Predicting Currency Crises with a Multiple Threshold-Variable Model

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

December 8, 2006

Abstract

Conventional threshold models contain only one threshold variable. Such models are of limited economic application. This paper provides the theoretical foundation for threshold models with multiple threshold variables. The new model is much more complicated than a model with a single threshold variable as several novel problems arise with an additional threshold variable. First, models with multiple threshold variables cannot be converted into change-point models in the manner of Tsay (1998) and Hansen (1999). Second, asymptotic joint distribution of the threshold estimators may be ill-behaved should the threshold variables be dependent. Third, having more threshold variables introduces the curse of dimensionality to the estimation.

In this paper, we establish the consistency of the threshold estimators. In particular, under certain conditions, we obtain the asymptotic joint distribution of threshold estimators and suggest a quick algorithm to estimate the threshold values. Tests for the number of threshold variables and their critical values are also developed. Asymptotic critical values of the LR type test for multiple threshold variables are computed. Simulations that

*Corresponding author. Department of Economics, Esther Lee Building, Room 930, the Chinese University of Hong Kong, Shatin, New Territories, Hong Kong. Phone: +(852)2609-8193; 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 also thanks Gilbert Lui and Haiqiang Chen for able research assistance. All remaining errors are our own.
support our asymptotic result are given. The model is applied to the study of currency crises. The major contribution of the empirical application is that it is the first study to provide clear estimates of the joint critical threshold values of multiple crisis indicators, which can be used by governments as guidelines in the regulation of short-term external borrowings and interest rate differentials.

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

**JEL Classification Number:** C33; C12; C13

1 Introduction

Threshold models have been developing very rapidly during the past decade. A common type of threshold model is the threshold autoregressive (TAR) model of Tong (1983), which uses the lag of the dependent variable as a threshold variable. Some extensions of the TAR model include the smooth transition threshold model (STAR) of Chan and Tong (1986), the functional-coefficient autoregressive (FAR) model of Chen and Tsay (1993) and the nested threshold autoregressive (NeTAR) models of Astatkie, Watts and Watt (1997). Hansen (1999) develops a threshold model for non-dynamic panels with individual fixed-effects and applies it to study whether financial constraints affect investment decisions. The threshold model has also been extended to incorporate multiple thresholds (Tsay, 1998; Hansen, 1999, 2000; Gonzalo and Pitarakis, 2002). There are a wide variety of applications of the threshold models. For example, Henry, Olekaln and Summers (2001) show evidence of threshold nonlinearity in the Australian real exchange rate. Dueker, Sola and Spagnolo (2003) propose a contemporaneous TAR model and apply it to the pricing of bonds. Other recent studies of threshold models include those of Ip, Wong, Li and Xie (1999), Caner and Hansen (2001, 2004) and Cheng (2004). Most of these studies focus on models with a single threshold variable only. Such models, however, have limited applications when two or more threshold variables are called for. For instance, it has long been observed that some economic variables, such as the foreign debt level and interest rate, cross a certain threshold value before currency crisis occurs (Lau and Yan, 2004). Studies in the literature of currency crises, such as those of Eichengreen, Rose and Wyplosz (1995), Sachs et al. (1996), Frankel and Rose (1996), Kaminsky, Lizondo and Reinhart (1997), Kaminsky (1998) and Edison (2000) also suggest that the occurrence of currency crises depends critically on the values of several threshold variables. However, none of these papers
has estimated and tested those important threshold values due to the lack of proper modelling techniques in the existing literature. To the authors’ 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 are available thus far.

In this paper, we will discuss 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 will drastically increase the complexity of the problem, because models with multiple threshold variables cannot be converted into change-point models in the manner of Tsay (1998) and Hansen (1999). The contributions of this paper are twofold. First, it provides modelling tools and a distributional theory for threshold models with multiple threshold variables. We will examine the asymptotic theory for estimators in this model. Secondly, it develops tests for the number of threshold variables.

The tests and estimation methods are applied to examine whether there are threshold effects in the currency crisis indicators as implied by the three generations of currency crisis models. The “first generation model” (Krugman, 1979; Flood and Garber, 1984) suggests that exogenous government budget deficits lay at the root of balance of payment crises. The empirical implication is that the pressure on the foreign exchange market becomes significantly higher once the fiscal deficit exceeds a certain threshold. The “second generation” model (Obstfeld, 1986) formulates the possibility of 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. The threat of an attack generates expectation-driven increases in interest rates and there is a strong incentive for the central bank to abandon the peg because devaluation allows the government to roll over the short-term public debt at a lower interest rate. The empirical implication of the second generation model is that, prior to an attack, we should observe a drastic increase in the domestic interest rate, but there is no reason to anticipate excessively expansionary monetary or fiscal policies. As a result, the relevant threshold variable is the differential between the domestic interest rate and the foreign interest rate. Nevertheless, Krugman (1999) observed that neither the first nor the second generation stories were able to explain the 1997 Asian crisis. For this reason, the third generation model was developed. The

\(^1\)Three related studies in this regard are those of Astatkie, Watts and Watt (1997), Xia and Li (1999) and Xia, Li and Tong (2004)
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 obligations in foreign currency exceed the amount of foreign currency to which it can have access at short notice. When governments implicitly guarantee the debts of financial systems, the problem of moral hazard arises, thereby encouraging overborrowing in short-term foreign currency. When authorities do not have adequate foreign reserves, the financial system is internationally illiquid and is highly vulnerable to speculative attacks. 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). Our paper takes the threshold variables implied by the three generations of currency crisis models and tests for the existence of threshold effects. If there is evidence of a threshold effect, we estimate the threshold values using panel data from 16 countries. Our finding of threshold effects in the indicators inspired by the currency crisis models reinforces the currency crisis literature.

The paper is organized as follows. Section 2 presents the model and the major assumptions. We examine the consistency and the asymptotic distribution of the threshold estimators in this section. In Section 3, a fast estimation algorithm is suggested. The model is extended to allow for panel data in Section 4. In Section 5, we examine the inference problem. An LR type test and its asymptotic distribution are studied. Section 6 provides experimental evidence to support our theory. Section 7 applies the estimation methods and tests to the study of currency crises. The last section concludes the paper and discusses the future direction of research. All proofs are relegated to the Appendix.

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 ‘\(\rightarrow p\)’ represents convergence in probability, ‘\(\rightarrow d\)’ 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 \(T \to \infty\) unless otherwise stated.

2 The Model

To begin, let us consider the following model:

\[
y_t = \beta'_t x_t + (\beta'_2 - \beta'_1) x_t \Psi (\gamma^0, Z_t) + \varepsilon_t, \tag{1}
\]

where \(\beta_1\) and \(\beta_2\) are the pre-shift and post-shift regression slope parameters respectively, with \(\beta_i = (\beta_{1i}, \beta_{2i}, \ldots, \beta_{Ki})'\) being a \(K \times 1\) vector of parameters.
vector of true parameters, \( i = 1, 2 \);

\( y_t \) is the dependent variable.

\( x_t \) is a \( K \) by 1 vector of covariates.

\((\varepsilon_1 \varepsilon_2 \ldots \varepsilon_T)^\top\) is a \( T \) by 1 vector of error term \( \varepsilon_t \) 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.

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

\( \gamma_0 = (\gamma_0^1, \ldots, \gamma_0^m) \in \prod_{j=1}^{m} [\gamma_j, \overline{\gamma}_j] \) is a vector of \( m \) true threshold parameters to be estimated.

The observations \( \{y_t, x_t, Z_t\}_{t=1}^{T} \) 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:

\[
\Psi(\gamma^0, Z_t) = \mathbf{1} (z_{1t} > \gamma_1^0, \ldots, z_{mt} > \gamma_m^0).
\]

(2)

In the scenario of currency crises, imposing such a threshold condition implies that the crisis will not be triggered until all of the threshold variables exceed the critical thresholds.

If the condition is that at least one threshold variable exceeds the critical value, then

\[
\Psi(\gamma^0, Z_t) = 1 - \mathbf{1} (z_{1t} \leq \gamma_1^0, \ldots, z_{mt} \leq \gamma_m^0).
\]

(3)

This case can be rewritten in the form of the previous case. Let \( w_{jt} = -z_{jt} \), then

\[
\Psi(\gamma^0, Z_t) = 1 - \Psi(-\gamma^0, W_t).
\]

As the second case can be incorporated into the first, we will focus on the first case in this paper. 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. We will derive the estimators of the structural and threshold parameters. We will also discuss the way to carry out statistical inferences in this model.

For notational simplicity, we let

\[
\Psi_t(\gamma^0) = \Psi(\gamma^0, Z_t) = \mathbf{1} (z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0).
\]

(4)

Define

\[
F(\gamma_1, \gamma_2) = \Pr(z_1 \leq \gamma_1, z_2 \leq \gamma_2)
\]

(5)
and
\[ F_i(\gamma_i) = \Pr (z_i \leq \gamma_i), \quad (i = 1, 2). \] (6)

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

(a) \( \frac{1}{T} \sum_{t=1}^{T} I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) \overset{p}{\to} \Pr (z_1 > \gamma_1, z_2 > \gamma_2) \overset{d}{=} F (\gamma_1, \gamma_2); \)

(b) \( \frac{1}{T} \sum_{t=1}^{T} I (z_{1t} > \gamma_1) \overset{p}{\to} \Pr (z_1 > \gamma_1) \overset{d}{=} F_1 (\gamma_1); \)

(c) \( \frac{1}{T} \sum_{t=1}^{T} I (z_{2t} > \gamma_2) \overset{p}{\to} \Pr (z_2 > \gamma_2) \overset{d}{=} F_2 (\gamma_2). \)

Define
\[ F_{i\gamma} = \frac{\partial}{\partial \gamma_i} F (\gamma_1, \gamma_2), \quad (i = 1, 2), \] (7)
\[ F_i^0 = F_i (\gamma_1^0, \gamma_2^0), \quad (i = 1, 2). \] (8)

Define the moment functionals:
\[ \overline{M}_i = \overline{M} (\gamma_1, \gamma_2) = E (x_t x_t' I (z_{1t} > \gamma_1, z_{2t} > \gamma_2)), \] (9)
\[ \overline{M}_0 = \overline{M} (\gamma_1^0, \gamma_2^0), \] (10)
\[ M = E (x_t x_t'), \] (11)
\[ D_\gamma = D (\gamma_1, \gamma_2) = E (x_t x_t' | z_{1t} = \gamma_1, z_{2t} = \gamma_2), \] (12)
\[ D = D (\gamma_1^0, \gamma_2^0), \] (13)
\[ V (\gamma_1, \gamma_2) = E (x_t x_t' \varepsilon_t^2 | z_{1t} = \gamma_1, z_{2t} = \gamma_2), \] (14)
\[ V = V (\gamma_1^0, \gamma_2^0), \] (15)
\[ G (\gamma_1, \gamma_2) = M^{-1} \overline{M} (\gamma_1, \gamma_2). \] (16)

We impose the following assumptions:

(A1) \( (x_t, Z_t, \varepsilon_t) \) is strictly stationary, ergodic and \( \rho \)-mixing, with \( \rho \)-mixing coefficients satisfying \( \sum_{j=1}^{\infty} \rho_j^2 < \infty; \)
\((A2)\) \(E(\varepsilon_t|Z_{t-1}) = 0;\)
\((A3)\) \(E|x_t|^4 < \infty\) and \(E|\varepsilon_t|^4 < \infty;\)
\((A4)\) For all \(\gamma \in \Gamma\) and \(i = 1, 2,\) \(E(|x_t|^4|Z_t = \gamma), F_i(\gamma_1, \gamma_2)\) are bounded;
\((A5)\) At \(\gamma = \gamma_0\) and \(i = 1, 2, F_{i\gamma}, D_\gamma\) and \(V(\gamma_1, \gamma_2)\) are continuous;
\((A6)\) \(\delta = \beta_2 - \beta_1 = cT^{-\alpha}\) where \(c \neq 0\) and \(0 < \alpha < \frac{1}{2};\)
\((A7)\) \(c'c, c'Vc, F_1^0\) and \(F_2^0\) are positive;
\((A8)\) \(M > M(\gamma_1, \gamma_2) > 0\) for all \(\gamma \in \Gamma.\)

\((A1)\) implies that all of the regressors are stationary and ergodic. It is needed 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 to have a non-degenerating distribution for the threshold estimators. \((A8)\) is the conventional full-rank condition which excludes multicollinearity.

Given \(\gamma = (\gamma_1, \gamma_2),\) the OLS estimators for \(\beta\) are

\[
\hat{\beta}_1(\gamma) = \sum_{t=1}^{T} y_t x'_t (1 - \Psi_t(\gamma)) \left( \sum_{t=1}^{T} x_t x'_t (1 - \Psi_t(\gamma)) \right)^{-1} \tag{17}
\]

and

\[
\hat{\beta}_2(\gamma) = \sum_{t=1}^{T} y_t x'_t \Psi_t(\gamma) \left( \sum_{t=1}^{T} x_t x'_t \Psi_t(\gamma) \right)^{-1} \tag{18}
\]

Define

\[
S_T(\gamma) = \sum_{t=1}^{T} \left( y_t - \hat{\beta}_1(\gamma) x_t - \left( \hat{\beta}_2(\gamma) - \hat{\beta}_1(\gamma) \right) x_t \Psi_t(\gamma) \right)^2, \tag{19}
\]

\[
\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2) = \arg \min_{(\gamma_1, \gamma_2) \in \Gamma_T} S_T(\gamma_1, \gamma_2), \tag{20}
\]

where

\[
\Gamma_T = \prod_{j=1}^{2} \left( \left[ \frac{\gamma_j}{\gamma_j} \right] \cap \{z_{j1}, ..., z_{jT} \} \right) \tag{21}
\]
The final structural estimators are then defined as

$$
\hat{\beta}_1 (\gamma_1, \gamma_2)
$$

and

$$
\hat{\beta}_2 (\gamma_1, \gamma_2).
$$

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

$$
M = E (x_t^2),
$$

(22)

$$
\mathcal{G} (\gamma_1, \gamma_2) = M^2 / M.
$$

(23)

Note from Appendix A1 that

$$
\hat{\beta}_1 (\gamma_1, \gamma_2) = \beta_1 + \delta \times \frac{\sum_{t=1}^T x_t^2 [I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2)]}{\sum_{t=1}^T x_t^2 (1 - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2))}
$$

+ o_p(1).

(24)

Similarly, we have

$$
\hat{\beta}_2 (\gamma_1, \gamma_2) = \beta_2 - \delta \times \frac{\sum_{t=1}^T x_t^2 [I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2)]}{\sum_{t=1}^T x_t^2 I (z_{1t} > \gamma_1, z_{2t} > \gamma_2)}
$$

+ o_p(1).

(25)

We can partition the space of $\gamma$ into four regions and discuss four separate cases:

**Case 1:** $\gamma_1 \leq \gamma_0^1, \gamma_2 \leq \gamma_0^2$

$$
\hat{\beta}_1 (\gamma_1, \gamma_2) \overset{p}{\rightarrow} \beta_1,
$$

(26)

$$
\hat{\beta}_2 (\gamma_1, \gamma_2) \overset{p}{\rightarrow} \beta_2 - \delta \left(1 - \frac{\mathcal{G} (\gamma_0^1, \gamma_0^2)}{\mathcal{G} (\gamma_1, \gamma_2)} \right).
$$

(27)
Case 2: \( \gamma_1 > \gamma_0^1, \gamma_2 \leq \gamma_0^2 \)

\[
\hat{\beta}_1 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_1 + \delta \frac{G(\gamma_0^1, \gamma_0^2) - G(\gamma_1, \gamma_2)}{1 - G(\gamma_1, \gamma_2)},
\]

\[
\hat{\beta}_2 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_2 - \delta \left(1 - \frac{G(\gamma_1, \gamma_2)}{G(\gamma_1, \gamma_2)}\right).
\]

Case 3: \( \gamma_1 \leq \gamma_0^1, \gamma_2 > \gamma_0^2 \)

\[
\hat{\beta}_1 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_1 + \delta \frac{G(\gamma_0^1, \gamma_0^2) - G(\gamma_0^1, \gamma_2)}{1 - G(\gamma_1, \gamma_2)},
\]

\[
\hat{\beta}_2 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_2 - \delta \left(1 - \frac{G(\gamma_0^1, \gamma_2)}{G(\gamma_1, \gamma_2)}\right).
\]

Case 4: \( \gamma_1 > \gamma_0^1, \gamma_2 > \gamma_0^2 \)

\[
\hat{\beta}_1 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_1 + \delta \frac{G(\gamma_0^1, \gamma_0^2) - G(\gamma_1, \gamma_2)}{1 - G(\gamma_1, \gamma_2)},
\]

\[
\hat{\beta}_2 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_2.
\]

Note that \( \hat{\beta}_i (\gamma_0^1, \gamma_0^2) \xrightarrow{p} \beta_i, \ i = 1, 2 \). This means the structural estimators can be consistently estimated if the threshold estimators are super-consistent.

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

We now study the behavior of the residual sum of squares. Let

\[
x_t (\gamma) = x_t \Psi_t (\gamma)
\]

and let \( X \) and \( X_\gamma \) be \( T \) by \( K \) matrices formed by stacking the vectors \( x_t' \) and \( x_t (\gamma)' \).

Thus, our model can be rewritten as

\[
Y = X\beta_1 + X_\gamma \delta + \varepsilon.
\]

The residual sum of squares can also be written as

\[
S_T (\gamma) = (Y - X\hat{\beta}_1 (\gamma) - X_\gamma \hat{\delta} (\gamma))' (Y - X\hat{\beta}_1 (\gamma) - X_\gamma \hat{\delta} (\gamma)) = Y' (I - P_\gamma) Y,
\]

where
\[ P_\gamma = \tilde{X}_\gamma \left( \tilde{X}_\gamma^t \tilde{X}_\gamma \right)^{-1} \tilde{X}_\gamma', \]

\[ \tilde{X}_\gamma = [X \ X_\gamma]. \]

As \( Y - X\beta_1 - X_\gamma\delta \) and \( X \) lies in the space spanned by \( P_\gamma \),

\[ S_T(\gamma) - \dot{\varepsilon}' \varepsilon = -\varepsilon' P_\gamma \varepsilon + 2\delta' X_0'(I - P_\gamma) \varepsilon + \delta' X_0'(I - P_\gamma) X_0 \delta \]

and

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

where \( X_0 = X_{\gamma_0} \).

We discuss four cases. From Appendix B, in each case, \( \frac{1}{T^{1-2\alpha}} (S_T(\gamma) - \dot{\varepsilon}' \varepsilon) \xrightarrow{p} b_i(\gamma), \) with \( b_i(\gamma_0) = 0 \), \( i = 1, 2, 3, 4 \).

**Case 1:** \( \gamma_1 \leq \gamma_1^0, \gamma_2 \leq \gamma_2^0 \)

\[ b_1(\gamma) = c' \left( \overline{M}_0 - \overline{M}\gamma_\gamma^{-1} \overline{M}_0 \right) c \geq 0. \]

\[ \frac{\partial}{\partial \gamma_1} b_1(\gamma) = c' \overline{M}_0 \overline{M}_\gamma^{-1} D_\gamma F_1 \overline{M}_\gamma^{-1} \overline{M}_0 c \leq 0, \]

\[ \frac{\partial}{\partial \gamma_2} b_1(\gamma) = c' \overline{M}_0 \overline{M}_\gamma^{-1} D_\gamma F_2 \overline{M}_\gamma^{-1} \overline{M}_0 c \leq 0. \]

**Case 2:** \( \gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0 \)

\[ b_2(\gamma) = c' \left( \overline{M}_0 - (\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2^0)) (M - \overline{M}_\gamma)^{-1} (\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2^0)) \right) \overline{M}_\gamma^{-1} \overline{M}(\gamma_1^0, \gamma_2^0) \overline{M}(\gamma_1^0, \gamma_2^0) c \]

\[ > 0. \]

**Case 3:** \( \gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0 \)

\[ b_3(\gamma) = c' \left( \overline{M}_0 - (\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2)) (M - \overline{M}_\gamma)^{-1} (\overline{M}_0 - \overline{M}(\gamma_1^0, \gamma_2)) \right) \overline{M}_\gamma^{-1} \overline{M}(\gamma_1^0, \gamma_2) \overline{M}(\gamma_1^0, \gamma_2) c \]

\[ > 0. \]

\[ 10 \]
Case 4: $\gamma_1 > \gamma_0^1$, $\gamma_2 > \gamma_0^2$

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

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

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

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.

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

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

and

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

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

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

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

$$b_4 (\gamma_1, \gamma_2) = c^2 \left( F (\gamma_1, \gamma_2) - F (\gamma_1, \gamma_2) \right) \frac{1 - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)}.$$
2.2 Asymptotic joint distribution of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ when $z_{1t}$ and $z_{2t}$ are independent

The threshold estimators are analogous to the change-point estimators in the structural-change model. As is well known, 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). To obtain a non-degenerate distribution, the usual practice is to let the magnitude of change go to zero at an appropriate rate.

In the threshold model, in order to obtain the distribution of the threshold estimators, we also let the change go to zero at a certain rate. The distribution of the threshold estimator has been obtained by Hansen (1999, 2000) for the single threshold variable case.

Tsay (1998) and Hansen (1999, 2000) argue that a threshold model can be transformed into a change-point model by re-indexing the threshold variable. Their approach, however, cannot be applied when there is more than one threshold variable. For cases with multiple thresholds, no study on the joint distribution of the threshold variable has ever been conducted. This paper attempts to explore such an area. From Appendix C, we have

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

\[
T^{1-2a} \frac{(c' Dc)^2}{c' Vc} \left( (\hat{\gamma}_1 - \gamma_1) F_{1}, (\hat{\gamma}_2 - \gamma_2) F_{2} \right) = (\hat{r}_1, \hat{r}_2)
\]

\[
\frac{\partial}{\partial (r_1, r_2) \in \mathbb{R}^2} \sum_{j=1}^{2} \left( -\frac{1}{2} |r_j| + W_j(r_j) \right),
\]

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

\[
F(\hat{r}_1, \hat{r}_2) (a_1, a_2) = \prod_{j=1}^{2} \left( 1 + \sqrt{\frac{a_j}{2\pi}} \exp \left( -\frac{a_j}{8} \right) + \frac{3}{2} \exp (a_j) \Phi \left( -\frac{3\sqrt{a_j}}{2} \right) - \frac{a_j + 5}{2} \Phi \left( -\frac{\sqrt{a_j}}{2} \right) \right),
\]

where $\Phi(\cdot)$ is the cdf of a standard normal distribution.
Thus, the joint density function, which is depicted in Figure 4b, can be found as

\[
f(\bar{r}_1, \bar{r}_2)(a_1, a_2) = \Pi_{j=1}^{2} \left( \frac{3}{2} \exp (a_j) \Phi \left( -\frac{3\sqrt{a_j}}{2} \right) - \frac{1}{2} \Phi \left( -\frac{\sqrt{a_j}}{2} \right) \right). \tag{39}
\]

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

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

\[
-T^{1-2\alpha} \frac{c'Dc}{c'Vc} (\hat{\gamma} - \gamma_0) \odot \frac{\partial F (\hat{\gamma}_1, \ldots, \hat{\gamma}_m)}{\partial \gamma} \max_{(r_1, \ldots, r_m) \in \mathbb{R}^m} \sum_{j=1}^{m} \left( -\frac{1}{2} |r_j| + W_j(r_j) \right). \tag{40}
\]

where \(\odot\) is the Hadamard product operator that multiplies on an element by element basis, and

\[
F(\bar{r}_1, \ldots, \bar{r}_m)(a_1, \ldots, a_m) = \Pi_{j=1}^{m} \left( 1 + \sqrt{\frac{a_j}{2\pi}} \exp \left( -\frac{a_j}{8} \right) + \frac{3}{2} \exp (a_j) \Phi \left( -\frac{3\sqrt{a_j}}{2} \right) - \frac{a_j + 5}{2} \Phi \left( -\frac{\sqrt{a_j}}{2} \right) \right), \tag{41}
\]

\[
f(\bar{r}_1, \ldots, \bar{r}_m)(a_1, \ldots, a_m) = \Pi_{j=1}^{m} \left( \frac{3}{2} \exp (a_j) \Phi \left( -\frac{3\sqrt{a_j}}{2} \right) - \frac{1}{2} \Phi \left( -\frac{\sqrt{a_j}}{2} \right) \right). \tag{42}
\]

It should be noted that if the threshold variables are dependent, the consistency and the joint distribution of the threshold estimators may not be easily obtained. For instance, if \(z_2 = -z_1\), we may not be able to partition the data into four groups according to the values of the two threshold variables, so the above distributional result will not hold.

### 3 A Fast Algorithm for the Estimation of \(\gamma\) when \(x_t, z_{1t},\) and \(z_{2t}\) are Independent

The estimation of the threshold values requires a global grid search, which is subject to the curse of dimensionality when the number of threshold variables increases. We propose a fast estimation method when the regressors and the threshold variables are all mutually independent.
To illustrate this, we consider the following simple model with a single regressor and two threshold variables:

\[ y_t = \beta_1 x_t + \delta x_t \Psi_t (\gamma^0) + \varepsilon_t. \]  

(43)

In this case, the asymptotic results are the same as the case in which all \( x_t = 1 \). From Appendix A2, it can be shown that

\[
\sup_{(\gamma_1, \gamma_2) \in \mathbb{R}^2} \left| \frac{1}{T} S_T (\gamma_1, \gamma_2) - g (\gamma_1, \gamma_2) \right| = o_p (1) .
\]

(44)

Let

\[ b_i (\gamma_1, \gamma_2) = T^{2a} (g (\gamma_1, \gamma_2) - \sigma^2) \quad \text{for } i = 1, 2, 3, 4. \]

(45)

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

**Case 1:** \( \gamma_1 \leq \gamma_1^0, \gamma_2 \leq \gamma_2^0 \)

\[ b_1 (\gamma_1, \gamma_2) = c^2 \mathcal{F}_1 (\gamma_1^0) \mathcal{F}_2 (\gamma_2^0) \left( 1 - \frac{\mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)}{\mathcal{F}_1 (\gamma_1^0) \mathcal{F}_2 (\gamma_2^0)} \right), \]

(46)

\[
\frac{\partial b_1 (\gamma_1, \gamma_2)}{\partial \gamma_1} = -c^2 \frac{\mathcal{F}_1 (\gamma_1^0)^2 \mathcal{F}_2 (\gamma_2^0)^2}{\mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)} H_1 (\gamma_1) \leq 0, \]

(47)

\[
\frac{\partial b_1 (\gamma_1, \gamma_2)}{\partial \gamma_2} = -c^2 \frac{\mathcal{F}_1 (\gamma_1^0)^2 \mathcal{F}_2 (\gamma_2^0)^2}{\mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)} H_2 (\gamma_2) \leq 0,
\]

(48)

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

**Case 2:** \( \gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0 \)

\[ b_2 (\gamma_1, \gamma_2) = c^2 \mathcal{F}_2 (\gamma_2^0) \left( \frac{\mathcal{F}_1 (\gamma_1^0)}{\mathcal{F}_2 (\gamma_2^0)} - \frac{\mathcal{F}_1 (\gamma_1)}{\mathcal{F}_2 (\gamma_2)} - \frac{(\mathcal{F}_1 (\gamma_1^0) - \mathcal{F}_1 (\gamma_1))^2}{1 - \mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)} \right), \]

(49)

\[
\frac{\partial b_2 (\gamma_1, \gamma_2)}{\partial \gamma_1} = c^2 \frac{\mathcal{F}_1 (\gamma_1^0)^2}{\mathcal{F}_2 (\gamma_2)} \left( \frac{1 - \mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)}{1 - \mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)} \right)^2 f_1 (\gamma_1) > 0, \]

(50)

\[
\frac{\partial b_2 (\gamma_1, \gamma_2)}{\partial \gamma_2} = -c^2 \mathcal{F}_2 (\gamma_2^0) \left( \frac{\mathcal{F}_1 (\gamma_1^0) - \mathcal{F}_1 (\gamma_1)}{1 - \mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)} + \frac{1}{\mathcal{F}_2 (\gamma_2)} \right) \frac{1 - \mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)}{1 - \mathcal{F}_1 (\gamma_1) \mathcal{F}_2 (\gamma_2)} \mathcal{F}_1 (\gamma_1) H_2 (\gamma_2)
\]

< 0.

(51)
Case 3: $\gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0$

$$b_3(\gamma_1, \gamma_2) = e^2 F_1 (\gamma_1^0)^2 \left( \frac{F_2 (\gamma_2)}{F_1 (\gamma_1)} - \frac{F_2 (\gamma_2^0)}{F_1 (\gamma_1^0)} - \frac{(F_2 (\gamma_2^0) - F_2 (\gamma_2))^2}{1 - F_1 (\gamma_1) F_2 (\gamma_2)} \right),$$

$$\frac{\partial b_3 (\gamma_1, \gamma_2)}{\partial \gamma_1} = -e^2 F_1 (\gamma_1^0)^2 \left( \frac{F_2 (\gamma_2) - F_2 (\gamma_2^0)}{1 - F_1 (\gamma_1) F_2 (\gamma_2)} + \frac{1}{F_1 (\gamma_1)} \right) \frac{1 - F_2 (\gamma_2) F_1 (\gamma_1) F_2 (\gamma_2)}{1 - F_1 (\gamma_1) F_2 (\gamma_2)} H_1 (\gamma_1) \quad \text{< 0,}$$

$$\frac{\partial b_3 (\gamma_1, \gamma_2)}{\partial \gamma_2} = e^2 F_1 (\gamma_1^0)^2 \left( \frac{F_2 (\gamma_2) - F_2 (\gamma_2^0)}{1 - F_1 (\gamma_1) F_2 (\gamma_2)} + \frac{1}{F_1 (\gamma_1)} \right)^2 F_1 (\gamma_1) f_2 (\gamma_2) > 0.$$  

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

$$b_4 (\gamma_1, \gamma_2) = e^2 (1 - F_1 (\gamma_1^0) F_2 (\gamma_2^0)) \left( \frac{1 - F_1 (\gamma_1^0) F_2 (\gamma_2^0)}{1 - F_1 (\gamma_1) F_2 (\gamma_2)} \right)^2,$$

$$\frac{\partial b_4 (\gamma_1, \gamma_2)}{\partial \gamma_1} = e^2 \left( \frac{1 - F_1 (\gamma_1^0) F_2 (\gamma_2^0)}{1 - F_1 (\gamma_1) F_2 (\gamma_2)} \right)^2 F_2 (\gamma_2) f_1 (\gamma_1) > 0,$$

$$\frac{\partial b_4 (\gamma_1, \gamma_2)}{\partial \gamma_2} = e^2 \left( \frac{1 - F_1 (\gamma_1^0) F_2 (\gamma_2^0)}{1 - F_1 (\gamma_1) F_2 (\gamma_2)} \right)^2 F_1 (\gamma_1) f_2 (\gamma_2) > 0.$$  

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

$$\text{Arg min}_b (\gamma_1, \gamma_2) = \gamma_1^0 \quad \forall \gamma_2$$  

and

$$\text{Arg min}_b (\gamma_1, \gamma_2) = \gamma_2^0 \quad \forall \gamma_1.$$  

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 a random value to another threshold estimate. This will drastically reduce the computational burden.
4 Model with Panel Data

Our model can be extended to incorporate panel data. We consider a balanced panel with \( n \) individuals over \( T \) periods. Following the lead of Hansen (1999), we assume that all individuals have the same threshold value for each threshold variable. Note that the model in the previous section can be a cross-sectional model or a time series model. 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

\[
\Psi_{it}(\gamma) = I(z_{1it} > \gamma_1, z_{2it} > \gamma_2).
\]

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

\[
y_{it} = \mu_i + \beta_1' x_{it} + \varepsilon_{it}, \quad \Psi_{it}(\gamma) = 0, \tag{60}
\]

\[
y_{it} = \mu_i + \beta_2' x_{it} + \varepsilon_{it}, \quad \Psi_{it}(\gamma) = 1. \tag{61}
\]

We impose the following assumptions:

- **(B1)** For each \( t \), \((x_{it}, Z_{it}, \varepsilon_{it})\) are i.i.d. across \( i \).
- **(B2)** For each \( i \), \( \varepsilon_{it} \) is i.i.d. over \( t \), is independent of \( \{x_{ij}, Z_{ij}\}_{j=1}^{T} \), and \( E(\varepsilon_{it}) = 0 \);
- **(B3)** For each \( j = 1, ..., k \), \( Pr(x_{it}^j = x_{it}^j = ... = x_{iT}^j) < 1 \), where \( x_{it}^j \) is the \( j \)th element of \( x_{it} \).
- **(B4)** \( E|x_{it}|^4 < \infty \) and \( E|\varepsilon_{it}|^4 < \infty \);
- **(B5)** \( \delta = cn^{-\alpha} \) where \( c \neq 0 \) and \( 0 < \alpha < \frac{1}{2} \);
- **(B6)** At \( \gamma = \gamma_0 \) and \( i = 1, 2 \), \( F_1(\gamma_1, \gamma_2), D_\gamma \) and \( V(\gamma_1, \gamma_2) \) are continuous;
- **(B7)** \( 0 < D < \infty \);
- **(B8)** For \( k > t \), \( f_{kit}(\gamma_1^0, \gamma_2^0 | \gamma_1, \gamma_2) < \infty \), where \( f_{kit}(\gamma_1^0, \gamma_2^0 | \gamma_1, \gamma_2) \) is the value of the conditional joint density of \( Z_{it} \) 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 in order for the data to be sortable.

Let

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

(62)

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

(63)

Averaging the above panel equation over $t$, we have

$$\bar{y}_i = \mu_i + \beta_{1i} \bar{x}_i + \delta_{1i} \bar{x}_i (\gamma) + \bar{\varepsilon}_i,$$

(64)

where

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

(65)

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

(66)

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

(67)

$$\bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^{T} \varepsilon_{it}.$$  

(68)

Taking the difference, we have

$$y_{it}^* = \beta_{1i} x_{it}^* + \delta_{1i} x_{it}^* (\gamma) + \varepsilon_{it}^*,$$  

(69)

where

$$y_{it}^* = y_{it} - \bar{y}_i,$$  

(70)

$$x_{it}^* = x_{it} - \bar{x}_i,$$  

(71)

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

(72)

$$\varepsilon_{it}^* = \varepsilon_{it} - \bar{\varepsilon}_i.$$  

(73)
The difference between our model and that of Hansen (1999) is that \( \gamma \) is a scalar in his model, whereas it is a vector in our model. Following Hansen (1999), we let

\[
\begin{align*}
y^*_i &= \begin{bmatrix} y^*_{i2} \\ \vdots \\ y^*_{iT} \end{bmatrix}, \\
 x^*_i &= \begin{bmatrix} x^*_{i2} \\ \vdots \\ x^*_{iT} \end{bmatrix}, \\
 x^*_i (\gamma) &= \begin{bmatrix} x^*_{i2} (\gamma) \\ \vdots \\ x^*_{iT} (\gamma) \end{bmatrix}, \\
 \varepsilon^*_i &= \begin{bmatrix} \varepsilon^*_{i2} \\ \vdots \\ \varepsilon^*_{iT} \end{bmatrix},
\end{align*}
\]

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

\[
\begin{align*}
Y^* &= \begin{bmatrix} y_1^* \\ \vdots \\ y_n^* \end{bmatrix}, \\
 X^* &= \begin{bmatrix} x_1^* \\ \vdots \\ x_n^* \end{bmatrix}, \\
 X^* (\gamma) &= \begin{bmatrix} x_1^* (\gamma) \\ \vdots \\ x_n^* (\gamma) \end{bmatrix}, \\
 \varepsilon^* &= \begin{bmatrix} \varepsilon_1^* \\ \vdots \\ \varepsilon_n^* \end{bmatrix}.
\end{align*}
\]

Thus, our model becomes

\[
Y^* = X^* \beta_1 + X^* (\gamma) \delta + \varepsilon^*. \tag{74}
\]

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

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

\[
\hat{\gamma} = (\hat{\gamma}_1, \hat{\gamma}_2) = \arg \min_{\gamma \in \Gamma_n} S_{nT} (\gamma_1, \gamma_2). \tag{75}
\]

\[
\Gamma_n = \Pi_{j=1}^2 \left( \left[ \gamma_{j1}, \gamma_{j2} \right] \cap \left( \cup_{i=1}^n \{ z_{ji1}, \ldots, z_{jiT} \} \right) \right) \tag{76}
\]

The final structural estimators are then defined as

\[
\hat{\beta}_1 (\hat{\gamma}_1, \hat{\gamma}_2)
\]

and

\[
\hat{\beta}_2 (\hat{\gamma}_1, \hat{\gamma}_2).
\]
and the residual variance is
\[
\hat{\sigma}^2 = \frac{1}{n (T - 1)} S_{nT} (\hat{\gamma}).
\] (77)

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
\[
F(0, 1, 1) = T \frac{S_T (-\infty, -\infty) - S_T (\hat{\gamma}_1, -\infty)}{S_T (\hat{\gamma}_1, -\infty)},
\] (78)
\[
F(0, 1, 2) = T \frac{S_T (-\infty, -\infty) - S_T (-\infty, \hat{\gamma}_2)}{S_T (-\infty, \hat{\gamma}_2)},
\] (79)

where
\(S_T (-\infty, -\infty)\) is the residual sum of squares from the regression without any threshold variable;
\(S_T (\hat{\gamma}_1, -\infty)\) is the residual sum of squares from the regression without the second threshold variable; and
\(S_T (-\infty, \hat{\gamma}_2)\) is the residual sum of squares from the regression without the first threshold variable.

For the notation \(F(\cdot, \cdot, \cdot)\), the first entry in the parenthesis stands for the value of \(m\) under the null hypothesis. The second entry represents the value of \(m\) under the alternative hypothesis. The last entry 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

\[
F(1, 2, 1) = T \frac{S_T(\gamma_1, -\infty) - S_T(\tilde{\gamma}_1, \tilde{\gamma}_2)}{S_T(\tilde{\gamma}_1, \tilde{\gamma}_2)},
\]

(80)

\[
F(1, 2, 2) = T \frac{S_T(-\infty, \gamma_2) - S_T(\tilde{\gamma}_1, \tilde{\gamma}_2)}{S_T(\tilde{\gamma}_1, \tilde{\gamma}_2)},
\]

(81)

where \( S_T(\gamma_1, \gamma_2) \) 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 we conclude that 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. This 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.

As the asymptotic distributions of the above tests are non-standard, we adopt the bootstrapping method of Hansen (1999). For the case of \( m = 0 \) against \( m = 1 \), the method suggested by Hansen (1999) is carried out \( R \) times if we have \( R \) potential candidates of threshold variables. For the tests in the next 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. After that, we draw a sample of size \( T \) with replacement from this empirical distribution and use the errors 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 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, if we are to test

\[
H_0 : m = 1 \\
H_1 : m = 2
\]

We estimate the threshold model with two threshold variables, take its LS residuals and draw the bootstrap \( \hat{\xi}_{it}^{bw} \) residuals from them \((i = 1, 2, \ldots, n; t = 1, 2, \ldots, T)\). Then we use the bootstrap residuals along with the estimated threshold model with one threshold variable to generate the bootstrap dependent variable.
\[ y_{it}^b = \beta_1' x_{it}^* + (\beta_2' - \beta_1') x_{it}^* \Psi (z_{it}, \gamma_1) + \varepsilon_{it}^b. \]  \hspace{1cm} (82)

Using the set of dependent and independent variables \{x_{it}^*, y_{it}^b\}, 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 \( S_{nT} (\tilde{\gamma}_{11}, \tilde{\gamma}_{2}) \). The sum of squared residuals under the null is \( S_{nT} (\tilde{\gamma}_{1}, -\infty) = \sum_{i=1}^{T} \sum_{t=1}^{n} \varepsilon_{it}^b \). The test statistic for testing two threshold variables under the alternative against the null that only the first threshold variable should appear in the model is

\[ F (1, 2, 1) = T \frac{S_{nT} (\tilde{\gamma}_{1}, -\infty) - S_{nT} (\tilde{\gamma}_{11}, \tilde{\gamma}_{2})}{S_{nT} (\tilde{\gamma}_{11}, \tilde{\gamma}_{2})}. \]  \hspace{1cm} (83)

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

\[ F (1, 2, 2) = T \frac{S_{nT} (-\infty, \tilde{\gamma}_{2}) - S_{nT} (\tilde{\gamma}_{11}, \tilde{\gamma}_{2})}{S_{nT} (\tilde{\gamma}_{11}, \tilde{\gamma}_{2})}. \]  \hspace{1cm} (84)

### 5.2 Testing the threshold values

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

\[ H_0 : \gamma = \gamma^0. \]

Under the assumption that \( \varepsilon_t \) is i.i.d. \( N (0, \sigma^2) \), we have

\[ LR_T (\gamma_1, \gamma_2) = T \frac{S_T (\gamma_1, \gamma_2) - S_T (\tilde{\gamma}_{1}, \tilde{\gamma}_{2})}{S_T (\tilde{\gamma}_{1}, \tilde{\gamma}_{2})}. \]  \hspace{1cm} (85)

\( H_0 \) is rejected for a large \( LR_T (\gamma_0^1, \gamma_0^2) \).

If the threshold variables are independent, we can show that

\[ LR_T (\gamma_0^1, \gamma_0^2) \xrightarrow{d} \eta^2 \xi, \]  \hspace{1cm} (86)

where

\[ \xi = \xi_1 + \xi_2, \]  \hspace{1cm} (87)

\[ \xi_1 = \max_{-\infty < r_1 < \infty} (-|r_1| + 2W_1 (r_1)), \]  \hspace{1cm} (88)
\[ \xi_2 = \max_{-\infty < r_2 < \infty} (-|r_2| + 2W_2(r_2)) \quad (89) \]

and

\[ \eta^2 = \frac{dVc}{\sigma^2e'Dc}. \quad (90) \]

The distribution of \( \xi_i \) \( (i = 1, 2) \) is

\[ \Pr (\xi_i \leq x) = \left(1 - e^{-\frac{1}{2}x}\right)^2, \quad (91) \]

\[ f_{\xi_i}(x) = \left(1 - e^{-\frac{1}{2}x}\right) e^{-\frac{1}{2}x}. \quad (92) \]

Thus,

\[ \Pr (\xi \leq x) = \Pr (\xi_1 + \xi_2 \leq x) \]

\[ = \int_{0}^{x} \Pr (\xi_1 \leq x - y) f_{\xi_2}(y) \, dy \]

\[ = 1 - (x + 5) e^{-x} - 2(x - 2) e^{-\frac{1}{2}x}, \quad (93) \]

The density function is given by

\[ f_{\xi}(x) = (x + 4) e^{-x} + (x - 4) e^{-\frac{1}{2}x}. \quad (94) \]

For \( m = 3 \), we have

\[ \Pr (\xi \leq x) = \Pr (\xi_1 + \xi_2 + \xi_3 \leq x) \]

\[ = \int_{0}^{x} \Pr (\xi_1 + \xi_2 \leq x - y) f_{\xi_3}(y) \, dy \]

\[ = \frac{1}{2} e^{-2x} \left( 62e^x + 14xe^x + 2e^{2x} + x^2e^x - 64e^{\frac{3}{2}x} + 16xe^{\frac{3}{2}x} - 2x^2e^{\frac{3}{2}x} \right) \]

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 a computer program, and the values are tabulated in Table A:
Table A: Asymptotic Critical Values

<table>
<thead>
<tr>
<th>Pr ($\xi \leq x$)</th>
<th>0.800</th>
<th>0.85</th>
<th>0.90</th>
<th>0.925</th>
<th>0.95</th>
<th>0.975</th>
<th>0.99</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 2$</td>
<td>8.33</td>
<td>9.13</td>
<td>10.21</td>
<td>10.96</td>
<td>11.98</td>
<td>13.68</td>
<td>15.85</td>
</tr>
<tr>
<td>$m = 3$</td>
<td>11.95</td>
<td>12.90</td>
<td>14.17</td>
<td>15.03</td>
<td>16.20</td>
<td>18.12</td>
<td>20.55</td>
</tr>
<tr>
<td>$m = 4$</td>
<td>15.47</td>
<td>16.54</td>
<td>17.96</td>
<td>18.92</td>
<td>20.21</td>
<td>22.32</td>
<td>29.46</td>
</tr>
<tr>
<td>$m = 5$</td>
<td>18.93</td>
<td>20.10</td>
<td>21.65</td>
<td>22.69</td>
<td>24.10</td>
<td>26.38</td>
<td>29.20</td>
</tr>
<tr>
<td>$m = 6$</td>
<td>22.34</td>
<td>23.61</td>
<td>25.28</td>
<td>26.39</td>
<td>27.90</td>
<td>30.32</td>
<td>33.33</td>
</tr>
<tr>
<td>$m = 7$</td>
<td>25.71</td>
<td>27.07</td>
<td>28.85</td>
<td>30.04</td>
<td>31.63</td>
<td>34.20</td>
<td>37.35</td>
</tr>
<tr>
<td>$m = 8$</td>
<td>29.06</td>
<td>30.50</td>
<td>32.38</td>
<td>33.63</td>
<td>35.31</td>
<td>38.00</td>
<td>41.31</td>
</tr>
<tr>
<td>$m = 9$</td>
<td>32.39</td>
<td>33.90</td>
<td>35.88</td>
<td>37.19</td>
<td>38.95</td>
<td>41.76</td>
<td>45.21</td>
</tr>
<tr>
<td>$m = 10$</td>
<td>35.70</td>
<td>37.28</td>
<td>39.35</td>
<td>40.72</td>
<td>42.55</td>
<td>45.48</td>
<td>49.06</td>
</tr>
</tbody>
</table>

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

$$MGF_{\xi}(t) = \left( \frac{1}{(1 - t)(1 - 2t)} \right)^m \quad \text{for } t < 0.5.$$  \hfill (95)

For the estimation of the nuisance $\eta^2$, we can extend the results of Hansen (2000). In our case for $m = 2$, it can be estimated via a polynomial regression with $(z_1, z_1^2, z_2, z_1 z_2)$ 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 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. $N(0, 1)$, $\gamma_0^1 = 0$, $\gamma_0^2 = 0$, $\beta_1 = 1$, $T = 1000$ (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 by GAUSS programs, which are available from the authors upon request.

**Experiment A.** This experiment studies the behavior of the residual sum of squares and the distribution of $\hat{\beta}_1(\hat{\gamma})$ and $\hat{\beta}_2(\hat{\gamma})$ 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 $\frac{1}{T} S_{T} (\gamma_{1}, \gamma_{2})$.

Figure 1b plots the 3D graph of $g (\gamma_{1}, \gamma_{2})$.

**FIGURES 1a and 1b HERE**

Figure 2a plots the distribution of $\frac{T}{2} \beta_{1} (\hat{\gamma}_{1}, \hat{\gamma}_{2}) - \beta_{1}$.

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

Figure 2c plots the 3D distribution of $T^{1/2} \left( \beta_{1} (\hat{\gamma}_{1}, \hat{\gamma}_{2}) - \beta_{1}, \beta_{2} (\hat{\gamma}_{1}, \hat{\gamma}_{2}) - \beta_{2} \right)$.

**FIGURES 2a – 2c HERE**

**Experiment B.** This experiment studies the distribution of $\hat{\gamma}$ for a shrinking break. Let $\delta = \beta_{2} - \beta_{1} = T^{-\frac{1}{8}}$.

In this case, we have

$$(\hat{\gamma}_{1}, \hat{\gamma}_{2}) = \arg \min_{(\gamma_{1}, \gamma_{2}) \in \Gamma} S_{T} (\gamma_{1}, \gamma_{2}) = \arg \min_{(\gamma_{1}, \gamma_{2}) \in \Gamma} \left[ S_{T} (\gamma_{1}, \gamma_{2}) - S_{T} (\gamma_{0}^{1}, \gamma_{0}^{2}) \right].$$

>From Appendix A3, for $\gamma_{1} = \gamma_{1}^{0} + \frac{u_{1}}{T^{1-2\alpha}}, \gamma_{2} = \gamma_{2}^{0} + \frac{u_{2}}{T^{1-2\alpha}}, [u_{1}, u_{2}] \in \mathbb{R}^{2}$, we have

$$S_{T} (\gamma_{1}, \gamma_{2}) - S_{T} (\gamma_{1}^{0}, \gamma_{2}^{0}) = \frac{d}{F_{1}^{0}} \left( v_{1} + 2T^{-\alpha} \sum_{t=1} |u_{1}|^{2T^{2\alpha}} \varepsilon_{t}^{0} \right) - \frac{d}{F_{2}^{0}} \left( v_{2} + 2T^{-\alpha} \sum_{t=1} |u_{2}|^{2T^{2\alpha}} \varepsilon_{t}^{b} \right),$$

where $\varepsilon_{t}^{0}$ and $\varepsilon_{t}^{b}$ are independent.

Let

$$r_{1} = -T_{1}^{0} u_{1},$$

$$r_{2} = -T_{2}^{0} u_{2}.$$  

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

$$T^{1-2\alpha} \left( f_{1} (\gamma_{1}^{0}) F_{2} (\gamma_{2}^{0}) (\hat{\gamma}_{1} - \gamma_{1}^{0}) ; f_{2} (\gamma_{2}^{0}) F_{1} (\gamma_{1}^{0}) (\hat{\gamma}_{2} - \gamma_{2}^{0}) \right) \rightarrow \arg \max \sum_{j=1}^{2} \left( -\frac{1}{2} |r_{j}| + W_{j} (r_{j}) \right).$$
Under this setting, we have

\[ f_1 (\gamma_0^0) = f_2 (\gamma_2^0) = f_1 (0) = \frac{1}{\sqrt{2\pi}}, \]

\[ T_1 (\gamma_1^0) = T_2 (\gamma_2^0) = T_1 (0) = \frac{1}{2}. \]

Figure 3a plots the distribution of \( T_1^{3/4} f_1 (\gamma_1^0) T_2 (\gamma_2^0) (\gamma_1 - \gamma_1^0) \) when \( \beta_2 - \beta_1 = T^{-1/8} \).

Figure 3b plots the distribution of \( T_1^{3/4} f_2 (\gamma_2^0) T_1 (\gamma_1^0) (\gamma_2 - \gamma_2^0) \) when \( \beta_2 - \beta_1 = T^{-1/8} \).

**FIGURES 3a and 3b HERE**

Figure 4a plots the 3D distribution of

\[ T_1^{1-2\alpha} (f_1 (\gamma_1^0) T_2 (\gamma_2^0) (\gamma_1 - \gamma_1^0), f_2 (\gamma_2^0) T_1 (\gamma_1^0) (\gamma_2 - \gamma_2^0)) \]

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

Figure 4b plots the joint density \( f_{(\tilde{\gamma}_1, \tilde{\gamma}_2)} (a_1, a_2) \), the definition of which is provided in eqn.(39).

**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 \( \delta = \beta_2 - \beta_1 = T^{-\frac{1}{4}}. \)

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

\[ LR_T (\gamma_1^0, \gamma_2^0) = T \frac{S_T (\gamma_1^0, \gamma_2^0) - S_T (\gamma_1, \gamma_2)}{S_T (\gamma_1, \gamma_2)}. \]  \hfill (96)

**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 \( (\gamma_1, \gamma_2) \) around the estimated thresholds such that \( LR_T (\gamma_1, \gamma_2) = 11.98. \)

**FIGURE 5b HERE**
7 Empirical Application

The empirical relevancy of the theory is studied through an application to the currency crisis models. The threshold model is particularly appropriate for this economic problem as all the currency crisis models in the literature suggest that there are significant threshold effects in the crisis indicators. The identification of the critical thresholds of the crisis indicators has important policy implications as they provide guidelines to the policy makers on the formulation of regulatory policies to minimize the stampede of currency crises.

So far, much of the empirical work on currency crises has been concerned with finding relevant crisis indicators: this work includes that of Kaminsky, Lizondo and Reinhart (1997), Kaminsky (1998) and Hali (2000), among others. While it is helpful to understand the relevancy of different crisis indicators, it is equally important for policy makers to be able to identify the critical thresholds of the indicators above which the economy becomes “too weak” 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. The selection of the threshold variables in this study is closely guided by the three generations of currency crisis models. The model is specified as follows:

\[ y_{it} = \mu_i + \beta_1^t x_{it} + (\beta_2^t - \beta_1^t) x_{it} \Psi (z_{it}, \gamma) + \varepsilon_{it}. \]  

(97)

Currency crises will not be triggered until all of the threshold variables exceed the critical thresholds. This means that

\[ \Psi (z_{it}, \gamma^0) = \Pi_{j=1}^m I (z_{jt} > \gamma^0_j). \]  

(98)

The number of threshold variables \((m)\) to be included in the model is determined by the tests that are 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. The rationale of this index is as follows.
Central banks can respond to a downward pressure in the foreign exchange market in three ways: they can let the exchange rate depreciate, or they can defend their currencies by running down their reserves or by raising the discount rate. 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:

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

(99)

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

$\alpha_1$, $\alpha_2$ and $\alpha_3$ 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 the pressure on the foreign exchange market becomes significantly higher once the fiscal deficit exceeds a certain threshold. The second generation model (Obstfeld, 1986) 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 (McKinnon and Huw, 1996; Chang and Velasco, 1998a and 1998b) indicates that short-term external liabilities relative to reserves is one crucial threshold variable in currency crises. Given these implications from the theories, an important empirical question is to test for the existence of threshold effects in the threshold variables and estimate the threshold values if threshold effects are found.
Table 1 summarizes the threshold variables $z_{it}$ that are implied by 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 the GDP. The interest rate differential is constructed as the difference between the 3-month domestic and 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 threshold variables are above their thresholds, the economy endogenously enters an unstable regime that accelerates the downward pressure in the foreign exchange market.

In estimating the thresholds, we include two fundamentals as the explanatory variables ($x_{it}$) in the regression. These variables include the real exchange rate and the ratio of domestic credit to GDP (Edwards, 1989; Dornbusch, Goldgajn and Valdés, 1995; Eichengreen, Rose and Wyplosz, 1995; Frankel and Rose, 1996; Sachs, Tornell and Velasco, 1996 and Lau and Yan, 2004), both in natural log. The real exchange rate index measures the change in the real exchange rate index 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 over-valued, and hence there is a stronger pressure for the real exchange rate to revert to the mean. Even though this measure of misalignment is only an indirect measure and does not control for long-run productivity changes, it is common in the literature because it is useful for identifying countries that have experienced extreme overvaluations.

The domestic credit variable is measured 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 and is dubbed the “lending boom effect” in the literature. As there are no internationally comparable ratios of non-performing loans to total assets, the ratio of domestic credit to GDP is employed 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 more 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.

7.3 Testing the number of threshold variables
In this section, we apply the tests described in Section 5.1 to identify the presence of threshold effects in the three threshold variables, motivated by the currency crisis theories. As the asymptotic distribution of the test statistic is non-standard and generally depends on the moments of the sample, the bootstrap procedure drawn from Hansen (1996, 1999) is used to approximate the sampling distribution of the test statistic. First, we estimate the model under the alternative hypothesis. Then, we group the regression residuals (after fixed-effect transformation) \( \hat{\varepsilon}_{it} \) by individual \( \hat{\varepsilon}_{it} = (\hat{\varepsilon}_{i1}, \hat{\varepsilon}_{i2}, \ldots, \hat{\varepsilon}_{iT}) \) and draw with replacement error sample of individual \( \hat{\varepsilon}_{it} (t = 1, 2, \ldots, T) \) from this empirical distribution \( \hat{\varepsilon}_{it}^* \). This gives the bootstrap errors. The bootstrap dependent variable \( \hat{y}_{it}^b \) is then generated under the null hypothesis, which depends on the LS estimates \( \hat{\beta} \) and \( \hat{\gamma} \) of the threshold model under the null. From the bootstrap sample \( \{x_{it}^*, \hat{y}_{it}^b\} \), 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 \( p = \frac{1}{B} \sum_{b=1}^{B} I \{ F^b > F_{\text{actual}} \} \) where \( F^b \) is the test statistic computed from one bootstrap sample, \( F_{\text{actual}} \) 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 and 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 for the fiscal deficits, short-term external liabilities and lending rate differentials variables 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 in which the fiscal deficit variable is dropped from the pair of fiscal deficit and short-term external liabilities and from the pair of fiscal

<table>
<thead>
<tr>
<th>Crisis models</th>
<th>( z_{it} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>First generation crisis model</td>
<td>ratio of fiscal deficits to GDP</td>
</tr>
<tr>
<td>Second generation crisis model</td>
<td>differentials between the domestic interest rate and foreign interest rate</td>
</tr>
<tr>
<td>Third generation crisis model</td>
<td>ratio of short-term external liabilities to foreign exchange reserves</td>
</tr>
</tbody>
</table>

Table 1: Threshold variables implied by the three generations of currency crisis models
Table 2: (a) Testing one threshold variable against no threshold variable

deficit and lending rate differential under the alternative. 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 variables can be dropped from the list of three cannot be 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. They are 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 to the interest rate differentials in practice and hence only one of the 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.
<table>
<thead>
<tr>
<th></th>
<th>$H_0 : m = 1$</th>
<th>$H_0 : m = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(fiscal deficit)</td>
<td>(short liabilities)</td>
</tr>
<tr>
<td></td>
<td>$H_1 : m = 2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(fiscal deficit, short liabilities)</td>
<td></td>
</tr>
<tr>
<td>$F$ p-value</td>
<td>27.1031</td>
<td>5.3490</td>
</tr>
<tr>
<td></td>
<td>0.0000**</td>
<td>0.9667</td>
</tr>
<tr>
<td></td>
<td>$H_0 : m = 1$</td>
<td>$H_0 : m = 1$</td>
</tr>
<tr>
<td></td>
<td>(fiscal deficit)</td>
<td>(lending rate diff.)</td>
</tr>
<tr>
<td></td>
<td>$H_1 : m = 2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(fiscal deficit, lending rate diff.)</td>
<td></td>
</tr>
<tr>
<td>$F$ p-value</td>
<td>62.0889</td>
<td>0.3865</td>
</tr>
<tr>
<td></td>
<td>0.0000**</td>
<td>1.0000</td>
</tr>
<tr>
<td></td>
<td>$H_0 : m = 1$</td>
<td>$H_0 : m = 1$</td>
</tr>
<tr>
<td></td>
<td>(short liabilities)</td>
<td>(lending rate diff.)</td>
</tr>
<tr>
<td></td>
<td>$H_1 : m = 2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(short liabilities, lending rate diff.)</td>
<td></td>
</tr>
<tr>
<td>$F$ p-value</td>
<td>91.0530</td>
<td>22.8329</td>
</tr>
<tr>
<td></td>
<td>0.0000**</td>
<td>0.0000**</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th></th>
<th>$H_0 : m = 2$</th>
<th>$H_0 : m = 2$</th>
<th>$H_0 : m = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(fiscal deficit,)</td>
<td>(fiscal deficit,</td>
<td>(short liabilities,</td>
</tr>
<tr>
<td></td>
<td>short liabilities)</td>
<td>lending diff.)</td>
<td>lending diff.)</td>
</tr>
<tr>
<td></td>
<td>$H_1 : m = 3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(fiscal deficit, short liabilities, lending rate diff.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$F$ p-value</td>
<td>96.1324</td>
<td>21.9379</td>
<td>1.1477</td>
</tr>
<tr>
<td></td>
<td>0.0000**</td>
<td>0.0153*</td>
<td>0.9866</td>
</tr>
</tbody>
</table>

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

In this section, we estimate the threshold values of the two threshold variables: namely, the ratio of short-term external liabilities to reserves and the lending rate differential. The threshold estimates are obtained by searching through values of $\gamma$ that equal the distinct values of the threshold variables in our sample. As it is undesirable for a threshold $\hat{\gamma}$ to be selected if too few observations fall into one or the other regime, we eliminate the smallest and largest 15 percent of each threshold variable when setting up the values of $\gamma$ to be searched for $\hat{\gamma}$ (see Hansen (1999) for a similar treatment). The $\hat{\gamma}$ that minimizes the sum of squared residuals is selected.

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. 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). The test statistics for testing the joint significance of the two threshold variables ($H_0: m = 0$ 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 are obtained using the bootstrap procedure discussed in Section 7.3 and they give strong evidence of threshold effects. When both threshold variables exceed the critical thresholds, the economy enters a zone of vulnerability and is likely to undergo an extreme downward adjustment in the foreign exchange market. In view of this, governments can use these threshold estimates to formulate regulatory policies to reduce the risk of having currency crises by taking preemptive measures to avoid the threshold variables from crossing these critical threshold values.

The coefficients of the ratio of domestic credit to GDP for both the Asian and Latin American countries are significantly positive when both threshold variables surpass the critical thresholds (that is, when $\Psi(z_{it}, \gamma) = 1$). This indicates that the vulnerability of the banking sector is a crucial factor in determining the exchange market pressure under this regime.
\[ y_{it} \equiv \text{exchange market pressure index (EMP}_{it} \text{)} \]
\[ z_{it} \equiv \{ \frac{\text{short term external liabilities}}{\text{reserves}}, \text{lending rate differentials} \} \]
\[ x_{it} \equiv \{ 1, \text{real exchange rate appreciation index,} \frac{\text{Domestic credit}}{\text{GDP}} \} \]

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>Asian countries</th>
<th>Latin American countries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(i)</td>
<td>(ii)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.1989</td>
<td>0.0512</td>
</tr>
<tr>
<td></td>
<td>(-0.9436)</td>
<td>(0.6371)</td>
</tr>
<tr>
<td>Real exch. rate appreciation</td>
<td>-2.7235</td>
<td>1.7293</td>
</tr>
<tr>
<td></td>
<td>(-1.8493)</td>
<td>(6.2783)**</td>
</tr>
<tr>
<td>Domestic credit GDP</td>
<td>0.0588</td>
<td>1.7464</td>
</tr>
<tr>
<td></td>
<td>(0.0442)</td>
<td>(6.0512)**</td>
</tr>
</tbody>
</table>

threshold estimates

<table>
<thead>
<tr>
<th></th>
<th>Asian countries</th>
<th>Latin American countries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short term external liabilities</td>
<td>3.1758</td>
<td>3.9851</td>
</tr>
<tr>
<td>Reserves</td>
<td>2.0566</td>
<td>21.8866</td>
</tr>
<tr>
<td>( F ) stat</td>
<td>33.8296</td>
<td>34.0269</td>
</tr>
<tr>
<td>p value</td>
<td>0.0000**</td>
<td>0.0000**</td>
</tr>
<tr>
<td>Observations</td>
<td>641</td>
<td>379</td>
</tr>
</tbody>
</table>

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
7.5 A Graphical Analysis of the Threshold Effects

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

\[
\text{Crisis}_{it} = \begin{cases} 
1 & \text{if } EMP_{it} > \mu_{EMP,it} + 3\sigma_{EMP,it} \\
0 & \text{otherwise}
\end{cases}
\]

where \( \mu_{EMP,it} \) and \( \sigma_{EMP,it} \) are the mean and standard deviation of the exchange market pressure index in country \( i \) and 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 lending rate differential started to go above the critical thresholds less than two years before the 1997 Indonesian and South Korean crises, and remained above the thresholds at the outbreak of the crises. Figure 6(f) indicates that the 1997 Philippine crisis occurred as soon as the short-term external liabilities exceeded the critical threshold, given that the lending rate differential had already surpassed the threshold a while ago. Figure 7(b) indicates that both the ratio of short-term external liabilities to reserves and lending rate differential started to go above the critical thresholds within 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 already gone above the critical threshold before that time.

Nevertheless, we do observe two false alarms in our sample. They occurred in the Philippines in 1990-92 and Mexico in 1999. One explanation is that the Philippine government has adopted a tight monetary policy, aggressively cut government spending and raised indirect taxes during this period to reduce the downward pressure in the foreign exchange market. Some analysts argue that this effectively steered the
<table>
<thead>
<tr>
<th>Countries</th>
<th>Crisis Episodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Argentina</td>
<td>2001Q4</td>
</tr>
<tr>
<td>2. Brazil</td>
<td>1998Q3-1999Q1, 2000Q2</td>
</tr>
<tr>
<td>3. Chile</td>
<td>1990Q4</td>
</tr>
<tr>
<td>5. Colombia</td>
<td>None</td>
</tr>
<tr>
<td>6. Hong Kong</td>
<td>None</td>
</tr>
<tr>
<td>7. Indonesia</td>
<td>1997Q3-1998Q2</td>
</tr>
<tr>
<td>8. S. Korea</td>
<td>1997Q4</td>
</tr>
<tr>
<td>10. Mexico</td>
<td>1994Q4</td>
</tr>
<tr>
<td>11. Philippines</td>
<td>1984Q1, 1997Q3</td>
</tr>
<tr>
<td>12. Singapore</td>
<td>1997Q3-Q4, 1998Q2</td>
</tr>
<tr>
<td>13. Taiwan</td>
<td>1997Q4</td>
</tr>
<tr>
<td>14. Thailand</td>
<td>1981Q3, 1997Q3-Q4</td>
</tr>
<tr>
<td>15. Uruguay</td>
<td>1994Q3-1995Q2</td>
</tr>
<tr>
<td>16. Venezuela</td>
<td>1994Q2</td>
</tr>
</tbody>
</table>

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

Economy away from the currency crisis despite the high values of the threshold variables (Bautista, 2000). For the case of Mexico, it was in the middle of a capital liberalization process during 1999, which significantly raised the amount of capital inflow into the country. This helped to lower the downward pressure in the foreign exchange market.
8 Conclusion

Threshold models capture nonlinear properties of a regression model, and thus have numerous applications in various academic disciplines. However, conventional threshold models only focus on the one threshold variable case and hence they have limited applications when two or more threshold variables are called for. So far, nothing is known about the distribution theory of the threshold estimators when there is more than one threshold variable. This paper studies a new kind of threshold model which has multiple threshold variables. We have derived the asymptotic properties of the OLS estimators of such a model. The results allow researchers to conduct estimation and inference in the existence of multiple threshold variables. Experimental evidence for our theory is provided.

We apply our model to the study of currency crises in 16 countries. We find overwhelming evidence that there are threshold effects generated by two threshold variables, the ratio of short-term external liabilities to reserves and the lending rate differential. The findings are consistent with the implications from the currency crisis models. The major contribution of our empirical study is that we provide clear estimates of the joint threshold values, which can be used by governments as guidelines in the regulation of short-term external borrowing and interest rate differentials.

The threshold models are highly useful but many interesting applications have not yet been explored. For instance, one may use the threshold model to estimate the critical age at which it is dangerous for a woman to get pregnant, to estimate the threshold temperature at which the gender of a reptile is determined, to estimate the year of schooling at which the income differential between people with low education and high education is maximized, and to estimate the tax rate over which a fiscal policy becomes ineffective. With the results obtained in this paper, the applicability of threshold models can be widely extended in the future.

References


Appendix A1: Asymptotic behavior of the OLS estimators
and \( \frac{1}{T} S_T (\gamma_1, \gamma_2) \) when \( x_t = 1 \)

Note that

\[
\hat{\beta}_1 (\gamma_1, \gamma_2) = \frac{\sum_{t=1}^{T} y_t (1 - \Psi_t (\gamma))}{\sum_{t=1}^{T} (1 - \Psi_t (\gamma))}
= \frac{\sum_{t=1}^{T} (1 - \Psi_t (\gamma))}{\sum_{t=1}^{T} (1 - \Psi_t (\gamma))}
= \beta_1 + \delta \frac{\sum_{t=1}^{T} \Psi_t (\gamma) (1 - \Psi_t (\gamma))}{\sum_{t=1}^{T} (1 - \Psi_t (\gamma))}
= \beta_1 + \delta \times \frac{\sum_{t=1}^{T} [I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) - I (z_{1t} > \max \{ \gamma_1, \gamma_2 \}, z_{2t} > \max \{ \gamma_1, \gamma_2 \})]}{\sum_{t=1}^{T} (1 - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2))}
+ o_p (1).
\]

Similarly, we have

\[
\hat{\beta}_2 (\gamma_1, \gamma_2) = \frac{\sum_{t=1}^{T} y_t \Psi_t (\gamma)}{\sum_{t=1}^{T} \Psi_t (\gamma)}
= \beta_2 - \delta \frac{\sum_{t=1}^{T} (1 - \Psi_t (\gamma)) \Psi_t (\gamma)}{\sum_{t=1}^{T} \Psi_t (\gamma)}
= \beta_2 - \delta \times \frac{\sum_{t=1}^{T} [I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) - I (z_{1t} > \max \{ \gamma_1, \gamma_2 \}, z_{2t} > \max \{ \gamma_1, \gamma_2 \})]}{\sum_{t=1}^{T} (1 - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2))}
+ o_p (1).
\]

\[
\frac{1}{T} S_T (\gamma_1, \gamma_2) = \sum_{t=1}^{T} \left( y_t - \hat{\beta}_1 - \hat{\beta}_2 - \hat{\Psi}_t (\gamma) \right)^2
= \frac{1}{T} \sum_{t=1}^{T} \left( \beta_1 + \delta \Psi_t (\gamma) + \varepsilon_t - \beta_1 - \hat{\Psi}_t (\gamma) \right)^2
= \frac{1}{T} \sum_{t=1}^{T} \left( \hat{\beta}_1 - \beta_1 \right)^2 + \delta \Psi_t (\gamma) - \hat{\Psi}_t (\gamma) \right)^2 + \frac{1}{T} \sum_{t=1}^{T} \varepsilon_t^2 + o_p (1).
\]

We discuss four cases:

**Case 1:** \( \gamma_1 \leq \gamma_0^1, \gamma_2 \leq \gamma_0^2 \)

\[
\hat{\beta}_1 (\gamma_1, \gamma_2) = \beta_1 + o_p (1),
\]

42
\[ \hat{\beta}_2 (\gamma_1, \gamma_2) = \beta_2 - \delta \left( 1 - \frac{\sum_{t=1}^{T} I (z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0)}{\sum_{t=1}^{T} I (z_{1t} > \gamma_1, z_{2t} > \gamma_2)} \right) + o_p (1) \]
\[ = \beta_2 - \delta \left( 1 - \frac{F (\gamma_1^0, \gamma_2^0)}{F (\gamma_1, \gamma_2)} \right) + o_p (1). \]

Note that

\[ \hat{\delta} = \hat{\beta}_2 (\gamma_1, \gamma_2) - \hat{\beta}_1 (\gamma_1, \gamma_2) = \delta \frac{F (\gamma_1^0, \gamma_2^0)}{F (\gamma_1, \gamma_2)} + o_p (1). \]

\[ \frac{1}{T} S_T \left( \gamma_1, \gamma_2 \right) \]
\[ = \delta^2 \frac{1}{T} \sum_{t=1}^{T} \left( \Psi_t (\gamma^0) - \frac{F (\gamma_1^0, \gamma_2^0)}{F (\gamma_1, \gamma_2)} \Psi_t (\gamma) \right)^2 + \sigma^2 + o_p (1) \]
\[ = \delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t (\gamma^0) + \delta^2 \left( \frac{F (\gamma_1^0, \gamma_2^0)}{F (\gamma_1, \gamma_2)} \right)^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t (\gamma) \]
\[ - 2 \delta^2 \frac{F (\gamma_1^0, \gamma_2^0)}{F (\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^{T} \Psi_t (\gamma^0) \Psi_t (\gamma) + \sigma^2 + o_p (1) \]
\[ \overset{p}{\longrightarrow} g (\gamma_1, \gamma_2), \]
where

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

\[ \frac{\partial g (\gamma_1, \gamma_2)}{\partial \gamma_1} = \delta^2 \left( \frac{F (\gamma_1^0, \gamma_2^0)}{F (\gamma_1, \gamma_2)} \right)^2 F_{1\gamma} \leq 0, \]

\[ \frac{\partial g (\gamma_1, \gamma_2)}{\partial \gamma_2} = \delta^2 \left( \frac{F (\gamma_1^0, \gamma_2^0)}{F (\gamma_1, \gamma_2)} \right)^2 F_{2\gamma} \leq 0. \]

**Case 2:** \( \gamma_1 > \gamma_1^0, \gamma_2 \leq \gamma_2^0 \)

\[ \hat{\beta}_1 (\gamma_1, \gamma_2) = \beta_1 + \delta \frac{\sum_{t=1}^{T} I (\gamma_1^0 < z_{1t} \leq \gamma_1, z_{2t} > \gamma_2^0)}{\sum_{t=1}^{T} (1 - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2))} + o_p (1) \]
\[ \overset{p}{\longrightarrow} \beta_1 + \delta \frac{F (\gamma_1^0, \gamma_2^0) - F (\gamma_1^0, \gamma_2^0)}{1 - F (\gamma_1, \gamma_2)}, \]

43
\begin{align*}
\hat{\beta}_2 (\gamma_1, \gamma_2) &= \beta_2 - \delta \frac{\sum_{t=1}^{T} [ I (z_{1t} > \gamma_1, z_{2t} > \gamma_2) - I (z_{1t} > \gamma_1, z_{2t} > \gamma_0) ]}{\sum_{t=1}^{T} I (z_{1t} > \gamma_1, z_{2t} > \gamma_2)} + o_p (1) \\
&= p_{\gamma_2} - \delta \left( 1 - \frac{F (\gamma_1, \gamma_2)}{F (\gamma_1, \gamma_2)} \right) .
\end{align*}

Note that

\begin{align*}
\hat{\delta} &= \frac{F (\gamma_1, \gamma_2) - F (\gamma_1, \gamma_2)}{(1 - F (\gamma_1, \gamma_2)) F (\gamma_1, \gamma_2)} \\
&= \frac{1}{T} \sum_{t=1}^{T} \left( \frac{F (\gamma_1, \gamma_2) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} + \Psi_t (\gamma) \right)^2 + \sigma^2 + o_p (1)
\end{align*}

\begin{align*}
&= \delta^2 \left( \frac{F (\gamma_1, \gamma_2) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} \right)^2 \frac{1}{T} \sum_{t=1}^{T} (1 - \Psi_t (\gamma)) + \delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t (\gamma) \\
&+ \delta^2 \left( \frac{F (\gamma_1, \gamma_2)}{F (\gamma_1, \gamma_2)} \right)^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t (\gamma) - 2 \frac{F (\gamma_1, \gamma_2)}{F (\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^{T} \Psi_t (\gamma) \\
&= \delta^2 \left( \frac{F (\gamma_1, \gamma_2) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} \right)^2 \frac{1}{T} \sum_{t=1}^{T} (1 - \Psi_t (\gamma)) + \delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t (\gamma) + \sigma^2 + o_p (1)
\end{align*}

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

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
\[
a = F(\gamma_1^0, \gamma_2^0),
b = F(\gamma_1, \gamma_2),
c = F(\gamma_1^0, \gamma_2^0).
\]

Using the fact that \(c > a, b > a\) and \(b + c - a < 1\), we have
\[
g(\gamma_1, \gamma_2) > g(\gamma_1^0, \gamma_2^0) = \sigma^2.
\]

**Case 3:** \(\gamma_1 \leq \gamma_1^0, \gamma_2 > \gamma_2^0\)
\[
\hat{\beta}_1(\gamma_1, \gamma_2) = \beta_1 + \delta \frac{\sum_{t=1}^{T} I(z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0) - I(z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0)}{\sum_{t=1}^{T} (1 - I(z_{1t} > \gamma_1, z_{2t} > \gamma_2))} + o_p(1)
\]
\[
\rightarrow \beta_1 + \delta \frac{F(\gamma_1^0, \gamma_2^0) - F(\gamma_1, \gamma_2)}{1 - F(\gamma_1, \gamma_2)}.
\]

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

Note that
\[
\hat{\delta} \rightarrow \delta \frac{F(\gamma_1^0, \gamma_2^0) - F(\gamma_1, \gamma_2) F(\gamma_1^0, \gamma_2^0)}{F(\gamma_1, \gamma_2) (1 - F(\gamma_1, \gamma_2))},
\]
\[
\frac{1}{T} S_T(\gamma_1, \gamma_2)
\]
\[
= \frac{\delta^2}{T} \sum_{t=1}^{T} \left( \frac{F(\gamma_1^0, \gamma_2^0) - F(\gamma_1, \gamma_2)}{1 - F(\gamma_1, \gamma_2)} + \Psi_t(\gamma^0) - \frac{F(\gamma_1^0, \gamma_2) - F(\gamma_1, \gamma_2) F(\gamma_1^0, \gamma_2^0) \Psi_t(\gamma)}{F(\gamma_1, \gamma_2) (1 - F(\gamma_1, \gamma_2))} \right)^2.
\]

45
\[
+ \sigma^2 + o_p(1)
\]
\[
= \delta^2 \frac{1}{T} \sum_{t=1}^{T} \left( \frac{F(\gamma_0, \gamma_2) - F(\gamma_1, \gamma_2)}{1 - F(\gamma_1, \gamma_2)} (1 - \Psi_t(\gamma)) + \Psi_t(\gamma) - \frac{F(\gamma_1, \gamma_2)}{F(\gamma_1, \gamma_2)} \Psi_t(\gamma) \right)^2
\]
\[
+ \sigma^2 + o_p(1)
\]
\[
= \delta^2 \left( \frac{F(\gamma_0, \gamma_2) - F(\gamma_1, \gamma_2)}{1 - F(\gamma_1, \gamma_2)} \right)^2 \frac{1}{T} \sum_{t=1}^{T} (1 - \Psi_t(\gamma)) + \delta^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t(\gamma)
\]
\[
+ \delta^2 \left( \frac{F(\gamma_0, \gamma_2)}{F(\gamma_1, \gamma_2)} \right)^2 \frac{1}{T} \sum_{t=1}^{T} \Psi_t(\gamma) + 2 \delta^2 \frac{F(\gamma_0, \gamma_2) - F(\gamma_1, \gamma_2)}{1 - F(\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^{T} (1 - \Psi_t(\gamma)) \Psi_t(\gamma)
\]
\[
- 2 \delta^2 \frac{F(\gamma_0, \gamma_2)}{F(\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^{T} \Psi_t(\gamma) \Psi_t(\gamma) + \sigma^2 + o_p(1)
\]
\[
= \delta^2 \left( \frac{F(\gamma_1, \gamma_2) - F(\gamma_0, \gamma_2)}{1 - F(\gamma_1, \gamma_2)} \right)^2 (1 - F(\gamma_1, \gamma_2)) + \delta^2 F(\gamma_0, \gamma_2)
\]
\[
+ \delta^2 \left( \frac{F(\gamma_0, \gamma_2)}{F(\gamma_1, \gamma_2)} \right)^2 F(\gamma_1, \gamma_2)
\]
\[
+ 2 \delta^2 \frac{F(\gamma_0, \gamma_2) - F(\gamma_1, \gamma_2)}{1 - F(\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^{T} (I(z_1 > \gamma_1, z_2 > \gamma_2) - I(z_1 > \gamma_1, z_2 > \gamma_2))
\]
\[
- 2 \delta^2 \frac{F(\gamma_0, \gamma_2)}{F(\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^{T} I(z_1 > \gamma_1, z_2 > \gamma_2) + \sigma^2 + o_p(1)
\]
\[
= \delta^2 \left( \frac{F(\gamma_1, \gamma_2) - F(\gamma_0, \gamma_2)}{1 - F(\gamma_1, \gamma_2)} \right)^2 + \delta^2 F(\gamma_0, \gamma_2) + \delta^2 \frac{F(\gamma_0, \gamma_2)^2}{F(\gamma_1, \gamma_2)}
\]
\[
- 2 \delta^2 F(\gamma_0, \gamma_2) \frac{1}{1 - F(\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^{T} F(\gamma_0, \gamma_2) + \sigma^2 + o_p(1)
\]
\[
\overset{p}{\rightarrow} g(\gamma_1, \gamma_2)
\]

where

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

Rewrite

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

where

\[
d = F(\gamma_1, \gamma_2).
\]

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

\[
g(\gamma_1, \gamma_2) > g(\gamma_1, \gamma_2) = \sigma^2.
\]
Case 4: $\gamma_1 > \gamma_0^0, \gamma_2 > \gamma_0^0$

$$\hat{\beta}_1 (\gamma_1, \gamma_2) = \beta_1 + \delta \frac{\sum_{t=1}^T [I (z_{1t} > \gamma_1^0, z_{2t} > \gamma_2^0) - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2)]}{\sum_{t=1}^T (1 - I (z_{1t} > \gamma_1, z_{2t} > \gamma_2))} + o_p (1)$$

$$= \beta_1 + \delta \frac{F (\gamma_1^0, \gamma_2^0) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} + o_p (1)$$

$$\hat{\beta}_2 (\gamma_1, \gamma_2) = \beta_2 + o_p (1)$$

Note that

$$\delta \rightarrow \delta \frac{1 - F (\gamma_1^0, \gamma_2^0)}{1 - F (\gamma_1, \gamma_2)},$$

$$\Psi_t (\gamma) \Psi_t (\gamma_0^0) = \Psi_t (\gamma).$$

$$\frac{1}{T} S_T (\gamma_1, \gamma_2)$$

$$= \delta^2 \frac{1}{T} \sum_{t=1}^T \left( - \frac{F (\gamma_1^0, \gamma_2^0) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} + \Psi_t (\gamma_0^0) - \frac{1 - F (\gamma_1^0, \gamma_2^0)}{1 - F (\gamma_1, \gamma_2)} \Psi_t (\gamma) \right)^2$$

$$+ \sigma^2 + o_p (1)$$

$$= \delta^2 \frac{1}{T} \left( - \frac{F (\gamma_1^0, \gamma_2^0) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} \right)^2 \frac{1}{T} \sum_{t=1}^T (1 - \Psi_t (\gamma))$$

$$+ \sigma^2 + o_p (1)$$

$$= \delta^2 \frac{1}{T} \sum_{t=1}^T \Psi_t (\gamma_0^0) + \delta^2 \frac{1}{T} \sum_{t=1}^T \Psi_t (\gamma)$$

$$- 2 \delta^2 \frac{F (\gamma_1^0, \gamma_2^0) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} \frac{1}{T} \sum_{t=1}^T (1 - \Psi_t (\gamma)) \Psi_t (\gamma_0^0)$$

$$- 2 \delta^2 \frac{1}{T} \sum_{t=1}^T \Psi_t (\gamma_0^0) \Psi_t (\gamma) + \sigma^2 + o_p (1)$$

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

$$+ \delta^2 F (\gamma_1^0, \gamma_2^0) + \delta^2 F (\gamma_1, \gamma_2)$$

$$- 2 \delta^2 \frac{F (\gamma_1^0, \gamma_2^0) - F (\gamma_1, \gamma_2)}{1 - F (\gamma_1, \gamma_2)} (F (\gamma_1^0, \gamma_2^0) - F (\gamma_1, \gamma_2))$$

$$- 2 \delta^2 \frac{1}{T} \sum_{t=1}^T \Psi_t (\gamma) + \sigma^2 + o_p (1)$$

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

$$\rightarrow g (\gamma_1, \gamma_2)$$

47
where

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

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

\[ \frac{\partial g(\gamma_1, \gamma_2)}{\partial \gamma_2} = -\delta^2 \left( \frac{1 - \mathcal{F}(\gamma_1^0, \gamma_2^0)}{1 - \mathcal{F}(\gamma_1, \gamma_2)} \right)^2 \mathcal{F}_{2\gamma} > 0. \]
Appendix A2: Asymptotic behavior of the OLS estimators
and $\frac{1}{T}ST (\gamma_1, \gamma_2)$ when $x_t = 1$ and threshold variables are independent

When threshold variables are independent, we have

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

Similarly, we have

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

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

$$
\hat{\beta}_1 (\gamma_1, \gamma_2) = \beta_1 + o_p (1),
$$

$$
\hat{\beta}_2 (\gamma_1, \gamma_2) = \beta_2 - \delta \left(1 - \frac{F_1 (\gamma_1^0) F_2 (\gamma_2^0)}{F_1 (\gamma_1) F_2 (\gamma_2)}\right) + o_p (1),
$$

$$
\hat{\delta} = \hat{\beta}_2 (\gamma_1, \gamma_2) - \hat{\beta}_1 (\gamma_1, \gamma_2) = \delta \frac{F_1 (\gamma_1^0) F_2 (\gamma_2^0)}{F_1 (\gamma_1) F_2 (\gamma_2)},
$$

$$
g (\gamma_1, \gamma_2) = \sigma^2 + \delta^2 F_1 (\gamma_1^0) F_2 (\gamma_2^0) \left(1 - \frac{F_1 (\gamma_1^0) F_2 (\gamma_2^0)}{F_1 (\gamma_1) F_2 (\gamma_2)}\right),
$$

$$
\frac{\partial g (\gamma_1, \gamma_2)}{\partial \gamma_1} = -\delta^2 \frac{F_1 (\gamma_1^0)^2 F_2 (\gamma_2^0)}{F_1 (\gamma_1) F_2 (\gamma_2)} H_1 (\gamma_1) \leq 0,
$$

$$
\frac{\partial g (\gamma_1, \gamma_2)}{\partial \gamma_2} = -\delta^2 \frac{F_1 (\gamma_1^0)^2 F_2 (\gamma_2^0)}{F_1 (\gamma_1) F_2 (\gamma_2)} H_2 (\gamma_2) \leq 0,
$$

where $H_1 (\gamma_1)$ and $H_2 (\gamma_2)$ are the hazard functions of $z_1$ and $z_2$ respectively.
Case 2: $\gamma_1 > \gamma_0^0, \gamma_2 \leq \gamma_2^0$

$$\hat{\beta}_1 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_1 + \delta \frac{(F_1(\gamma_0^0) - F_1(\gamma_1)) F_2(\gamma_2^0)}{1 - F_1(\gamma_1) F_2(\gamma_2)},$$

$$\hat{\beta}_2 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_2 - \delta \left(1 - \frac{F_2(\gamma_2^0)}{F_2(\gamma_2)}\right),$$

$$\hat{\delta} \xrightarrow{p} \frac{F_2(\gamma_2^0) [1 - F_1(\gamma_1) F_2(\gamma_0^0)]}{F_2(\gamma_2) [1 - F_1(\gamma_1) F_2(\gamma_2)]}.$$

$$g (\gamma_1, \gamma_2) = \sigma^2 + \sigma^2 F_2(\gamma_2^0) \left[ -\frac{(F_1(\gamma_0^0) - F_1(\gamma_1))^2}{1 - F_1(\gamma_1) F_2(\gamma_2)} + \frac{F_1(\gamma_0^0)}{F_2(\gamma_2)} - \frac{F_1(\gamma_1)}{F_2(\gamma_2)} \right],$$

$$\frac{\partial g (\gamma_1, \gamma_2)}{\partial \gamma_1} = \sigma^2 F_2(\gamma_2^0) \left( \frac{F_1(\gamma_0^0) - F_1(\gamma_1)}{1 - F_1(\gamma_1) F_2(\gamma_2)} - \frac{1}{F_2(\gamma_2)} \right)^2 F_2(\gamma_2) f_1(\gamma_1) > 0,$$

$$\frac{\partial g (\gamma_1, \gamma_2)}{\partial \gamma_2} = -\sigma^2 F_2(\gamma_2^0) \left[ \frac{F_1(\gamma_0^0) - F_1(\gamma_1)}{1 - F_1(\gamma_1) F_2(\gamma_2)} + \frac{1}{F_2(\gamma_2)} \right] \times$$

$$\frac{1 - F_1(\gamma_1) F_2(\gamma_2)}{1 - F_1(\gamma_1) F_2(\gamma_2)} \frac{F_1(\gamma_1) H_2(\gamma_2)}{1 - F_1(\gamma_1) F_2(\gamma_2)} < 0.$$  

Case 3: $\gamma_1 \leq \gamma_0^0, \gamma_2 > \gamma_2^0$

$$\hat{\beta}_1 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_1 + \delta \frac{F_1(\gamma_1) (F_2(\gamma_2^0) - F_2(\gamma_2))}{1 - F_1(\gamma_1) F_2(\gamma_2)},$$

$$\hat{\beta}_2 (\gamma_1, \gamma_2) \xrightarrow{p} \beta_2 - \delta \left(1 - \frac{F_1(\gamma_0^0)}{F_1(\gamma_1)}\right),$$

$$\hat{\delta} \xrightarrow{p} \delta \frac{F_1(\gamma_1) [1 - F_1(\gamma_1) F_2(\gamma_2)]}{F_1(\gamma_1) [1 - F_1(\gamma_1) F_2(\gamma_2)]}.$$

$$g (\gamma_1, \gamma_2) = \sigma^2 + \sigma^2 F_1(\gamma_0^0) \left[ -\frac{(F_2(\gamma_2^0) - F_2(\gamma_2))^2}{1 - F_1(\gamma_1) F_2(\gamma_2)} + \frac{F_2(\gamma_2^0)}{F_1(\gamma_0^0)} - \frac{F_2(\gamma_2)}{F_1(\gamma_1)} \right],$$

50
\[
\frac{\partial g(\gamma_1, \gamma_2)}{\partial \gamma_1} = -\delta^2 F_1(\gamma_1)^2 \left( \frac{F_2(\gamma_2) - F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)} + \frac{1}{F_1(\gamma_1)} \right) \\
\times \frac{1 - F_1(\gamma_1)F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)} F_2(\gamma_2) H_1(\gamma_1)
\]

\[
< 0,
\]

\[
\frac{\partial g(\gamma_1, \gamma_2)}{\partial \gamma_2} = \delta^2 F_1(\gamma_1)^2 \left( \frac{F_2(\gamma_2) - F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)} - \frac{1}{F_1(\gamma_1)} \right)^2 F_1(\gamma_1) f_2(\gamma_2) > 0.
\]

**Case 4:** \( \gamma_1 > \gamma_0, \gamma_2 > \gamma_0 \)

\[
\begin{align*}
\hat{\beta}_1(\gamma_1, \gamma_2) &\xrightarrow{p} \beta_1 + \delta \frac{F_1(\gamma_1)F_2(\gamma_2) - F_1(\gamma_1)F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)}, \\
\hat{\beta}_2(\gamma_1, \gamma_2) &= \beta_2 + o_p(1), \\
\hat{\delta} &\xrightarrow{p} \delta \frac{1 - F_1(\gamma_1)F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)},
\end{align*}
\]

\[
g(\gamma_1, \gamma_2) = \sigma^2 + \delta^2 \left( 1 - F_1(\gamma_1)F_2(\gamma_2) \right) \left( 1 - \frac{F_1(\gamma_1)F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)} \right),
\]

\[
\frac{\partial g(\gamma_1, \gamma_2)}{\partial \gamma_1} = \delta^2 \left( \frac{1 - F_1(\gamma_1)F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)} \right)^2 F_2(\gamma_2) f_1(\gamma_1) > 0,
\]

\[
\frac{\partial g(\gamma_1, \gamma_2)}{\partial \gamma_2} = \delta^2 \left( \frac{1 - F_1(\gamma_1)F_2(\gamma_2)}{1 - F_1(\gamma_1)F_2(\gamma_2)} \right)^2 F_1(\gamma_1) f_2(\gamma_2) > 0.
\]
Appendix A3: Distribution of $\hat{\gamma}_1$ and $\hat{\gamma}_2$ when $x_t = 1$

Define

$$(\hat{\gamma}_1, \hat{\gamma}_2) = \arg\min S_T (\gamma_1, \gamma_2) = \arg\min [S_T (\gamma_1, \gamma_2) - S_T (\gamma_1^0, \gamma_2^0)].$$

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

$$T (\gamma_1, \gamma_2) = T (\gamma_1^0, \gamma_2^0) + (\gamma_1 - \gamma_1^0) T_1^0 + (\gamma_2 - \gamma_2^0) T_2^0 + o(1),$$

$$T (\gamma_1^0, \gamma_2^0) = T (\gamma_1^0, \gamma_2^0) + (\gamma_1 - \gamma_1^0) T_1^0 + o(1),$$

$$T (\gamma_1^0, \gamma_2^0) = T (\gamma_1^0, \gamma_2^0) + (\gamma_2 - \gamma_2^0) T_2^0 + o(1).$$

Thus,

$$S_T (\gamma_1, \gamma_2) - S_T (\gamma_1^0, \gamma_2^0)$$

$$= \sum_{t=1}^{T} \left( \delta \Psi_t (\gamma) + \varepsilon_t - \delta \Psi_t (\gamma) \right)^2 - \sum_{t=1}^{T} \left( \delta \Psi_t (\gamma) + \varepsilon_t - \delta \Psi_t (\gamma) \right)^2$$

$$+ o_p (1)$$

$$= \delta \sum_{t=1}^{T} \left( \delta \Psi_t (\gamma) + 2 \varepsilon_t - \delta \Psi_t (\gamma) \right) \left( \Psi_t (\gamma) - \Psi_t (\gamma) \right) + o_p (1)$$

$$= \delta^2 \sum_{t=1}^{T} \left( \Psi_t (\gamma) - \Psi_t (\gamma) \right)^2 + 2 \delta \sum_{t=1}^{T} \varepsilon_t \left( \Psi_t (\gamma) - \Psi_t (\gamma) \right) + o_p (1).$$

In the neighborhood of the true thresholds, where $\gamma_1 = \gamma_1^0 + \frac{\nu_2}{T^{1-2\alpha}}$, $\gamma_2 = \gamma_2^0 + \frac{\nu_2}{T^{1-2\alpha}}$, all estimators can be approximated by the true values, so we have the following:

**Case 1:** $\nu_1 \leq 0$, $\nu_2 \leq 0$

$$S_T (\gamma_1, \gamma_2) - S_T (\gamma_1^0, \gamma_2^0)$$

$$= T^{-2\alpha} T \left( F (\gamma_1, \gamma_2) - F (\gamma_1^0, \gamma_2^0) \right)$$

$$+ 2 T^{-\alpha} \sum_{(x_1 < x_1^0 \text{ and } x_2^0 < z_2)} \left( \gamma_1 - \gamma_1^0 \right) E_1^0 + \left( \gamma_2 - \gamma_2^0 \right) E_2^0$$

$$+ o_p (1)$$

$$= T^{-2\alpha} \left( (\gamma_1 - \gamma_1^0) F_1^0 + (\gamma_2 - \gamma_2^0) F_2^0 \right)$$

$$+ 2 T^{-\alpha} \sum_{(x_1 < x_1^0 \text{ and } x_2^0 < z_2)} \left( \gamma_1 - \gamma_1^0 \right) E_1^0 + \left( \gamma_2 - \gamma_2^0 \right) E_2^0$$

$$= \frac{d}{v_1} F_1^0 + \frac{d}{v_2} F_2^0$$

$$+ 2 T^{-\alpha} \sum_{t=1}^{T} \left( \gamma_1 < x_1^0 \text{ and } x_2^0 < z_2 \right) E_1^0 + \left( \gamma_2 < x_2^0 \text{ and } \gamma_1^0 < x_1 \right) E_2^0$$

$$= \frac{d}{v_1} F_1^0 + 2 T^{-\alpha} \sum_{t=1}^{T} \left( \gamma_1 < x_1^0 \text{ and } x_2^0 < z_2 \right) E_1^0 + \left( \gamma_2 < x_2^0 \text{ and } \gamma_1^0 < x_1 \right) E_2^0$$

$$= \frac{d}{v_1} F_1^0 + 2 T^{-\alpha} \sum_{t=1}^{T} \left( \gamma_1 < x_1^0 \text{ and } x_2^0 < z_2 \right) E_1^0 + \left( \gamma_2 < x_2^0 \text{ and } \gamma_1^0 < x_1 \right) E_2^0,$$

where $E_1^0$ and $E_2^0$ are independent.
Case 2: \( v_1 > 0, v_2 \leq 0 \)
\[
S_T (\gamma_1, \gamma_2) - S_T (\gamma_1^0, \gamma_2^0) = T^{-2\alpha} T \left( \overline{F} (\gamma_1, \gamma_2) - \overline{F} (\gamma_1^0, \gamma_2^0) \right) + 2T^{-\alpha} \sum_{\gamma_1^0 < \gamma_2 < \gamma_1 \text{ or } \gamma_1^0 < \gamma_2} \varepsilon_t + o_p (1)
\]
\[
= T^{-2\alpha} T \left( (\gamma_2 - \gamma_1^0) \overline{T}_2^0 - (\gamma_1 - \gamma_1^0) \overline{T}_1^0 \right) + 2T^{-\alpha} \sum_{\gamma_1^0 < \gamma_2 < \gamma_1} \varepsilon_t + 2T^{-\alpha} \sum_{\gamma_2 < \gamma_1^0} \varepsilon_t + o_p (1) \]
\[
d = v_2 \overline{T}_2^0 - v_1 \overline{T}_1^0 + 2T^{-\alpha} \sum_{t=1}^{T} \left( T \Pr (\gamma_1^0 < z_{1t} < \gamma_1) \right) \varepsilon_t + 2T^{-\alpha} \sum_{t=1}^{T} \varepsilon_t + o_p (1)
\]
\[
d = -v_1 \overline{T}_1^0 + 2T^{-\alpha} \sum_{t=1}^{T} -v_1 T^{2\alpha} \overline{T}_1^0 \varepsilon_t + v_2 \overline{T}_2^0 + 2T^{-\alpha} \sum_{t=1}^{T} v_2 T^{2\alpha} \overline{T}_2^0 \varepsilon_t.
\]

Case 3: \( v_1 \leq 0, v_2 > 0 \)
\[
S_T (\gamma_1, \gamma_2) - S_T (\gamma_1^0, \gamma_2^0) = T^{-2\alpha} T \left( \overline{F} (\gamma_1, \gamma_2) - \overline{F} (\gamma_1^0, \gamma_2^0) \right) + 2T^{-\alpha} \sum_{\gamma_1^0 < \gamma_2 < \gamma_1} \varepsilon_t + o_p (T^{-2\alpha}) \]
\[
d = v_1 \overline{T}_1^0 - v_2 \overline{T}_2^0 + 2T^{-\alpha} \sum_{t=1}^{T} \left( T \Pr (\gamma_1^0 < z_{1t} < \gamma_1) \right) \varepsilon_t + 2T^{-\alpha} \sum_{t=1}^{T} \varepsilon_t + o_p (T^{-2\alpha}) \]
\[
d = v_1 \overline{T}_1^0 + 2T^{-\alpha} \sum_{t=1}^{T} -v_1 T^{2\alpha} \overline{T}_1^0 \varepsilon_t - v_2 \overline{T}_2^0 + 2T^{-\alpha} \sum_{t=1}^{T} v_2 T^{2\alpha} \overline{T}_2^0 \varepsilon_t.
\]

Similarly, we have

Case 4: \( v_1 > 0, v_2 > 0 \)
\[
S_T (\gamma_1, \gamma_2) - S_T (\gamma_1^0, \gamma_2^0) = -v_1 \overline{T}_1^0 + 2T^{-\alpha} \sum_{t=1}^{T} -v_1 T^{2\alpha} \overline{T}_1^0 \varepsilon_t - v_2 \overline{T}_2^0 + 2T^{-\alpha} \sum_{t=1}^{T} v_2 T^{2\alpha} \overline{T}_2^0 \varepsilon_t.
\]

Let
\[
r_1 = -\overline{T}_1^0 v_1,
\]
\[
r_2 = -\overline{T}_2^0 v_2.
\]

We have
\[-T^{1-2\alpha} \left((\hat{\gamma}_1 - \gamma^0_1) F^0_1, (\hat{\gamma}_2 - \gamma^0_2) F^0_2\right)\]

\[\xrightarrow{d} \arg \max_{-\infty < r_1 < \infty, -\infty < r_2 < \infty} \left(-\frac{1}{2} |r_1| + W_1(r_1) - \frac{1}{2} |r_2| + W_2(r_2)\right).\]

When \(z_1\) and \(z_2\) are independent, we have

\[\mathcal{F}_1^0 = -f_1(\gamma^0_1) \mathcal{F}_2(\gamma^0_2)\]

and

\[\mathcal{F}_2^0 = -f_2(\gamma^0_2) \mathcal{F}_1(\gamma^0_1).\]

Thus,

\[T^{1-2\alpha} (f_1(\gamma^0_1) \mathcal{F}_2(\gamma^0_2) (\hat{\gamma}_1 - \gamma^0_1), f_2(\gamma^0_2) \mathcal{F}_1(\gamma^0_1) (\hat{\gamma}_2 - \gamma^0_2))\]

\[\xrightarrow{d} \arg \max_{(r_1, r_2) \in \mathbb{R}^2} \left(-\frac{1}{2} |r_1| + W_1(r_1) - \frac{1}{2} |r_2| + W_2(r_2)\right).\]
Appendix B: Asymptotic behavior of $S_T (\gamma_1, \gamma_2)$ for vector $x_t$.

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

$$S_T (\gamma) - \varepsilon' \varepsilon = Y'(I - P_\gamma) Y - \varepsilon' \varepsilon$$

$$= (X\beta_1 + X_0 \delta + \varepsilon)'(I - P_\gamma) (X\beta_1 + X_0 \delta + \varepsilon) - \varepsilon' \varepsilon$$

$$= -\varepsilon' P_\gamma \varepsilon + \beta_1' X'(I - P_\gamma) X \beta_1 + 2\delta' X_0' (I - P_\gamma) \varepsilon + \delta' X_0' (I - P_\gamma) X_0 \delta + 2\beta_1' X'(I - P_\gamma) X_0 \delta$$

$$= -\varepsilon' P_\gamma \varepsilon + 2\delta' X_0' (I - P_\gamma) \varepsilon + \delta' X_0' (I - P_\gamma) X_0 \delta.$$

Let $\delta = \frac{\varepsilon}{T^{\alpha_2}}$.

$$\frac{1}{T^{1-2\alpha_2}} (S_T (\gamma) - \varepsilon' \varepsilon) = \frac{1}{T} c' (X_0' (I - P_\gamma) X_\gamma) c + o_p (1).$$

The projection $P_\gamma$ can be written as the projection onto $[X - X_{\gamma_1}, X_{\gamma_2}]$ where $X - X_\gamma$ is a matrix whose $t^{th}$ row is $x_t' (1 - \Psi_t (\gamma))$. Observe that $(X - X_{\gamma_1})' X_{\gamma_2} = 0$.

$$P_\gamma = (X - X_{\gamma_1}, X_{\gamma_2}) \left( \begin{array}{cc} (X - X_{\gamma_1})' (X - X_{\gamma_2}) & 0 \\ 0 & X_{\gamma_1}' X_{\gamma_2} \end{array} \right)^{-1} (X - X_{\gamma_1}, X_{\gamma_2})'$$

$$= (X - X_{\gamma_1}) [(X - X_{\gamma_1})' (X - X_{\gamma_1})]^{-1} (X - X_{\gamma_1})' + X_{\gamma_1} (X_{\gamma_1}' X_{\gamma_2})^{-1} X_{\gamma_2}.$$

$$X_0' P_\gamma X_0 = X_0' (X - X_{\gamma_1}) [(X - X_{\gamma_1})' (X - X_{\gamma_1})]^{-1} (X - X_{\gamma_1})' X_0 + X_0' X_{\gamma_1} (X_{\gamma_1}' X_{\gamma_2})^{-1} X_{\gamma_2} X_0.$$

We discuss four cases:

**Case 1: $\gamma_1 \leq \gamma_0, \gamma_2 \leq \gamma_0$**

In this case, we have

$$X_{\gamma_1}' X_0 = X_0' X_0;$$

$$(X - X_{\gamma_1})' X_0 = 0;$$

$$X_0' P_\gamma X_0 = X_0' X_0 (X_{\gamma_1}' X_{\gamma_2})^{-1} X_{\gamma_2} X_0.$$

Thus, we have

$$\frac{1}{T^{1-2\alpha_2}} (S_T (\gamma) - \varepsilon' \varepsilon)$$

$$= \frac{1}{T} c' X_0' (I - P_\gamma) X_0 c + o_p (1)$$

$$= c' \left( M_T (\gamma_0) - M_T (\gamma_0) M_T^{-1} (\gamma) M_T (\gamma_0) \right) c$$

$$\overset{p}{\to} c' \left( M_0 - M_0 M_0^{-1} M_0 \right) c$$

$$\equiv b_1 (\gamma).$$

As

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

55
\[ \frac{\partial}{\partial \gamma_2} \mathcal{M}(\gamma_1, \gamma_2) = D_s F_{2\gamma} \leq 0. \]
Thus,
\[ \frac{\partial}{\partial \gamma_1} b_1(\gamma) = c' \mathcal{M}_0 \mathcal{M}_1^{-1} D_s F_{1\gamma} \mathcal{M}_1^{-1} \mathcal{M}_0 c \leq 0, \]
\[ \frac{\partial}{\partial \gamma_2} b_1(\gamma) = c' \mathcal{M}_0 \mathcal{M}_1^{-1} D_s F_{2\gamma} \mathcal{M}_1^{-1} \mathcal{M}_0 c \leq 0. \]

**Case 2:** \( \gamma_1 > \gamma_0^1, \gamma_2 \leq \gamma_0^2 \) In this case, we have
\[ \frac{1}{T} X' \gamma X_0 \xrightarrow{p} \mathcal{M}(\gamma_1, \gamma_0^0); \]
\[ \frac{1}{T} (X - X_\gamma)' X_0 \xrightarrow{p} \mathcal{M}_0 - \mathcal{M}(\gamma_1, \gamma_0^0); \]
\[ \frac{1}{T} X'_0 \mathcal{P}_\gamma X_0 \xrightarrow{p} (\mathcal{M}_0 - \mathcal{M}(\gamma_1, \gamma_0^0)) (M - \mathcal{M}_\gamma)^{-1} \left( \mathcal{M}_0 - \mathcal{M}(\gamma_1, \gamma_0^0) \right) \]
\[ + \mathcal{M}(\gamma_1, \gamma_0^0) M_\gamma^{-1} \mathcal{M}(\gamma_1, \gamma_0^0). \]
Then, we have
\[ \frac{1}{T} (S_T(\gamma) - \varepsilon' \varepsilon) \xrightarrow{p} \mathcal{L}_c \left( \mathcal{M}_0 - \mathcal{M}(\gamma_1, \gamma_0^0) \right) (M - \mathcal{M}_\gamma)^{-1} \left( \mathcal{M}_0 - \mathcal{M}(\gamma_1, \gamma_0^0) \right) \]
\[ \equiv b_2(\gamma) > b_2(\gamma_0) = 0, \]
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:** \( \gamma_1 \leq \gamma_0^1, \gamma_2 > \gamma_0^2 \) In this case, we have
\[ \frac{1}{T} X' \gamma X_0 \xrightarrow{p} \mathcal{M}(\gamma_0^1, \gamma_2) \]
\[ \frac{1}{T} (X - X_\gamma)' X_0 \xrightarrow{p} \mathcal{M}_0 - \mathcal{M}(\gamma_0^1, \gamma_2); \]
\[ \frac{1}{T} X'_0 \mathcal{P}_\gamma X_0 \xrightarrow{p} (\mathcal{M}_0 - \mathcal{M}(\gamma_0^1, \gamma_2)) (M - \mathcal{M}_\gamma)^{-1} \left( \mathcal{M}_0 - \mathcal{M}(\gamma_0^1, \gamma_2) \right) \]
\[ + \mathcal{M}(\gamma_0^1, \gamma_2) M_\gamma^{-1} \mathcal{M}(\gamma_0^1, \gamma_2). \]
Then, we have
\[ \frac{1}{T} (S_T(\gamma) - \varepsilon' \varepsilon) \xrightarrow{p} \mathcal{L}_c \left( \mathcal{M}_0 - \mathcal{M}(\gamma_1, \gamma_0^0) \right) (M - \mathcal{M}_\gamma)^{-1} \left( \mathcal{M}_0 - \mathcal{M}(\gamma_1, \gamma_0^0) \right) \]
\[ \equiv b_3(\gamma) > b_3(\gamma_0) = 0, \]
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: $\gamma_1 > \gamma_0^1, \gamma_2 > \gamma_0^2$ 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

$$\frac{1}{T^{1-2\alpha}} (S_T (\gamma) - \varepsilon' \varepsilon) 
\xrightarrow{P} c' \left( M_0 - M_\gamma - (M_0 - M_\gamma) (M - M_\gamma)^{-1} (M_0 - M_\gamma) \right) c
= c' \left( M_0 - M_\gamma - (M_0 - M + M - M_\gamma) (M - M_\gamma)^{-1} (M_0 - M + M - M_\gamma) \right) c
= c' \left( M - M_0 - (M - M_0) (M - M_\gamma)^{-1} (M - M_0) \right) c
\equiv b_4 (\gamma).$$

Thus,

$$\frac{\partial}{\partial \gamma_1} b_4 (\gamma) = -c' \left( (M - M_0) (M - M_\gamma)^{-1} D_\gamma F_{1\gamma} (M - M_\gamma)^{-1} (M - M_0) \right) c > 0,$$

$$\frac{\partial}{\partial \gamma_2} b_4 (\gamma) = -c' \left( (M - M_0) (M - M_\gamma)^{-1} D_\gamma F_{2\gamma} (M - M_\gamma)^{-1} (M - M_0) \right) c > 0.$$

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 $\hat{\gamma}_1$ and $\hat{\gamma}_2$ for vector $x_t$

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

$$\hat{\beta}_1 = (X'X - X'_\gamma X_\gamma)^{-1} (X - X_\gamma)'Y'$$

$$= \beta_1 + (X'X - X'_\gamma X_\gamma)^{-1} (X'_0 X_0 - X_\gamma X_0) \delta + (X'X - X'_\gamma X_\gamma)^{-1} (X - X_\gamma)' \varepsilon$$

$$\hat{\beta}_1 - \hat{\beta}_1^0 = (X'X - X'_\gamma X_\gamma)^{-1} \left( X'_0 X_0 - X'_\gamma X_\gamma \right) \delta$$

$$+ (X'X - X'_\gamma X_\gamma)^{-1} (X - X_\gamma)' \varepsilon - (X'X - X'_\gamma X_\gamma)^{-1} (X - X_\gamma)' \varepsilon$$

$$= (X'X - X'_\gamma X_\gamma)^{-1} \left( X'_0 X_0 - X'_\gamma X_\gamma \right) \delta + (X'X - X'_\gamma X_\gamma)^{-1} (X_0 - X_\gamma)' \varepsilon$$

$$+ o_p \left( \frac{1}{T^{1-\alpha}} \right)$$

$$= O_p \left( \frac{1}{T^{1-\alpha}} \right) + O_p \left( \frac{1}{T^{1-\alpha}} \right) + o_p \left( \frac{1}{T^{1-\alpha}} \right)$$

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

$$\hat{\delta} = (X'_\gamma X_\gamma)^{-1} X'_\gamma Y - \hat{\beta}_1 ,$$

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

$$S_T (\gamma_1, \gamma_2) = \left( Y - X\hat{\beta}_1 - X_\gamma \hat{\delta} \right)' \left( Y - X\hat{\beta}_1 - X_\gamma \hat{\delta} \right)$$

$$= Y'Y - 2Y' \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} \right) + \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} \right)' \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} \right) .$$

In the neighborhood of the true thresholds, where $\gamma_1 = \gamma_1^0 + \frac{\nu_1}{T^{1-2\alpha}}$, $\gamma_2 = \gamma_2^0 + \frac{\nu_2}{T^{1-2\alpha}}$, we have:

$$S_T (\gamma_1, \gamma_2) - S_T (\gamma_1^0, \gamma_2^0)$$

$$= 2Y' \left( X\hat{\beta}_1^0 + X_0 \hat{\delta}_0 \right) - 2Y' \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} \right) + \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} \right)' \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} \right)$$

$$- \left( X\hat{\beta}_1^0 + X_0 \hat{\delta}_0 \right)' \left( X\hat{\beta}_1^0 + X_0 \hat{\delta}_0 \right)$$

$$= -2 (X\beta_1 + X_0 \delta + \varepsilon)' \left( X \left( \hat{\beta}_1 - \hat{\beta}_1^0 \right) + X_\gamma \hat{\delta} - X_0 \hat{\delta}_0 \right)$$

$$+ \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} + X\hat{\beta}_1^0 + X_0 \hat{\delta}_0 \right)' \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} - X\hat{\beta}_1^0 - X_0 \hat{\delta}_0 \right)$$

$$= -2 \left( X\beta_1 + X_0 \delta + \varepsilon \right)' \left( X, \hat{\delta} - X_0 \hat{\delta}_0 \right) + \left( X\hat{\beta}_1 + X_\gamma \hat{\delta} + X\hat{\beta}_1^0 + X_0 \hat{\delta}_0 \right)' \left( X, \hat{\delta} - X_0 \hat{\delta}_0 \right) .$$
\[ + o_p(1) \]
\[ = -2 \varepsilon' (X_\gamma - X_0) \delta + (X \tilde{\beta}_1 + X \tilde{\beta}_1^0 - 2X \beta_1 + X \gamma \delta - X_0 \delta)' (X_\gamma - X_0) \delta + o_p(1) \]
\[ = -2 \varepsilon' (X_\gamma - X_0) \delta + \delta' (X_\gamma - X_0)' (X_\gamma - X_0) \delta + (\tilde{\beta}_1 + \tilde{\beta}_1^0 - 2 \beta_1)' X' (X_\gamma - X_0) \delta + o_p(1) \]
\[ = -2 \varepsilon' (X_\gamma - X_0) \delta + \delta' (X_\gamma - X_0)' (X_\gamma - X_0) \delta + o_p(1) \]
\[ = -2T^{-\alpha} \sum_{t=1}^T \varepsilon_t (\Psi_t (\gamma) - \Psi_t (\gamma^0)) + \sum_{t=1}^T (\varepsilon_t)^2 |\Psi_t (\gamma) - \Psi_t (\gamma^0)| + o_p(1) \]

Now, using

\[ \overline{M} (\gamma_1, \gamma_2) = \overline{M}_0 + (\gamma_1 - \gamma_1^0) DF_1^0 + (\gamma_2 - \gamma_2^0) DF_2^0 + o(1) \]

**Case 1: \( v_1 \leq 0, v_2 \leq 0 \)** In this case, we have

\[ X'_\gamma X_0 = X'_0 X_0; \]
\[ S_T (\gamma_1, \gamma_2) = S_T (\gamma_1^0, \gamma_2^0) \]
\[ = 2 \varepsilon' (X_0 - X_\gamma) \delta + \delta' (X_0' X_0 - X'_\gamma X_\gamma) \delta + o_p(1) \]
\[ = 2 \varepsilon' (X_0 - X_\gamma) cT^{-\alpha} + c' (\overline{M}_0 - \overline{M}_\gamma) cT^{1-2\alpha} + o_p(1) \]
\[ = -2T^{-\alpha} \sum (\gamma_1 < z_{1l} < \gamma_1^0 \text{ and } \gamma_1^0 < z_{2l}) \text{ or } (\gamma_2 < z_{2l} < \gamma_2^0 \text{ and } \gamma_2^0 < z_{1l}) \in \varepsilon_t x_t^T c \]
\[ cT^{-\alpha} \left( (\gamma_1 - \gamma_1^0) DF_1^0 + (\gamma_2 - \gamma_2^0) DF_2^0 \right) c + o_p(1) \]
\[ = -2T^{-\alpha} \sum (\gamma_1 < z_{1l} < \gamma_1^0 \text{ and } \gamma_1^0 < z_{2l}) \text{ or } (\gamma_2 < z_{2l} < \gamma_2^0 \text{ and } \gamma_2^0 < z_{1l}) \in \varepsilon_t x_t c - c' \left( v_1 DF_1^0 + v_2 DF_2^0 \right) c + o_p(1) \]
\[ = -d' Dcv_1 F_1^0 - 2T^{-\alpha} \sum (\gamma_1 < z_{1l} < \gamma_1^0 \text{ and } \gamma_1^0 < z_{2l}) \varepsilon_t x_t + \gamma_1^0 < z_{1l} \varepsilon_t x_t + o_p(1) \]

Note that
\[ T^{-\alpha} \sum (\gamma_1 < z_{1l} < \gamma_1^0 \text{ and } \gamma_2^0 < z_{2l}) \in \varepsilon_t x_t \text{ converge in distribution to } B_1 (v), \text{ which is a vector Brownian motion with covariance matrix } E \left( B_1 (1) B_1 (1) \right) = -V_1^0, \]
\[ T^{-\alpha} \sum (\gamma_2 < z_{2l} < \gamma_2^0 \text{ and } \gamma_2^0 < z_{1l}) \varepsilon_t x_t \text{ converge in distribution to } B_2 (v_2), \text{ which is a vector Brownian motion with covariance matrix } E \left( B_2 (1) B_2 (1) \right) = -V_2^0, \]

where \( B_1 \) and \( B_2 \) are independent.

Thus, in the neighborhood of the true threshold values, the above is equal in distribution to
\[ d' Dcv_1 F_1^0 - 2d' B_1 (v) - d' Dcv_2 F_2^0 - 2d' B_2 (v_2). \]

We apply the same arguments to the following cases:
Case 2: $v_1 > 0$, $v_2 \leq 0$ We have
\[
S_T \left( \gamma_1, \gamma_2 \right) - S_T \left( \gamma_1^0, \gamma_2^0 \right) = -2 \varepsilon' \left( X_\gamma - X_0 \right) c T^{-\alpha} + T^{1-2a} c' \left( \bar{M}_\gamma + \bar{M}_0 - 2 \bar{M} \left( \gamma_1, \gamma_2 \right) \right)' c + o_p(1)
\]
\[
= -2 \varepsilon' \left( X_\gamma - X_0 \right) c T^{-\alpha} + T^{1-2a} c' \left( \bar{M}_0 + (\gamma_1 - \gamma_0) D \bar{F}_1^0 + (\gamma_2 - \gamma_0^0) D \bar{F}_2^0 + \bar{M}_0 - 2 \left( \bar{M}_0 + (\gamma_1 - \gamma_0^0) D \bar{F}_1^0 \right) \right)' c + o_p(1)
\]
\[
= -2 \varepsilon' \left( X_\gamma - X_0 \right) c T^{-\alpha} + T^{1-2a} c' \left( \bar{M}_0 + (\gamma_1 - \gamma_0^0) D \bar{F}_1^0 \right)' c + o_p(1)
\]
\[
= -2 \varepsilon' \left( X_\gamma - X_0 \right) c T^{-\alpha} + c' \left( v_2 D \bar{F}_2^0 - v_1 D \bar{F}_1^0 \right)' c + o_p(1)
\]
\[
d = -c' D v_1 \bar{F}_1^0 + 2 c' B_1 (v_1) + c' D v_2 \bar{F}_2^0 - 2 c' B_2 (v_2).
\]

Similarly, we have

Case 3: $v_1 \leq 0$, $v_2 > 0$
\[
S_T \left( \gamma_1, \gamma_2 \right) - S_T \left( \gamma_1^0, \gamma_2^0 \right) \]
\[
d = -c' D v_1 \bar{F}_1^0 - 2 c' B_1 (v_1) + c' D v_2 \bar{F}_2^0 + 2 c' B_2 (v_2).
\]

Case 4: $v_1 > 0$, $v_2 > 0$
\[
S_T \left( \gamma_1, \gamma_2 \right) - S_T \left( \gamma_1^0, \gamma_2^0 \right) \]
\[
d = c' D v_1 \bar{F}_1^0 + 2 c' B_1 (v_1) + c' D v_2 \bar{F}_2^0 + 2 c' B_2 (v_2).
\]

Making the change of variables
\[
v_1 = -\frac{c' V c}{\left( c' D c \right)^2} r_1,
\]
\[
v_2 = -\frac{c' V c}{\left( c' D c \right)^2} r_2.
\]

In general,
\[
S_T \left( \gamma_1, \gamma_2 \right) - S_T \left( \gamma_1^0, \gamma_2^0 \right) \]
\[
d = c' D c |v_1| \bar{F}_1^0 + 2 c' D c |v_1| \bar{F}_2^0 + 2 c' D c |v_2| \bar{F}_2^0 + 2 c' D c |v_2| \bar{F}_2^0
\]
\[
= c' V c \left( r_1 \right) + 2 \frac{\sqrt{c' V c}}{c' D c} c' B_1 \left( r_1 \right) + c' V c \left( r_2 \right) + 2 \frac{\sqrt{c' V c}}{c' D c} c' B_2 \left( r_2 \right)
\]
\[
= c' V c \left( r_1 \right) + 2 c' V c \left( r_1 \right) W_1 \left( r_1 \right) + c' V c \left( r_2 \right) + 2 c' V c \left( r_2 \right) W_2 \left( r_2 \right)
\]
\[
= 2 \frac{c' V c}{c' D c} \left( r_1 + W_1 \left( r_1 \right) + r_2 + W_2 \left( r_2 \right) \right).
\]

60
We have
\[
-T^{1-2\alpha} \frac{(c'Dc)^2}{c'Vc} \left( F^0_1 (\hat{\gamma}_1 - \gamma_1^0), F^0_2 (\hat{\gamma}_2 - \gamma_2^0) \right)
\]
\[
\overset{d}{\rightarrow} \arg \min_{(r_1, r_2) \in \mathbb{R}^2} \left( 2c'Vc \left( \frac{|r_1|}{2} + W_1(r_1) + \frac{|r_2|}{2} + W_2(r_2) \right) \right)
\]
\[
\overset{d}{=} \arg \max_{(r_1, r_2) \in \mathbb{R}^2} \left( -\frac{1}{2} |r_1| + W_1(r_1) - \frac{1}{2} |r_2| + W_2(r_2) \right).
\]

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
\[
\Pr \left( \arg \max \sum_{j=1}^{2} \left( -\frac{1}{2} |r_j| + W_j(r_j) \right) \leq a_1, \arg \max \sum_{j=1}^{2} \left( -\frac{1}{2} |r_j| + W_j(r_j) \right) \leq a_2 \right)
\]
\[
= \Pr \left( \arg \max \left( -\frac{1}{2} |r_1| + W_1(r_1) \right) \leq a_1, \arg \max \left( -\frac{1}{2} |r_2| + W_2(r_2) \right) \leq a_2 \right)
\]
\[
= \Pi_{j=1}^{2} \Pr \left( \arg \max \left( -\frac{1}{2} |r_j| + W_j(r_j) \right) \leq a_j \right)
\]
\[
\overset{\text{def}}{=} \Pi_{j=1}^{2} F_{r_j}(a_j).
\]

According to Bhattacharya and Brockwell (1976) and Hansen (2000), for \(a_1 > 0\) and \(a_2 > 0\), the above joint distribution equals
\[
F_{(\tilde{r}_1, \tilde{r}_2)}(a_1, a_2) = \Pi_{j=1}^{2} F_{\tilde{r}_j}(a_j)
\]
\[
= \Pi_{j=1}^{2} \left( 1 + \sqrt{\frac{a_j}{2\pi}} \exp \left( -\frac{a_j}{8} \right) + \frac{3}{2} \exp (a_j) \Phi \left( -\frac{3\sqrt{a_j}}{2} \right) - \frac{a_j + 5}{2} \Phi \left( -\frac{\sqrt{a_j}}{2} \right) \right),
\]

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

Thus, using the fact that \(\phi(x) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{x^2}{2} \right)\), the joint density function is
\[
f_{(\tilde{r}_1, \tilde{r}_2)}(a_1, a_2) = \Pi_{j=1}^{2} f_{\tilde{r}_j}(a_j)
\]
\[
= \Pi_{j=1}^{2} \left( \frac{3}{2} \exp (a_j) \Phi \left( -\frac{3\sqrt{a_j}}{2} \right) - \frac{1}{2} \Phi \left( -\frac{\sqrt{a_j}}{2} \right) \right).
\]

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

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Sources and Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ratio of fiscal deficits to GDP</td>
<td>Fiscal deficit is taken from IFS line 80 and GDP is taken from IFS line 99B.</td>
</tr>
<tr>
<td>2. Ratio of short-term external liabilities to foreign exchange reserves</td>
<td>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.</td>
</tr>
<tr>
<td>3. Lending rate differential</td>
<td>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.</td>
</tr>
<tr>
<td>4. Real exchange rate appreciation index</td>
<td>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.</td>
</tr>
<tr>
<td>5. Ratio of domestic credit to GDP</td>
<td>The domestic credit data is taken from IFS line 32.ZF and the GDP data is from IFS line 99B.</td>
</tr>
</tbody>
</table>
Figure 1a: $S_T (\gamma_1, \gamma_2) / T$

Figure 1b: $g (\gamma_1, \gamma_2)$
Figure 2a: Distribution of $T^{1/2} \left( \hat{\beta}_1 (\hat{\gamma}_1, \hat{\gamma}_2) - \beta_1 \right)$

Figure 2b: Distribution of $T^{1/2} \left( \hat{\beta}_2 (\hat{\gamma}_1, \hat{\gamma}_2) - \beta_2 \right)$
Figure 2c: Joint Distribution of $T^{1/2} \left( \hat{\beta}_1 (\hat{\gamma}_1, \hat{\gamma}_2) - \beta_1, \hat{\beta}_2 (\hat{\gamma}_1, \hat{\gamma}_2) - \beta_2 \right)$
Figure 3a: Distribution of $T^{3/4} f_1 (\gamma_1^0) \overline{F}_2 (\gamma_2^0) (\gamma_1 - \gamma_1^0)$

Figure 3b: Distribution of $T^{3/4} f_2 (\gamma_2^0) \overline{F}_1 (\gamma_1^0) (\gamma_2 - \gamma_2^0)$
Figure 4a: Joint Distribution of
\[ T^{1-2\alpha} \left( f_1(\gamma_1^0) \overline{F}_2(\gamma_2^0) (\gamma_1 - \gamma_1^0), f_2(\gamma_2^0) \overline{F}_1(\gamma_1^0) (\gamma_2 - \gamma_2^0) \right) \]

Figure 4b: Joint Density of \( f_{(\gamma_1, \gamma_2)}(a_1, a_2) \)
Figure 5a: Distribution of the $LR_T$ Statistic

Figure 5b: 95% Confidence Region for $\gamma_1$ and $\gamma_2$
Figure 6: Threshold Effects and Exchange Market Pressure Index of selected Asian Countries

(a): China

(b): Hong Kong

Lending rate differential

Short-term external liabilities

Reserves

Excess over threshold

Short liabilities/reserves (avg. of 4 quarterly lags)

Exchange market pressure index (right axis)
Figure 6: Threshold Effects and Exchange Market Pressure Index of selected Asian Countries (Continued)

(c): **Indonesia**

(d): **S. Korea**

<table>
<thead>
<tr>
<th>Short-term external liabilities</th>
<th>Reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lending rate differential

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

(e): Malaysia

(f): Philippines

<table>
<thead>
<tr>
<th>Short-term external liabilities</th>
<th>Reserves</th>
</tr>
</thead>
</table>

Lending rate differential

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

(g): Singapore

(h): Taiwan

Short-term external liabilities
Reserves

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

(i): Thailand

<table>
<thead>
<tr>
<th>Short-term external liabilities</th>
<th>Reserves</th>
</tr>
</thead>
</table>

![Graph showing short-term external liabilities and reserves over time]

Lending rate differential

![Graph showing lending rate differential over time]
Figure 7: Threshold Effects and Exchange Market Pressure Index of selected Latin American Countries

(a): Argentina

(b): Brazil

Short-term external liabilities

Reserves

Lending rate differential

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

(c): Chile

(d): Colombia

<table>
<thead>
<tr>
<th>Short-term external liabilities</th>
<th>Reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>Colombia</td>
</tr>
</tbody>
</table>

Lending rate differential

<table>
<thead>
<tr>
<th>Lending rate differential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
</tr>
</tbody>
</table>
Figure 7: Threshold Effects and Exchange Market Pressure Index of selected Latin American Countries (Continued)

(e): Mexico

(f): Uruguay

<table>
<thead>
<tr>
<th>Short-term external liabilities</th>
<th>Reserves</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mexico</td>
<td>Uruguay</td>
</tr>
</tbody>
</table>

Lending rate differential

76
Figure 7: Threshold Effects and Exchange Market Pressure Index of selected Latin American Countries (Continued)

(g): Venezuela

<table>
<thead>
<tr>
<th>Short-term external liabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reserves</td>
</tr>
</tbody>
</table>

Lending rate differential