An Optimal Design of Early Warning Systems: A Bayesian Quickest Change Detection Approach

Li, Haixi

University of Wisconsin-Madison

March 2012

An Optimal Design of Early Warning Systems for Financial Disruptions: A Bayesian Quickest Change Detection Approach

Haixi Li[*,†]
Department of Economics, University of Wisconsin-Madison

March 10, 2012

Abstract

This paper proposed a new optimal design of Early Warning Systems (EWS) to detect early warning signals of an impending financial crisis. The problem of EWS was formulated from a policy maker's perspective. Hence the probability threshold was obtained by minimizing the policy maker's welfare loss. This paper employed the state-of-the-art Bayesian Quickest Change Detection (BQCD) as the methodology to detect the early warning signals as soon as possible. We showed that the BQCD method outperformed the Logit model used in traditional EWS models based on results of simulation exercise and the out-of-sample predictions of the 1997 Asian financial crises. We found that not only early warning signals were stronger prior to a crisis, but also stronger warning signals appeared more frequently. The BQCD method was sensitive to the increase in frequency, hence out-performed the traditional Logit-EWS model.

Keywords: early warning system, financial crisis, monetary policy, Bayesian quickest change detection, optimal stopping
JEL code: E5, F3

1 Introduction

1.1 Motivation

In the past few decades, we saw a large number of financial disruptions often with devastating economic and social consequences. As a result, international organizations begun to develop

[*]Department of Economics, University of Wisconsin-Madison, 1180 Observatory Dr. Madison, WI 53706-1393, USA. Tel: 1-608-2620200. Email: hli26@wisc.edu

[†]I am grateful to my main adviser Steven Durlauf for his advice and inspiration. I'd like to thank William Brock and Noah Williams for their help and advice. The project benefited from suggestions from seminar participants at UW-Madison. All mistakes are of the author. The latest version of this paper can be found in the following link: `http://dl.dropbox.com/u/22521981/Early_Warning_System_Working_Edition.pdf`

Early Warning System (EWS) models with the aim of anticipating whether and when a country may be affected by a financial crisis. EWS models can have substantial value to policy makers by allowing them to detect symptoms of financial disruptions sufficiently in advance to take preemptive measures to reduce risks of financial crises. As stated in IMF (2010), the primary purpose of EWS is to identify vulnerabilities that predisposes an economy to a crisis, so that preemptive policies can be implemented. Hence such an alarm system constitutes a crucial part of measures taken by authorities to prevent financial crises and facilitate financial stability.

Two main approaches are developed since the early contributions of EWS models by Kaminsky et al. (1998) and Berg and Pattillo (1999): the leading indicators approach and the discrete-dependent-variable approach.

The leading indicators approach developed by Kaminsky and Reinhart (1999) and Kaminsky et al. (1998) considers various indicators and transforms them into binary signals. For example, when an indicator falls above a given threshold, this particular indicator will flash a red light. The information from different indicators is aggregated into a composite measure of probabilities of an impending crisis. The level of the threshold which is predetermined is of great importance, since the lower the threshold, the more signals will this indicator send, hence the cost of false alarm will increase.

Berg and Pattillo (1999) first popularized the discrete-dependent-variable approach. It assumes the probability of an impending financial crisis depends on a non-linear function of the indicators, i.e. $Pr\,(Y=1) = F\,(X\beta)$. Then, one needs to specify a threshold T beyond which the predicted probability can be interpreted as issuing an alarm of an impending financial crisis. The key issue is how to determine the threshold. The lower it is, the more alarms will be issued, hence, the cost of false alarm (Type 1 error) will increase. On the other hand, if it is chosen too high, it will be more likely to miss crisis signals (Type 2 error). The trade-off problem has been considered in the EWS literature. Criterion such as Noise-to-Signal Ratio has been used. But this criterion is based on statistical measures, which do not represent the cost of false alarm and delay detection. As Bussiere and Fratzscher (2008) and Candelon et al. (2010b) pointed out that the threshold T is chosen with no explicit reasons in standard EWS models.

In order to model costs of the two types of mistakes, the policy maker's problem needs to be considered in the optimal design of EWS. In other words, the problem of EWS should be framed as the policy maker's problem. Gramlich et al. (2010) pointed out that a critical element of the optimal design of EWS model is the tight correspondence between the outcome of EWS model and the objective of its user. Davis and Karim (2008) argued that it is important to consider policy maker's objective in the design of EWS model and determination of the thresholds since there is trade-off between the two types of mistakes.

Bussiere and Fratzscher (2008) first explored issues of the optimal design of EWS models from a policy maker's perspective, since both the outcome of an EWS model and the policy reaction are discrete events. The authors proposed a simple objective function of the policy maker to demonstrate the necessity of incorporating policy maker's preference in the optimal design of EWS. The objective function took the following form: $L(T) = \theta p(T) + (1 - \theta) q(T)$, where p is the probability of a missed crisis and q is the probability of issuing an alarm but the crisis did not occur, θ is the relative cost of missing a crisis. The authors also listed three critical components of the optimal design of EWS: the degree of risk aversion θ of missing a crisis, the forecast horizon of the model, and the probability threshold T for issuing crisis alarms. θ is of fundamental importance, once it is determined, the threshold T is uniquely determined, and it depends on the relative cost of a crisis and the policy maker's preference.

This paper is considered as a continuation of Bussiere and Fratzscher (2008)'s effort. We will frame the problem of early warning system models as the policy maker's problem. The policy maker will extract information from standard Logit-EWS models about the likelihood of an impending financial crisis, and determine the timing to take preemptive actions. To determine the optimal timing of preemptive actions, the policymaker will have to strike a balance between costs of the two types of mistakes: false alarm and delay detection.

As we will argue later, to detect the vulnerability of the economy and determine the timing of preemptive action to be taken is inherently a quickest change detection problem which has been studied intensively in engineering and mathematics. We are going to adopt this method to solve the policy maker's problem.

1.2 Quickest Change Detection

Quickest change detection deals with the design and analysis of techniques for quickest detection of a change in the state of observed stochastic processes. More thorough review and list of references can be found in Polunchenko and Tartakovsky (2011) and Zacks (1991). In many problems, when the state representing the pattern of behavior of observed stochastic processes undergoes a change, one is interested in detecting the change as soon as it happens.

In general, observations arrive sequentially, and if it is believed that the observations are generated from the normal state, it is let to continue. If it is believed that the state has changed, it is one's interest to detect the change as soon as possible, so that an appropriate response can be made in a timely manner. The decision needs to be made in real time with available data. However, any detection policy will give rise to two types of mistakes: false alarm and delay detection. The optimal detection policy will need to strike a balance between costs of the two types of mistakes.

Quickest change detection has applications in many fields, such as signal processing, automatic control and so on. It started in the 1920-1930's designed to deal with quality control issues. One of pioneering work is Shewhart (1931), who popularized the Shewhart's charts. More efficient sequential detection procedures were developed later in the 1950-1960's after the emergence of Sequential Analysis(Wald, 1947). This leads to a large amount of literature on both theory and practice of sequential change-point detection.

The Bayesian Quickest Change Detection problem was first proposed and solved by Shiryaev (1963, 1978). Consider a random sequence $\{Z_k, k = 1, 2, ...\}$ with a random structural break time θ, conditional on θ, $\{Z_k, k = 1, 2, ...\}$ is an independent sequence with $Z_1...Z_{\theta-1}$ being i.i.d. with marginal distribution Q_0 and $Z_\theta, Z_{\theta+1}, ...$ being i.i.d with marginal distribution Q_1. The objective is to find a stopping time T that solves the following problem:

$$\inf_{T \in \mathcal{T}} \left\{ P\left(T < \theta\right) + cE\left(T - \theta\right)^+ \right\}. \tag{1.1}$$

$P\left(T < \theta\right)$ is the frequency of false alarm and $E\left(T - \theta\right)^+$ is the expected length of delay detection. The optimal stopping time T^* must strike a balance between these two conflicting objectives. The constant $c > 0$ is the relative weight assigned to the expected length of delay detection.

The monetary policy maker's problem when facing an impending financial crisis, is inherently a quickest change detection problem. The behavior of the early warning signals X changes prior to a crisis, indicating vulnerability of the economy. The policymaker needs to detection the change as soon as it happens and take preemptive action to attenuate the risk of an impending financial crisis.

2 The Model

2.1 The Stochastic Process

There are two regimes of the economy, the economy starts with the tranquil regime then jumps into a financial crisis at some random and unobservable time ν. Assume there is a D-dimensional controlled stochastic process ξ_t with u_t as the control. The process is given by

$$d\xi_t = k_\gamma\left(\xi_t, u_t\right)dt + \sigma dW_t$$

with

$$\gamma = \begin{cases} 0, & t \leq \nu \\ 1, & t > \nu \end{cases}.$$

There is a structural break in the controlled stochastic process at time ν, ν is a random variable with probability measure ϕ as the prior:

$$\begin{aligned} \phi\left(\nu = 0\right) &= \pi \\ \phi\left(\nu \geq s\right) &= \left(1 - \pi\right)e^{-\lambda s},\ \lambda > 0. \end{aligned}$$

As assumed in the EWS literature, There is a collection of early warning signals, whose behaviors will start to change h periods before ν, where h is deterministic. Even though the reasons that caused a financial crisis might be different from case to case, it might be possible to identify a common pattern of the behaviors of the early warning signals that is detectable prior to a financial crisis. Once the common pattern of the early warning signals starts to show the vulnerability of the economy, it is in the policy maker's interest to detect the change in the pattern as soon as it appears and take preemptive action. Hence, it is assumed that the joint distribution of the early warning signals will change at random and unobservable time $\theta = \nu - h$. More formally, the policy maker is conducting a hypothesis test sequentially*:

- H0: no financial crisis will happen within the following h periods
- H1: a financial crisis will happen within the following h periods.

It can be interpreted that at random time θ, the behaviors of the early warning signals will start to support the alternative instead of the null. The policy maker will conduct the hypothesis test sequentially. Once the alternative is accepted, the policy maker will take preemptive actions to prevent the crisis. The policy maker will choose a stopping time T ex-ante to determine the timing to take preemptive actions optimally.

2.2 The Sequential Decision Rule

It is conventional wisdom that commitment monetary policy making is superior to discretionary monetary policy making. It has been standard to assume the central bank is committed to a policy rule once for all, assuming the model of the economy never change. How to specify a commitment policy rule when the environment of the economy changes is an important question.

In this paper, we assume the policy maker is committed to a sequential decision rule $(T, u_{T,t})$, which is specified as follows

$$u_{T,t} = \begin{cases} u_{0,t} & \text{before T declares stopping} \\ u_{1,t} & \text{after T declares stopping} \end{cases}, \tag{2.2}$$

where

$$u_{1,t} = \begin{cases} u'_{1,t} & \text{after T declares stopping, before T+h} \\ u''_{1,t} & \text{after T+h} \end{cases},$$

Where $u_{0,t}$ is the policy rule that is appropriate when the economy is in the tranquil period, once stopping time T declares that there is enough warning of a pending financial crisis, policy

*More detailed discussion can be found in Subsection 2.6

maker will take temporary preemptive policy rule $u'_{1,t}$. After the preemptive action is taken, policy rule $u''_{2,t}$ will be adopted which is appropriate in the regime of post crisis.

The central bank is committed to a simple rule $u_{0,t}$ first, but the life span of this commitment is a random variable, the timing at which the central bank decides to switch to a different policy regime is determined by newly arrived information and a principle that is chosen at the beginning of the commitment optimally, i.e. the stopping time T. The principle that the central bank is committed to specifies a rule (T) that determines that should the central bank continue the current commitment or switch to a new policy regime, given the information available at the time.

2.3 The Probability Measure

Consider the following measurable spaces

1. $(\Phi, \mathcal{J}) = \left(C[0,\infty)^D \times (\mathcal{R}^+)^2, \mathcal{B}\left(C[0,\infty)^D \times (\mathcal{R}^+)^2\right)\right)$,
2. $(\Lambda, \mathcal{G}) = \left(C[0,\infty)^D \times \mathcal{R}^+, \mathcal{B}\left(\left(C[0,\infty)^D\right) \times \mathcal{R}^+\right)\right)$,

where $C[0,\infty)$ is the space of continuous functions. We consider the following filtration $\mathcal{G}_t = \sigma\{T \leq s, \xi_s, s \leq t\}$ and $\mathcal{J}_t = \sigma\{k, l, \xi_s, s \leq t\}$. $\mathcal{J}_t$ is the information sets when the policy maker has full information: they know the structure break happens at l, and policy regime switches at k. $\mathcal{G}_t$ is the information of the policy maker at t.

Without loss of generality, assume $\zeta = 1$. For any given $k \in \mathcal{R}^+$and $l \in \mathcal{R}^+$, here k represents the realization of stopping time T and l represents the realization of structure break time θ. When $k \leq l$, define function

$$K_-^{k,l}(s) = k_0(\xi_s, u_{0t}) 1_{\{s<k\}} + k_0(\xi_s, u_{1t}) 1_{\{k \leq s \leq l+h\}} + k_1(\xi_s, u_{1t}) 1_{\{l+h<s\}};$$

when $l + h \geq k > l$, define function

$$K_+^{k,l}(s) = k_0(\xi_s, u_{0t}) 1_{\{s<k\}} + k_0(\xi_s, u_{1t}) 1_{\{k \leq s \leq l+h\}} + k_1(\xi_s, u_{1t}) 1_{\{l+h<s\}}.$$

Given a probability space $(\Omega, \mathcal{F}, P_0)$, and a D-dimensional Brownian motion

$$W = \left\{W_t = \left(W_t^{(1)}, ..., W_t^{(D)}\right), \mathcal{F}_t, 0 \leq t < \infty\right\}$$

defined on it with $P_0(W_0 = 0) = 1$. We can define the following Radon-Nikondým derivative and define the corresponding probability measure.

When $k \leq l$, let

$$\frac{dP_-^{k,l}}{dP_0} \equiv \exp\left\{-\int_0^t \sum_{d=1}^D K_-^{k,l,(d)}(s)\, dW_s^{(d)} - \frac{1}{2}\int_0^t \left\|K_-^{k,l}(s)\right\|^2 ds\right\}; \tag{2.3}$$

when $l+h \geq k > l$, define

$$\frac{dP_{+}^{k,l}}{dP_0} \equiv \exp\left\{-\int_0^t \sum_{d=1}^{D} K_{+}^{k,l,(d)}(s)\, dW_s^{(d)} - \frac{1}{2}\int_0^t \left\| K_{-}^{k,l}(s)\right\|^2 ds\right\}. \tag{2.4}$$

For any k and l, define

$$\frac{dP^{k,l}}{dP_0} \equiv \frac{dP_{-}^{k,l}}{dP_0} 1_{\{k \leq l\}} + \frac{dP_{+}^{k,l}}{dP_0} 1_{\{l+h \geq k > l\}}.$$

The next thing we need is to average out l and k, with prior ϕ and probability measure P_θ to be defined below.

For a given stopping time T, define

$$T_n(\omega) = \begin{cases} T(\omega), & if\, T(\omega) = \infty \\ \frac{m}{2^n}, & if\, \frac{m-1}{2^n} \leq T(\omega) < \frac{m}{2^n} \end{cases}.$$

We have that T_n is a stopping time and $\lim_{n\to\infty} T_n = T^*$. Define probability measure

$$P_{\theta,n}(T_n(\omega) = k|l) \equiv P^{k,l}(\omega \in \Lambda : T_n(\omega) = k),$$

given a stopping time T, and correspondingly defined T_n, $P_{\theta,n}$ tells us the probability of $T_n = k$, when structure break happens at l.

For any event $A \in \Lambda$,

$$E_\theta\left(1_{\{A\}}\right) \equiv \lim_{n\to\infty} \sum_{m=0}^{\infty} \left\{P_{\theta,n}\left(T_n(\omega) = \frac{m}{2^n}|\theta\right) P^{\frac{m}{2^n},l}(A)\right\}.$$

$P^{k,l}(A)$ is the probability of observing event A, given structure break happens at l and policy switch at k. $E_\theta\left(1_{\{A\}}\right)$ tells us, given stopping time T, and structure break happens at l, the probability of observing event A. Compare to $P^{k,l}(A)$, we just average out the randomness in stopping time T.

$$P_\theta(A) = E_\theta\left\{1_{\{A\}}\right\} \tag{2.5}$$

$P_\theta(A)$ is the probability of A conditional on θ.

Now we can use the prior distribution for θ to average out l, for any event $A \in \Lambda$

$$P(A) = E_\pi\left(E_\theta\left(1_{\{A\}}\right)\right). \tag{2.6}$$

*For more details, please see Problem 2.24 of Karatzas and Shreve (1991).

E_π is defined with respect to prior distribution ϕ, expectation operator E is defined with respect to probability measure P.

The existence of probability measure P under which the process ξ_t is Brownian motion is guaranteed by Novikov condition*

$$\begin{aligned} E\left\{\exp\left(\frac{1}{2}\int_0^t \left\|K_-^{k,l}(s)\right\|^2 ds\right)\right\} &< \infty; \\ E\left\{\exp\left(\frac{1}{2}\int_0^t \left\|K_+^{k,l}(s)\right\|^2 ds\right)\right\} &< \infty. \end{aligned}$$

2.4 The Objective Function

Now consider k and l are both deterministic, k is the time that policy maker choose to preempt, l is the time that the behavior of the early warning signals start to change.

The pattern of early warning signals will change at l indicating the vulnerability of the economy, the policy maker tries to detect the change, and will switch policy regime once an alarm is triggered. Once an alarm is issued, the policy maker will take temporary preemptive action, which is a temporary policy rule that will last h periods, then he will switch to a policy rule that is appropriate in the second regime. To do so, the policymaker will make two types of mistakes: false alarm, that the policy maker respond before l, i.e. $k \le l$; delayed detection, that the policy maker respond after l, but before $l+h$, i.e. $l < k \le l+h$. l is unobservable in case $k \le l$, since in general, once preemptive action is taken, the transition of the economy from regime 1 to regime 2 might not be observable. If the policymaker fail to take preemptive actions before $l+h$, then the financial disruption will manifest itself, and it will be observable when it happens.

$$v(k,l) = v_-(k,l)\,1_{\{k\le l\}} + v_+(k,l)\,1_{\{l+h\ge k>l\}}$$

More precisely, in the domain of false alarm, when $k \le l$

$$\begin{aligned} v_-(k,l) &= E_-\left\{\int_0^k e^{-\rho t} L(\xi_t,u_{0t})\,dt + \int_k^{k+h} e^{-\rho t} L\left(\xi_t,u'_{1t}\right)dt\right\} \\ &\quad + E_-\left\{\int_{k+h}^{l+h} e^{-\rho t} L\left(\xi_t,u''_{1t}\right)dt + \int_{l+h}^{\infty} e^{-\rho t} L\left(\xi_t,u''_{1t}\right)dt\right\} \\ d\xi_t &= k_0(\xi_t,u_{0t})\,dt + \sigma dW_t,\; t\le k \\ d\xi_t &= k_0(\xi_t,u_{1t})\,dt + \sigma dW_t,\; k<t\le k+h \\ d\xi_t &= k_0\left(\xi_t,u''_{1t}\right)dt + \sigma dW_t,\; k+h<t\le l+h \\ d\xi_t &= k_1\left(\xi_t,u''_{1t}\right)dt + \sigma dW_t,\; l+h<t. \end{aligned}$$

*More details about this condition can be found in page 199 of Karatzas and Shreve (1991)

In the domain of delayed detection, when $l+h \geq k > l$

$$v_{+}(k,l) = E_{+}\left\{\int_0^k e^{-\rho t} L(\xi_t, u_{0t})\,dt + \int_k^{l+h} e^{-\rho t} L\left(\xi_t, u'_{1t}\right)dt + \int_{l+h}^{\infty} e^{-\rho t} L\left(\xi_t, u''_{1t}\right)dt\right\}$$

$$d\xi_t = k_0(\xi_t, u_{0t})\,dt + \sigma dW_t,\ t \leq k$$

$$d\xi_t = k_0\left(\xi_t, u'_{1t}\right)dt + \sigma dW_t,\ k < t \leq l+h$$

$$d\xi_t = k_1\left(\xi_t, u''_{1t}\right)dt + \sigma dW_t,\ l+h < t.$$

The policy maker's problem can be written as

$$\inf_{T\in\mathcal{T}} E\{v(T,\theta) - v(\theta,\theta)\}$$

Proposition 1. *The policy maker's objective function $E\{v(T,\theta) - v(\theta,\theta)\}$ can be approximated by*

$$E\left\{-v'_{-}(\theta,\theta)(\theta-T)^{+} + v'_{+}\left(\theta',\theta\right)h1_{\{T>\theta\}}\right\}.$$

Proof. Please see Appendix B. □

The policy maker's problem can be approximated as

$$\inf_{T\in\mathcal{T}} E\left\{-v'_{-}(\theta,\theta)(\theta-T)^{+} + v'_{+}\left(\theta',\theta\right)h1_{\{T>\theta\}}\right\}.$$

We can further simplify the objective:

$$E\left\{\left(v'_{+}\left(\theta^{+},\theta\right)\left[(\theta-T)^{+} + h1_{\{T>\theta\}}\right]\right)\right\}.$$

Define probability distribution function

$$\tilde{P}_\theta = \frac{v'_{+}(\theta^{+},\theta)}{K}P_\theta,$$

where $K = v'_{+}(\theta^{+},\theta)\,P_\theta(\Omega)$. Based on $\tilde{P}_\theta$, we can define expectation $\tilde{E}$, hence the problem can be converted to

$$\inf_{T\in\mathcal{T}}\left\{\tilde{E}\left[c(\theta-T)^{+}\right] + \tilde{P}(T>\theta)\right\}, \tag{2.7}$$

where

$$c = -\frac{v'_{-}(\theta,\theta)}{v'_{+}(\theta^{+},\theta)}.$$

c is the relative cost ratio, it is the ratio of the speeds of increase in losses associated with the two types of mistakes. When $c > 1$, that means the increase in loss of false alarm is faster than that of delay detection, hence higher weight should be put on the expected length of false alarm.

Proposition 2. *At any time t, given deterministic k and l, if the expected period loss function can be written as*

$$E\left\{L_t\left(k,l\right)\right\} = \mathbb{L}\left(1_{\{t\geq k\}}, 1_{\{t\geq l\}}\right),$$

then $c = -\frac{v'_-(\theta,\theta)}{v'_+(\theta^+,\theta)}$ is a constant.

Proof. Please see Appendix D. □

2.5 Discrete Time Version

In this subsection, we formulate the policy maker's problem in discrete time analogous to that of continuous time. The objective in continuous time has the same form as in discrete time

$$\inf_{T\in\mathcal{T}}\left\{\tilde{E}\left[c\left(\theta - T\right)^+\right] + \tilde{P}\left(T > \theta\right)\right\},$$

where

$$c = -\frac{v_-\left(\theta,\theta\right) - v_-\left(\theta-1,\theta\right)}{v_+\left(\theta+1,\theta\right) - v_-\left(\theta,\theta\right)}, \tag{2.8}$$

we simply use growth rate to replace time derivative.

Assume geometric prior distribution

$$p\left(\theta = k\right) = \begin{cases} \pi & \text{if } k = 0 \\ \left(1-\pi\right)\rho\left(1-\rho\right) & \text{if } k = 1, 2, ... \end{cases}. \tag{2.9}$$

Define $\pi_t = \tilde{P}\left(\theta < t|\mathcal{F}_t\right)$, which is the posterior probability of structure break happening before t.

Proposition 3. *Assume $\tilde{E}\left(\theta\right)$ and $\tilde{E}\left(T\right)$ are finite, then*

$$\tilde{E}\left[c\left(\theta - T\right)^+\right] + \tilde{P}\left(T > \theta\right) = \tilde{E}\left\{c\theta - \sum_{m=0}^{T-1} c\left(1-\pi_m\right) + \pi_{T-1}\right\}$$

Proof. Please see Appendix C. □

The next Proposition asserts that the relative cost ratio c is a constant under certain conditions in discrete time.

Proposition 4. *At any time t, given deterministic k and l, if the expected period loss function can be written as*

$$E\left\{L_t\left(k,l\right)\right\} = \mathbb{L}\left(1_{\{t\geq k\}}, 1_{\{t\geq l\}}\right),$$

then $c = -\frac{v_-(\theta,\theta)-v_-(\theta-1,\theta)}{v_+(\theta+1,\theta)-v_+(\theta,\theta)}$ is a constant.

Proof. Please see Appendix E. □

2.6 The Evolution of Posterior Probability

There are a collection of early warning signals X, the behavior of these early warning signs are different prior to a financial crisis. This one of fundamental assumption in EWS literature that the behavior of some particular variables is discernibly different in some periods before a crisis from that in the tranquil periods. Our job is to detect the change in this behavior. Consider the following two alternatives

- H0: no financial crisis will happen within the following h periods
- H1: a financial crisis will happen within the following h periods.

Define dependent variable $Y = 0$ when the null is true, and $Y = 1$ when the alternative is true. We will run a Logit model on historical data and obtain conditional probability:

$$\begin{aligned}\Pr\left(Y_i = 1|X_i\right) &= \frac{1}{1-\exp\left(-X_i\beta\right)} \\ \Pr\left(Y_i = 0|X_i\right) &= 1-\Pr\left(Y_i = 1|X_i\right).\end{aligned}$$

The likelihood ratio is defined as the probability of observing X_i given the alternative is true divided by the probability of observing X_i given the null is true. Hence, by this definition

$$\begin{aligned}L\left(X_i\right) &= \frac{\Pr\left(X_i|Y_i = 1\right)}{\Pr\left(X_i|Y_i = 0\right)} \\ &= \frac{\Pr\left(Y_i = 1|X_i\right)\Pr\left(Y_i = 0\right)}{\Pr\left(Y_i = 0|X_i\right)\Pr\left(Y_i = 1\right)}\end{aligned}$$

The unconditional probability $\Pr\left(Y_i = 0\right)$ will be taken as the prior π.

The posterior probability $\pi_k = \tilde{P}\left(\theta \leq k|\mathcal{F}_k\right),\ k = 0, 1, ..$ consists a sequence $\{\pi_k\}$ evolves according to the recursion

$$\pi_k = \frac{L\left(X_k\right)\left[\pi_{k-1} + \rho\left(1-\pi_{k-1}\right)\right]}{L\left(X_k\right)\left[\pi_{k-1} + \rho\left(1-\pi_{k-1}\right)\right] + \left(1-\rho\right)\left(1-\pi_{k-1}\right)}.$$

2.7 Solution Method

The policy maker's problem has been converted to

$$\inf_{T\in\mathcal{T}} \tilde{E}\left\{c\theta - \sum_{m=0}^{T-1} c\left(1-\pi_m\right) + \pi_{T-1}\right\}, \tag{2.10}$$

the prior is given by Equation 2.9.

The solution of the discrete time version is similar to section 4.3 of Shiryaev (1978).

Let $g(\pi) = \pi - c(1-\pi)$, and define operator $\mathcal{Q}$ similar to equation (4.128) in Shiryaev (1978)

$$\mathcal{Q}^1 g(\pi) = \min\left\{g(\pi), \pi - c(1-\pi) + E_\pi g(\pi')\right\}, \pi \in [0,1].$$

Then the value function given π is

$$v(\pi) = \lim_{n\to\infty} \mathcal{Q}^n g(\pi)$$

$$v(\pi) = \min\left\{g(\pi), \pi - c(1-\pi) + E_\pi v(\pi')\right\},$$

and the stopping time

$$T^* = \inf\{n \geq 0 : v(\pi_n) = \pi_n - c(1-\pi_n)\} \tag{2.11}$$

is the optimal stopping time.

3 Quantitative Study

In this paper, we borrow Zampolli (2006)'s model to study currency crises. Zampolli (2006) models large adjustments in asset prices with a Markov regime-switching model. There are two regimes of the economy: bubble regime and no-bubble regime. When the economy is the the bubble regime, exchange rate will experience sustained deviations from fundamentals. When the economy is in the no-bubble regime, exchange rate fluctuates around its fundamentals. When the economy switches from bubble regime to no-bubble regime, the exchange rate collapses and a currency crisis happens. The evolution over time of the two regimes is described by a Markov chain.

In this paper, we borrow the idea of two regimes of the economy to model currency crisis, in order to address the timing issue, we follow the construction of Shiryaev (1963)'s Bayesian Quickest Change Detection problem. We assume the economy starts with the bubble regime, and then the jumps into no-bubble regime at some random time ν.

Zampolli (2006) modified Ball (1999)'s small open economy model:

$$\begin{aligned} y_{t+1} &= \alpha y_t - \beta(i_t - \pi_t) - \chi a_t + \eta_t & (3.12)\\ \pi_{t+1} &= \delta\pi_t + \gamma y_t - f(a_t - a_{t-1}) + \varepsilon_t & (3.13)\\ a_t &= \rho_\tau a_{t-1} + \kappa(i_t - \pi_t) + \upsilon_t. & (3.14) \end{aligned}$$

where

$$\begin{cases} \rho_0 > 1 & \text{if}\, t \leq \nu, \\ \rho_1 = 0 & \text{if}\, t > \nu. \end{cases}$$

Equation 3.12 is the open-economy IS curve, where y_t is output gap, $i_t - \pi_t$ is real interest rate. Equation 3.13 is the open-economy Phillips curve, where π_t is inflation, a_t is real exchange rate. Equation 3.14 is the reduced form equation that relate the real exchange rate with real interest rate and transitive shocks. When $t \leq \nu$, the real exchange rate exhibits bubble like behavior, this can allow the real exchange rate to grow away from it fundamental value, when the regime switches, the real exchange rate collapses towards its fundamental value.

The policy rule can be specified as

$$i_t = f_y y_t + f_\pi \pi_t + f_a a_{t-1}.$$

The objective of the policy maker is given by

$$\sum_{t=0}^{\infty} \beta^t \left[\operatorname{var}(\pi_t) + \lambda \operatorname{var}(y_t) \right].$$

The economy starts with bubble regime, at time ν, the real exchange rate will experience abrupt reversion to its fundamental value. The policymaker is monitoring a set of indicators, whose joint distribution will change at $\theta = \nu - h$, and determine the timing to take preemptive action to reduce to risk of the pending financial disruption. Preemptive action is necessary in this model since the policy effect is delayed.

Three elementary policy rules in the sequential decision rule is determined by the following method. u_{0t} is the optimal policy rule in regime 1, u''_{1t} is the optimal policy rule in regime 2, u'_{1t} is the optimal policy rule given u_{0t} and u''_{1t} and the timing of policy is set at one period before the time of regime switch. u_{0t} and u'_{1t} are solved with the method provided by Zampolli (2006) with the transition matrix

$$P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where initial state set at 1 and 2 respectively.

With this model and the parameter values taken from Zampolli (2006), we can compute $c = 1.7857$ according to Equation 2.8. The optimal stopping 2.10 can be solved with the method presented is Subsection 2.7. The result can be presented in Figure 1. As is shown in Figure 1, π^* is determined by the optimality condition 2.11. π^* separates the space of π into two regions: continuation region and stopping region. The optimal stopping time is defined as the first time π_t jumps from continuation region into stopping region.

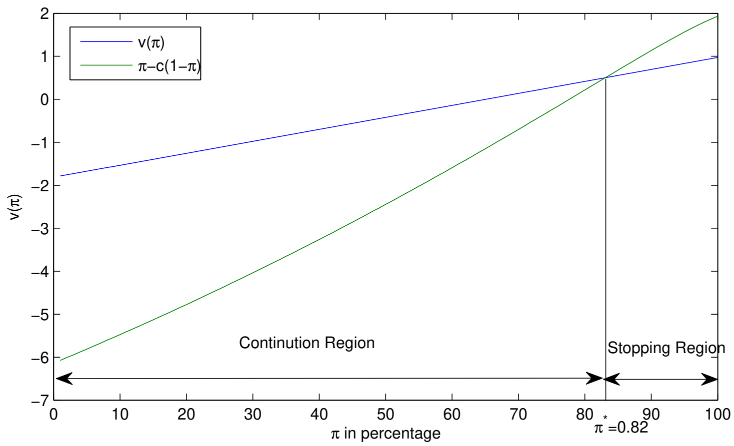

Figure 1: Optimal Stopping

3.1 Data

Most of the data was extracted from IFS database, some are extracted from national banks of the countries under analysis via Datastream*. It consists of monthly observations from Jan 1986 through Dec 1998 for 15 emerging countries†.

Following Candelon et al. (2010a), variables of interest include the annual growth rate of international reserves, annual growth rate of imports, the annual growth rate of exports, the ratio of M2 to foreign reserves, and the annual growth rate of M2 to foreign reserves, the annual growth rate of M2 multiplier, the annual growth rate of domestic credit over GDP, real interest rate and real exchange rate overvaluation. To reduce the impact of extreme values, all variables are dampen using the formula $f(x_t) = \text{sign}(x_t)\ln(1+|x_t|)$.

3.2 Definition of Currency Crises

Following Candelon et al. (2010a), we use the KLR modified pressure index

$$\text{KLRm}_{n,t} = \frac{\triangle e_{n,t}}{e_{n,t}} - \frac{\sigma_e}{\sigma_r}\frac{\triangle r_{n,t}}{r_{n,t}} + \frac{\sigma_e}{\sigma_i}\triangle i_{n,t},$$

where $e_{n,t}$ denotes the exchange rate, $r_{n,t}$ denotes the foreign reserves of country n in period t, while $i_{n,t}$ denotes the interest rate of country n at time t. σ_x denotes the standard deviation of the growth rate of variable x, where x denotes the corresponding variable. A currency crisis

*Data is provided by Candelon et al. (2010a).

†Argentina, Brazil, Chile, Indonesia, Israel, Malaysia, Mexico, Morocco, Peru, Philippines, South Korea, Turkey, Thailand, Uruguay and Venezuela.

is defined as the pressure index exceed two standard deviation above the mean:

$$\text{Crisis}_{n,t} = \begin{cases} 1, & \text{if } \text{KLRm}_{n,t} > 2\sigma + \mu \\ 0, & \text{otherwise.} \end{cases}$$

Using such definition is appropriate if one views crises from the standpoint of a policymaker, who is interested in both successful and unsuccessful speculative attacks.

3.3 Repetition of Similar Early Warning Signals

Given an estimation of the Logit model $\Pr(Y|X_t) = F(\beta X_t)$ from historical data, suppose we have received similar early warning signals continuously, that means we have received similar observations X for several consecutive periods. This will give us the same estimated probabilities for a number of consecutive periods.

Consider the following imaginary scenarios demonstrated in Figure 2. In Scenario 1, the estimated probability remains at 20%, except a spike of 40%. In scenario 2, the estimated probability remains at 40%. In scenario 1, the 40% estimated probability may be generated by some random events. In scenario 2, the probability remains high at 40% consecutively. The repetition of similar early warning signals may indicate some fundamental problems with the economy. The alarming signal should be stronger as the repetition goes on. With the traditional Logit-EWS models, the repetition is not taken in consideration so that the two scenarios can not be distinguished. If the threshold is set above 40%, the warning signals will be missed in both scenarios, no matter how many time the early warning signals are repeated. If the threshold is set lower that 40%, then alarm will be issued in both scenarios. Neither way is desirable.

The BQCD method can appropriately account for the repetition of early warning signals. As can be seen in Figure 2, in Scenario 2, the posterior probability labeled by Bayesian Quickest Change Detection is getting higher as long as we see the same warning signals repeatedly. An alarm will eventually be issued. If the threshold is set higher, more repetitions will be needed. Bayesian updating will enhance repetitive warning signals, which will make detecting the warning signs easier. This is one of the reason why Bayesian Quickest Change Detection method can perform better that traditional Logit-EWS models. One might ask why is being sensitive to repetition important. Notice the aim of both methods is to detect the change of the underlying data generating process of warning signals. If the data generating process has changed, not only we will observe data that supports the alternative, but we will observe that more often. The state-of-the-art BQCD method is keen to manifest the increase in the frequency of data that supports the alternative. This ability does not render the BQCD method perform marginally better, but it is of fundamental importance to the task of quickest change detection.

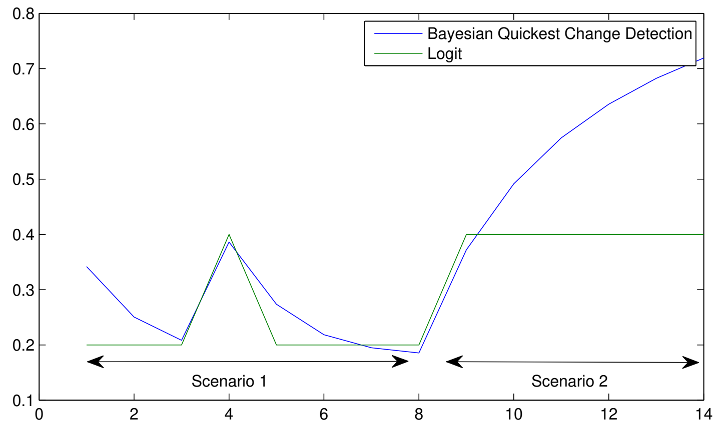

Figure 2: A Demonstration

4 Simulation Exercise

We will first resort to simulation exercises to compare the performance between the Logit-EWS model and BQCD model. In this exercise, we simulate a random independent sequence $\{Z_k; k = 1, 2, ..., 200\}$ with $Z_1, Z_2, ..., Z_{100}$ being i.i.d. with marginal distribution $\mathcal{N}(0, 0.9)$, and $\{Z_{101}, ..., Z_{200}\}$ being i.i.d. with marginal distribution $\mathcal{N}(0.12, 0.9)$. The marginal distribution of Z_k change at 101. We create independent variable $Y_1 = 0, ..., Y_{100} = 0, Y_{101} = 1, ..., Y_{200} = 1$. We estimate the Logit model $\Pr(Y = 1) = F(Z\beta)$. With these simulations, we compare the performance of the traditional EWS model and the BQCD model to identify the structural break.

To evaluate the performances of the two methods, given a threshold, we record the length of false alarm or the occurrence of delay detection. Notice, for each simulation, only one type of mistake can be made. For each threshold given, we repeat the simulations a large number of times and calculate the average length of false alarm and the proportion of delay detection. The results are recorded in Table 1. As can be seen, the traditional EWS model results in lengthy false alarms, where for BQCD, the average lengths of false alarms are much smaller for all levels of thresholds.

Figure 3 shows us why BQCD model performs better. From the top panel of Figure 3, we can see the signal is too noisy to make any decision sequentially. For a wide range of thresholds, lengthy false alarms are inevitable. As can be seen from the lower panel of Figure 3, before 100, BQCD model filters out noisy signals, so the posterior probability remains low; after 100, the Logit model starts to show higher probability of the occurrence of the structural break. We can see from the top penal, from 100-110, the estimated probability remains over 80 percent for about 10 consecutive observations. The traditional EWS model would miss that

Table 1: Summary of Simulation Results

	Logit-EWS			BQCD		
	Average Length of False Alarm	Proportion of Delay Detection	Welfare[a]	Average Length of False Alarm	Proportion of Delay Detection	Welfare
Threshold=0.5	98	0	175	17.14	0.72	31.36
Threshold=0.7	94.52	0	168.78	2.09	0.96	4.7
Threshold=0.8	82.53	0	147.37	1.26	0.98	3.23
Threshold=0.9	35.33	0.36	63.45	0	1	1
Threshold=0.95	18	0.8	32.94	0	1	1

[a]Welfare is weighted average of the Average Length of False Alarm and the Proportion of Delay Detection according to Equation 2.7.

information, the BQCD model will pick that up quickly. So with a wide range of thresholds, the BQCD model will issue alarms with delays less than 10 observations. With the welfare measure proposed earlier in this paper, the BQCD model outperforms the traditional EWS model quite convincingly in this simulation exercise.

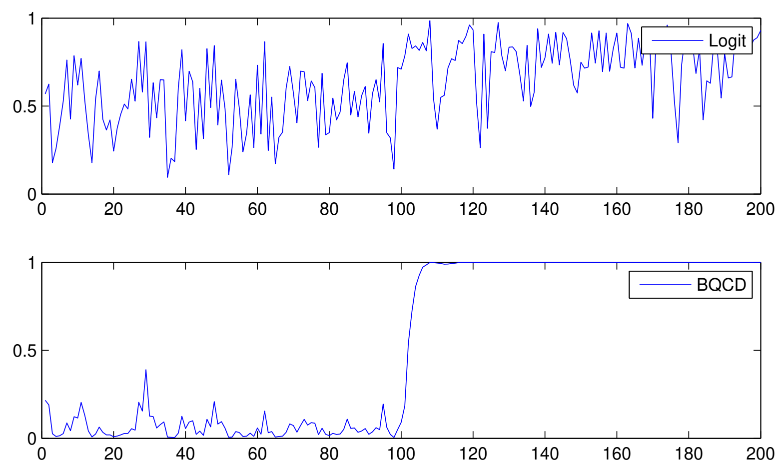

Figure 3: Simulation Result

5 Results

This section shows the results of this study. We evaluate the out-of-sample performance of the Logit model and the BQCD method to predict 1997 Asian financial crises. The Logit model is estimate with data cover from Jan 1986 through Apr 1995. The out-of-sample cover from

May 1995 through Dec 1998. As it can be see in Figure 4, for appropriately chosen threshold, both models can predict the crisis within 24 months in advance. But for the Logit model, if the threshold is set above certain level, say 60%, then the warning signals in Malaysia and Philippines will be entirely missed. If the threshold is set above 70%, then the warning signals will be missed for all five countries under analysis. The reason for this is the repetition of strong warning signals are not accounted. Just as we showed the in Figure 2, if the threshold is set too high, the warning signals will be missed, no matter how many time the warning signals are repeated, they will be missed altogether. On the other hand, this is not the case for BQCD model, since the repetition of strong warning signals are accounted by Bayesian updating, the posterior probability is driven up to almost 100% as it is closer to the financial crises. In this case, no matter how high the threshold is, it is almost impossible to miss the warning signal prior to the crises in all five countries under analysis. Given the uncertainty involved in setting the threshold, this is a desirable feature. In some cases, alarm can be issued earlier than the Logit model given the same threshold in certain countries.

We summarize some statistics in Table 2. We are particularly interested in the number of months in advance of the actual crisis that the alarm is issued. As can be seen from the table, when the thresholds are lower than 50%, both models can predict the crisis, but BQCD model will trigger the alarm earlier for all five countries. Most importantly, if the threshold is set above 60%, the Logit-EWS model will miss warning signals in Malaysia. Once the threshold is set above 80%, the Logit-EWS model will fail to issue alarms in all five countries in the analysis. In this regard, the BQCD model out-performs the Logit-EWS model. As we argue before, the BQCD method is keen to detect the increase in the data that support the alternative. This is very important to the design of EWS, since we are trying to detect the vulnerability of the economy that will lead to speculative attacks. When the economy has indeed became more vulnerable, the correctly chosen early warning signals will not only become stronger, but stronger warning signals will also appear more frequently. This is indeed what has happened before the 1997 Asian financial crises. As can be seen from the five Asian countries in the analysis, prior to the crises, stronger warning signals do tend to appear more often. To be able to detect the increase in the frequency is of fundamental importance.

6 Conclusion

In this paper, we consider the optimal design of Early Warning Systems form a policy maker's perspective. We assume the behavior of a collection of early warning signals changes prior to a financial crisis. The policy maker's problem is to detect the change as soon as possible and take preemptive actions. We employ the state-of-the-art Bayesian Quickest Change Detection method to solve the policy maker's problem. Since the problem of EWS is framed from the policy maker's perspective, the probability threshold is determined by minimizing the welfare

Table 2: Comparisons of Out-of-Sample Predictions Between Logit-EWS and BQCD

Threshold	0.5		0.6		0.7		0.8		0.9	
	L-EWS	BQCD	L-EWS	BQCD	L-EWS	BQCD	L-EWS	BQCD	L-EWS	BQCD
Indonesia	8[a]	9	2	8	0[b]	7	0	7	0	6
Korea	5	12	4	5	1	5	0	5	0	4
Malaysia	1	4	0	4	0	3	0	2	0	1
Philippines	5	5	1	4	0	4	0	4	0	3
Thailand	7	7	6	7	1	6	0	6	0	5
Average[c]	5.2	7.4	2.6	5.6	0.4	5	0	4.8	0	3.8

[a]This means the alarm is issued 8 months before the crisis.
[b]Zeros indicate failures to issue alarms before the crisis.
[c]The average is takes across countries.

loss of the policy maker. We also argue that prior to a financial crisis, not only the warning signals are getting stronger, but also stronger warning signals will appear more frequently. The increase in frequency of early warning signals has been ignored in previous EWS models. Increase in the frequency of the early warning signals contains important information about the well-being of the economy: if the same warning signals appears more frequently, it is a stronger indication that there are some fundamental problems with the economy. The BQCD method can pick this information up appropriately. We compare the performance of the Logit-EWS model and the BQCD model with simulation exercises and out-of-sample prediction of the 1997 Asian financial crises, the results show that BQCD method outperform the traditional EWS model.

A Appendix

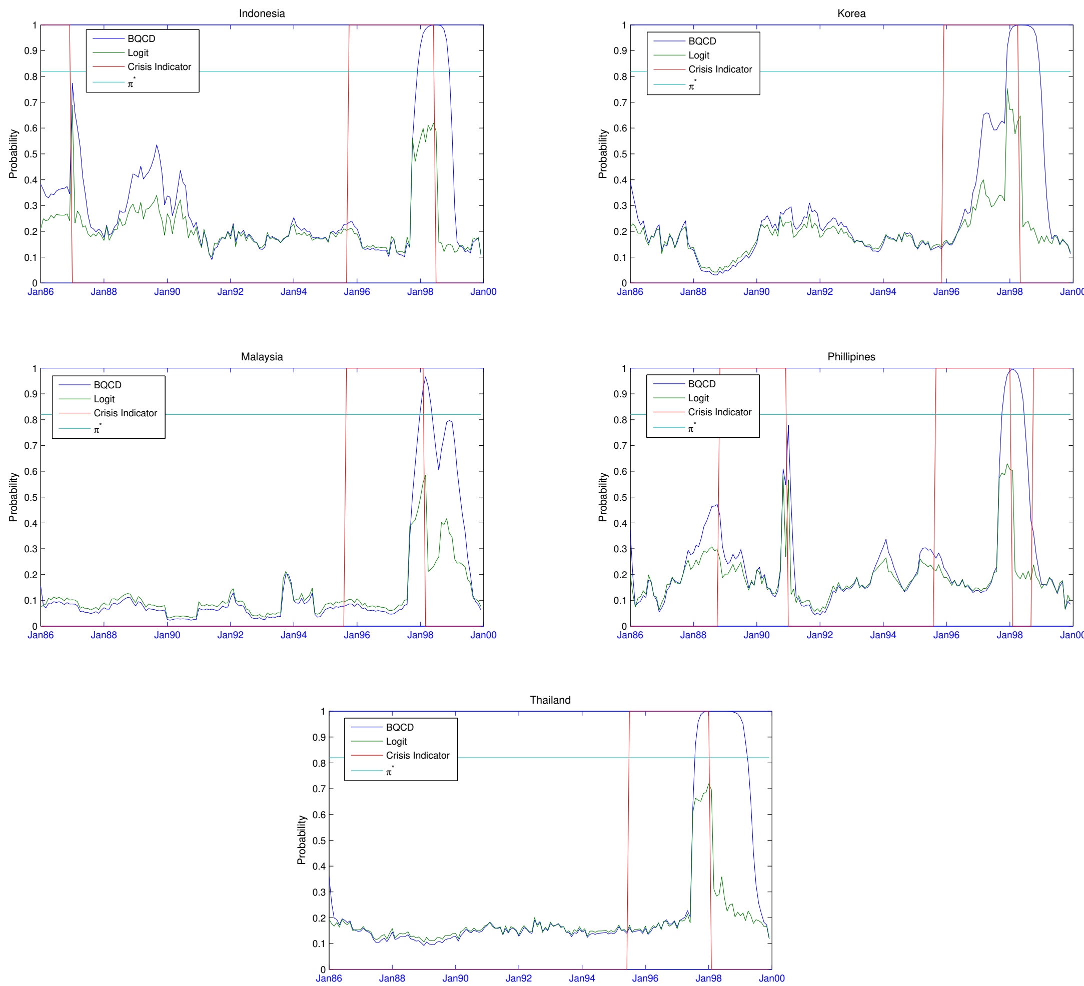

Figure 4: Predicted Probability of Crisis-Out-of-Sample with $\pi^* = 0.82$

B Appendix

Given deterministic k and l, let

$$v(k,l) = v_{-}(k,l)\, 1_{\{k \leq l\}} + v_{+}(k,l)\, 1_{\{k>l\}}$$

More precisely, in the domain of false alarm, when $k \leq l$

$$
\begin{aligned}
v_-(k,l) &= E_-\left\{\int_0^k e^{-\rho t}L(\xi_t,u_{0t})\,dt + \int_k^{k+h} e^{-\rho t}L(\xi_t,u_{1t})\,dt\right\} \\
&\quad + E_-\left\{\int_{k+h}^{l+h} e^{-\rho t}L(\xi_t,u_{2t})\,dt + \int_{l+h}^{\infty} e^{-\rho t}L(\xi_t,u_{2t})\,dt\right\} \\
d\xi_t &= k_0(\xi_t,u_{0t})\,dt + \sigma dW_t,\ t \le k \\
d\xi_t &= k_0(\xi_t,u_{1t})\,dt + \sigma dW_t,\ k < t \le k+h \\
d\xi_t &= k_0(\xi_t,u_{2t})\,dt + \sigma dW_t,\ k+h < t \le l+h \\
d\xi_t &= k_1(\xi_t,u_{2t})\,dt + \sigma dW_t,\ l+h < t
\end{aligned}
$$

In the domain of delayed detection, when $l+h \ge k > l$

$$
\begin{aligned}
&v_+(k,l) = E_+\left\{\int_0^k e^{-\rho t}L(\xi_t,u_{0t})\,dt + \int_k^{l+h} e^{-\rho t}L(\xi_t,u_{1t})\,dt + \int_{l+h}^{\infty} e^{-\rho t}L(\xi_t,u_{1t})\,dt\right\} \\
&\quad d\xi_t = k_0(\xi_t,u_{0t})\,dt + \sigma dW_t,\ t \le k \\
&\quad d\xi_t = k_0(\xi_t,u_{1t})\,dt + \sigma dW_t,\ k < t \le l+h \\
&\quad d\xi_t = k_1(\xi_t,u_{2t})\,dt + \sigma dW_t,\ l+h < t
\end{aligned}
$$

Where E_- and E_+ are defined with respect to probability measure constructed through Radon-Nikodým derivative (2.3) and (2.4) respectively. Next we need to linearized $v_+(k,l)$ around $k = l+\delta$, since $v_+(k,l)$ is defined on $k > l$, $v_-(l,l)$ is not defined. We need to choose an approximation, let $\delta > 0$ be an arbitrarily small positive number, then $v_+(k,l)$ can be linearized on $l^+ = l+\delta$, for $k \le i^-$.

Given $\theta = l$ and $\delta > 0$ for $l+h \ge k \ge l^+$, and let $l^+ = l+\delta$, we have*

$$
v_+(k,l) - v_-(l,l) = v_+(k,l) - v_+\left(l^+,l\right) + v_+\left(l^+,l\right) - v_-(l,l)
$$

$$
\begin{aligned}
v_+(k,l) - v_+\left(l^+,l\right) &\approx v'_+\left(l^+,l\right)\left(k-l^+\right)^+ \\
v_+\left(l^+,l\right) - v_-(l,l) &\approx v'_+\left(l^+,l\right)\delta,
\end{aligned}
$$

hence we will have

$$
v_+(k,l) - v_-(l,l) \approx v'_+\left(l^+,l\right)\left(k-l^+\right)^+ - v'_+\left(l^+,l\right)\delta,
$$

for $k \le l$, we have

$$
v_-(k,l) - v_-(l,l) \approx v'_-(l,l)(k-l)^+ .
$$

*$x^- = -\inf\{x,0\}$ and $x^+ = \sup\{x,0\}$.

Hence,

$$\begin{aligned} v(k,l) - v_-(l,l) &= [v_-(k,l) - v_-(l,l)]\, 1_{\{k\le l\}} + [v_+(k,l) - v_-(l,l)]\, 1_{\{k>l\}} && (2.15)\\ &\approx -v'_-(l,l)(l-k)^+ + v'_+\left(l^+,l\right)\left(k-l^+\right)^+ + v'_+\left(l^+,l\right)\delta 1_{\{k>l\}} && (2.16)\\ &= -v'_-(l,l)(l-k)^+ + v'_+\left(l^+,l\right)(k-l)^+ && (2.17)\end{aligned}$$

When $h=1$, then the above equation is equal to

$$v(k,l) - v_-(l,l) \approx -v'_-(l,l)(l-k)^+ + v'_+\left(l^+,l\right) 1_{\{k>l\}}.$$

where $v'_-(l,l)$ denote the left derivative of $v_-(k,l)$ with respect to k evaluated at l.

Given θ, and a stopping time $T\in\mathcal{T}$, where $\mathcal{T}$ is the class of all stopping times, define a sequence of random times by

$$T_n(\omega) = \begin{cases} T(\omega); & \textit{on } \{\omega;\, T(\omega)=+\infty\} \\ \frac{m}{2^n}; & \textit{on } \left\{\omega;\, \frac{m-1}{2^n} \le T(\omega) < \frac{m}{2^n}\right\} \end{cases},$$

T_n is a stopping time and $\lim_{n\to\infty} T_n = T$*.

Without loss of generality, assume $h=1$. Conditional on $\theta=l$ and given $\delta>0$, we have the following:

$$\begin{aligned} E_\theta\{v(T,l) - v(l,l)\} &= \lim_{n\to\infty}\sum_{m=0}^{\infty} P_{\theta,n}\left(T_n = \frac{m}{2^n}|l\right)\left[v\left(\frac{m}{2^n},l\right) - v(l,l)\right]\\ &\approx \lim_{n\to\infty}\sum_{m=0}^{\infty} P_{\theta,n}\left(T_n = \frac{m}{2^n}|l\right)\left[-v'_-(l,l)\left(l-\frac{m}{2^n}\right)^+ + v'_+\left(l^+,l\right)1_{\left\{\frac{m}{2^n}>l\right\}}\right]\\ &= E_\theta\left[-v'_-(l,l)(l-T)^+ + v'_+\left(l^+,l\right)1_{\{T>l\}}\right]\end{aligned}$$

where E_θ is expectation conditional on θ, and P_θ is the probability distribution given θ, defined by (2.5).

$$\begin{aligned} E_\pi\{E_\theta[v(T,l) - v(l,l)]\} &\approx E_\pi\left\{E_\theta\left[-v'_-(l,l)(l-T)^+ + v'_+\left(l^+,l\right)1_{\{T>l\}}\right]\right\}\\ E\{v(T,\theta) - v(\theta,\theta)\} &\approx E\left\{-v'_-(\theta,\theta)(\theta-T)^+ + v'_+\left(\theta^+,\theta\right)1_{\{T>\theta\}}\right\} && (2.18)\end{aligned}$$

here E is the expectation under probability measure P, defined by (2.6). The left side of the approximation is the policymaker's objective as a function of stopping time, which can be approximated by the left hand side of the approximation.

*Please see Problem 2.24 in Karatzas and Shreve (1991)

C Appendix

$$\inf_{T\in\mathcal{T}} E\left\{c\left(\theta-T\right)^{+}+1_{\{T>\theta\}}\right\}$$

Let $E\left\{\left(\theta-T\right)^{+}\right\}=E\left\{C_T\right\}$

$$\begin{aligned}
C_k &= \sum_{m=k}^{\infty}\left(m-k\right)P\left(\theta=m|\mathcal{F}_k\right)\\
&= \sum_{m=k}^{\infty}P\left(\theta>m|\mathcal{F}_k\right)\\
&= \sum_{m=k}^{\infty}\left[1-P\left(\theta\le m|\mathcal{F}_k\right)\right]\\
&= \sum_{m=0}^{\infty}\left[1-P\left(\theta\le m|\mathcal{F}_k\right)\right]-\sum_{m=1}^{k-1}\left[1-P\left(\theta\le m|\mathcal{F}_k\right)\right]\\
&= \sum_{m=0}^{\infty}P\left(\theta>m|\mathcal{F}_k\right)-\sum_{m=0}^{k-1}\left[1-P\left(\theta\le m|\mathcal{F}_k\right)\right]\\
&\quad+\sum_{m=0}^{k-1}\left[1-P\left(\theta\le m|\mathcal{F}_m\right)\right]-\sum_{m=0}^{k-1}\left[1-P\left(\theta\le m|\mathcal{F}_m\right)\right]\\
&= \sum_{m=0}^{\infty}mP\left(\theta=m|\mathcal{F}_k\right)-\sum_{m=0}^{k-1}\left[1-P\left(\theta\le m|\mathcal{F}_m\right)\right]+N_k
\end{aligned}$$

we can show that

$$E\left\{c\left(\theta-T\right)^{+}\right\}=E\left\{c\theta-\sum_{m=0}^{T-1}c\left(1-\pi_m\right)\right\}$$

Since we have

$$E\left\{1_{\{T>\theta\}}\right\}$$

Let

$$\begin{aligned}
C_k &= E\left\{1_{\{k>m\}}\right\}\\
&= \sum_{m=0}^{k-1}P\left(\theta=m|\mathcal{F}_k\right)\\
&= P\left(\theta\le k-1|\mathcal{F}_k\right)+P\left(\theta\le k-1|\mathcal{F}_{k-1}\right)-P\left(\theta\le k-1|\mathcal{F}_{k-1}\right)\\
&= \pi_{k-1}+P\left(\theta\le k-1|\mathcal{F}_k\right)-P\left(\theta\le k-1|\mathcal{F}_{k-1}\right)
\end{aligned}$$

The rest is easy to show that

$$E\left\{1_{\{T>\theta\}}\right\} = E\left\{\pi_{T-1}\right\}$$

$$E\left\{c\theta - \sum_{m=0}^{T-1} c\left(1-\pi_m\right) + \pi_{T-1}\right\}$$

D Appendix

Given the condition in Lemma 2 is satisfied, the period loss function can then be expressed as

$$\mathbb{L}\left(1_{\{t\geq k\}}, 1_{\{t\geq l\}}\right) = E_\theta\left(L_t\left(k,l\right)\right),$$

For $k \geq l^+$, where $l^+ = l + \delta$,

$$\begin{aligned}
v_+\left(k,l\right) - v_+\left(l^+,l\right) &= E\left\{\int_0^\infty e^{-\rho t} L_t\left(k,l\right) dt - \int_0^\infty e^{-\rho t} L_t\left(l,l\right) dt\right\} \\
&= \int_0^\infty e^{-\rho t}\mathbb{L}\left(1_{\{t\geq k\}}, 1_{\{t\geq l\}}\right) dt - \int_0^\infty e^{-\rho t}\mathbb{L}\left(1_{\{t\geq l\}}, 1_{\{t\geq l\}}\right) dt \\
&= \int_k^{l+h} e^{-\rho t}\left[\mathbb{L}\left(0,1\right) - \mathbb{L}\left(1,1\right)\right] dt \\
&= \frac{e^{-\rho(l+h-k)}}{-\rho}\left[\mathbb{L}\left(0,1\right) - \mathbb{L}\left(1,1\right)\right]
\end{aligned}$$

$$\begin{aligned}
v'_+\left(l^+,l\right) &= \lim_{k\to l^+} \frac{e^{-\rho(l+h-k)}}{-\rho\left(l+h-k\right)}\left[\mathbb{L}\left(0,1\right) - \mathbb{L}\left(1,1\right)\right] \\
&= \frac{e^{-\rho(\delta+h)}}{-\rho\left(\delta+h\right)}\left(\mathbb{L}\left(0,1\right) - \mathbb{L}\left(1,1\right)\right),
\end{aligned}$$

hence $v'_+\left(l^+,l\right)$ is a constant. Similar result can be shown for $v'_-\left(l,l\right)$, hence $c = -\frac{v'_-(\theta,\theta)}{v'_+(\theta^+,\theta)}$ is a constant.

E Appendix

Period loss function $L\left(1_{\{t\geq k\}}, 1_{\{t\geq l\}}\right)$, then

$$\begin{aligned}
v_-\left(l,l\right) - v_-\left(l-1,l\right) &= \sum_{t=0}^{\infty} \beta^t\left\{L\left(1_{\{t\geq l\}}, 1_{\{t\geq l\}}\right) - L\left(1_{\{t\geq l-1\}}, 1_{\{t\geq l\}}\right)\right\} \\
&= \beta^{l-1}\left\{\underbrace{L(0,0)}_{t=l-1} - \underbrace{L(1,0)}_{t=l-1}\right\}
\end{aligned}$$

$$\begin{aligned} v_+(l+1,l) - v_+(l,l) &= \sum_{t=0}^{\infty} \beta^t \left\{ L_t\left(1_{\{t \geq l+1\}}, 1_{\{t \geq l\}}\right) - L_t\left(1_{\{t \geq l\}}, 1_{\{t \geq l\}}\right) \right\} \\ &= \beta^l \left\{ L(0,1) - L(1,1) \right\} \end{aligned}$$

Finally, we have

$$\begin{aligned} c &= -\frac{v_-(l,l) - v_-(l-1,l)}{v_+(l+1,l) - v_+(l,l)} \\ &= -\beta^{-1} \frac{L(0,0) - L(1,0)}{L(0,1) - L(1,1)}. \end{aligned}$$

Clearly, c is not a function of l.

F Appendix

Table 3: Logit Maximum Likelihood Estimates

Variables	Coefficient	t-statistics	t-probability
Constant	-1.218744	-6.424941	0.000000
Growth of International Reserve	-1.059737	-4.916337	0.000001
Growth of Imports	0.005356	0.016094	0.987162
Growth of Exports	-0.529670	-1.348524	0.177673
Growth of Domestic Credit Over GDP	1.862253	6.982468	0.000000
Real Interest Rate	0.078776	1.894862	0.058283
M2 to Reserve	0.211484	12.094560	0.000000
Growth of M2 to Reserve	-0.000751	-0.020199	0.983887
Real Exchange Rate	-3.677564	-5.164558	0.000000
Real Exchange Rate Over-evaluation	3.655495	5.252078	0.000000
Exchange Market Pressure Index	1.667829	4.906118	0.000001
McFadden R-squared	0.2171		
Estrella R-squared	0.2463		
Log-Likelihood	-759.5282		
Nobs, Nvars	1680, 11		

References

Ball, Laurence, "Policy Rules for Open Economies," in John Taylor, ed., *Monetary Policy Rules*, The University of Chicago Press, 1999.

Berg, Andrew and Catherine Pattillo, "Predicting currency crises: The indicators approach and an alternative," *Journal of International Money and Finance*, 1999, *18* (4), 561 – 586.

Bussiere, Matthieu and Marcel Fratzscher, "Low probability, high impact: Policy making and extreme events," *Journal of Policy Modeling*, 2008, *30* (1), 111–121.

Candelon, Bertrand, Elena-Ivona Dumitrescu, and Christophe Hurlin, "Currency Crises Early Warning Systems: why they should be Dynamic," Research Memoranda 047, Maastricht : METEOR, Maastricht Research School of Economics of Technology and Organization 2010.

_ , _ , and _ , "How to evaluate an Early Warning System? Towards a United Statistical Framework for Assessing Financial Crises Forecasting Methods," Research Memoranda 046, Maastricht : METEOR, Maastricht Research School of Economics of Technology and Organization 2010.

Davis, E. Philip and Dilruba Karim, "Comparing early warning systems for banking crises," *Journal of Financial Stability*, June 2008, *4* (2), 89–120.

Gramlich, Dieter, Gavin L. Miller, Mikhail V. Oet, and Stephen J. Ong, "Early Warning Systems for Systemic Banking Risk: Critical Review and Modeling Implications," *Banks and Bank System*, 2010, *5*, 199–211.

IMF, "The IMF-FSB Early Warning Exercise: Design and Methodological Toolkit," International Monetary Fund 2010.

Kaminsky, Graciela L. and Carmen M. Reinhart, "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems," *American Economic Review*, June 1999, *89* (3), 473–500.

Kaminsky, Graciela, Saul Lizondo, and Carmen M. Reinhart, "Leading Indicators of Currency Crises," *IMF Staff Papers*, March 1998, *45* (1), 1–48.

Karatzas, Ioannis and Steven E. Shreve, *Brownian motion and stochastic calculus*, second ed., Vol. 113 of *Graduate Texts in Mathematics*, New York: Springer-Verlag, 1991.

Polunchenko, Aleksey S and Alexander G Tartakovsky, "State-of-the-Art in Sequential Change-Point Detection," *Sequential Analysis*, 2011, *29* (2), 34.

Shewhart, W A, *Economic Control of Quality of Manufactured Product*, Vol. 27, D. Van Nostrand Company, Inc, 1931.

Shiryaev, A. N., "On Optimum Methods in Quickest Detection Problems," *Theory of Probability and its Applications*, 1963, *8* (1), 22–46.

Shiryaev, Albert N., *Optimal Stopping Rules*, Springer-Verlag, 1978.

Wald, Abraham, *Sequential Analysis*, New York: John Wiley & Sons Inc., 1947.

Zacks, S., "Detection and change-point problems," in "Handbook of sequential analysis," Vol. 118 of *Statist. Textbooks Monogr.*, New York: Dekker, 1991, pp. 531–562.

Zampolli, Fabrizio, "Optimal monetary policy in a regime-switching economy: The response to abrupt shifts in exchange rate dynamics," *Journal of Economic Dynamics and Control*, 2006, *30* (9-10), 1527–1567.