dependent variable y_{it}^{b*} is then generated based on the estimates $\widehat{\beta}$ and $\widehat{\gamma}$ of the threshold model under the null hypothesis. From the bootstrap sample $\left\{x_{it}^{*}, y_{it}^{b*}\right\}$, the test statistic is calculated. This procedure is repeated a large number of times and the p-value of the test statistic is calculated as $p = \frac{1}{B}\sum_{b=1}^{B} I\left\{F^{b} > F^{actual}\right\}$ where F^{b} is the test statistic computed from one bootstrap sample, F^{actual} is the test statistic computed from the actual data, and B is the number of bootstrap replications. In this paper, 300 bootstrap replications are used for each of the tests. The null hypothesis is rejected if the p-value is smaller than the desired significance level.

The test statistics and p-values for testing zero against one, one against two, and two against three threshold variables are performed. The results are reported in Tables 2(a), 2(b) and 2(c). The tests for zero against one threshold variable are all highly significant, with p-values of 0.000, 0.014 and 0.000 when the threshold variable is fiscal deficits, short-term external liabilities and the lending rate differentials, respectively.

The tests for one against two threshold variables are statistically significant for almost all of the cases, with p-values close to 0, except for the cases where the fiscal deficit variable is dropped from the pair of fis-

	$H_0 : m = 0$		
	$H_1 : m = 1$ (fiscal deficit)	$H_1 : m = 1$ (short liability)	$H_1 : m = 1$ (lending rate diff.)
F	51.7909	30.0459	50.1684
p-value	0.0000**	0.0140*	0.0000**

Note: The numbers in parentheses are the p-values. "**" means the test statistic is significant at the 5% level and "*" means the test statistic is significant at the 1% level. 300 bootstrap replications are used for each of the test.

Table 2: (a) Testing one threshold variable against no threshold variable

cal deficit and short-term external liabilities, and from the pair of fiscal deficit and the lending rate differential under the alternative hypothesis. The p-values for these two cases are 0.9667 and 1. When testing two against three threshold variables, the null hypothesis that the fiscal deficit variable can be dropped from the list of three is not rejected and the p-value is 0.9866. Based on these results, we conclude that there is strong evidence for two threshold variables in the regression relationship, namely, the ratio of short-term external liabilities to reserves and the lending rate differential. For the remainder of the paper we work with a threshold model with these two threshold variables.

An explanation for the absence of a threshold effect in the fiscal deficit variable is that fiscal deficits are often closely related in practice to interest rate differentials; hence, only one of these two needs to be included as the threshold variable. One reason for this is that large fiscal deficits are commonly financed by excessively expansionary monetary policies, which drive up the risk premium of the domestic currency and widen the interest rate differential. In addition, if a large fiscal deficit is accompanied by a high public debt, the government can only roll over its short-term public debt by offering a higher domestic interest rate, which results in a larger interest rate differential.

	$H_0 : m = 1$ (fiscal deficit)	$H_0 : m = 1$ (short liabilities)
	$H_1 : m = 2$ (fiscal deficit, short liabilities)	
F	27.1031	5.3490
p-value	0.0000**	0.9667
	$H_0 : m = 1$ (fiscal deficit)	$H_0 : m = 1$ (lending rate diff.)
	$H_1 : m = 2$ (fiscal deficit, lending rate diff.)	
F	62.0889	0.3865
p-value	0.0000**	1.0000
	$H_0 : m = 1$ (short liabilities)	$H_0 : m = 1$ (lending rate diff.)
	$H_1 : m = 2$ (short liabilities, lending rate diff.)	
F	91.0530	22.8329
p-value	0.0000**	0.0000**

Table 2: (b) Testing two threshold variables against one threshold variable

	$H_0 : m = 2$ (fiscal deficit, short liabilities)	$H_0 : m = 2$ (fiscal deficit, lending diff.)	$H_0 : m = 2$ (short liabilities, lending diff.)
	$H_1 : m = 3$ (fiscal deficit, short liabilities, lending rate diff.)		
F	96.1324	21.9379	1.1477
p-value	0.0000**	0.0153*	0.9866

Table 2: (c) Testing three threshold variables against two threshold variables

7.4 Estimation Results

In this section, we estimate the threshold values for the ratio of short-term external liabilities to reserves and the lending rate differential. To allow for different thresholds for countries in different geographical regions, we divide the sample countries into the Asian and Latin American regions. The estimation results are represented in Table 3. The point estimates of the ratio of short-term external liabilities to reserves for the Asian and Latin American countries are 3.1758 and 3.9851 respectively. The point estimates for the lending rate differential for the Asian and Latin American countries are 2.0566 and 21.8866 percentage points (or 205.66 and 2188.66 basis points) respectively. The test statistics for $H_0 : m = 0$ against $H_1 : m = 2$ are highly significant for countries in both regions. The test statistic, along with the p-value, are 33.8296 and 0.0000 for the Asian countries and 34.0269 and 0.0000 for the Latin American countries. The p-values obtained using the bootstrap procedure discussed in Section 7.3 give strong evidence of threshold effects. When both threshold variables exceed their critical thresholds, the risk of currency depreciation is imminent. These threshold estimates can be used to formulate regulatory policies to reduce the risk of a currency crisis.

The coefficients of the ratio of domestic credit to GDP for both the Asian and the Latin American countries are significantly positive when both threshold variables surpass the critical thresholds (that is, when $\Psi(\mathbf{z}_{it}, \boldsymbol{\gamma}) = 1$). This indicates that the vulnerability of the banking sector is a crucial factor in determining the exchange market pressure under this regime.

7.5 A Graphical Analysis of the Threshold Effects

We analyse how well these threshold values can be used to distinguish the normal regime from the crisis regime in foreign exchange markets. Given (1) the threshold estimates of 3.1758 for the short-term external liability variable and 2.0566 for the lending rate differential variable for Asian countries, and (2) the estimates of 3.9851 and 21.8866 for the Latin American countries, we define crisis episodes as extreme values of the exchange market pressure index,

$$\begin{aligned} Crisis_{it} &= 1 \quad \text{if } EMP_{it} > \mu_{EMP,it} + 3\sigma_{EMP,it} \\ &= 0 \quad \text{otherwise} \end{aligned}$$

$\mathbf{y}_{it} \equiv$ exchange market pressure index (EMP_{it}) $\mathbf{z}_{it} \equiv \{\frac{\text{short term external liabilities}}{\text{reserves}}$, lending rate differentials$\}$ $\mathbf{x}_{it} \equiv \{1$, real exchange rate appreciation index, $\frac{\text{Domestic credit}}{\text{GDP}}\}$				
	Asian countries		Latin American countries	
Explanatory variables	(i) $\Psi(\mathbf{z}_{it}, \boldsymbol{\gamma}) = 0$	(ii) $\Psi(\mathbf{z}_{it}, \boldsymbol{\gamma}) = 1$	(iii) $\Psi(\mathbf{z}_{it}, \boldsymbol{\gamma}) = 0$	(iv) $\Psi(\mathbf{z}_{it}, \boldsymbol{\gamma}) = 1$
Constant	-0.1989	0.0512	0.2281	0.3692
	(-0.9436)	(0.6371)	(1.5579)	(1.1598)
Real exch. rate appreciation	-2.7235	1.7293	-0.6257	3.8072
	(-1.8493)	(6.2783)**	(-1.7762)	(1.4069)
$\frac{\text{Domestic credit}}{\text{GDP}}$	0.0588	1.7464	1.0528	3.4369
	(0.0442)	(6.0512)**	(1.7287)	(3.0245)**
threshold estimates				
$\frac{\text{Short term external liabilities}}{\text{Reserves}}$	3.1758		3.9851	
Lending rate differentials	2.0566		21.8866	
F stat	33.8296		34.0269	
p value	0.0000**		0.0000**	
Observations	641		379	

Note: The numbers in parentheses are the t-statistics. "**" means that the t statistic is significant at the 5% level and "*" means that the t statistic is significant at the 1% level.

Table 3: Estimates of the threshold models with two threshold variables

where $\mu_{EMP,it}$ and $\sigma_{EMP,it}$ are, respectively, the mean and standard deviation of the exchange market pressure index for country i at time t. The dates of the crisis episodes in the sample are reported in Table 4.

The threshold effects are illustrated in Figures 6 and 7, which show the values of the threshold variables (represented by the bars in the figures), the critical thresholds (the dashed lines), and the exchange market pressure index (the solid lines) of the Latin American and Asian countries. The crisis episodes are shaded in grey. The figures indicate that the threshold variables perform reasonably well in predicting the regime shifts. For instance, Figures 6(c) and 6(d) show that the ratio of short-term external liabilities to reserves and the lending rate differentials in Indonesia and South Korea started to exceed the critical thresholds less than two years before the 1997 financial crisis, and remained above the thresholds at the outbreak of the crisis. Figure 6(f) indicates that the 1997 currency crisis in the Philippines occurred as soon as the short-term external liabilities exceeded the critical threshold, given that the lending rate differential had already surpassed the threshold prior to the crisis. Figure 7(b) indicates that both the ratio of short-term external liabilities to reserves and the lending rate differential started to go above the critical thresholds less than one year prior to the Brazilian crisis of 2000, and remained above the thresholds throughout the crisis. Figure 7(g) shows that the 1994 Venezuelan crisis broke out as soon as the short-term external liabilities moved above the critical threshold, given that the lending rate differential had, at that time, already gone above the critical threshold.

Nevertheless, we observe two false alarms in our sample. These occurred in the Philippines in 1990-92 and in Mexico in 1999. One explanation is that the Philippine government adopted a tight monetary policy, aggressively cut government spending and raised indirect taxes during that period to reduce the downward pressure in the foreign exchange market. This effectively steered the economy away from a currency crisis, despite the high values of the threshold variables (Bautista, 2000). Mexico was in the middle of a capital liberalization process during 1999, which significantly raised the amount of capital inflow into the country. This helped lower the downward pressure in the foreign exchange market.

Countries	Crisis Episodes
1. Argentina	2001Q4
2. Brazil	1998Q3-1999Q1, 2000Q2
3. Chile	1990Q4
4. China	1992Q3-1993Q2
5. Colombia	None
6. Hong Kong	None
7. Indonesia	1997Q3-1998Q2
8. S. Korea	1997Q4
9. Malaysia	1997Q3-Q4, 1998Q2
10. Mexico	1994Q4
11. Philippines	1984Q1, 1997Q3
12. Singapore	1997Q3-Q4, 1998Q2
13. Taiwan	1997Q4
14. Thailand	1981Q3, 1997Q3-Q4
15. Uruguay	1994Q3-1995Q2
16. Venezuela	1994Q2

Note: Crisis episodes that occurred within one year of each other in the same country are considered as one continuous episode.

Table 4: Dates of Crisis Episodes

8 Conclusion

Threshold regression models have been widely studied in various academic disciplines. However, conventional threshold models only contain one threshold variable and thus have limited applications when two or more threshold variables are needed. Thus far, little is known about the estimation and inference in models with multiple threshold variables. This paper presents a new model that allows for more than one multiple threshold variable. The asymptotic estimation and testing theories are derived for this model. These asymptotic results are supported by simulation evidence. We apply our model to the study of currency crises in 16 countries. The selection of the threshold variables in this study is closely guided by the premises of the three generations of currency crisis models. The first generation model suggests that the pressure on the foreign exchange market becomes significantly higher once the fiscal deficit exceeds a certain threshold. The second generation model suggests that an economy switches from a good equilibrium (a non-crisis equilibrium) to a bad equilibrium (a crisis equilibrium) once the expectation-driven increases in domestic interest rates relative to the foreign interest rates exceed a certain threshold. The third generation model indicates that short-term external liabilities relative to reserves is one crucial threshold variable in currency crises. We find overwhelming evidence of threshold effects in the ratio of short-term external liabilities to reserves and in the lending rate differential. Our finding provides strong empirical support for the currency crisis models in the literature. More importantly, our estimates of the joint threshold values can be adopted as policy guidelines in the regulation of short-term external borrowing and interest rate differentials. Pre-emptive measures might be taken to prevent these variables from crossing the critical threshold values, in order to prevent a currency crisis from happening. Finally, it should be mentioned that the applicability of our new model extends beyond the scope of economics. It can also serve as a foundation for further studies.

References

1. Astatkie, T., D. G. Watts, and W. E. Watt (1997). Nested Threshold Autoregressive (NeTAR) Models. International Journal of Forecasting 13, 105-116.

2. Bai, J. (1997). Estimating Multiple Breaks one at a time. Econometric Theory 13, 315-52.

3. Bai, J. and P. Perron (1998). Estimating and Testing Linear Models with Multiple Structural Changes. Econometrica 66(1), 47-78.

4. Bautista, C. (2000). Boom-bust Cycles and Crisis Periods in the Philippines, a Regime-switching Analysis. Working paper, the University of the Philippines.

5. Bhattacharya, P. K, and P. J. Brockwell (1976). The Minimum of an Additive Process with Applications to Signal Estimation and Storage Theory. Z. Wahrschein. Verw. Gebiete, 37, 51-75.

6. Billingsley, P. (1968). Convergence of Probability Measures. New York: John Wiley and Sons.

7. Calvo, G. and G. Mendoza (2000). Capital Markets Crises and Economic Collapse in Emerging Markets: An Informational-Frictions Approach. American Economic Review Papers and Proceedings 90(2), 59-64.

8. Chan, K. S. and H. Tong (1986). On Estimating Thresholds in Autoregressive Models. Journal of Time Series Analysis 7, 179-190.

9. Chan, K. S. and R. S. Tsay (1998). Limiting Properties of the Least Squares Estimator of a Continuous Threshold Autoregressive Model. Biometrika 85, 413-426.

10. Chang, R. and A. Velasco (1998a). Financial Fragility and the Exchange Rate Regime, NBER WP6469.

11. Chang, R. and A. Velasco (1998b). Financial Crises in Emerging Markets: A Canonical Model, Federal Reserve Bank of Atlanta, working paper.

12. Chen, H., T. T. L. Chong and J. Bai (2012). Theory and Applications of TAR Model with two Threshold Variables. Econometric Reviews 31(2), 142–170.

13. Chen, R. and S. Tsay (1993). Functional-coefficient Autoregressive Models. Journal of the American Statistical Association 88, 298-308.

14. Chong, T. T. L. (2001). Structural Change in AR(1) Models. Econometric Theory 17, 87-155.

15. Christiano, L. J. (1992). Searching for a Break in GNP. Journal of Business & Economic Statistics 10, 237-250.

16. Dornbusch, R., I. Goldgajn and R. O. Valdĕs (1999). Currency Crises and Collapses. Brookings Papers on Economic Activity, No.2, 219-95.

17. Durlauf, S. N. and P. A. Johnson (1995). Multiple Regimes and Cross-country Growth Behavior. Journal of Applied Econometrics 10, 365-384.

18. 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.

19. Edwards, S. (1989). Real Exchange Rates, Devaluations, and Adjustment: Exchange Rate Policy in Developing Countries, Cambridge, Massachusetts: MIT Press.

20. Eichengreen, B., Rose, A. and C. Wyplosz (1995). Exchange Market Mayhem: The Antecedents and Aftermath of Speculative Attacks. *Economic Policy* 21, 249-312.

21. Eichengreen, B., Rose, A., and C. Wyplosz (1996). Contagious Currency Crises, *NBER* WP5681.

22. Flood, R. and P. Garber (1984). Collapsing Exchange Rate Regimes: Some Linear Examples. *Journal of International Economics* 17, 1-13.

23. Frankel, J. A. and A. Rose (1996). Currency Crashes in Emerging Markets: Empirical Indicators, *NBER* WP5437.

24. Garcia, R. and P. Perron (1994). An Analysis of the Real Interest rate under Regime Shifts. *Review of Economics and Statistics* 78, 111-125.

25. Goldstein, M., G. Kaminsky and C. Reinhart (2000). Assessing Financial Vulnerability – An Early Warning System for Emerging Markets, *Institute for International Economics* 51, 145-168.

26. Gonzalo, J. and J. Pitarakis (2002). Estimation and Model Selection Based Inference in Single and Multiple Threshold Models. *Journal of Econometrics* 110, 319-352.

27. Hali, J. E. (2000). Do Indicators of Financial Crises Work? An evaluation of the early warning system. Board of Governors of Federal Reserve System, International Finance Discussion Paper #675.

28. Hansen, B. E. (1999). Threshold Effect in Non-dynamic Panels: Estimation, Testing, and Inference. *Journal of Econometrics* 93, 345-368.

29. Hansen, B. E. (2000). Sample Splitting and Threshold Estimation. *Econometrica* 68, 575-603.

30. Henry, Ó., N. Olekaln and P. M. Summers (2001). Exchange Rate Instability: a Threshold Autoregressive Approach. *Economic Record* 77, 160-166.

31. 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.

32. Kaminsky, G. L., S. Lizondo and C. M. Reinhart (1997). Leading Indicators of Currency Crises. IMF Staff Papers 45, 1-48.

33. Kakkar, V. and I. Yan (2012). Real Exchange Rates and Productivity: Evidence From Asia. *Journal of Money Credit and Banking* 44(2-3), 301-322.

34. Krugman, P. (1979). A Model of Balance of Payments Crises. *Journal of Money, Credit and Banking*, 311-325.

35. Krugman, P. (1999). Balance Sheets, the Transfer Problem, and Financial Crisis. *Paper prepared for the volume in honor of Robert Flood.*

36. Kumhof, M. , S. Li and I. Yan (2007). Balance of Payments Crises under Inflation Targeting. *Journal of International Economics* 72(1), 242-264.

37. Mankiw, N. G., D. Romer and D. N. Weil (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* 107(2), 407-437.

38. McKinnon, I. Ronald. and P. Huw (1996). Credible Economic Liberalizations and Overborrowing. *American Economic Review* 87(2), 189-193.

39. Mishkin, F. S. (1990). What does the Term Structure Tell us about Future Inflation? *Journal of Monetary Economics* 25, 77-95.

40. Obstfeld, M. (1986). Rational and Self-Fulfilling Balance of Payment Crises. *American Economic Review* 76(1), 72-81.

41. Pollard, D. (1984), *Convergence of Stochastic Processes*, Springer-Verlag Inc., New York.

42. Potter, S. M. (1995). A Nonlinear Approach to US GNP. *Journal of Applied Econometrics* 2, 109-125.

43. Qiu, P., Asano, C. and X. Li (1991). Estimation of Jump Regression Function. *Bulletin of Informatics and Cybernetic*s, 24, No. 3-4, 197-212.

44. Quintos, C., Fan, Z. and P. C. B. Phillips (2001). Structural Change Tests in Tail Behavior and the Asian Crisis. *Review of Economic Studies* 68, 633-663.

45. Roley, V. V. and S. M. Wheatley (1996). Shifts in the Interest-rate Response to Money Announcements: What can We say about when they occur? *Journal of Business & Economics Statistics* 14, 135-138.

46. Sachs, J., A. Tornell and A. Velasco (1996). Financial Crises in Emerging Markets: The Lessons from 1995. *Brookings Papers on Economic Activity No.1,* 147-215.

47. Tong, H. (1983). *Threshold Models in Nonlinear Time Series Analysis: Lecture Notes in Statistic*s 21. Berlon: Springer.

48. Tsay, R. S. (1998). Testing and Modeling Multivariate Threshold Models. *Journal of the American Statistical Association* 93, 1188-1202.

49. Xia, Y. and W. K. Li (1999). On Single-index Coefficient Regression Model, *Journal of the American Statistical Association* 94, 1275-1285.

Appendix A1: Asymptotic behavior of the OLS estimators and $\frac{1}{T}S_T\left(\gamma_{1,}\gamma_2\right)$ when $x_t=1$

Note that

$$
\begin{aligned}
\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) &= \frac{\sum_{t=1}^T y_t\left(1-\Psi_t\left(\gamma\right)\right)}{\sum_{t=1}^T\left(1-\Psi_t\left(\gamma\right)\right)} \\
&= \frac{\sum_{t=1}^T\left(\beta_1\left(1-\Psi_t\left(\gamma^0\right)\right)+\beta_2\Psi_t\left(\gamma^0\right)+\varepsilon_t\right)\left(1-\Psi_t\left(\gamma\right)\right)}{\sum_{t=1}^T\left(1-\Psi_t\left(\gamma\right)\right)} \\
&= \beta_1+\delta\frac{\sum_{t=1}^T\Psi_t\left(\gamma^0\right)\left(1-\Psi_t\left(\gamma\right)\right)}{\sum_{t=1}^T\left(1-\Psi_t\left(\gamma\right)\right)}+o_p\left(1\right) \\
&= \beta_1+\delta\times \\
&\quad \frac{\sum_{t=1}^T\left[I\left(z_{1t}>\gamma_1^0,z_{2t}>\gamma_2^0\right)-I\left(z_{1t}>\max\left\{\gamma_1^0,\gamma_1\right\},z_{2t}>\max\left\{\gamma_2^0,\gamma_2\right\}\right)\right]}{\sum_{t=1}^T\left(1-I\left(z_{1t}>\gamma_1,z_{2t}>\gamma_2\right)\right)} \\
&\quad +o_p\left(1\right).
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) &= \frac{\sum_{t=1}^T y_t\Psi_t\left(\gamma\right)}{\sum_{t=1}^T\Psi_t\left(\gamma\right)} \\
&= \beta_2-\delta\frac{\sum_{t=1}^T\left(1-\Psi_t\left(\gamma^0\right)\right)\Psi_t\left(\gamma\right)}{\sum_{t=1}^T\Psi_t\left(\gamma\right)}+o_p\left(1\right) \\
&= \beta_2-\delta\times \\
&\quad \frac{\sum_{t=1}^T\left[I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2\right)-I\left(z_{1t}>\max\left\{\gamma_1^0,\gamma_1\right\},z_{2t}>\max\left\{\gamma_2^0,\gamma_2\right\}\right)\right]}{\sum_{t=1}^T I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2\right)} \\
&\quad +o_p\left(1\right).
\end{aligned}
$$

$$
\begin{aligned}
&\frac{1}{T}S_T\left(\gamma_{1,}\gamma_2\right)=\sum_{t=1}^T\left(y_t-\widehat{\beta}_1-\widehat{\delta}\Psi_t\left(\gamma\right)\right)^2 \\
&=\frac{1}{T}\sum_{t=1}^T\left(\beta_1+\delta\Psi_t\left(\gamma^0\right)+\varepsilon_t-\widehat{\beta}_1-\widehat{\delta}\Psi_t\left(\gamma\right)\right)^2 \\
&=\frac{1}{T}\sum_{t=1}^T\left(-\left(\widehat{\beta}_1-\beta_1\right)+\delta\Psi_t\left(\gamma^0\right)-\widehat{\delta}\Psi_t\left(\gamma\right)\right)^2+\frac{1}{T}\sum_{t=1}^T\varepsilon_t^2+o_p\left(1\right).
\end{aligned}
$$

We discuss four cases:

Case 1: $\gamma_1\leq\gamma_1^0,\gamma_2\leq\gamma_2^0$

$$\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right)=\beta_1+o_p\left(1\right),$$

$$\widehat{\beta}_2(\gamma_1,\gamma_2) = \beta_2 - \delta\left(1 - \frac{\sum_{t=1}^{T} I(z_{1t} > \gamma_1^0,\ z_{2t} > \gamma_2^0)}{\sum_{t=1}^{T} I(z_{1t} > \gamma_1,\ z_{2t} > \gamma_2)}\right) + o_p(1)$$
$$= \beta_2 - \delta\left(1 - \frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\right) + o_p(1).$$

Note that

$$\widehat{\delta} = \widehat{\beta}_2(\gamma_1,\gamma_2) - \widehat{\beta}_1(\gamma_1,\gamma_2) = \delta\frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)} + o_p(1).$$

$$\begin{aligned}
&\frac{1}{T}S_T(\gamma_1,\gamma_2)\\
&= \delta^2\frac{1}{T}\sum_{t=1}^{T}\left(\Psi_t(\gamma^0) - \frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\Psi_t(\gamma)\right)^2 + \sigma^2 + o_p(1)\\
&= \delta^2\frac{1}{T}\sum_{t=1}^{T}\Psi_t(\gamma^0) + \delta^2\left(\frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\right)^2\frac{1}{T}\sum_{t=1}^{T}\Psi_t(\gamma)\\
&-2\delta^2\frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\frac{1}{T}\sum_{t=1}^{T}\Psi_t(\gamma^0)\Psi_t(\gamma) + \sigma^2 + o_p(1)\\
&= \delta^2\overline{F}(\gamma_1^0, \gamma_2^0) + \delta^2\left(\frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\right)^2\overline{F}(\gamma_1, \gamma_2) - 2\delta^2\frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\overline{F}(\gamma_1^0, \gamma_2^0)\\
&+\sigma^2 + o_p(1)\\
&\xrightarrow{p} g(\gamma_1, \gamma_2),
\end{aligned}$$
where

$$g(\gamma_1, \gamma_2) = \sigma^2 + \delta^2\overline{F}(\gamma_1^0, \gamma_2^0)\left(1 - \frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\right) \geq g(\gamma_1^0, \gamma_2^0) = \sigma^2,$$

$$\frac{\partial g(\gamma_1, \gamma_2)}{\partial \gamma_1} = \delta^2\left(\frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\right)^2\overline{F}_{1\gamma} \leq 0,$$

$$\frac{\partial g(\gamma_1, \gamma_2)}{\partial \gamma_2} = \delta^2\left(\frac{\overline{F}(\gamma_1^0, \gamma_2^0)}{\overline{F}(\gamma_1, \gamma_2)}\right)^2\overline{F}_{2\gamma} \leq 0.$$

Case 2: $\gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0$

$$\widehat{\beta}_1(\gamma_1,\gamma_2) = \beta_1 + \delta\frac{\sum_{t=1}^{T} I(\gamma_1^0 < z_{1t} \leq \gamma_1, z_{2t} > \gamma_2^0)}{\sum_{t=1}^{T}(1 - I(z_{1t} > \gamma_1,\ z_{2t} > \gamma_2))} + o_p(1)$$
$$\xrightarrow{p} \beta_1 + \delta\frac{\overline{F}(\gamma_1^0, \gamma_2^0) - \overline{F}(\gamma_1, \gamma_2^0)}{1 - \overline{F}(\gamma_1, \gamma_2)},$$

$$\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right)=\beta_2-\delta\frac{\sum_{t=1}^{T}\left[I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2\right)-I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2^0\right)\right]}{\sum_{t=1}^{T}I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2\right)}+o_p\left(1\right)$$

$$\xrightarrow{p}\beta_2-\delta\left(1-\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\right).$$

Note that

$$\widehat{\delta}\xrightarrow{p}\delta\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)-\overline{F}\left(\gamma_1^0,\gamma_2^0\right)\overline{F}\left(\gamma_1,\gamma_2\right)}{\left(1-\overline{F}\left(\gamma_1,\gamma_2\right)\right)\overline{F}\left(\gamma_1,\gamma_2\right)}$$

$$\begin{aligned}
&\frac{1}{T}S_T\left(\gamma_{1,}\gamma_2\right)\\
&=\delta^2\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}\dfrac{\overline{F}\left(\gamma_1,\gamma_2^0\right)-\overline{F}\left(\gamma_1^0,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}+\Psi_t\left(\gamma^0\right)\\ -\dfrac{\overline{F}\left(\gamma_1,\gamma_2^0\right)-\overline{F}\left(\gamma_1^0,\gamma_2^0\right)\overline{F}\left(\gamma_1,\gamma_2\right)}{\left(1-\overline{F}\left(\gamma_1,\gamma_2\right)\right)\overline{F}\left(\gamma_1,\gamma_2\right)}\Psi_t\left(\gamma\right)\end{pmatrix}^2+\sigma^2+o_p\left(1\right)\\
&=\delta^2\frac{1}{T}\sum_{t=1}^{T}\left(-\frac{\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\left(1-\Psi_t\left(\gamma\right)\right)+\Psi_t\left(\gamma^0\right)-\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\Psi_t\left(\gamma\right)\right)^2\\
&+\sigma^2+o_p\left(1\right)\\
&=\delta^2\left(\frac{\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\right)^2\frac{1}{T}\sum_{t=1}^{T}\left(1-\Psi_t\left(\gamma\right)\right)+\delta^2\frac{1}{T}\sum_{t=1}^{T}\Psi_t\left(\gamma^0\right)\\
&+\delta^2\left(\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\right)^2\frac{1}{T}\sum_{t=1}^{T}\Psi_t\left(\gamma\right)-2\frac{\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\delta^2\frac{1}{T}\sum_{t=1}^{T}\left(1-\Psi_t\left(\gamma\right)\right)\Psi_t\left(\gamma^0\right)\\
&-2\delta^2\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\frac{1}{T}\sum_{t=1}^{T}\Psi_t\left(\gamma^0\right)\Psi_t\left(\gamma\right)+\sigma^2+o_p\left(1\right)\\
&=\delta^2\left(\frac{\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\right)^2\left(1-\overline{F}\left(\gamma_1,\gamma_2\right)\right)+\delta^2\overline{F}\left(\gamma_1^0,\gamma_2^0\right)+\delta^2\left(\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\right)^2\overline{F}\left(\gamma_1,\gamma_2\right)\\
&-2\frac{\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\delta^2\frac{1}{T}\sum_{t=1}^{T}\left(I\left(z_{1t}>\gamma_1^0,z_{2t}>\gamma_2^0\right)-I\left(z_{1t}>\gamma_1,z_{2t}>\gamma_2^0\right)\right)\\
&-2\delta^2\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\frac{1}{T}\sum_{t=1}^{T}I\left(z_{1t}>\gamma_1,z_{2t}>\gamma_2^0\right)+\sigma^2+o_p\left(1\right)\\
&=\delta^2\frac{\left(\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)\right)^2}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}+\delta^2\overline{F}\left(\gamma_1^0,\gamma_2^0\right)+\delta^2\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)^2}{\overline{F}\left(\gamma_1,\gamma_2\right)}\\
&-2\frac{\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\delta^2\left(\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)\right)-2\delta^2\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\overline{F}\left(\gamma_1,\gamma_2^0\right)\\
&+\sigma^2+o_p\left(1\right)\\
&\xrightarrow{p}g\left(\gamma_1,\gamma_2\right),
\end{aligned}$$

where

$$g(\gamma_1,\gamma_2)=\sigma^2+\delta^2\left[\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\frac{\left(\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2^0\right)\right)^2}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}-\frac{\overline{F}\left(\gamma_1,\gamma_2^0\right)^2}{\overline{F}\left(\gamma_1,\gamma_2\right)}\right].$$

Rewrite

$$g(\gamma_1,\gamma_2)=\sigma^2+\delta^2\left[(c-a)\frac{1-b-c+a}{1-b}+a\frac{b-a}{b}\right],$$

where

$$a=\overline{F}\left(\gamma_1,\gamma_2^0\right),$$

$$b=\overline{F}\left(\gamma_1,\gamma_2\right),$$

$$c=\overline{F}\left(\gamma_1^0,\gamma_2^0\right).$$

Using the fact that $c>a$, $b>a$ and $b+c-a<1$, we have

$$g(\gamma_1,\gamma_2)>g\left(\gamma_1^0,\gamma_2^0\right)=\sigma^2.$$

Case 3: $\gamma_1\le\gamma_1^0,\gamma_2>\gamma_2^0$

$$\begin{aligned}\widehat{\beta}_1\left(\gamma_1,\gamma_2\right)&=\beta_1+\delta\frac{\sum_{t=1}^T\left[I\left(z_{1t}>\gamma_1^0,\ z_{2t}>\gamma_2^0\right)-I\left(z_{1t}>\gamma_1^0,\ z_{2t}>\gamma_2\right)\right]}{\sum_{t=1}^T\left(1-I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2\right)\right)}+o_p(1)\\&\xrightarrow{p}\beta_1+\delta\frac{\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1^0,\gamma_2\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)},\end{aligned}$$

$$\begin{aligned}\widehat{\beta}_2\left(\gamma_1,\gamma_2\right)&=\beta_2-\delta\frac{\sum_{t=1}^T\left[I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2\right)-I\left(z_{1t}>\gamma_1^0,\ z_{2t}>\gamma_2\right)\right]}{\sum_{t=1}^T I\left(z_{1t}>\gamma_1,\ z_{2t}>\gamma_2\right)}+o_p(1)\\&\xrightarrow{p}\beta_2-\delta\left(1-\frac{\overline{F}\left(\gamma_1^0,\gamma_2\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)}\right).\end{aligned}$$

Note that

$$\widehat{\delta}\xrightarrow{p}\delta\frac{\overline{F}\left(\gamma_1^0,\gamma_2\right)-\overline{F}\left(\gamma_1,\gamma_2\right)\overline{F}\left(\gamma_1^0,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)\left(1-\overline{F}\left(\gamma_1,\gamma_2\right)\right)}.$$

$$\begin{aligned}&\frac{1}{T}S_T\left(\gamma_1,\gamma_2\right)\\&=\delta^2\frac{1}{T}\sum_{t=1}^T\left(\frac{\overline{F}\left(\gamma_1^0,\gamma_2\right)-\overline{F}\left(\gamma_1^0,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}+\Psi_t\left(\gamma^0\right)-\frac{\overline{F}\left(\gamma_1^0,\gamma_2\right)-\overline{F}\left(\gamma_1,\gamma_2\right)\overline{F}\left(\gamma_1^0,\gamma_2^0\right)}{\overline{F}\left(\gamma_1,\gamma_2\right)\left(1-\overline{F}\left(\gamma_1,\gamma_2\right)\right)}\Psi_t\left(\gamma\right)\right)^2\end{aligned}$$

$$
\begin{aligned}
&+\sigma^2 + o_p(1) \\
&= \delta^2 \frac{1}{T} \textstyle\sum_{t=1}^{T} \left(\dfrac{\overline{F}(\gamma_1^0,\gamma_2) - \overline{F}(\gamma_1^0,\gamma_2^0)}{1-\overline{F}(\gamma_1,\gamma_2)} (1-\Psi_t(\gamma)) + \Psi_t(\gamma^0) - \dfrac{\overline{F}(\gamma_1^0,\gamma_2)}{\overline{F}(\gamma_1,\gamma_2)} \Psi_t(\gamma) \right)^2 \\
&+\sigma^2 + o_p(1) \\
&= \delta^2 \left(\frac{\overline{F}(\gamma_1^0,\gamma_2) - \overline{F}(\gamma_1^0,\gamma_2^0)}{1-\overline{F}(\gamma_1,\gamma_2)} \right)^2 \frac{1}{T} \textstyle\sum_{t=1}^{T} (1-\Psi_t(\gamma)) + \delta^2 \dfrac{1}{T} \textstyle\sum_{t=1}^{T} \Psi_t(\gamma^0) \\
&+\delta^2 \left(\frac{\overline{F}(\gamma_1^0,\gamma_2)}{\overline{F}(\gamma_1,\gamma_2)} \right)^2 \frac{1}{T} \textstyle\sum_{t=1}^{T} \Psi_t(\gamma) + 2\delta^2 \dfrac{\overline{F}(\gamma_1^0,\gamma_2) - \overline{F}(\gamma_1^0,\gamma_2^0)}{1-\overline{F}(\gamma_1,\gamma_2)} \dfrac{1}{T} \textstyle\sum_{t=1}^{T} (1-\Psi_t(\gamma)) \Psi_t(\gamma^0) \\
&-2\delta^2 \frac{\overline{F}(\gamma_1^0,\gamma_2)}{\overline{F}(\gamma_1,\gamma_2)} \frac{1}{T} \textstyle\sum_{t=1}^{T} \Psi_t(\gamma^0) \Psi_t(\gamma) + \sigma^2 + o_p(1) \\
&= \delta^2 \left(\frac{\overline{F}(\gamma_1^0,\gamma_2) - \overline{F}(\gamma_1^0,\gamma_2^0)}{1-\overline{F}(\gamma_1,\gamma_2)} \right)^2 \left(1-\overline{F}(\gamma_1,\gamma_2)\right) + \delta^2 \overline{F}(\gamma_1^0,\gamma_2^0) \\
&+\delta^2 \left(\frac{\overline{F}(\gamma_1^0,\gamma_2)}{\overline{F}(\gamma_1,\gamma_2)} \right)^2 \overline{F}(\gamma_1,\gamma_2) \\
&+2\delta^2 \frac{\overline{F}(\gamma_1^0,\gamma_2) - \overline{F}(\gamma_1^0,\gamma_2^0)}{1-\overline{F}(\gamma_1,\gamma_2)} \frac{1}{T} \textstyle\sum_{t=1}^{T} \left(I(z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0) - I(z_{1t} > \gamma_1^0, z_{2t} > \gamma_2) \right) \\
&-2\delta^2 \frac{\overline{F}(\gamma_1^0,\gamma_2)}{\overline{F}(\gamma_1,\gamma_2)} \frac{1}{T} \textstyle\sum_{t=1}^{T} I(z_{1t} > \gamma_1^0, z_{2t} > \gamma_2) + \sigma^2 + o_p(1) \\
&= \delta^2 \frac{\left(\overline{F}(\gamma_1^0,\gamma_2) - \overline{F}(\gamma_1^0,\gamma_2^0)\right)^2}{1-\overline{F}(\gamma_1,\gamma_2)} + \delta^2 \overline{F}(\gamma_1^0,\gamma_2^0) + \delta^2 \frac{\overline{F}(\gamma_1^0,\gamma_2)^2}{\overline{F}(\gamma_1,\gamma_2)} \\
&-2\delta^2 \frac{\left(\overline{F}(\gamma_1^0,\gamma_2) - \overline{F}(\gamma_1^0,\gamma_2^0)\right)^2}{1-\overline{F}(\gamma_1,\gamma_2)} - 2\delta^2 \frac{\overline{F}(\gamma_1^0,\gamma_2)}{\overline{F}(\gamma_1,\gamma_2)} \overline{F}(\gamma_1^0,\gamma_2) + \sigma^2 + o_p(1) \\
&\xrightarrow{p} g(\gamma_1,\gamma_2),
\end{aligned}
$$

where

$$
g(\gamma_1,\gamma_2) = \sigma^2 + \delta^2 \left[\overline{F}\left(\gamma_1^0,\gamma_2^0\right) - \frac{\left(\overline{F}(\gamma_1^0,\gamma_2^0) - \overline{F}(\gamma_1^0,\gamma_2)\right)^2}{1-\overline{F}(\gamma_1,\gamma_2)} - \frac{\overline{F}(\gamma_1^0,\gamma_2)^2}{\overline{F}(\gamma_1,\gamma_2)} \right].
$$

Rewrite

$$
g(\gamma_1,\gamma_2) = \sigma^2 + \delta^2 \left[(c-d) \frac{1-b-c+d}{1-b} + d \frac{b-d}{b} \right]
$$

where

$$
d = \overline{F}\left(\gamma_1^0,\gamma_2\right).
$$

Use the facts that $c > d$, $b > d$ and $b + c - d < 1$, we have

$$
g(\gamma_1,\gamma_2) > g\left(\gamma_1^0,\gamma_2^0\right) = \sigma^2.
$$

Case 4: $\gamma_1 > \gamma_1^0, \gamma_2 > \gamma_2^0$

$$\begin{aligned}
\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) &= \beta_1 + \delta \frac{\sum_{t=1}^{T}\left[I\left(z_{1t} > \gamma_1^0,\ z_{2t} > \gamma_2^0\right) - I\left(z_{1t} > \gamma_1,\ z_{2t} > \gamma_2\right)\right]}{\sum_{t=1}^{T}\left(1 - I\left(z_{1t} > \gamma_1,\ z_{2t} > \gamma_2\right)\right)} + o_p(1) \\
&= \beta_1 + \delta \frac{\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} + o_p(1)
\end{aligned}$$

$$\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) = \beta_2 + o_p(1)$$

Note that

$$\widehat{\delta} \xrightarrow{p} \delta \frac{1 - \overline{F}\left(\gamma_1^0, \gamma_2^0\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)},$$

$$\Psi_t(\gamma)\,\Psi_t\left(\gamma^0\right) = \Psi_t(\gamma)\,.$$

$$\begin{aligned}
&\frac{1}{T} S_T\left(\gamma_{1,}\gamma_2\right) \\
&= \delta^2 \frac{1}{T} \sum_{t=1}^{T} \left(-\frac{\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} + \Psi_t\left(\gamma^0\right) - \frac{1 - \overline{F}\left(\gamma_1^0, \gamma_2^0\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} \Psi_t(\gamma) \right)^2 \\
&+\sigma^2 + o_p(1) \\
&= \delta^2 \frac{1}{T} \sum_{t=1}^{T} \left(-\frac{\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} \left(1 - \Psi_t(\gamma)\right) + \Psi_t\left(\gamma^0\right) - \Psi_t(\gamma) \right)^2 \\
&+\sigma^2 + o_p(1) \\
&= \delta^2 \left(\frac{\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} \right)^2 \frac{1}{T} \sum_{t=1}^{T} \left(1 - \Psi_t(\gamma)\right) \\
&+\delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t\left(\gamma^0\right) + \delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t(\gamma) \\
&-2\delta^2 \frac{\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} \frac{1}{T} \sum_{t=1}^{T} \left(1 - \Psi_t(\gamma)\right) \Psi_t\left(\gamma^0\right) \\
&-2\delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t\left(\gamma^0\right) \Psi_t(\gamma) + \sigma^2 + o_p(1) \\
&= \delta^2 \left(\frac{\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} \right)^2 \left(1 - \overline{F}\left(\gamma_1, \gamma_2\right)\right) \\
&+\delta^2 \overline{F}\left(\gamma_1^0, \gamma_2^0\right) + \delta^2 \overline{F}\left(\gamma_1, \gamma_2\right) \\
&-2\delta^2 \frac{\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} \left(\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)\right) \\
&-2\delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t(\gamma) + \sigma^2 + o_p(1) \\
&= \delta^2 \left(\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right) - \frac{\left(\overline{F}\left(\gamma_1^0, \gamma_2^0\right) - \overline{F}\left(\gamma_1, \gamma_2\right)\right)^2}{1 - \overline{F}\left(\gamma_1, \gamma_2\right)} \right) + \sigma^2 + o_p(1) \\
&\xrightarrow{p} g\left(\gamma_1, \gamma_2\right)
\end{aligned}$$

where

$$g\left(\gamma_1,\gamma_2\right)=\sigma^2+\delta^2\left(\overline{F}\left(\gamma_1^0,\gamma_2^0\right)-\overline{F}\left(\gamma_1,\gamma_2\right)\right)\frac{1-\overline{F}\left(\gamma_1^0,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}>g\left(\gamma_1^0,\gamma_2^0\right)=\sigma^2,$$

$$\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_1}=-\delta^2\left(\frac{1-\overline{F}\left(\gamma_1^0,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\right)^2\overline{F}_{1\gamma}>0,$$

$$\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_2}=-\delta^2\left(\frac{1-\overline{F}\left(\gamma_1^0,\gamma_2^0\right)}{1-\overline{F}\left(\gamma_1,\gamma_2\right)}\right)^2\overline{F}_{2\gamma}>0.$$

Appendix A2: Asymptotic behavior of the OLS estimators and $\frac{1}{T}S_T\left(\gamma_{1,}\gamma_2\right)$ when $x_t = 1$ and threshold variables are independent

When threshold variables are independent, we have

$$\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) = \beta_1 + \delta \times \frac{\sum_{t=1}^{T}\left[I\left(z_{1t} > \gamma_1^0\right)I\left(z_{2t} > \gamma_2^0\right) - I\left(z_{1t} > \max\left\{\gamma_1^0, \gamma_1\right\}\right)I\left(z_{2t} > \max\left\{\gamma_2^0, \gamma_2\right\}\right)\right]}{\sum_{t=1}^{T}\left(1 - I\left(z_{1t} > \gamma_1\right)I\left(z_{2t} > \gamma_2\right)\right)} + o_p(1).$$

Similarly, we have

$$\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) = \beta_2 - \delta \times \frac{\sum_{t=1}^{T}\left[I\left(z_{1t} > \gamma_1\right)I\left(z_{2t} > \gamma_2\right) - I\left(z_{1t} > \max\left\{\gamma_1^0, \gamma_1\right\}\right)I\left(z_{2t} > \max\left\{\gamma_2^0, \gamma_2\right\}\right)\right]}{\sum_{t=1}^{T}I\left(z_{1t} > \gamma_1\right)I\left(z_{2t} > \gamma_2\right)} + o_p(1)$$

Case 1: $\gamma_1 \leq \gamma_1^0, \gamma_2 \leq \gamma_2^0$

$$\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) = \beta_1 + o_p(1),$$

$$\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) = \beta_2 - \delta\left(1 - \frac{\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right) + o_p(1),$$

$$\widehat{\delta} = \widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) - \widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) \xrightarrow{p} \delta\frac{\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)},$$

$$g\left(\gamma_1, \gamma_2\right) = \sigma^2 + \delta^2\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\left(1 - \frac{\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right),$$

$$\frac{\partial g\left(\gamma_1, \gamma_2\right)}{\partial \gamma_1} = -\delta^2\frac{\overline{F}_1\left(\gamma_1^0\right)^2\overline{F}_2\left(\gamma_2^0\right)^2}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}H_1\left(\gamma_1\right) \leq 0,$$

$$\frac{\partial g\left(\gamma_1, \gamma_2\right)}{\partial \gamma_2} = -\delta^2\frac{\overline{F}_1\left(\gamma_1^0\right)^2\overline{F}_2\left(\gamma_2^0\right)^2}{\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}H_2\left(\gamma_2\right) \leq 0,$$

where $H_1\left(\gamma_1\right)$ and $H_2\left(\gamma_2\right)$ are the hazard functions of z_1 and z_2 respectively.

Case 2: $\gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0$

$$\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) \xrightarrow{p} \beta_1 + \delta\frac{\left(\overline{F}_1\left(\gamma_1^0\right) - \overline{F}_1\left(\gamma_1\right)\right)\overline{F}_2\left(\gamma_2^0\right)}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)},$$

$$\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) \xrightarrow{p} \beta_2 - \delta\left(1 - \frac{\overline{F}_2\left(\gamma_2^0\right)}{\overline{F}_2\left(\gamma_2\right)}\right),$$

$$\widehat{\delta} \xrightarrow{p} \delta\frac{\overline{F}_2\left(\gamma_2^0\right)\left[1 - \overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2\right)\right]}{\overline{F}_2\left(\gamma_2\right)\left[1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)\right]},$$

$$g\left(\gamma_1,\gamma_2\right) = \sigma^2 + \delta^2\overline{F}_2\left(\gamma_2^0\right)^2\left[-\frac{\left(\overline{F}_1\left(\gamma_1^0\right) - \overline{F}_1\left(\gamma_1\right)\right)^2}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)} + \frac{\overline{F}_1\left(\gamma_1^0\right)}{\overline{F}_2\left(\gamma_2^0\right)} - \frac{\overline{F}_1\left(\gamma_1\right)}{\overline{F}_2\left(\gamma_2\right)}\right],$$

$$\begin{aligned}\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_1} &= \delta^2\overline{F}_2\left(\gamma_2^0\right)^2\left(\frac{\overline{F}_1\left(\gamma_1^0\right) - \overline{F}_1\left(\gamma_1\right)}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)} - \frac{1}{\overline{F}_2\left(\gamma_2\right)}\right)^2\overline{F}_2\left(\gamma_2\right)f_1\left(\gamma_1\right)\\ &> 0,\end{aligned}$$

$$\begin{aligned}\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_2} &= -\delta^2\overline{F}_2\left(\gamma_2^0\right)^2\left[\frac{\overline{F}_1\left(\gamma_1^0\right) - \overline{F}_1\left(\gamma_1\right)}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)} + \frac{1}{\overline{F}_2\left(\gamma_2\right)}\right]\times\\ &\quad\frac{1 - \overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2\right)}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\overline{F}_1\left(\gamma_1\right)H_2\left(\gamma_2\right)\\ &< 0.\end{aligned}$$

Case 3: $\gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0$

$$\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) \xrightarrow{p} \beta_1 + \delta\frac{\overline{F}_1\left(\gamma_1^0\right)\left(\overline{F}_2\left(\gamma_2^0\right) - \overline{F}_2\left(\gamma_2\right)\right)}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)},$$

$$\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) \xrightarrow{p} \beta_2 - \delta\left(1 - \frac{\overline{F}_1\left(\gamma_1^0\right)}{\overline{F}_1\left(\gamma_1\right)}\right),$$

$$\widehat{\delta} \xrightarrow{p} \delta\frac{\overline{F}_1\left(\gamma_1^0\right)\left[1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2^0\right)\right]}{\overline{F}_1\left(\gamma_1\right)\left[1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)\right]},$$

$$g\left(\gamma_1,\gamma_2\right) = \sigma^2 + \delta^2\overline{F}_1\left(\gamma_1^0\right)^2\left[-\frac{\left(\overline{F}_2\left(\gamma_2^0\right) - \overline{F}_2\left(\gamma_2\right)\right)^2}{1 - \overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)} + \frac{\overline{F}_2\left(\gamma_2^0\right)}{\overline{F}_1\left(\gamma_1^0\right)} - \frac{\overline{F}_2\left(\gamma_2\right)}{\overline{F}_1\left(\gamma_1\right)}\right],$$

$$
\begin{aligned}
\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_1} &= -\delta^2\overline{F}_1\left(\gamma_1^0\right)^2\left(\frac{\overline{F}_2\left(\gamma_2^0\right)-\overline{F}_2\left(\gamma_2\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}+\frac{1}{\overline{F}_1\left(\gamma_1\right)}\right) \\
&\quad\times\frac{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\overline{F}_2\left(\gamma_2\right)H_1\left(\gamma_1\right) \\
&< 0,
\end{aligned}
$$

$$
\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_2} = \delta^2\overline{F}_1\left(\gamma_1^0\right)^2\left(\frac{\overline{F}_2\left(\gamma_2^0\right)-\overline{F}_2\left(\gamma_2\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}-\frac{1}{\overline{F}_1\left(\gamma_1\right)}\right)^2\overline{F}_1\left(\gamma_1\right)f_2\left(\gamma_2\right) > 0.
$$

Case 4: $\gamma_1 > \gamma_1^0, \gamma_2 > \gamma_2^0$

$$
\widehat{\beta}_1\left(\gamma_{1,}\gamma_2\right) \xrightarrow{p} \beta_1+\delta\frac{\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)},
$$

$$
\widehat{\beta}_2\left(\gamma_{1,}\gamma_2\right) = \beta_2+o_p\left(1\right),
$$

$$
\widehat{\delta} \xrightarrow{p} \delta\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)},
$$

$$
g\left(\gamma_1,\gamma_2\right) = \sigma^2+\delta^2\left(1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\right)\left(1-\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right),
$$

$$
\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_1} = \delta^2\left(\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right)^2\overline{F}_2\left(\gamma_2\right)f_1\left(\gamma_1\right) > 0,
$$

$$
\frac{\partial g\left(\gamma_1,\gamma_2\right)}{\partial\gamma_2} = \delta^2\left(\frac{1-\overline{F}_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)}{1-\overline{F}_1\left(\gamma_1\right)\overline{F}_2\left(\gamma_2\right)}\right)^2\overline{F}_1\left(\gamma_1\right)f_2\left(\gamma_2\right) > 0.
$$

Appendix A3: Distribution of $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ when $x_t = 1$

Define

$$(\widehat{\gamma}_1, \widehat{\gamma}_2) = \arg\min S_T\left(\gamma_{1,}\gamma_2\right) = \arg\min\left[S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma_{1,}^0\gamma_2^0\right)\right].$$

To derive the limiting distribution of $\widehat{\gamma}$ for a shrinking break, we let $\delta = T^{-\alpha}$, $0 < \alpha < \frac{1}{2}$, we have

$$\overline{F}(\gamma_1, \gamma_2) = \overline{F}(\gamma_1^0, \gamma_2^0) + (\gamma_1 - \gamma_1^0)\overline{F}_1^0 + (\gamma_2 - \gamma_2^0)\overline{F}_2^0 + o(1),$$
$$\overline{F}(\gamma_1, \gamma_2^0) = \overline{F}(\gamma_1^0, \gamma_2^0) + (\gamma_1 - \gamma_1^0)\overline{F}_1^0 + o(1),$$
$$\overline{F}(\gamma_1^0, \gamma_2) = \overline{F}(\gamma_1^0, \gamma_2^0) + (\gamma_2 - \gamma_2^0)\overline{F}_2^0 + o(1).$$

Thus,

$$\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma_{1,}^0\gamma_2^0\right)\\
&= \textstyle\sum_{t=1}^T \left(\delta\Psi_t(\gamma^0) + \varepsilon_t - \widehat{\delta}\Psi_t(\gamma)\right)^2 - \sum_{t=1}^T \left(\delta\Psi_t(\gamma^0) + \varepsilon_t - \widehat{\delta}^0\Psi_t(\gamma^0)\right)^2\\
&+o_p(1)\\
&= \delta\textstyle\sum_{t=1}^T (\delta\Psi_t(\gamma^0) + 2\varepsilon_t - \delta\Psi_t(\gamma))(\Psi_t(\gamma^0) - \Psi_t(\gamma)) + o_p(1)\\
&= \delta^2\textstyle\sum_{t=1}^T (\Psi_t(\gamma^0) - \Psi_t(\gamma))^2 + 2\delta\sum_{t=1}^T \varepsilon_t(\Psi_t(\gamma^0) - \Psi_t(\gamma)) + o_p(1)
\end{aligned}$$

In the neighborhood of the true thresholds, where $\gamma_1 = \gamma_1^0 + \dfrac{\upsilon_1}{T^{1-2\alpha}}$, $\gamma_2 = \gamma_2^0 + \dfrac{\upsilon_2}{T^{1-2\alpha}}$, all estimators can be approximated by the true values, so we have the following:

Case 1: $\upsilon_1 \leq 0$, $\upsilon_2 \leq 0$

$$\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma_{1,}^0\gamma_2^0\right)\\
&= T^{-2\alpha}T\left(\overline{F}(\gamma_1, \gamma_2) - \overline{F}(\gamma_1^0, \gamma_2^0)\right)\\
&+2T^{-\alpha}\textstyle\sum_{\left(\gamma_1<z_{1t}<\gamma_1^0 \text{ and } \gamma_2^0<z_{2t}\right) \text{ or } \left(\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1^0<z_{1t}\right) \text{or } \left(\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1<z_{1t}<\gamma_1^0\right)} \varepsilon_t\\
&+o_p(1)\\
&= T^{1-2\alpha}\left((\gamma_1 - \gamma_1^0)\overline{F}_1^0 + (\gamma_2 - \gamma_2^0)\overline{F}_2^0\right)\\
&+2T^{-\alpha}\textstyle\sum_{\left(\gamma_1<z_{1t}<\gamma_1^0 \text{ and } \gamma_2^0<z_{2t}\right) \text{ or } \left(\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1^0<z_{1t}\right)} \varepsilon_t + o_p(1)\\
&\overset{d}{=} \upsilon_1\overline{F}_1^0 + \upsilon_2\overline{F}_2^0\\
&+2T^{-\alpha}\textstyle\sum_{t=1}^{T\Pr\left(\gamma_1<z_{1t}<\gamma_1^0 \text{ and } \gamma_2^0<z_{2t}\right)} \varepsilon_t^a + 2T^{-\alpha}\sum_{t=1}^{T\Pr\left(\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1^0<z_{1t}\right)} \varepsilon_t^b\\
&\overset{d}{=} \upsilon_1\overline{F}_1^0 + 2T^{-\alpha}\textstyle\sum_{t=1}^{-\upsilon_1T^{2\alpha}\overline{F}_1^0} \varepsilon_t^a + \upsilon_2\overline{F}_2^0 + 2T^{-\alpha}\sum_{t=1}^{-\upsilon_2T^{2\alpha}\overline{F}_2^0} \varepsilon_t^b,
\end{aligned}$$

where ε_t^a and ε_t^b are independent.

Case 2: $\upsilon_1 > 0,\ \upsilon_2 \leq 0 \quad S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma^0_{1,}\gamma^0_2\right)$

$$
\begin{aligned}
&= T^{-2\alpha}T\left(\overline{F}\left(\gamma^0_1,\gamma_2\right) - \overline{F}\left(\gamma_1,\gamma^0_2\right)\right) \\
&+2T^{-\alpha}\sum\nolimits_{\left(\gamma^0_1<z_{1t}<\gamma_1 \text{ and } \gamma^0_2<z_{2t}\right) \text{ or } \left(\gamma_2<z_{2t}<\gamma^0_2 \text{ and } \gamma_1<z_{1t}\right) \text{ or } \left(\gamma_2<z_{2t}<\gamma^0_2 \text{ and } \gamma^0_1<z_{1t}<\gamma_1\right)} \varepsilon_t \\
&+o_p(1) \\
&= T^{-2\alpha}T\left(\left(\gamma_2-\gamma^0_2\right)\overline{F}^0_2 - \left(\gamma_1-\gamma^0_1\right)\overline{F}^0_1\right) \\
&+2T^{-\alpha}\sum\nolimits_{\gamma^0_1<z_{1t}<\gamma_1 \text{ and } \gamma^0_2<z_{2t}} \varepsilon_t + 2T^{-\alpha}\sum\nolimits_{\gamma_2<z_{2t}<\gamma^0_2 \text{ and } \gamma_1<z_{1t}} \varepsilon_t + o_p(1) \\
&\overset{d}{=} \upsilon_2\overline{F}^0_2 - \upsilon_1\overline{F}^0_1 \\
&+2T^{-\alpha}\sum\nolimits_{t=1}^{T\Pr\left(\gamma^0_1<z_{1t}<\gamma_1 \text{ and } \gamma^0_2<z_{2t}\right)} \varepsilon^a_t + 2T^{-\alpha}\sum\nolimits_{t=1}^{T\Pr\left(\gamma_2<z_{2t}<\gamma^0_2 \text{ and } \gamma_1<z_{1t}\right)} \varepsilon^b_t \\
&\overset{d}{=} -\upsilon_1\overline{F}^0_1 + 2T^{-\alpha}\sum\nolimits_{t=1}^{-\upsilon_1 T^{2\alpha}\overline{F}^0_1}\varepsilon^a_t + \upsilon_2\overline{F}^0_2 + 2T^{-\alpha}\sum\nolimits_{t=1}^{-\upsilon_2 T^{2\alpha}\overline{F}^0_2}\varepsilon^b_t.
\end{aligned}
$$

Case 3: $\upsilon_1 \leq 0,\ \upsilon_2 > 0$

$$
\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma^0_{1,}\gamma^0_2\right) \\
&= T^{-2\alpha}T\left(\overline{F}\left(\gamma_1,\gamma^0_2\right) - \overline{F}\left(\gamma^0_1,\gamma_2\right)\right) \\
&+2T^{-\alpha}\sum\nolimits_{\left(\gamma_1<z_{1t}<\gamma^0_1 \text{ and } \gamma_2<z_{2t}\right) \text{ or } \left(\gamma^0_2<z_{2t}<\gamma_2 \text{ and } \gamma^0_1<z_{1t}\right) \text{ or } \left(\gamma_1<z_{2t}<\gamma^0_1 \text{ and } \gamma^0_2<z_{1t}<\gamma_2\right)} \varepsilon_t \\
&+o_p(1) \\
&= T^{1-2\alpha}\left(\left(\gamma_1-\gamma^0_1\right)\overline{F}^0_1 - \left(\gamma_2-\gamma^0_2\right)\overline{F}^0_2\right) \\
&+2T^{-\alpha}\sum\nolimits_{\left(\gamma_1<z_{1t}<\gamma^0_1 \text{ and } \gamma_2<z_{2t}\right) \text{ or } \left(\gamma^0_2<z_{2t}<\gamma_2 \text{ and } \gamma^0_1<z_{1t}\right)} \varepsilon_t + O_p\left(T^{-2\alpha}\right) \\
&\overset{d}{=} \upsilon_1\overline{F}^0_1 - \upsilon_2\overline{F}^0_2 \\
&+2T^{-\alpha}\sum\nolimits_{t=1}^{T\Pr\left(\gamma_1<z_{1t}<\gamma^0_1 \text{ and } \gamma_2<z_{2t}\right)} \varepsilon^a_t + 2T^{-\alpha}\sum\nolimits_{t=1}^{T\Pr\left(\gamma^0_2<z_{2t}<\gamma_2 \text{ and } \gamma^0_1<z_{1t}\right)} \varepsilon^b_t \\
&\overset{d}{=} \upsilon_1\overline{F}^0_1 + 2T^{-\alpha}\sum\nolimits_{t=1}^{-\upsilon_1 T^{2\alpha}\overline{F}^0_1}\varepsilon^a_t - \upsilon_2\overline{F}^0_2 + 2T^{-\alpha}\sum\nolimits_{t=1}^{-\upsilon_2 T^{2\alpha}\overline{F}^0_2}\varepsilon^b_t.
\end{aligned}
$$

Similarly, we have

Case 4: $\upsilon_1 > 0,\ \upsilon_2 > 0$

$$
\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma^0_{1,}\gamma^0_2\right) \\
&\overset{d}{=} -\upsilon_1\overline{F}^0_1 + 2T^{-\alpha}\sum\nolimits_{t=1}^{-\upsilon_1 T^{2\alpha}\overline{F}^0_1}\varepsilon^a_t - \upsilon_2\overline{F}^0_2 + 2T^{-\alpha}\sum\nolimits_{t=1}^{-\upsilon_2 T^{2\alpha}\overline{F}^0_2}\varepsilon^b_t.
\end{aligned}
$$

Let

$$r_1 = -\overline{F}^0_1\upsilon_1,$$

$$r_2 = -\overline{F}^0_2\upsilon_2.$$

We have

$$
\begin{aligned}
&-T^{1-2\alpha}\left(\left(\widehat{\gamma}_1-\gamma_1^0\right)\overline{F}_1^0,\left(\widehat{\gamma}_2-\gamma_2^0\right)\overline{F}_2^0\right)\\
&\xrightarrow{d}\arg\max_{-\infty<r_1<\infty,-\infty<r_2<\infty}\left(-\frac{1}{2}\left|r_1\right|+W_1\left(r_1\right)-\frac{1}{2}\left|r_2\right|+W_2\left(r_2\right)\right).
\end{aligned}
$$

When z_1 and z_2 are independent, we have

$$
\overline{F}_1^0=-f_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)
$$

and

$$
\overline{F}_2^0=-f_2\left(\gamma_2^0\right)\overline{F}_1\left(\gamma_1^0\right).
$$

Thus,

$$
\begin{aligned}
&T^{1-2\alpha}\left(f_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\left(\widehat{\gamma}_1-\gamma_1^0\right),f_2\left(\gamma_2^0\right)\overline{F}_1\left(\gamma_1^0\right)\left(\widehat{\gamma}_2-\gamma_2^0\right)\right)\\
&\xrightarrow{d}\arg\max_{(r_1,r_2)\in R^2}\left(-\frac{1}{2}\left|r_1\right|+W_1\left(r_1\right)-\frac{1}{2}\left|r_2\right|+W_2\left(r_2\right)\right).
\end{aligned}
$$

Appendix B: Asymptotic behavior of $S_T\left(\gamma_{1,}\gamma_2\right)$ for vector x_t.

Let $X_0 = X_{\gamma_0}$. As $Y - X\beta_1 - X_\gamma\boldsymbol{\delta}$ and X lies in the space spanned by $P_\gamma = \widetilde{X}_\gamma\left(\widetilde{X}'_\gamma\widetilde{X}_\gamma\right)^{-1}\widetilde{X}'_\gamma$,

$S_T(\gamma) - \varepsilon'\varepsilon$
$= Y'(I - P_\gamma)Y - \varepsilon'\varepsilon$
$= (X\beta_1 + X_0\boldsymbol{\delta} + \varepsilon)'(I - P_\gamma)(X\beta_1 + X_0\boldsymbol{\delta} + \varepsilon) - \varepsilon'\varepsilon$
$= -\varepsilon' P_\gamma\varepsilon + \beta_1' X'(I - P_\gamma)X\beta_1 + 2\boldsymbol{\delta}'X_0'(I - P_\gamma)\varepsilon + \boldsymbol{\delta}'X_0'(I - P_\gamma)X_0\boldsymbol{\delta}$
$+2\beta_1'X'(I - P_\gamma)X_0\boldsymbol{\delta} + 2\beta_1'X'(I - P_\gamma)\varepsilon$
$= -\varepsilon'P_\gamma\varepsilon + 2\boldsymbol{\delta}'X_0'(I - P_\gamma)\varepsilon + \boldsymbol{\delta}'X_0'(I - P_\gamma)X_0\boldsymbol{\delta}.$

Let $\delta = \dfrac{c}{T^\alpha}$,

$$\frac{1}{T^{1-2\alpha}}(S_T(\gamma) - \varepsilon'\varepsilon) = \frac{1}{T}c'(X_0'(I - P_\gamma)X_0)c + o_p(1).$$

The projection P_γ can be written as the projection onto $[X - X_\gamma, X_\gamma]$ where $X - X_\gamma$ is a matrix whose t^{th} row is $x_t'(1 - \Psi_t(\gamma))$. Observe that $(X - X_\gamma)'X_\gamma = 0$.

$$P_\gamma = (X - X_\gamma, X_\gamma)\begin{pmatrix}(X - X_\gamma)'(X - X_\gamma) & 0 \\ 0 & X_\gamma'X_\gamma\end{pmatrix}^{-1}(X - X_\gamma, X_\gamma)'$$
$$= (X - X_\gamma)\left[(X - X_\gamma)'(X - X_\gamma)\right]^{-1}(X - X_\gamma)' + X_\gamma\left(X_\gamma'X_\gamma\right)^{-1}X_\gamma'.$$

$$X_0'P_\gamma X_0 = X_0'(X - X_\gamma)\left[(X - X_\gamma)'(X - X_\gamma)\right]^{-1}(X - X_\gamma)'X_0 + X_0'X_\gamma\left(X_\gamma'X_\gamma\right)^{-1}X_\gamma'X_0.$$

We discuss four cases:

Case 1: $\gamma_1 \leq \gamma_1^0, \gamma_2 \leq \gamma_2^0$ In this case, we have

$X_\gamma'X_0 = X_0'X_0$;
$(X - X_\gamma)'X_0 = 0$;
$X_0'P_\gamma X_0 = X_0'X_0\left(X_\gamma'X_\gamma\right)^{-1}X_0'X_0$.

Thus, we have

$$\frac{1}{T^{1-2\alpha}}(S_T(\gamma) - \varepsilon'\varepsilon)$$
$$= \frac{1}{T}c'X_0'(I - P_\gamma)X_0c + o_p(1)$$
$$= c'\left(\overline{M}_T(\gamma_0) - \overline{M}_T(\gamma_0)\overline{M}_T^{-1}(\gamma)\overline{M}_T(\gamma_0)\right)c$$
$$\xrightarrow{p} c'\left(\overline{M}_0 - \overline{M}_0\overline{M}_\gamma^{-1}\overline{M}_0\right)c$$
$$\equiv b_1(\gamma).$$

As

$$\frac{\partial}{\partial\gamma_1}\overline{M}(\gamma_1, \gamma_2) = D_\gamma\overline{F}_{1\gamma} \leq 0,$$

$$\frac{\partial}{\partial \gamma_2} \overline{M}(\gamma_1, \gamma_2) = D_\gamma \overline{F}_{2\gamma} \leq 0.$$

Thus,

$$\frac{\partial}{\partial \gamma_1} b_1(\gamma) = c' \overline{M}_0 \overline{M}_\gamma^{-1} D_\gamma \overline{F}_{1\gamma} \overline{M}_\gamma^{-1} \overline{M}_0 c \leq 0,$$

$$\frac{\partial}{\partial \gamma_2} b_1(\gamma) = c' \overline{M}_0 \overline{M}_\gamma^{-1} D_\gamma \overline{F}_{2\gamma} \overline{M}_\gamma^{-1} \overline{M}_0 c \leq 0.$$

Case 2: $\gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0$ In this case, we have

$$\frac{1}{T} X_\gamma' X_0 \xrightarrow{p} \overline{M}(\gamma_1, \gamma_2^0);$$

$$\frac{1}{T}(X - X_\gamma)' X_0 \xrightarrow{p} \overline{M}_0 - \overline{M}(\gamma_1, \gamma_2^0);$$

$$\frac{1}{T} X_0' P_\gamma X_0 \xrightarrow{p} \left(\overline{M}_0 - \overline{M}(\gamma_1, \gamma_2^0)\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - \overline{M}(\gamma_1, \gamma_2^0)\right)$$
$$+\overline{M}(\gamma_1, \gamma_2^0) \overline{M}_\gamma^{-1} \overline{M}(\gamma_1, \gamma_2^0).$$

Then, we have

$$\frac{1}{T^{1-2\alpha}}\left(S_T(\gamma) - \varepsilon'\varepsilon\right)$$
$$\xrightarrow{p} c' \begin{pmatrix} \overline{M}_0 - \left(\overline{M}_0 - \overline{M}(\gamma_1, \gamma_2^0)\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - \overline{M}(\gamma_1, \gamma_2^0)\right) \\ -\overline{M}(\gamma_1, \gamma_2^0) \overline{M}_\gamma^{-1} \overline{M}(\gamma_1, \gamma_2^0) \end{pmatrix} c$$
$$\equiv b_2(\gamma) > b_2(\gamma_0) = 0,$$

as $X_0'(I - P_\gamma) X_0 = X_0'(I - P_\gamma)'(I - P_\gamma) X_0$ which is positive semi-definite.

Case 3: $\gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0$ In this case, we have

$$\frac{1}{T} X_\gamma' X_0 \xrightarrow{p} \overline{M}(\gamma_1^0, \gamma_2)$$

$$\frac{1}{T}(X - X_\gamma)' X_0 \xrightarrow{p} \overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2).$$

$$\frac{1}{T} X_0' P_\gamma X_0 \xrightarrow{p} \left(\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2)\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2)\right)$$
$$+\overline{M}(\gamma_1^0, \gamma_2) \overline{M}_\gamma^{-1} \overline{M}(\gamma_1^0, \gamma_2).$$

Then we have

$$\frac{1}{T^{1-2\alpha}}\left(S_T(\gamma) - \varepsilon'\varepsilon\right)$$
$$\xrightarrow{p} c' \begin{pmatrix} \overline{M}_0 - \left(\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2)\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2)\right) \\ -\overline{M}(\gamma_1^0, \gamma_2) \overline{M}_\gamma^{-1} \overline{M}(\gamma_1^0, \gamma_2) \end{pmatrix} c$$
$$\equiv b_3(\gamma) > b_3(\gamma_0) = 0,$$

as $X_0'(I - P_\gamma) X_0 = X_0'(I - P_\gamma)'(I - P_\gamma) X_0$, which is positive semi-definite.

Case 4: $\gamma_1 > \gamma_1^0, \gamma_2 > \gamma_2^0$ In this case, we have

$X_\gamma' X_0 = X_\gamma' X_\gamma;$

$(X - X_\gamma)' X_0 = X_0' X_0 - X_\gamma' X_\gamma.$

Then, we have

$$\frac{1}{T^{1-2\alpha}} \left(S_T(\gamma) - \varepsilon'\varepsilon\right)$$

$$\xrightarrow{p} c' \left(\overline{M}_0 - \overline{M}_\gamma - \left(\overline{M}_0 - \overline{M}_\gamma\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - \overline{M}_\gamma\right)\right) c$$

$$= c' \left(\overline{M}_0 - \overline{M}_\gamma - \left(\overline{M}_0 - M + M - \overline{M}_\gamma\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(\overline{M}_0 - M + M - \overline{M}_\gamma\right)\right) c$$

$$= c' \left(M - \overline{M}_0 - \left(M - \overline{M}_0\right)\left(M - \overline{M}_\gamma\right)^{-1}\left(M - \overline{M}_0\right)\right) c$$

$$\equiv b_4(\gamma).$$

Thus,

$$\frac{\partial}{\partial \gamma_1} b_4(\gamma) = -c' \left(\left(M - \overline{M}_0\right)\left(M - \overline{M}_\gamma\right)^{-1} D_\gamma \overline{F}_{1\gamma} \left(M - \overline{M}_\gamma\right)^{-1}\left(M - \overline{M}_0\right)\right) c$$
$$> 0,$$

$$\frac{\partial}{\partial \gamma_2} b_4(\gamma) = -c' \left(\left(M - \overline{M}_0\right)\left(M - \overline{M}_\gamma\right)^{-1} D_\gamma \overline{F}_{2\gamma} \left(M - \overline{M}_\gamma\right)^{-1}\left(M - \overline{M}_0\right)\right) c$$
$$> 0.$$

As all of the four functions are minimized at the true thresholds, and it can be shown that $b_i(\gamma) \neq b_i(\gamma_0)$ iff $\gamma \neq \gamma_0$ for $i = 1, 2, 3, 4$, the threshold estimators are consistent.

Appendix C: Distribution of $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ for vector $\mathbf{x}_t$

To derive the limiting distribution of $\widehat{\gamma}$ for a shrinking break, we let $\delta = cT^{-\alpha}$, $\alpha < \frac{1}{2}$, and we have

$$\widehat{\beta}_1 = \left(X'X - X_\gamma' X_\gamma\right)^{-1}(X - X_\gamma)' Y'$$
$$= \beta_1 + \left(X'X - X_\gamma' X_\gamma\right)^{-1}\left(X_0' X_0 - X_\gamma X_0\right)\delta + \left(X'X - X_\gamma' X_\gamma\right)^{-1}(X - X_\gamma)'\varepsilon$$

$$\begin{aligned}
\widehat{\beta}_1 - \widehat{\beta}_1^0 &= \left(X'X - X_\gamma' X_\gamma\right)^{-1}\left(X_0' X_0 - X_\gamma' X_0\right)\delta \\
&\quad + \left(X'X - X_\gamma' X_\gamma\right)^{-1}(X - X_\gamma)'\varepsilon - \left(X'X - X_0' X_0\right)^{-1}(X - X_0)'\varepsilon \\
&= \left(X'X - X_\gamma' X_\gamma\right)^{-1}\left(X_0' X_0 - X_\gamma' X_0\right)\delta + \left(X'X - X_\gamma' X_\gamma\right)^{-1}(X_0 - X_\gamma)'\varepsilon \\
&\quad + o_p\left(\frac{1}{T^{1-\alpha}}\right) \\
&= O_p\left(\frac{1}{T^{1-\alpha}}\right) + O_p\left(\frac{1}{T^{1-\alpha}}\right) + o_p\left(\frac{1}{T^{1-\alpha}}\right) \\
&= O_p\left(\frac{1}{T^{1-\alpha}}\right),
\end{aligned}$$

$$\widehat{\delta} = \left(X_\gamma' X_\gamma\right)^{-1} X_\gamma' Y - \widehat{\beta}_1,$$

$$\widehat{\delta} - \delta = O_p\left(\frac{1}{T^{1-\alpha}}\right),$$

$$\begin{aligned}
S_T(\gamma_1, \gamma_2) &= \left(Y - X\widehat{\beta}_1 - X_\gamma \widehat{\boldsymbol{\delta}}\right)'\left(Y - X\widehat{\beta}_1 - X_\gamma \widehat{\boldsymbol{\delta}}\right) \\
&= Y'Y - 2Y'\left(X\widehat{\beta}_1 + X_\gamma \widehat{\boldsymbol{\delta}}\right) + \left(X\widehat{\beta}_1 + X_\gamma \widehat{\boldsymbol{\delta}}\right)'\left(X\widehat{\beta}_1 + X_\gamma \widehat{\boldsymbol{\delta}}\right).
\end{aligned}$$

In the neighborhood of the true thresholds, where $\gamma_1 = \gamma_1^0 + \frac{\upsilon_1}{T^{1-2\alpha}}$, $\gamma_2 = \gamma_2^0 + \frac{\upsilon_2}{T^{1-2\alpha}}$, we have:

$$\begin{aligned}
&S_T(\gamma_1, \gamma_2) - S_T(\gamma_1^0, \gamma_2^0) \\
&= 2Y'\left(X\widehat{\beta}_1^0 + X_0\widehat{\boldsymbol{\delta}}^0\right) - 2Y'\left(X\widehat{\beta}_1 + X_\gamma\widehat{\boldsymbol{\delta}}\right) + \left(X\widehat{\beta}_1 + X_\gamma\widehat{\boldsymbol{\delta}}\right)'\left(X\widehat{\beta}_1 + X_\gamma\widehat{\boldsymbol{\delta}}\right) \\
&- \left(X\widehat{\beta}_1^0 + X_0\widehat{\boldsymbol{\delta}}^0\right)'\left(X\widehat{\beta}_1^0 + X_0\widehat{\boldsymbol{\delta}}^0\right) \\
&= -2\left(X\beta_1 + X_0\boldsymbol{\delta} + \boldsymbol{\varepsilon}\right)'\left(X\left(\widehat{\beta}_1 - \widehat{\beta}_1^0\right) + X_\gamma\widehat{\boldsymbol{\delta}} - X_0\widehat{\boldsymbol{\delta}}^0\right) \\
&+ \left(X\widehat{\beta}_1 + X_\gamma\widehat{\boldsymbol{\delta}} + X\widehat{\beta}_1^0 + X_0\widehat{\boldsymbol{\delta}}^0\right)'\left(X\widehat{\beta}_1 + X_\gamma\widehat{\boldsymbol{\delta}} - X\widehat{\beta}_1^0 - X_0\widehat{\boldsymbol{\delta}}^0\right) \\
&= -2\left(X\beta_1 + X_0\boldsymbol{\delta} + \boldsymbol{\varepsilon}\right)'\left(X_\gamma\widehat{\boldsymbol{\delta}} - X_0\widehat{\boldsymbol{\delta}}^0\right) + \left(X\widehat{\beta}_1 + X_\gamma\widehat{\boldsymbol{\delta}} + X\widehat{\beta}_1^0 + X_0\widehat{\boldsymbol{\delta}}^0\right)'\left(X_\gamma\widehat{\boldsymbol{\delta}} - X_0\widehat{\boldsymbol{\delta}}^0\right)
\end{aligned}$$

$+o_p(1)$
$= -2\boldsymbol{\varepsilon}'(X_\gamma - X_0)\boldsymbol{\delta} + \left(X\widehat{\beta}_1 + X\widehat{\beta}_1^0 - 2X\beta_1 + X_\gamma\boldsymbol{\delta} - X_0\boldsymbol{\delta}\right)'(X_\gamma - X_0)\boldsymbol{\delta}$
$+o_p(1)$
$= -2\boldsymbol{\varepsilon}'(X_\gamma - X_0)\boldsymbol{\delta} + \boldsymbol{\delta}'(X_\gamma - X_0)'(X_\gamma - X_0)\boldsymbol{\delta} + \left(\widehat{\beta}_1 + \widehat{\beta}_1^0 - 2\beta_1\right)' X'(X_\gamma - X_0)\boldsymbol{\delta}$
$+o_p(1)$
$= -2\boldsymbol{\varepsilon}'(X_\gamma - X_0)\boldsymbol{\delta} + \boldsymbol{\delta}'(X_\gamma - X_0)'(X_\gamma - X_0)\boldsymbol{\delta} + o_p(1)$
$= -2T^{-\alpha}\sum_{t=1}^{T} c'x_t\varepsilon_t(\Psi_t(\gamma) - \Psi_t(\gamma^0)) + \sum_{t=1}^{T}(c'x_t)^2|\Psi_t(\gamma) - \Psi_t(\gamma^0)|$
$+o_p(1)$.
Now, using

$$\overline{M}(\gamma_1, \gamma_2) = \overline{M}_0 + \left(\gamma_1 - \gamma_1^0\right) D\overline{F}_1^0 + \left(\gamma_2 - \gamma_2^0\right) D\overline{F}_2^0 + o(1).$$

Case 1: $\upsilon_1 \leq 0,\ \upsilon_2 \leq 0$ In this case, we have
$X_\gamma' X_0 = X_0' X_0;$
$S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma_{1,}^0\gamma_2^0\right)$
$= 2\boldsymbol{\varepsilon}'(X_0 - X_\gamma)\delta + \delta'\left(X_0'X_0 - X_\gamma'X_\gamma\right)\delta + o_p(1)$
$= 2\boldsymbol{\varepsilon}'(X_0 - X_\gamma)cT^{-\alpha} + c'\left(\overline{M}_0 - \overline{M}_\gamma\right)cT^{1-2\alpha} + o_p(1)$
$= -2T^{-\alpha}\sum_{\left(\gamma_1<z_{1t}<\gamma_1^0 \text{ and } \gamma_2^0<z_{2t}\right) \text{ or } \left(\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1^0<z_{1t}\right) \text{ or } \left(\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1<z_{1t}<\gamma_1^0\right)} \varepsilon_t x_t' c$
$-T^{1-2\alpha}c'\left(\left(\gamma_1 - \gamma_1^0\right)D\overline{F}_1^0 + \left(\gamma_2 - \gamma_2^0\right)D\overline{F}_2^0\right)c + o_p(1)$
$= -2T^{-\alpha}\sum_{\left(\gamma_1<z_{1t}<\gamma_1^0 \text{ and } \gamma_2^0<z_{2t}\right) \text{ or } \left(\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1^0<z_{1t}\right)} \varepsilon_t x_t' c - c'\left(\upsilon_1 D\overline{F}_1^0 + \upsilon_2 D\overline{F}_2^0\right)c$
$+o_p(1)$
$= -c'Dc\upsilon_1\overline{F}_1^0 - 2T^{-\alpha}\sum_{\gamma_1<z_{1t}<\gamma_1^0 \text{ and } \gamma_2^0<z_{2t}} c'x_t\varepsilon_t - c'Dc\upsilon_2\overline{F}_2^0$
$-2T^{-\alpha}\sum_{\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1^0<z_{1t}} c'x_t\varepsilon_t + o_p(1)$.
Note that
$T^{-\alpha}\sum_{\gamma_1<z_{1t}<\gamma_1^0 \text{ and } \gamma_2^0<z_{2t}} \varepsilon_t x_t'$ converge in distribution to $B_1(\upsilon)$, which is a vector Brownian motion with covariance matrix $E\left(B_1(1)B_1(1)'\right) = -V\overline{F}_1^0$,
$T^{-\alpha}\sum_{\gamma_2<z_{2t}<\gamma_2^0 \text{ and } \gamma_1^0<z_{1t}} \varepsilon_t x_t'$ converge in distribution to $B_2(\upsilon_2)$, which is a vector Brownian motion with covariance matrix $E\left(B_2(1)B_2(1)'\right) = -V\overline{F}_2^0$, where B_1 and B_2 are independent.

Thus, in the neighborhood of the true threshold values, the above is equal in distribution to
$\stackrel{d}{=} -c'Dc\upsilon_1\overline{F}_1^0 - 2c'B_1(\upsilon) - c'Dc\upsilon_2\overline{F}_2^0 - 2c'B_2(\upsilon_2)$.

We apply the same arguments to the following cases:

Case 2: $\upsilon_1 > 0,\ \upsilon_2 \leq 0$ We have

$$
\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma^0_{1,}\gamma^0_2\right)\\
&= -2\boldsymbol{\varepsilon}'\left(X_\gamma - X_0\right)cT^{-\alpha} + T^{1-2\alpha}c'\left(\overline{M}_\gamma + \overline{M}_0 - 2\overline{M}\left(\gamma_1, \gamma^0_2\right)\right)'c + o_p\left(1\right)\\
&= -2\boldsymbol{\varepsilon}'\left(X_\gamma - X_0\right)cT^{-\alpha}\\
&+T^{1-2\alpha}c'\left(\overline{M}_0 + \left(\gamma_1 - \gamma^0_1\right)D\overline{F}^0_1 + \left(\gamma_2 - \gamma^0_2\right)D\overline{F}^0_2 + \overline{M}_0 - 2\left(\overline{M}_0 + \left(\gamma_1 - \gamma^0_1\right)D\overline{F}^0_1\right)\right)'c\\
&+o_p\left(1\right)\\
&= -2\boldsymbol{\varepsilon}'\left(X_\gamma - X_0\right)cT^{-\alpha} + T^{1-2\alpha}c'\left(\left(\gamma_2 - \gamma^0_2\right)D\overline{F}^0_2 - \left(\gamma_1 - \gamma^0_1\right)D\overline{F}^0_1\right)'c\\
&+o_p\left(1\right)\\
&= -2\boldsymbol{\varepsilon}'\left(X_\gamma - X_0\right)cT^{-\alpha} + c'\left(\upsilon_2 D\overline{F}^0_2 - \upsilon_1 D\overline{F}^0_1\right)'c + o_p\left(1\right)\\
&\overset{d}{=} -c'Dc\upsilon_1\overline{F}^0_1 + 2c'B_1\left(\upsilon_1\right) + c'Dc\upsilon_2\overline{F}^0_2 - 2c'B_2\left(\upsilon_2\right).
\end{aligned}
$$

Similarly, we have

Case 3: $\upsilon_1 \leq 0,\ \upsilon_2 > 0$

$$
\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma^0_{1,}\gamma^0_2\right)\\
&\overset{d}{=} -c'Dc\upsilon_1\overline{F}^0_1 - 2c'B_1\left(\upsilon_1\right) + c'Dc\upsilon_2\overline{F}^0_2 + 2c'B_2\left(\upsilon_2\right).
\end{aligned}
$$

Case 4: $\upsilon_1 > 0,\ \upsilon_2 > 0$

$$
\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma^0_{1,}\gamma^0_2\right)\\
&\overset{d}{=} c'Dc\upsilon_1\overline{F}^0_1 + 2c'B_1\left(\upsilon_1\right) + c'Dc\upsilon_2\overline{F}^0_2 + 2c'B_2\left(\upsilon_2\right).
\end{aligned}
$$

Making the change of variables

$$
\upsilon_1 = -\frac{c'Vc}{\left(c'Dc\right)^2\overline{F}^0_1}r_1,
$$

$$
\upsilon_2 = -\frac{c'Vc}{\left(c'Dc\right)^2\overline{F}^0_2}r_2.
$$

In general,

$$
\begin{aligned}
&S_T\left(\gamma_{1,}\gamma_2\right) - S_T\left(\gamma^0_{1,}\gamma^0_2\right)\\
&\overset{d}{=} c'Dc\left|\upsilon_1\right|\overline{F}^0_1 + 2c'B_1\left(\upsilon_1\right) + c'Dc\left|\upsilon_2\right|\overline{F}^0_2 + 2c'B_2\left(\upsilon_2\right)\\
&\overset{d}{=} \frac{c'Vc}{c'Dc}\left|r_1\right| + 2\frac{\sqrt{c'Vc}}{c'Dc\sqrt{-\overline{F}^0_1}}c'B_1\left(-r_1\right) + \frac{c'Vc}{c'Dc}\left|r_2\right| + 2\frac{\sqrt{c'Vc}}{c'Dc\sqrt{-\overline{F}^0_2}}c'B_2\left(-r_2\right)\\
&\overset{d}{=} \frac{c'Vc}{c'Dc}\left|r_1\right| + 2\frac{c'Vc}{c'Dc}W_1\left(-r_1\right) + \frac{c'Vc}{c'Dc}\left|r_2\right| + 2\frac{c'Vc}{c'Dc}W_2\left(-r_2\right)\\
&\overset{d}{=} 2\frac{c'Vc}{c'Dc}\left(\frac{\left|r_1\right|}{2} + W_1\left(r_1\right) + \frac{\left|r_1\right|}{2} + W_2\left(r_2\right)\right).
\end{aligned}
$$

We have

$$
\begin{aligned}
&-T^{1-2\alpha}\frac{(c'Dc)^2}{c'Vc}\left(\overline{F}_1^0\left(\widehat{\gamma}_1-\gamma_1^0\right),\overline{F}_2^0\left(\widehat{\gamma}_2-\gamma_2^0\right)\right)\\
&\xrightarrow{d} \underset{(r_1,r_2)\in R^2}{\arg\min}\left(2\frac{c'Vc}{c'Dc}\left(\frac{|r_1|}{2}+W_1(r_1)+\frac{|r_2|}{2}+W_2(r_2)\right)\right)\\
&\overset{d}{=} \underset{(r_1,r_2)\in R^2}{\arg\max}\left(-\frac{1}{2}|r_1|+W_1(r_1)-\frac{1}{2}|r_2|+W_2(r_2)\right).
\end{aligned}
$$

To find the close-form joint distribution, note that the selection of r_1 does not depend on the choice of r_2 and vice versa, so we have

$$
\begin{aligned}
&\Pr\left(\underset{r_1\in R}{\arg\max}\textstyle\sum_{j=1}^2\left(-\frac{1}{2}|r_j|+W_j(r_j)\right)\le a_1,\ \underset{r_2\in R}{\arg\max}\textstyle\sum_{j=1}^2\left(-\frac{1}{2}|r_j|+W_j(r_j)\right)\le a_2\right)\\
&=\Pr\left(\underset{r_1\in R}{\arg\max}\left(-\tfrac{1}{2}|r_1|+W_1(r_1)\right)\le a_1,\ \underset{r_2\in R}{\arg\max}\left(-\tfrac{1}{2}|r_2|+W_2(r_2)\right)\le a_2\right)\\
&=\Pi_{j=1}^2\Pr\left(\underset{r_j\in R}{\arg\max}\left(-\tfrac{1}{2}|r_j|+W_j(r_j)\right)\le a_j\right)\\
&\overset{def}{=}\Pi_{j=1}^2F_{\widehat{r}_j}(a_j).
\end{aligned}
$$

According to Bhattacharya and Brockwell (1976), for $a_1 > 0$ and $a_2 > 0$, the above joint distribution equals

$$
\begin{aligned}
&F_{(\widehat{r}_1,\widehat{r}_2)}(a_1,a_2)\\
&=\Pi_{j=1}^2F_{\widehat{r}_j}(a_j)\\
&=\Pi_{j=1}^2\left(1+\sqrt{\frac{a_j}{2\pi}}\exp\left(-\frac{a_j}{8}\right)+\frac{3}{2}\exp(a_j)\Phi\left(-\frac{3\sqrt{a_j}}{2}\right)-\frac{a_j+5}{2}\Phi\left(-\frac{\sqrt{a_j}}{2}\right)\right),
\end{aligned}
$$

where $\Phi(\cdot)$ is the cdf of a standard normal distribution.

Thus, using the fact that $\phi(x)=\dfrac{1}{\sqrt{2\pi}}\exp\left(-\dfrac{x^2}{2}\right)$, the joint density function is

$$
\begin{aligned}
&f_{(\widehat{r}_1,\widehat{r}_2)}(a_1,a_2)\\
&=\Pi_{j=1}^2f_{\widehat{r}_j}(a_j)\\
&=\Pi_{j=1}^2\left(\frac{3}{2}\exp(a_j)\Phi\left(-\frac{3\sqrt{a_j}}{2}\right)-\frac{1}{2}\Phi\left(-\frac{\sqrt{a_j}}{2}\right)\right).
\end{aligned}
$$

For cases where some of the $a_j < 0$, we can replace those items in the above expression by $F_{\widehat{r}_j}(a_j)=1-F_{\widehat{r}_j}(-a_j)$ and $f_{\widehat{r}_j}(a_j)=f_{\widehat{r}_j}(-a_j)$.

Appendix D: data description

The sample data consists of quarterly data from 1982 Q1 through 2001 Q4 of the following economies: Argentina, Brazil, Chile, Colombia, Mexico, Uruguay and Venezuela in Latin America, and Mainland China, Hong Kong, Indonesia, South Korea, Malaysia, the Philippines, Singapore, Taiwan and Thailand in Asia.

The primary data sources are International Financial Statistics (IFS), and the websites of both the Asian Development Bank (ADB) and the Bank of International Settlements (BIS). The following table gives the sources and definitions of the variables:

Predictors	Sources and Definitions
1. Ratio of fiscal deficits to GDP	Fiscal deficit is taken from IFS line 80 and GDP is taken from IFS line 99B.
2. Ratio of short-term external liabilities to foreign exchange reserves	The short-term external debt data is obtained from the Asian Development Bank (ADB) website and the Bank of International Settlements (BIS) website. The cumulative portfolio liabilities data is constructed as the cumulative sum of the portfolio liabilities flow data obtained from IFS line 78BGD. The import data is from IFS line 98C. The foreign exchange reserve data is from IFS line 1L.
3. Lending rate differential	The lending rate differential is constructed as the difference between the 3-month domestic lending rate and that of the US. The lending interest rate is taken from IFS line 60P.
4. Real exchange rate appreciation index	The exchange rate data is obtained from IFS line ..AE..ZF. The exchange rate of China before 1994 Q1 is the swap rate obtained from Global Financial Data. The nominal exchange rate is deflated by the Wholesale Price Index (WPI), which is taken from IFS line 63..ZF, and then the real exchange rate is normalized to 1986 Q1=1.
5. Ratio of domestic credit to GDP	The domestic credit data is taken from IFS line 32.ZF and the GDP data is from IFS line 99B.

Figure 1a: $S_T\left(\gamma_{1,}\gamma_2\right)/T$

Figure 1b: $g\left(\gamma_{1,}\gamma_2\right)$

Figure 2a: Distribution of $T^{1/2}\left(\widehat{\beta}_1\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_1\right)$

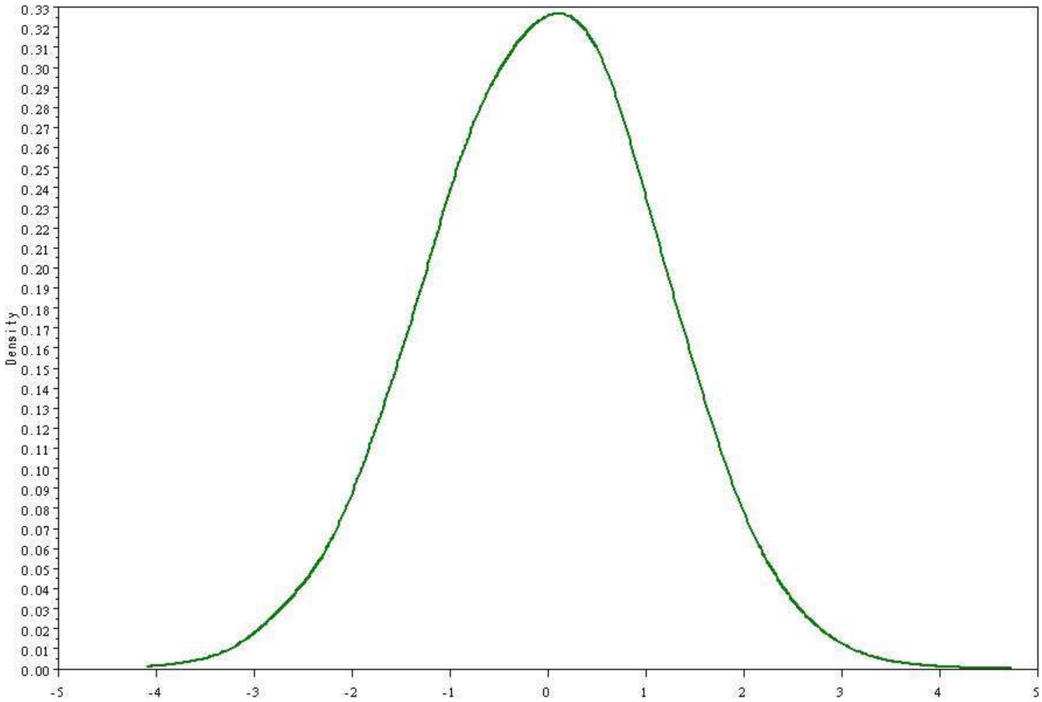


Figure 2b: Distribution of $T^{1/2}\left(\widehat{\beta}_2\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_2\right)$

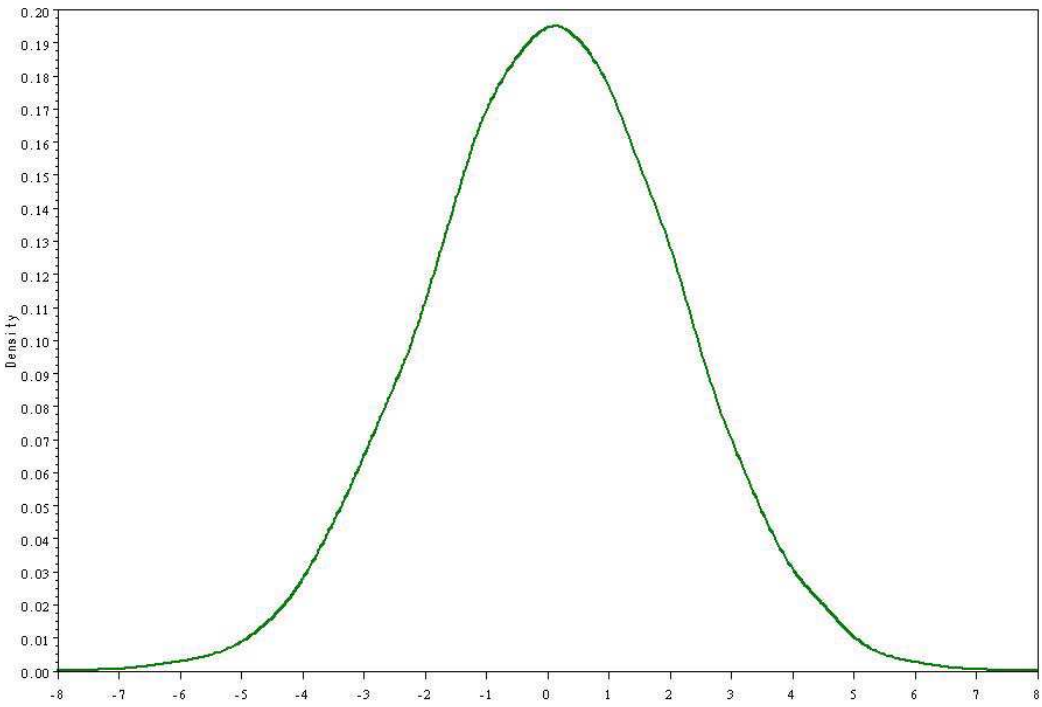


Figure 2c: Joint Distribution of $T^{1/2}\left(\widehat{\beta}_1\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_1,\widehat{\beta}_2\left(\widehat{\gamma}_{1,}\widehat{\gamma}_2\right)-\beta_2\right)$

Figure 3a: Distribution of $T^{3/4} f_1\left(\gamma_1^0\right) \overline{F}_2\left(\gamma_2^0\right)\left(\widehat{\gamma}_1-\gamma_1^0\right)$

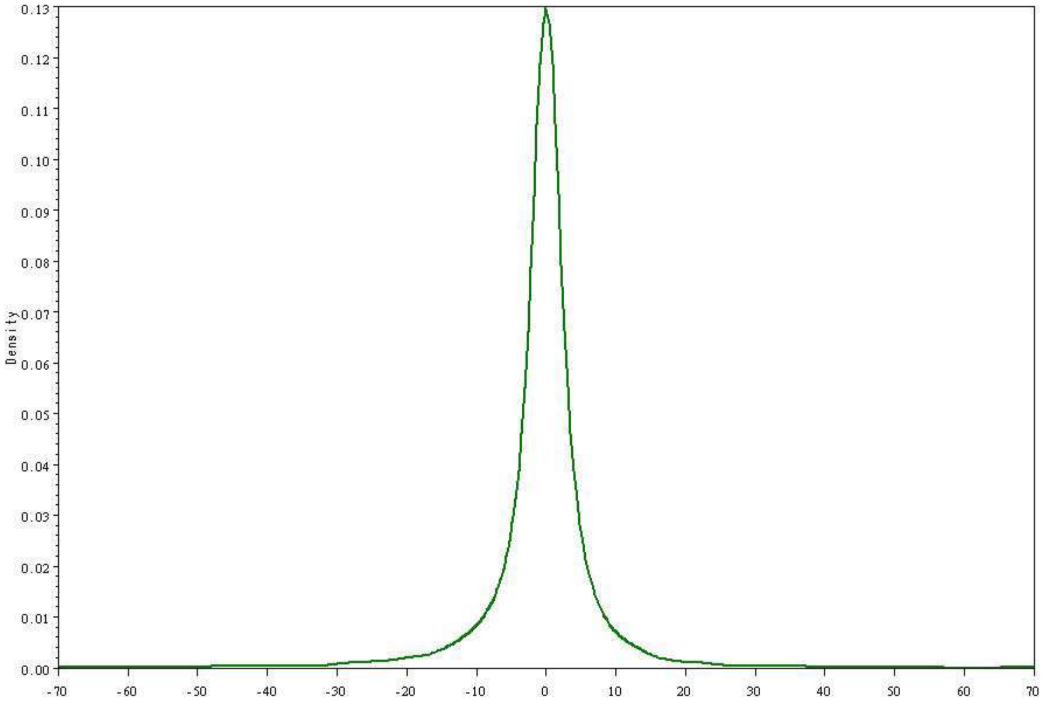


Figure 3b: Distribution of $T^{3/4} f_2\left(\gamma_2^0\right) \overline{F}_1\left(\gamma_1^0\right)\left(\widehat{\gamma}_2-\gamma_2^0\right)$

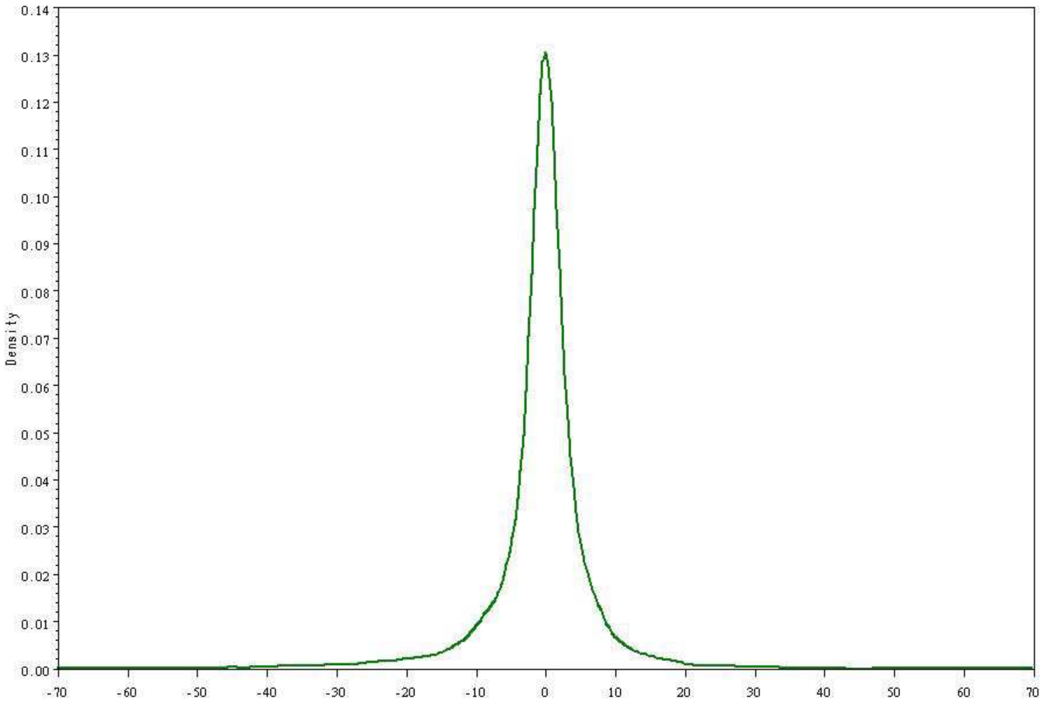


Figure 4a: Joint Distribution of
$T^{1-2\alpha}\left(f_1\left(\gamma_1^0\right)\overline{F}_2\left(\gamma_2^0\right)\left(\widehat{\gamma}_1-\gamma_1^0\right), f_2\left(\gamma_2^0\right)\overline{F}_1\left(\gamma_1^0\right)\left(\widehat{\gamma}_2-\gamma_2^0\right)\right)$

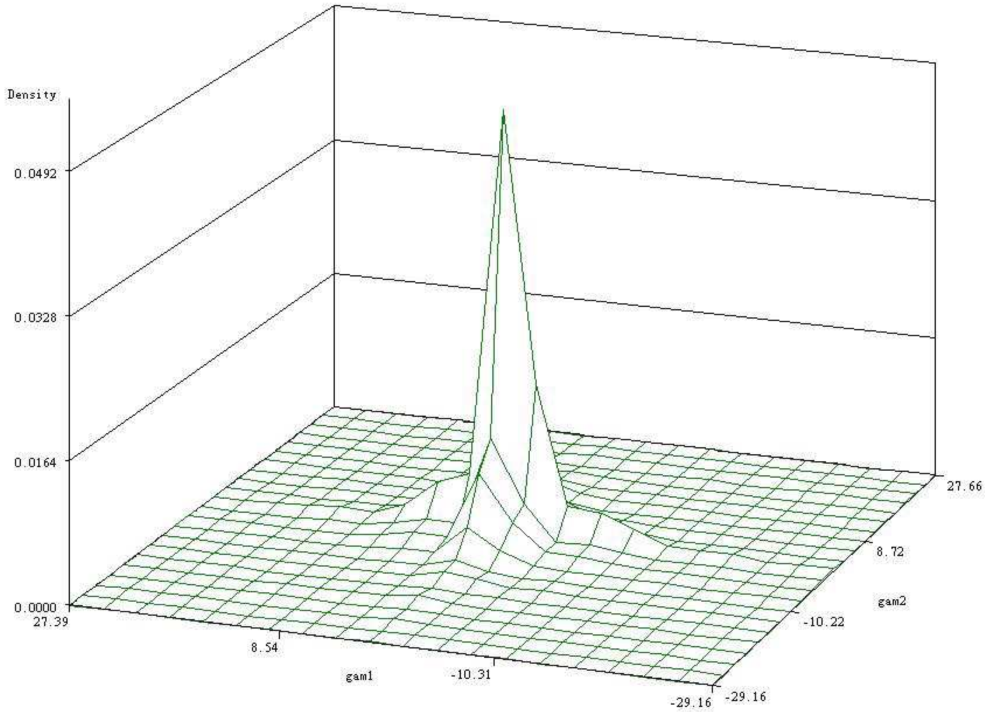


Figure 4b: Joint Density of $f_{(\widehat{r}_1,\widehat{r}_2)}\left(a_1, a_2\right)$

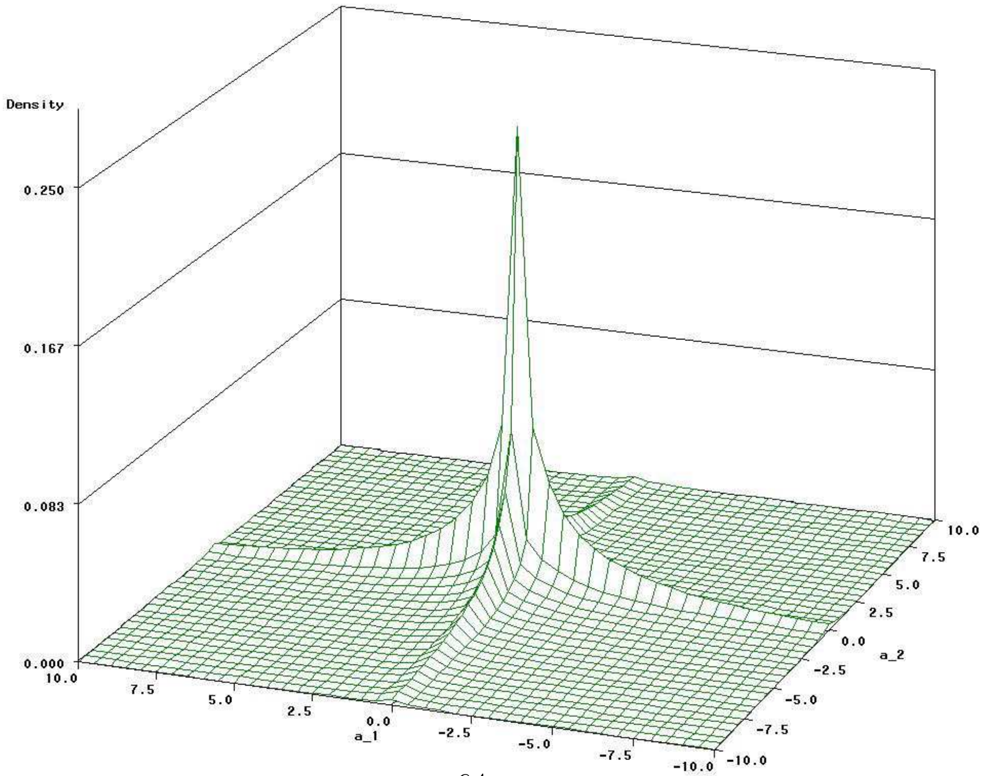


Figure 5a: Distribution of the LR_T Statistic

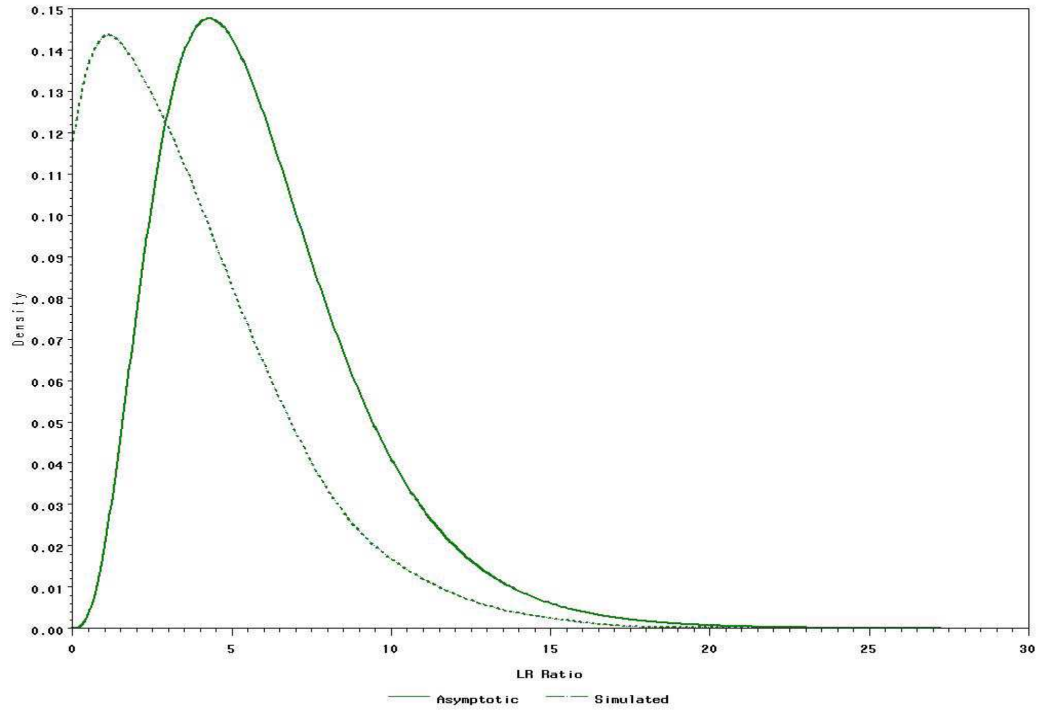


Figure 5b: 95% Confidence Region for γ_1 and γ_2

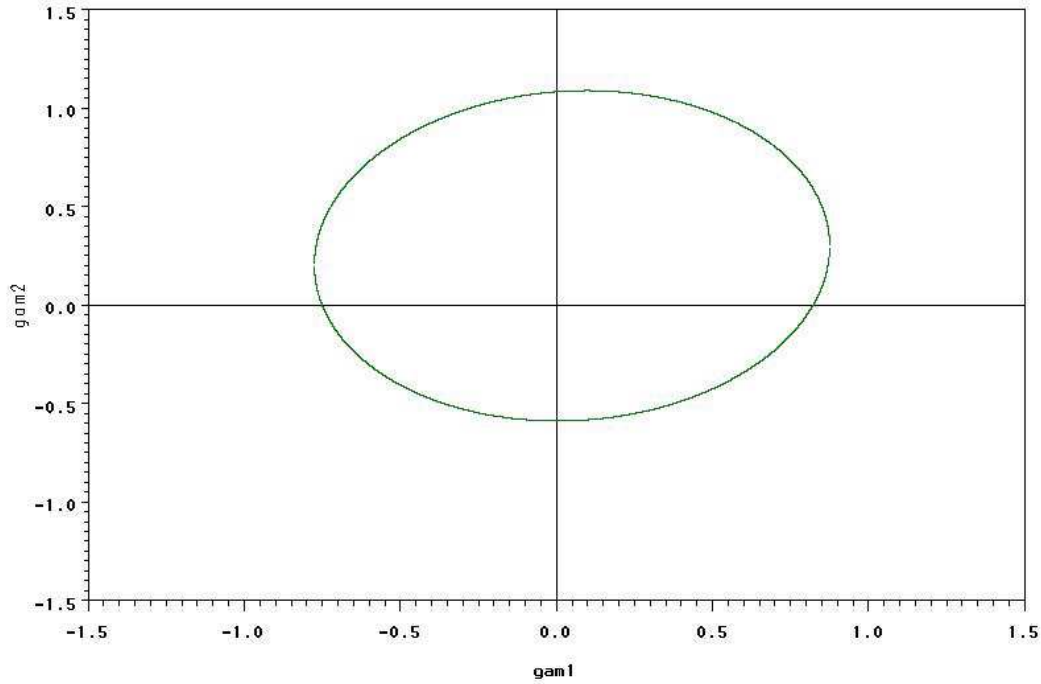
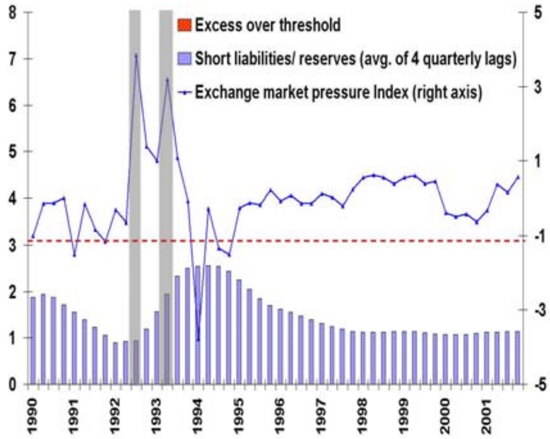


Figure 6: Threshold Effects and Exchange Market Pressure Index of selected Asian Countries

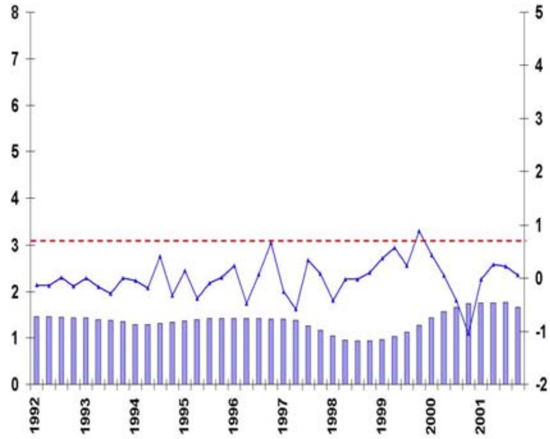
(a): **China**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$

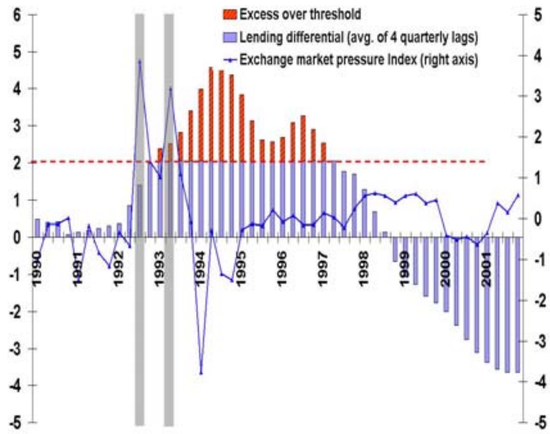


(b): **Hong Kong**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$



Lending rate differential



Lending rate differential

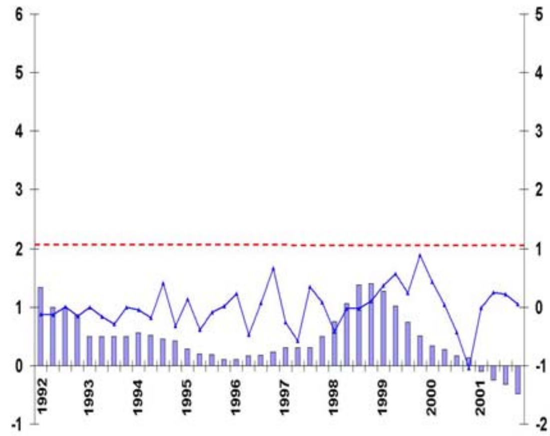
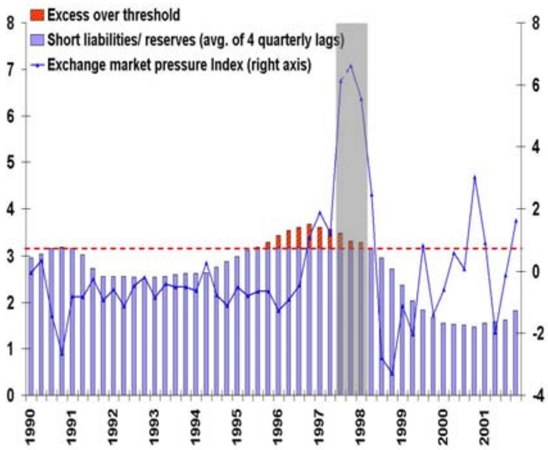


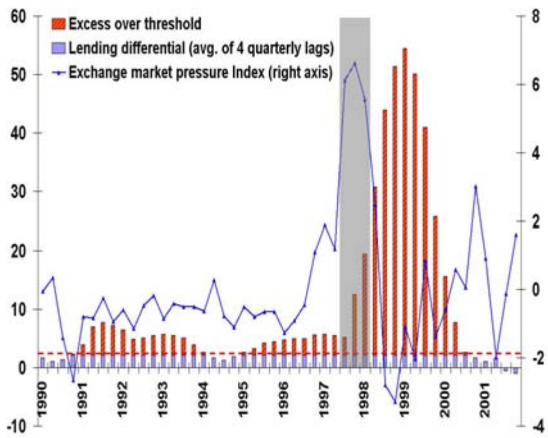
Figure 6: Threshold Effects and Exchange Market Pressure Index of selected Asian Countries (Continued)

(c): **Indonesia**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$

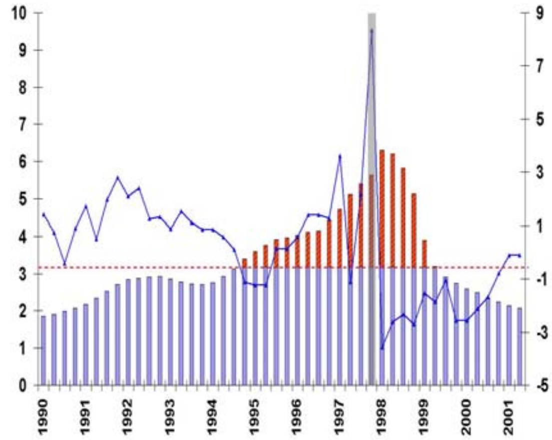


Lending rate differential



(d): **S. Korea**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$



Lending rate differential

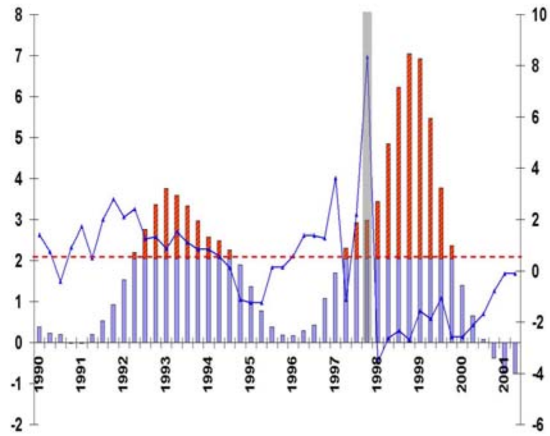
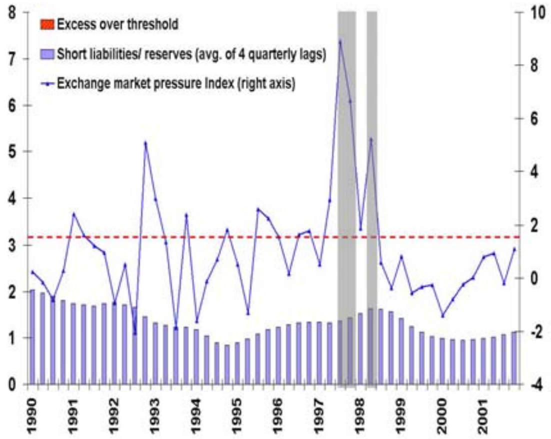


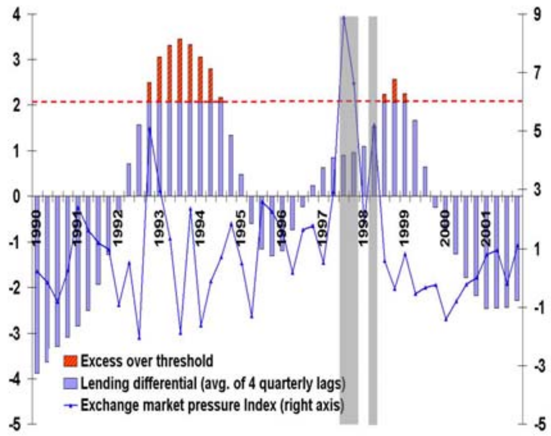
Figure 6: Threshold Effects and Exchange Market Pressure Index of selected Asian Countries (Continued)

(e): **Malaysia**

Short-term external liabilities / Reserves

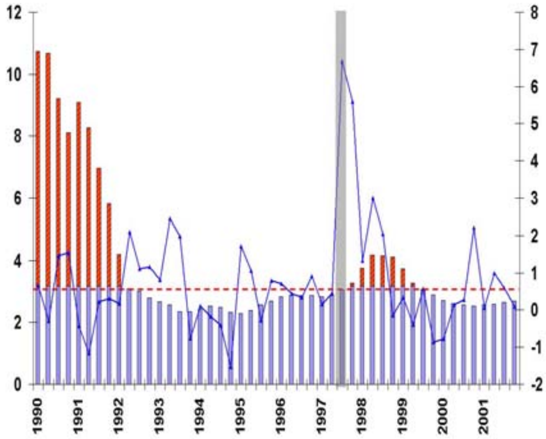


Lending rate differential



(f): **Philippines**

Short-term external liabilities / Reserves



Lending rate differential

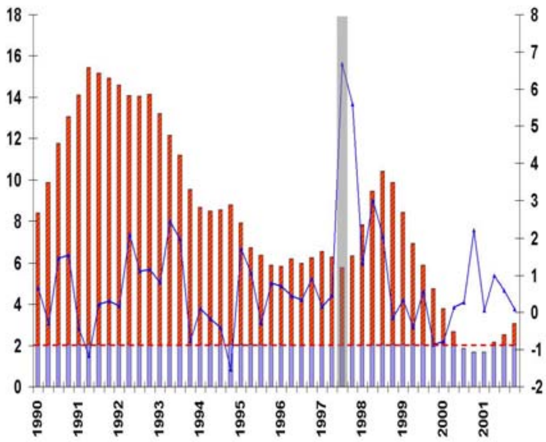
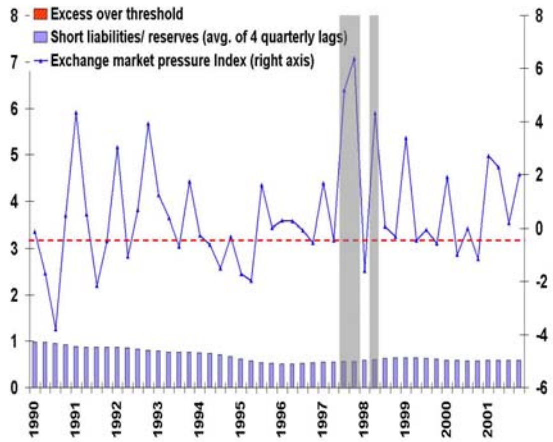


Figure 6: Threshold Effects and Exchange Market Pressure Index of selected Asian Countries (Continued)

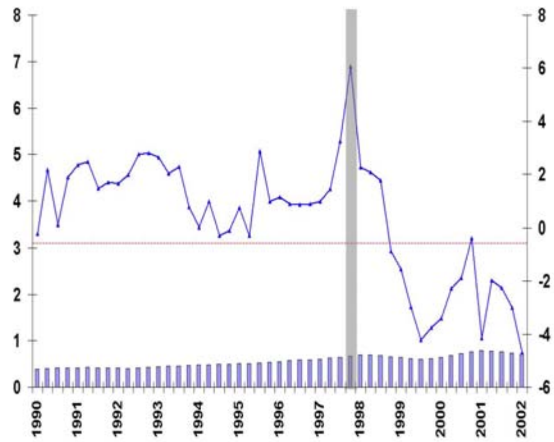
(g): **Singapore**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$

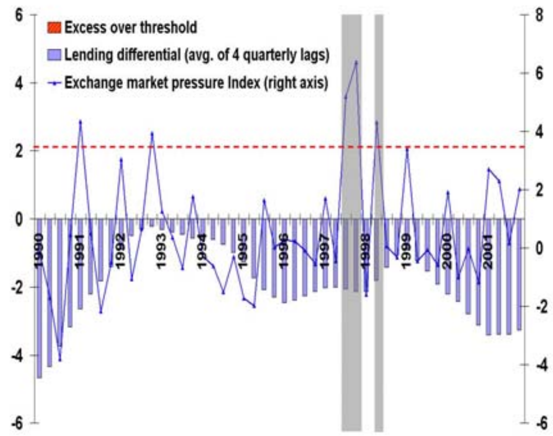


(h): **Taiwan**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$



Lending rate differential



Lending rate differential

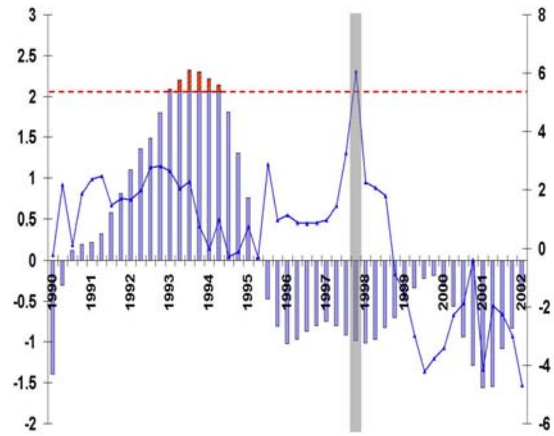


Figure 6: Threshold Effects and Exchange Market Pressure Index of selected Asian Countries (Continued)

(i): **Thailand**

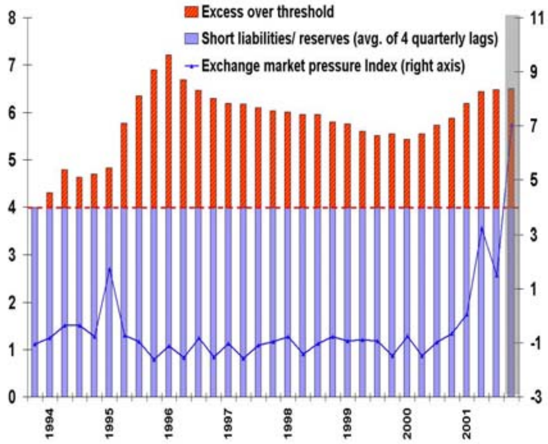
$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$

Lending rate differential

Figure 7: Threshold Effects and Exchange Market Pressure Index of selected Latin American Countries

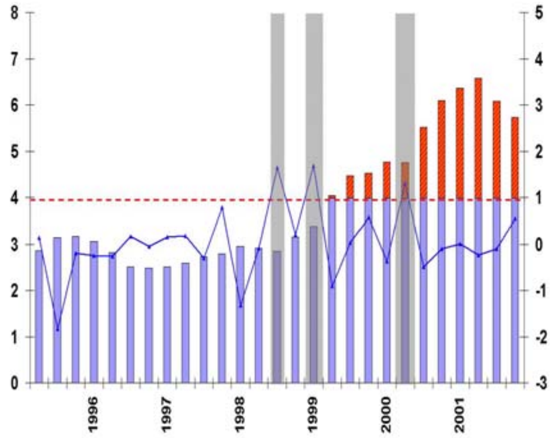
(a): **Argentina**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$

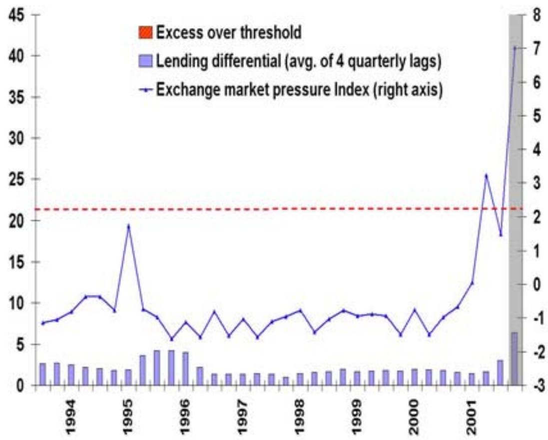


(b): **Brazil**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$



Lending rate differential



Lending rate differential

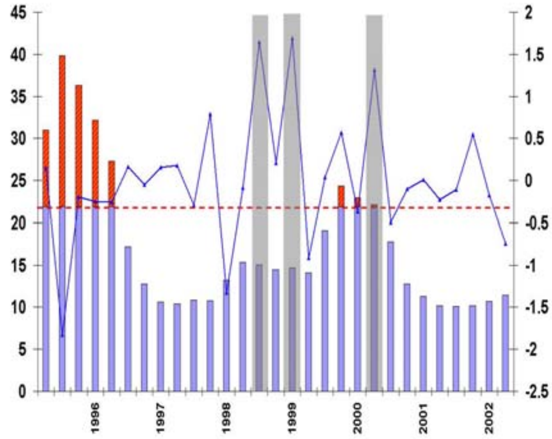
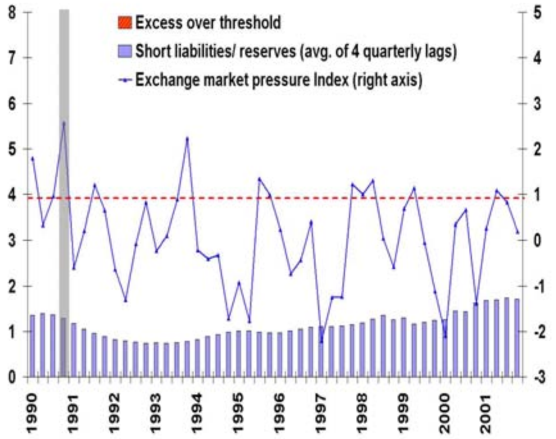


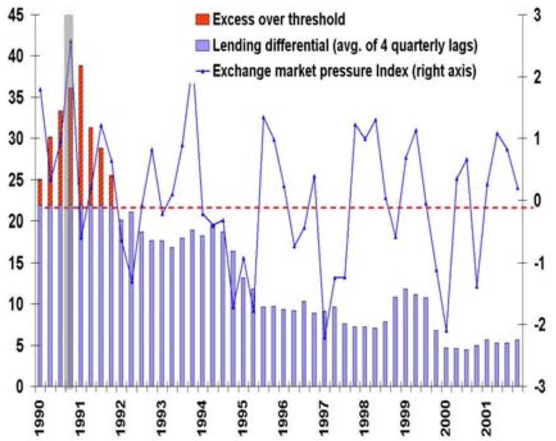
Figure 7: Threshold Effects and Exchange Market Pressure Index of selected Latin American Countries (Continued)

(c): **Chile**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$

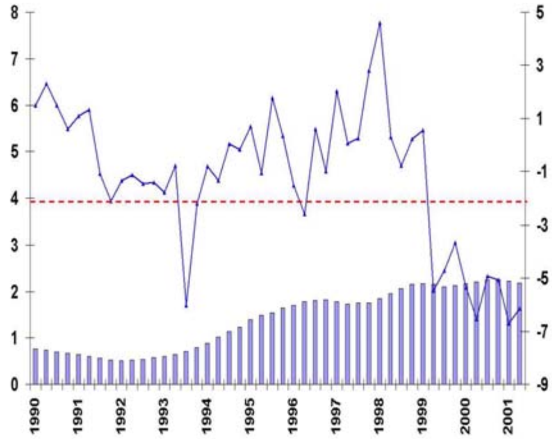


Lending rate differential



(d): **Colombia**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$



Lending rate differential

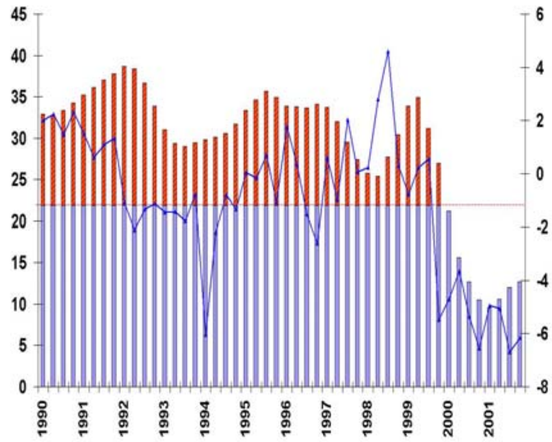
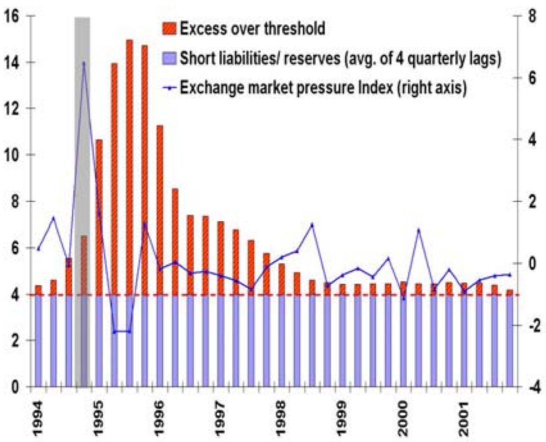


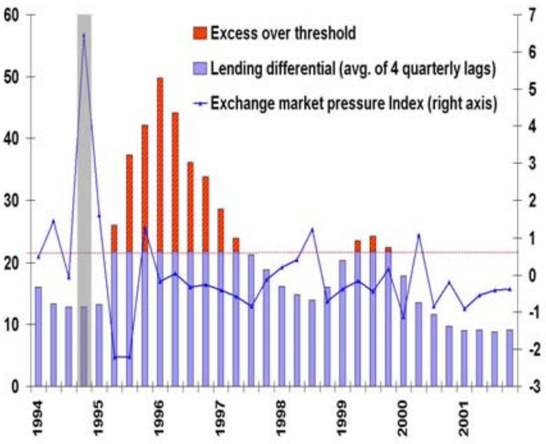
Figure 7: Threshold Effects and Exchange Market Pressure Index of selected Latin American Countries (Continued)

(e): **Mexico**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$

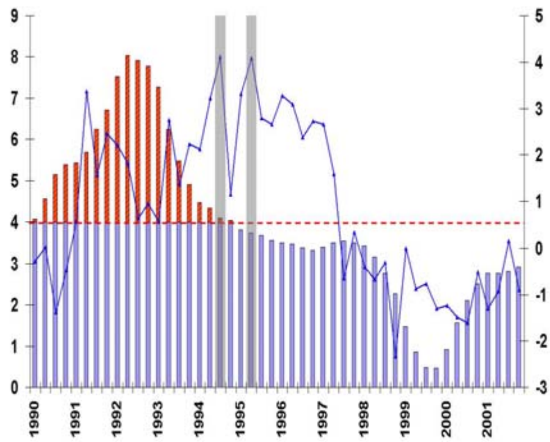


Lending rate differential



(f): **Uruguay**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$



Lending rate differential

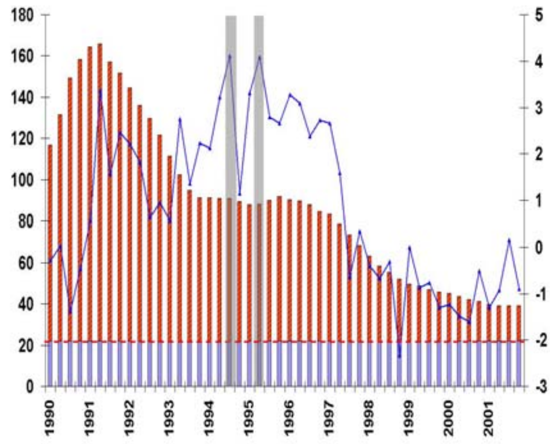
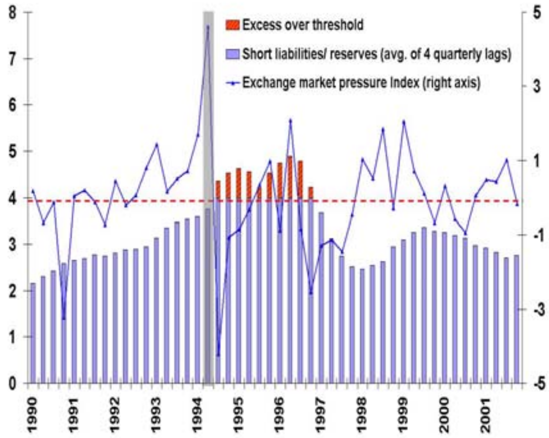


Figure 7: Threshold Effects and Exchange Market Pressure Index of selected Latin American Countries (Continued)

(g): **Venezuela**

$\frac{\text{Short-term external liabilities}}{\text{Reserves}}$



Lending rate differential