THREE ESSAYS ON MONETARY POLICY UNDER UNCERTAINTY

by

Haixi Li

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Economics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2012

Date of final oral examination: 06/01/2012

The dissertation is approved by the following members of the Final Oral Committee:
Steven Durlauf, Professor, Economics
William Brock, Professor, Economics
Noah Williams, Professor, Economics
Kenichi Fukushima, Assistant Professor, Economics
Xiaodong Du, Assistant Professor, Agricultural and Applied Economics

Acknowledgment

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.

Table of Contents

Abstract

This dissertation consists of three self-contained essays on monetary policy making under uncertainty.

In the first chapter, I propose a new optimal design of Early Warning Systems (EWS) to detect early warning signals of an impending financial crisis. The problem of EWS is formulated from a policy maker's perspective. Hence the probability threshold is obtained by minimizing the policy maker's welfare loss. This paper employs the state-of-the-art Bayesian Quickest Change Detection (BQCD) as the methodology to detect the early warning signals as soon as possible. We show that the BQCD method outperforms the Logit model used in traditional EWS models based on results of simulation exercise and the out-of-sample prediction of the 1997 Asian financial crises.

In the second chapter, I investigated the timing of U.S. monetary policy response to the 2007 financial crisis. In the presence of financial turmoil, monetary policy needs to deviate from standard Taylor rule to credit-spread adjusted Taylor rule. The timing of the switching is crucial, since switching too early will introduce unnecessary noise into the system, and switching too late will result in further deterioration of the economy. The optimal timing will strike a balance between the two types of mistakes. The LIBOR-OIS spread jumped significantly on Aug 9, 2007, indicating an unusual increase in distress of the financial sector, but it was not until Aug 17, 2007, that the Federal Reserve responded with a 50 basis points of reduction of the primary credit rate after an unscheduled meeting. Assuming that the policy maker knew the intensity of the coming financial crisis, I found that the optimal timing should be Aug 9, 2007, contradicting to the observed response time of Aug 17, 2007. I further assumed that the policy maker was uncertain about the intensity of the financial crisis, and that monetary policy responded to more severe financial crisis more intensively. Hence, if the policy maker chose the wrong policy regime, there would be costs. In order to increase the accuracy of choosing the right policy regime, waiting for more information may be desirable. I found that for certain specifications of the intensity of the financial crisis, the optimal timing of the Federal Reserve's policy response should be Aug 15, 2007. I concluded that uncertainty about the intensity of the financial crisis played an important role in the timing decision of the policy maker.

In the third chapter, I investigated the issue of international monetary policy coordination under uncertainty. The consensus is that international monetary policy coordination is welfare improving, but some argue that the improvement is not significant

quantitatively. This paper studied the role of model uncertainty in international monetary policy coordination, and found that considering model uncertainty can enhance welfare gain of coordination. This was because without coordination, due to information asymmetry, domestic policy maker would be more concerned about model misspecification of the foreign economy. Under coordination, with full information sharing between domestic policy maker and the foreign policy maker, this concern would vanish. Assuming that the domestic policy maker adopted robust control to deal with model misspecification, coordination would enhance welfare gains.

Chapter 1

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

1.1 Introduction

1.1.1 Motivation

In the past few decades, we saw a large number of financial disruptions of different sorts often with devastating economic and social consequences. As a result, in the 90s, international organizations begun to develop Early Warning System (EWS) models with the aim of anticipating whether and when a country may be affected by a financial crisis. More so, the 2007 financial crisis has renewed the calls for EWS. 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 early warning system constitutes a crucial part of integrated measures taken by authorities to prevent financial crises and facilitate financial stability.

Two main approaches have 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, and 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. $Pr\left(Y=1\right)=F\left(X\beta\right)$. 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. Criteria such as the Noise-to-Signal Ratio have been used. But this criterion is based on statistical measures, which do not represent the costs of false alarm and delay detection. As Bussiere and Fratzscher (2008) and Candelon et al. (2010b) pointed out 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 the 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 threshold 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: $L\left(T\right)=\theta p\left(T\right)+\left(1-\theta\right)q\left(T\right)$, 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. Parameter 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.

In this paper, I 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 these two types of mistakes: false alarm and delay detection. As I 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. I are going to adopt this method to solve the policy maker's problem. I will also show the new design of EWS with BQCD method outperform the Logit-EWS model.

1.1.2 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 the 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 led to a large volume 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:

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

$P(T < \theta)$ is the frequency of false alarm and $E(T - \theta)^+$ 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 policy-maker needs to detect the change as soon as it happens and take preemptive actions to attenuate the risk of an impending financial crisis.

1.2 The Model

There are two regimes of the economy: the normal regime and the regime with a financial crisis. The economy starts with the normal regime and jumps into a financial crisis at some random time ν. Before ν, the pattern of the behavior of a collection of early warning signals will change. The policy maker's problem is to monitor the collection of early warning signals, and detect the change in the pattern of the behavior and determine the timing to take preemptive actions.

1.2.1 The Stochastic Process

Assume there is a D-dimensional controlled stochastic process ξ_t with u_t as the control. The process is given by

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

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:

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

As assumed in the EWS literature, there is a collection of early warning signals among process ξ_t, whose behaviors will 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 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 a random and unobservable time $\theta = \nu - h$. More formally, the policy maker is conducting a hypothesis test sequentially:[1]

- 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.

1.2.2 The Sequential Decision Rule

In this subsection, I will specify the policy maker's response to the evolution of information.

It is conventional wisdom that commitment monetary policy making is superior to discretionary monetary policy making. It has been standard to assume that the central

[1] More detailed discussion can be found in Subsection 1.2.6

bank is committed to a policy rule once for all, assuming that the model of the economy does not change. How to specify a commitment policy rule when the environment of the economy changes is an important question.

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

$$u_{T,t} = \begin{cases} u_{0,t} & \text{before T declares stopping} \\ u_{1,t} & \text{after T declares stopping} \end{cases}, \tag{1.2}$$

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 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.

1.2.3 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. I 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 $P_0(W_0 = 0) = 1$. I can define the following Radon-Nikondým derivative and define the corresponding probability measure.

When $k \leq l$, let

$$\frac{dP_-^{k,l}}{dP_0} \equiv \exp\left\{ -\int_0^t \sum_{d=1}^D K_-^{k,l,(d)}(s)\, dW_s^{(d)} - \frac{1}{2}\int_0^t \left\| K_-^{k,l}(s) \right\|^2 ds \right\}; \tag{1.3}$$

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

$$\frac{dP_+^{k,l}}{dP_0} \equiv \exp\left\{ -\int_0^t \sum_{d=1}^D K_+^{k,l,(d)}(s)\, dW_s^{(d)} - \frac{1}{2}\int_0^t \left\| K_-^{k,l}(s) \right\|^2 ds \right\}. \tag{1.4}$$

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 I 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

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

I have that T_n is a stopping time and $\lim_{n\to\infty} T_n = T$[2]. Define probability measure

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

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 $A \in \Lambda$,

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

$P^{k,l}\left(A\right)$ 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 $P^{k,l}\left(A\right)$, I just average out the randomness in stopping time T.

$$P_\theta\left(A\right) = E_\theta\left\{1_{\{A\}}\right\} \tag{1.5}$$

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

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

$$P\left(A\right) = E_\pi\left(E_\theta\left(1_{\{A\}}\right)\right). \tag{1.6}$$

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[3]

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

1.2.4 The Objective Function

In this subsection, I will define the objective function of the policy maker. I will then convert this objective function into an objective function that is of similar form of Equa-

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

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

tion 1.1, then I can solve it as an optimal stopping problem.

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 regime of post crisis. To do so, the policymaker will make two types of mistakes: false alarm, that the policy maker respond before l, i.e. $k \leq l$; delayed detection, that the policy maker respond after l, but before $l+h$, i.e. $l < k \leq l+h$. l is unobservable in case $k \leq l$, 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.

The policy maker's objective can be defined as

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

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

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

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

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

The expectation E_- and E_+ are defined by probability measures defined by Equation 1.3 and 1.4.

The policy maker's problem can be written as

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

The policy maker's objective function $E\left\{v(T,\theta) - v(\theta,\theta)\right\}$ can be approximated by

$$
E\left\{-v'_{-}(\theta,\theta)(\theta - T)^{+} + v'_{+}\left(\theta',\theta\right) h 1_{\{T>\theta\}}\right\}.
$$

Please see A.2.

The policy maker's problem can be approximated as

$$
\inf_{T \in \mathcal{T}} E\left\{-v'_{-}(\theta,\theta)(\theta - T)^{+} + v'_{+}\left(\theta',\theta\right) h 1_{\{T>\theta\}}\right\}.
$$

I can further simplify the objective:

$$
E\left\{\left(v'_{+}\left(\theta^{+},\theta\right)\left[c(\theta - T)^{+} + h 1_{\{T>\theta\}}\right]\right)\right\}.
$$

Define probability distribution function

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

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

$$
\inf_{T \in \mathcal{T}} \left\{\tilde{E}\left[c(\theta - T)^{+}\right] + \tilde{P}(T > \theta)\right\}, \tag{1.7}
$$

where

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

c is the relative cost ratio, it is the ratio of the speed 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.

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.

Please see A.4.

1.2.5 Discrete Time Version

In this subsection, I 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

$$c = -\frac{v_-\left(\theta,\theta\right) - v_-\left(\theta - 1,\theta\right)}{v_+\left(\theta + 1,\theta\right) - v_-\left(\theta,\theta\right)}, \tag{1.8}$$

I simply use growth rate to replace time derivative.

Assume geometric prior distribution

$$p\left(\theta = k\right) = \begin{cases} \pi & \text{if } k = 0 \\ \left(1-\pi\right)\rho\left(1-\rho\right)^{k-1} & \text{if } k = 1, 2, ... \end{cases}. \tag{1.9}$$

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

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

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

Please see A.3.

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

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.

Please see A.5.

1.2.6 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. I 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 estimated from the full sample which is equal to 0.2702.

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

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

1.2.7 Solution Method

The policy maker's problem has been converted to

$$\inf_{T\in\mathcal{T}} \tilde{E}\left\{c\theta - \sum_{m=0}^{T-1} c\left(1-\pi_m\right) + \pi_{T-1}\right\}, \tag{1.10}$$

the prior is given by Equation 1.9.

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

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

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

Then the value function given π is

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

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

and the stopping time

$$T^* = \inf\left\{n \geq 0 : v\left(\pi_n\right) = \pi_n - c\left(1-\pi_n\right)\right\} \tag{1.11}$$

is the optimal stopping time.

1.3 Quantitative Study

In this paper, I employ 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, I borrow the idea of two regimes of the economy to model currency crisis, in order to address the timing issue, I follow the construction of Shiryaev (1963)'s Bayesian Quickest Change Detection problem. I 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:

$$\begin{aligned} y_{t+1} &= \alpha y_t - \beta\left(i_t - \pi_t\right) - \chi a_t + \eta_t & (1.12) \\ \pi_{t+1} &= \delta\pi_t + \gamma y_t - f\left(a_t - a_{t-1}\right) + \varepsilon_t & (1.13) \\ a_t &= \rho_\tau a_{t-1} + \kappa\left(i_t - \pi_t\right) + \upsilon_t. & (1.14) \end{aligned}$$

where

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

Equation 1.12 is the open-economy IS curve, where y_t is output gap, $i_t - \pi_t$ is real interest rate. Equation 1.13 is the open-economy Phillips curve, where π_t is inflation, a_t is real exchange rate. Equation 1.14 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

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

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 are 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

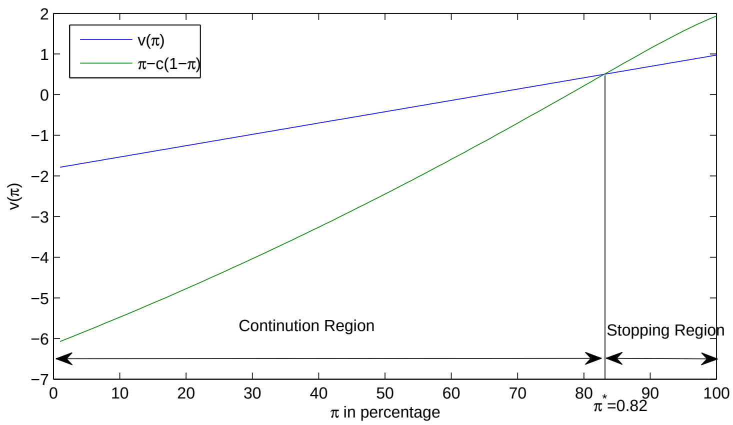


Figure 1.1: The Optimal Stopping Rule

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), I can compute $c = 1.7857$ according to Equation 1.8. The optimal stopping 1.10 can be solved with the method presented is Subsection 1.2.7. The result can be presented in Figure 1.1. As is shown in Figure 1.1, π^* is determined by the optimality condition 1.11. 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.

1.3.1 Data

Most of the data are extracted from IFS database, some are extracted from national banks of the countries under analysis via Datastream.[4] It consists of monthly observations from Jan 1986 through Dec 1998 for 15 emerging countries.[5]

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

[4]Data is provided by Candelon et al. (2010a).

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

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 $f(x_t) = \textbf{sign}(x_t) \ln(1 + |x_t|)$.

1.3.2 Definition of Currency Crises

Following Candelon et al. (2010a), I 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 is defined as the pressure index exceed two standard deviation above the mean:

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

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

1.3.3 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 I have received similar early warning signals continuously, that means I 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 1.2. 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 1.2, in Scenario 2, the posterior probability labeled by Bayesian Quickest Change Detection is getting higher as long as I 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 being sensitive to repetition is 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 I will observe data that supports the alternative, but I 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.

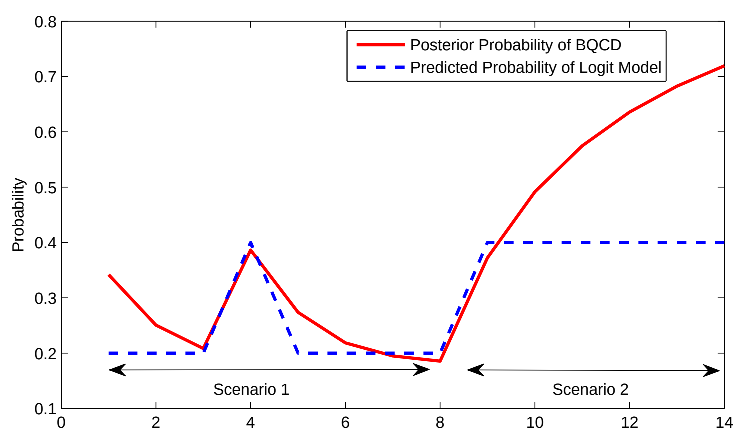


Figure 1.2: A Demonstration

1.4 Simulation Exercise

I will first resort to simulation exercises to compare the performance between the Logit-EWS model and BQCD model. In this exercise, I 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. I create independent variable $Y_1 = 0, ..., Y_{100} = 0, Y_{101} = 1, ..., Y_{200} = 1$. I estimate the Logit model $\Pr(Y = 1) = F(Z\beta)$. With these simulations, I 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, I 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, I 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 1.1. 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.

Table 1.1: 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 1.7.

Figure 1.3 shows us why BQCD model performs better. From the top panel of Figure 1.3, I 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 1.3, 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. I 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 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.

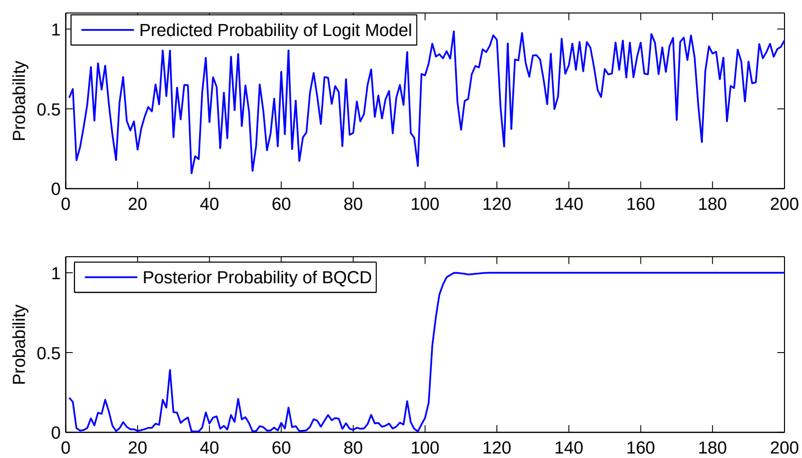


Figure 1.3: Simulation Result

1.5 Results

This section shows the results of this study. I 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 can be see in Figure A.1, for an 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 a 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 is not accounted for. Just as I showed in Figure 1.2, 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.

I summarize some statistics in Table 1.2. I 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 I 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 I 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.

Table 1.2: 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.

1.6 Conclusion

In this paper, I consider the optimal design of Early Warning Systems from a policy maker's perspective. I 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. I 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 loss of the policy maker. I 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. I 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.

Chapter 2

Did the Federal Reserve Respond Too Late to the 2007 Financial Crisis?

2.1 Introduction

2.1.1 Motivation

On Aug 9, 2007, the spread between 3-month LIBOR (London Inter-bank Offered Rate) and OIS (Overnight Index Swaps) jumped unusually by 21 basis points, indicating an unusual increase in distress of the interbank lending market. This event took both market participants and central banks by surprise. Taylor and Williams (2008) refer to this highly unusual event as a "black swan" in the money market; Cecchetti (2008) takes this day as the beginning of the financial crisis.

This event can be categorized as a low-probability extreme event. Other examples include a large and sudden depreciation of a currency, or an oil price hike. How to deal with such low probability events has been a real challenge to the central banks. Freedman (2002) comments on Svensson (2002):

> There are two points about this paper I'd like to make. The first relates to the mean versus mode debate, and I'd like to tie that to the asset price issue. One of the real challenges to central banks is how to deal with small probability cases. If we have a situation where there is, say, a 10 percent probability of an asset price collapse and a 90 percent probability that it is not going to happen, do you then go ahead and focus on what one would call

> the mode, which is the appropriate path for policy in the 90 percent case and say, "if the 10 percent case happens, we'll try and deal with it later."? Or do we try to deal with the mean? In that case if the small probability outcome doesn't happen, you are going to have an interest rate path that is not appropriate. But even if it does happen, you will not have moved interest rates enough to deal with the collapse of asset prices in any case. So, it is very much an open question of how to deal with such a situation.

Nickell (2002) prefers a wait-and-see approach:

> The important question here is should we move interest rates up today in order to forestall the potential inflationary consequences of a significant exchange rate fall......? In my view, the answer is simply no. If, and when, such a fall in the exchange rate comes about, then is the time to make appropriate adjustment, if any, in interest rates. To act pre-emptively on this front is difficult because of the huge uncertainties involved in forecasting exchange rate movements and......

Svensson (2004) shares the same point of view under certain conditions and points out the option value of waiting:

> Clearly, if the extreme event in question is such that there is time to react to it in case it materializes, there is no point in trying to pre-empt the event. Instead, there is an option value of waiting to see if the event materializes and only then take the appropriate action.

Poole (2007) also hints the option value of waiting in monetary policy making:

> When new information arrives, most of the time the central bank can wait for the market to respond and the passage of time to clarify what is happening....
>
> ...The market understands, I believe, that the Fed will act in due time, if and when evidence accumulates that action would be appropriate.

Mishkin (2010) envisions a framework of a systematic approach to manage the macroeconomic risks posed by financial disruptions. In his opinion, the systematic approach of policy response needs to be timely and decisive. First, timely action is crucial since waiting too long will result in further deterioration of the macro economy and increase the cost of easing that would be needed eventually; second, decisive action needs to be taken in response to financial disruption, even though it might lead to less inertia and gradualism than otherwise would be typical.

How should central banks respond to low-probability extreme events? On the normative side, this is very much an open question; on the positive side, what happened in Aug, 2007 provides a natural experiment to study the Federal Reserve's policy response to the low-probability extreme event.

Despite the unusual jump in the LIBOR-OIS spread on Aug 9, 2007 and the panicky demand from the market to lower key interest rates, it was not until Aug 17, 2007 that the Board voted to reduce the primary credit rate by 50 basis points after an unscheduled meeting. John (2007) wrote an article in the Financial Times commenting on the Federal Reserve's decision with the title: "Bernanke acts too late to help Nikkei". On the other hand, Mishkin (2010) argues that the Federal Reserve's response is timely and decisively. Mizen (2008) also states that "On August 17, 2007, the Fed extended its normal lending period to 30 days and cut the interest rate offered to banks at the discount window by 50 basis points, acting swiftly and decisively."

Did the Federal Reserve respond too late? How do we evaluate the merits of such criticism?

The seemingly delayed response on Aug 17, 2007, reveals that there is some interesting dynamics in the timing decision of the Federal Reserve. The option value of waiting might help to explain the discrepancy between Aug 9, and Aug 17, 2007.

From the above discussions, one can derive two main ingredients needed for a systematic approach of policy responses to financial disruptions. First, the timing is a major component of policy response that needs to be chosen optimally according to certain criteria; second, switching from one policy regime that is appropriate in normal time to another that is appropriate in the times of financial distress, needs to be decisive. In order to address the issues above, this paper studies the wait-and-see policy approach proposed above with mathematical tools developed in quickest change detection literature (Poor and Hadjiliadis, 2009; Shiryaev, 1978; Dayanik et al., 2008). With this framework, we model the option value of waiting and its effect on the timing decision of the Federal Reserve. To the best of our knowledge, the timing issue of a policy regime switch and the corresponding option value of waiting have not been rigorously studied.

It is assumed that there is an abrupt change in the model of the economy at some random time in the future which is not observable to the policy maker. There are two regimes of the economy: the normal regime, and the regime with a financial crisis. The economy jumps from the normal regime into regime with financial crisis at some random time θ. There are two sets of parameter values for each regime of the economy; when the regime switch happens, the parameter values of the model change, including the mean of the credit spread. Corresponding to the two regimes of the economy, there are two monetary policy regimes: the standard Taylor rule for the normal regime, and the credit-spread adjusted Taylor rule for the regime with a financial crisis. In the spirit of the wait-and-see approach, given that there is a change in the structure of the economy, the policy regime will switch to adapt to that change at a time that is optimally chosen. We assume that the central bank is committed to a sequential decision rule (T, u_{Tt}), similar to that in Poor and Hadjiliadis (2009, Chapter 4), where T is a stopping time that determines when the policy regime switches and u_{Tt} specifies the policy rules in two regimes[1].

The main results of this paper can be summarized as follows. In the case of a known alternative, that is the policy maker is certain about the intensity of the coming financial crisis, to determine the time of the policy regime switch when the structure of the economy changes at some random, unobservable time in the future, the policy maker will make two types of mistakes: false alarm, when the policy regime switches before the structure of the economy changes, and delayed detection, when the policy regime switches after that. The optimal timing will need to strike a balance between these two mistakes. This paper finds that the optimal timing should be Aug 9, 2007, contradicting to the observed response time of Aug 17, 2007.

We further assume that the policy maker is uncertain about the intensity of the financial crisis, and monetary policy responds to more severe financial crises more intensively. Hence, if the policy maker chooses the wrong policy regime, there will be costs. In order to reduce the costs, and increase the accuracy of choosing the right policy regime, waiting for more information about the financial crisis may be desirable. We employ the mathematical tool of Bayesian sequential change diagnosis to study this problem. We find that the optimal timing of the Federal Reserve's policy response should be Aug 15, 2007 for certain specifications of the intensity of the financial crisis. We conclude that uncertainty about the intensity of the financial crisis plays an important role in the timing decision of the policy maker.

2.1.2 LIBOR-OIS Spread

Commercial banks need funds to carry out business or satisfy regulations of the central bank. They can finance themselves from the interbank market, where they can borrow through unsecured short-term loans from other banks. The interest rates on these

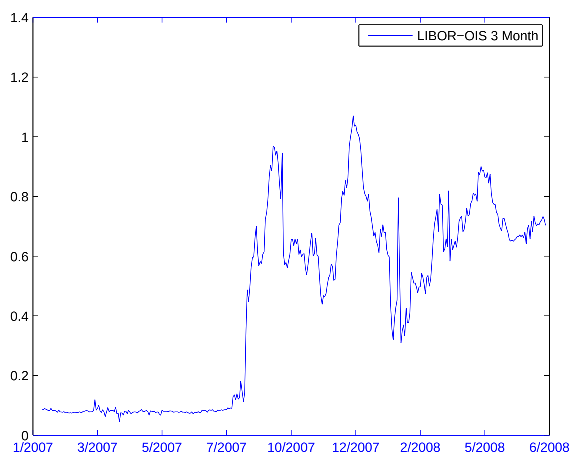
[1] By policy rule, we mean a simple rule, such as a Taylor rule, where the coefficients of the rule are constants in each policy regime.

unsecured short-term loans are negotiated individually between banks, LIBOR is the average of such interest rates. In times of high uncertainty, banks will charge high interest rates on unsecured loans, hence LIBOR, as an average, will increase.

The OIS rate is closely related to the average overnight interest rate expected to prevail until maturity, and it involves minimal counterparty risk since no money exchanges hands until maturity. The LIBOR-OIS spread is a commonly used indicator of the credit spread, that in turn serves as an indicator of distress in the financial sector. In the context of what happened in the summer of 2007, market participants were uncertain about the extent and distribution of subprime mortgage related losses, which makes it difficult to evaluate credit worthiness of each other. Counterparty risk was higher, hence the LIBOR-OIS spread jumped.

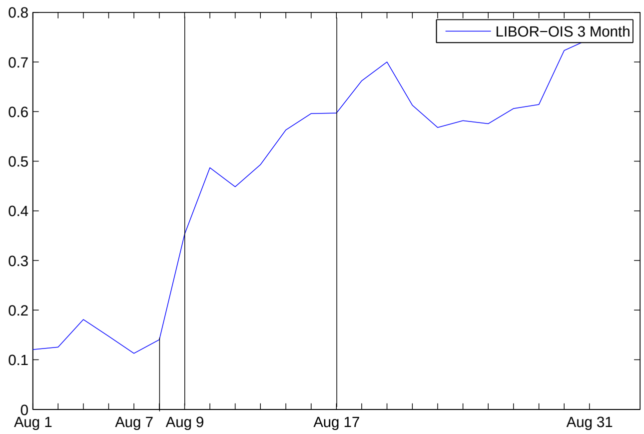
Figure 2.1 shows the 3-Month LIBOR-OIS spread from Jan 2007 to Jun 2008.

Figure 2.1: The 3-Month LIBOR-OIS Spread



The next figure shows the 3-Month LIBOR-OIS spread in August, 2007. Notice the sudden jump on Aug 9, 2007.

Figure 2.2: The 3-Month LIBOR-OIS Spread in August, 2007



2.1.3 A Narrative Account of the Federal Reserve's Response

The following is a brief narrative account of the Federal Reserve's response in August 2007.

- On August 7, 2007, after a scheduled meeting, the FOMC voted to maintain its target for the federal funds rate at 5.25%, which is the same level as June 28, 2007.
- On August 10, 2007, the Federal Reserve Board stated that it "will provide reserves as necessary...to promote trading in the federal funds markets at rates close to the FOMC's target rate of 5.25%.
- On August 17, 2007, after an unscheduled meeting, the Federal Board voted to reduce the primary credit rate by 50 basis points to 5.75 percent[2].

The policy response on August 17, 2007 has been described as "acting swiftly and decisively"(Mizen, 2008; Mishkin, 2010). In this paper, it is assumed that the Federal Reserve changed its policy rule in response to the financial crisis on Aug 17, 2007.

2.1.4 Shiryaev's Problem and On-line Quality Control

Quickest change detection has applications in many fields, such as signal processing, automatic control and so on. The objective is to detect the change in the process in

[2]The complete press release can be found in Appendix B.11.

real time as quickly as possible once it occurs. One of early applications is on-line quality control. On-line quality control procedures are used when decisions are to be made sequentially as new data arrives. Consider a production process can jump from a state of in control to a state of out of control at some random, unobservable time. It is necessary to detect the disorder as soon as it happens with as few false alarms as possible, and then take the appropriate action. The optimal rule to detect the disorder will need to strike a balance between false alarm and delay detection.

Shiryaev (1963, 1978) proposed and solved the following problem. 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:

$$\inf_{T \in \mathcal{T}} \left\{ P\left(T < \theta\right) + cE\left(T - \theta\right)^+ \right\}. \tag{2.1}$$

$P\left(T < \theta\right)$ 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.

It is not clear how to determine the value of c given a particular problem. In order to solve the optimal timing in our problem, the functional form of objective (2.1) and the value of c needs to be determined in our context.

2.2 A Known Alternative

In order to solve this problem, we need to study the objective of the quickest detection problem. Especially, the interpretation and the determination of the value of c in Equation (2.1), the weight assigned to the expected length of delay detection. We are going to address these issues in continuous time. Time needs to be continuous, since time derivative needs to be taken at some point in the derivation.

2.2.1 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 propose a sequential decision rule (T, u_{Tt}), where T is a stopping time and u_{Tt} is the terminal decision rule

$$u_{Tt} = \begin{cases} u_{0t} & before\, T\, declares\, stopping \\ u_{1t} & after\, T\, declares\, stopping \end{cases}. \tag{2.1}$$

In our problem, u_{0t} is the Taylor rule (2.18), and u_{1t} is the credit-spread adjusted Taylor rule (2.19).

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 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.

2.2.2 The Process and the Probability Measure

The period loss function is given as $L(\xi_t, u_t)$, the D-dimensional controlled stochastic process ξ_t with u_t as the control is given by

$$d\xi_t = k_\gamma(\xi_t, u_t)\, dt + \zeta dW_t, \tag{2.2}$$

alternatively

$$\xi_t = \xi_0 + \int_0^t k_\gamma(\xi_s, u_s)\, ds + \zeta \int_0^t dW_s$$

with

$$\gamma = \begin{cases} 0, & t < \theta \\ 1, & t \geq \theta \end{cases}.$$

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

$$\begin{aligned} \phi(\theta = 0) &= \pi \\ \phi(\theta \geq s) &= (1-\pi)\, e^{-\lambda s},\ \lambda > 0. \end{aligned} \tag{2.3}$$

The exponential distribution is a common choice of prior in quickest detection literature, since it is commonly used as distribution for life spans or arrival times. Among the D-dimensional process ξ_t, there is one signal process, and it is assumed that this signal process contains all the information available about θ. The decision maker will monitor this signal process and determine whether the structure break has happened or not. In this study, this signal process is the credit spread, an indicator of distress in the financial sector. The Federal Reserve will monitor this process to infer whether a financial crisis has happened, and determine when to switch policy rule.

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 $\mathcal{G}_t = \sigma\{T \le s, \xi_s, s \le t\}$ and $\mathcal{J}_t = \sigma\{k, i, \xi_s, s \le t\}$. $\mathcal{J}_t$ is the information sets when the policy maker has full information: they know the structure break happens at i, 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 $i \in \mathcal{R}^+$, here k represents the realization of stopping time T and i represents the realization of structure break time θ. When $k < i$, define function

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

when $k \ge i$, define function

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

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

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

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 $k < i$, let

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

when $k \geq i$, define

$$\frac{dP_{+}^{k,i}}{dP_0} \equiv \exp\left\{-\int_0^t \sum_{d=1}^{D} K_{+}^{k,i,(d)}(s)\, dW_s^{(d)} - \frac{1}{2}\int_0^t \left\|K_{-}^{k,i}(s)\right\|^2 ds\right\}. \tag{2.5}$$

For any k and i, define

$$\frac{dP^{k,i}}{dP_0} \equiv \frac{dP_{-}^{k,i}}{dP_0} 1_{\{k<i\}} + \frac{dP_{+}^{k,i}}{dP_0} 1_{\{k \geq i\}}.$$

The next thing we need is to average out i 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 $\lim_{n\to\infty} T_n = T$[3]. Define probability measure

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

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 i.

For any event $A \in \Lambda$,

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

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

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

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

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

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

$$P(A) = E_{\pi}\left(E_{\theta}\left(1_{\{A\}}\right)\right). \tag{2.7}$$

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[4]

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

2.2.3 Derivation of the Objective

Given the period loss function $L(\xi_t, u_{\cdot t})$ and the stochastic process

$$\begin{aligned} d\xi_t &= k_{\gamma}(\xi_t, u_{Tt})\,dt + \sigma dW_t \\ \gamma &= \begin{cases} 0, & t < \theta \\ 1, & t \geq \theta \end{cases}. \end{aligned}$$

The objective can be further defined as the following.

For $T < \theta$, the value of a given sequential decision rule (T, u_{Tt}) is given by

$$\begin{aligned} V_{-}(T,\theta) &= E_{\theta}\left\{\int_0^T e^{-\rho t} L(\xi_t, u_{0t})\,dt + \int_T^{\theta} e^{-\rho t} L(\xi_t, u_{1t})\,dt + \int_{\theta}^{\infty} e^{-\rho t} L(\xi_t, u_{1t})\,dt\right\} \\ d\xi_t &= k_0(\xi_t, u_{0t})\,dt + \sigma dW_t,\ t < T \\ d\xi_t &= k_0(\xi_t, u_{1t})\,dt + \sigma dW_t,\ T \leq t < \theta \\ d\xi_t &= k_1(\xi_t, u_{1t})\,dt + \sigma dW_t,\ \theta \leq t, \end{aligned}$$

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

for $T \geq \theta$, the value if given as

$$\begin{aligned}
V_{+}(T,\theta) &= E_{\theta}\left\{\int_{0}^{\theta} e^{-\rho t}L(\xi_t,u_{0t})\,dt + \int_{\theta}^{T} e^{-\rho t}L(\xi_t,u_{0t})\,dt + \int_{T}^{\theta} e^{-\rho t}L(\xi_t,u_{1t})\,dt\right\} \\
d\xi_t &= k_0(\xi_t,u_{0t})\,dt + \sigma dW_t,\; t \leq \theta \\
d\xi_t &= k_1(\xi_t,u_{0t})\,dt + \sigma dW_t,\; \theta < t \leq T \\
d\xi_t &= k_1(\xi_t,u_{1t})\,dt + \sigma dW_t,\; T < t.
\end{aligned}$$

The expectation is a conditional expectation given θ, then given u_{Tt}, the value of the given stopping time T is

$$V(T) = E_{\pi}\left\{V_{-}(T,\theta)\,1_{\{T<\theta\}} + V_{+}(T,\theta)\,1_{\{T\geq\theta\}}\right\}. \tag{2.8}$$

The opportunity cost of any stopping time T can be derived from the original objective (2.8) as

$$\begin{aligned}
R(T) &= E_{\pi}\left\{V(T,\theta) - V_{+}(\theta,\theta)\right\}, && (2.9)\\
&= E_{\pi}\left\{\left[V_{-}(T,\theta) - V_{+}(\theta,\theta)\right]1_{\{T<\theta\}} + \left[V_{+}(T,\theta) - V_{+}(\theta,\theta)\right]1_{\{T\geq\theta\}}\right\}, && (2.10)
\end{aligned}$$

$V_{+}(\theta,\theta)$ is the expected welfare loss of stopping perfectly in the event of delay detection.

Following Curdia and Woodford (2009b), the terminal decision rule u_{Tt} represents two policy regimes, which can be defined as

$$\begin{aligned}
u_{0t} = &\quad \phi_y y_t + \phi_x X_t, && before\, T\, declares\, stopping \\
u_{1t} = &\quad \phi_y y_t + \phi_x X_t - \phi_\omega \omega_t, && after\, T\, declares\, stoppint.
\end{aligned}$$

The period loss function has the following properties by assumption

$$\begin{aligned}
E_{\theta}\left(L(\xi_t,u_{0t})\right) &> E_{\theta}\left(L(\xi_t,u_{1t})\right),\; when\, t < \theta, \\
E_{\theta}\left(L(\xi_t,u_{0t})\right) &< E_{\theta}\left(L(\xi_t,u_{1t})\right),\; when\, t \geq \theta.
\end{aligned}$$

This is true since as argued by Taylor (2008) and Curdia and Woodford (2009a,b) that the credit-spread adjusted Taylor rule is more appropriate in the era of financial crisis.

The objective function in (2.8) can be approximated by

$$E\left\{-v_{-}'\left(\theta^{-},\theta\right)(\theta - T)^{+} + v_{+}'(\theta,\theta)(T-\theta)^{+} - v_{-}\left(\theta^{-},\theta\right)\delta 1_{\{T<\theta\}}\right\},$$

given an arbitrarily small number $\delta > 0$ and $\theta^- \equiv \theta - \delta$.

please see Appendix B.1.

Then the problem

$$\inf_{T\in\mathcal{T}} R(T)$$

can be approximated by

$$\inf_{T\in\mathcal{T}} E\left\{-v'_-\left(\theta^-,\theta\right)(\theta-T)^+ + v'_+(\theta,\theta)(T-\theta)^+ - v_-\left(\theta^-,\theta\right)\delta 1_{\{T<\theta\}}\right\}.$$

We need the following lemma to further simplify the objective.

At any time t, given k and i, if the expected period loss function can be written as

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

then

$$c = -\frac{v'_+(\theta,\theta)}{v'_-(\theta^-,\theta)} \tag{2.11}$$

is a constant.

Please see Appendix A.4.

The condition in Proposition 2.2.3 says that the expected period loss function depends on k and i, and the way it depends on k and i is only through two indicator functions $1_{\{t\geq k\}}$ and $1_{\{t\geq i\}}$. Given the condition in Proposition 2.2.3 is satisfied, we have

$$\inf_{T\in\mathcal{T}} E\left(-v'_-(\theta,\theta)\right)\left[(\theta-T)^+ + c(T-\theta)^+ + \delta 1_{\{T<\theta\}}\right], \tag{2.12}$$

further, we can define a new conditional 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

$$\inf_{T\in\mathcal{T}} \tilde{E}\left[(\theta-T)^+ + c(T-\theta)^+ + \delta 1_{\{T<\theta\}}\right] \tag{2.13}$$

Compare (2.13) with (2.1), where $c = -\frac{v'_+(\theta,\theta)}{v'_-(\theta^-,\theta)}$.

The solution to this problem is provided in Appendix B.8

c is the relative weight assigned to the expected length of delay detection, 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 delay detection is faster than that of false alarm, hence higher weight should be put on the expected length of delay detection. Now, there is a clear interpretation for c, and a method to compute given a particular problem.

2.2.4 Discrete Version

Assuming geometric prior distribution

$$p(\theta = k) = \begin{cases} \pi & if\, k = 0 \\ (1-\pi)\,\rho\,(1-\rho)^{k-1} & if\, k = 1, 2, ... \end{cases}$$

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

Assume $\tilde{E}(\theta)$ and $\tilde{E}(T)$ are finite, then

$$\tilde{E}\left\{(\theta - 1 - T)^+ + c(T-\theta)^+ + 1_{\{T<\theta\}}\right\} = \tilde{E}\left\{\theta + \sum_{m=0}^{T-1}\left[c\pi_m - (1-\pi_m)\right]\right\}. \tag{2.14}$$

Please see Appendix B.4

Similar to Proposition 2.2.3, we have the following proposition for discrete time.

At any time t, given k and i, if the expected period loss function can be written as

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

then

$$c = -\frac{v_+(\theta+1,\theta) - v_+(\theta,\theta)}{v_-(\theta-2,\theta) - v_-(\theta-1,\theta)} \tag{2.15}$$

is a constant.

Please see Appendix B.7.

2.2.5 Solution Method

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

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

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

Then the value function given π is

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

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

and the stopping time

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

is the optimal stopping time.

2.3 Quantitative Study

The model of the economy is based on Curdia and Woodford (2009a,b). The authors extend the basic New Keynesian model to incorporate financial frictions, and evaluate the performance of the credit-spread adjusted Taylor rule which incorporates a response to some indicator of the distress in the financial sector as proposed by Taylor (2008). Curdia and Woodford (2009b) find that the credit-spread adjusted Taylor rule by introducing a contemporaneous response to the size of the credit spread can improve policy performance upon that of the standard Taylor rule.

2.3.1 The Model

2.3.1.1 The Credit Spread

In Curdia and Woodford (2009a,b)'s model, there are two interest rates: the borrowing rate i_t^b and the lending rate i_t^d. The difference between two interest rates is driven by the credit spread ω_t. The relation between two interest rates and the credit spread is determined by the following rule

$$1 + i_t^b = \left(1 + i_t^d\right)(1 + \omega_t),$$

where the credit spread ω_t follows the following process

$$\omega_t = \chi_\gamma + \sigma_\gamma \epsilon_t \tag{2.16}$$

$$\gamma = \begin{cases} 0 & t < \theta \\ 1 & t \geq \theta \end{cases}. \tag{2.17}$$

χ_γ is the loss rate which means in order to originate a quantity of loans b_t that will be repaid in the following period, it is necessary to make additional $\chi_\gamma b_t$ amount of loans that will not be repaid(Curdia and Woodford, 2009b). In this paper, χ_γ is indexed by $\gamma \in \{0, 1\}$, and γ changes from 0 to 1 after random time θ. ω_t is equal to χ_γ plus a random noise, so that the time of the structure break is not observable. In this paper, the LIBOR-OIS spread is used as the indicator for credit spread.

In this study, it is assumed that the Federal Reserve is aware of the null. Assuming normal distribution, the specifications of the null (mean and standard deviation) are estimated form data ranging from Jan 1, 2007 through Aug 6, 2007. The specification of the alternative is rather a complicated issue. Intuitively, the alternative is the Fed's perspective about what is going to happen. One approach is to specify it with data from Aug 7, 2007 through Aug 1, 2008. This is to assume the policy makers know exactly what is going to happen, they just do not know when. This is a strong assumption and it will be relaxed in the following section with unknown alternatives.

2.3.1.2 Monetary Policy

The monetary policy rule in regime 1 is the standard Taylor rule

$$i_t^d = \phi_x X_t + \phi_y y_t, \tag{2.18}$$

and the policy rule in regime 2 will be the credit-spread adjusted Taylor rule proposed by Curdia and Woodford (2009a)

$$i_t^d = \phi_x X_t + \phi_y y_t - \phi_\omega \omega_t. \tag{2.19}$$

Taylor (2008) also suggested a similar rule in the era of financial crisis.

2.3.1.3 The Economy

Consider the following welfare loss function

$$\sum_{t=0}^{\infty} \beta^t \left[var\left(X_t\right) + \alpha var\left(Y_t\right) \right]$$

where X_t is inflation and Y_t is welfare-relevant output gap.

The simplified log-linear system of equilibrium conditions given by Curdia and

Woodford (2009b) and Curdia and Woodford (2009a, page 46) are given by

$$\begin{aligned}
i_t^{avg} &= i_t^d + \pi_b 1_{t\geq i}\omega_t, \\
\Omega_t &= \omega_t + \delta E_t\Omega_{t+1}, \\
Y_t &= E_tY_{t+1} - \bar{\sigma}\left(i_t^{avg} - E_tX_{t+1}\right) - \bar{\sigma}s_\Omega\Omega_t + \bar{\sigma}\left(s_\Omega + \psi_\Omega\right)E_t\Omega_{t+1}, \\
X_t &= \beta E_tX_{t+1} + \kappa Y_t + \xi\left(s_\Omega + \pi_b - \gamma_b\right)\Omega_t + u_t, \\
i_t^d &= \phi_x X_t + \phi_y Y_t - \phi_\omega 1_{t\geq k}\omega_t, \\
\omega_t &= \chi_0 + \left(\chi_1 - \chi_0\right)1_{t\geq i} + \left(\sigma_0 + \left(\sigma_1 - \sigma_1\right)1_{\{t\geq i\}}\right)\epsilon_t.
\end{aligned} \tag{2.20}$$

In this paper, we abstract away from government spending, and the real cost of originating loans, and the credit spread is an exogenous shock, so that we can abstract away from private indebtedness. We assume that before the financial crisis, the average interest rate does not respond to interest rate spread; this is similar to the standard New Keynesian model in the sense that only one interest rate affects the economy. After the financial crisis, both the lending rate and the borrowing rate will affect the economy through the average interest rate. The interpretation of i should be the time when financial sector is in so much trouble, that it's variation will have significant impact on the aggregate economy which is symbolized by a jump in the mean of credit spread. Through equations of (2.20), one can see that the additive noise of the credit spread is introduced into the system through average interest rate i_t^{ave} after the beginning of the financial crisis, i.e. $t \geq \theta$. The monetary policy rule with $\phi_\omega > 0$ can dampen the impact of this noise on the system. But if the policy rule changes before financial crisis starts ($k < \theta$), then the policy rule introduces the additive noise of the credit spread into the system when the credit spread does not have any impact on the system through other channels, and this causes higher welfare loss compared to the case when $k = \theta$. On the other hand, if $k > \theta$, the financial crisis has already happened, if the policy regime switches later than θ, there will be higher welfare loss compare to the case when $k = \theta$.

The model expressed by equation (2.20) is solved with Blanchard and Kahn (1980)'s method. The detailed solution can be found in Appendix B.5. Let $x_t = [\Omega_t, Y_t, \pi_t]'$, we have

$$var\left(x_t\right) = \left(\Lambda_{11}^{-1}\Lambda_{12}\right)var\left(u_t\right)\left(\Lambda_{11}^{-1}\Lambda_{12}\right)' + \left(\Lambda_{11}^{-1}J_1^{-1}D_1\right)var\left(f_t\right)\left(\Lambda_{11}^{-1}J_1^{-1}D_1\right)'.$$

The period loss function can be written as

$$\mathbb{L} = H'var\left(x_t\right)H, \tag{2.21}$$

where $H = [0, 1, \alpha^{1/2}]$.

The entries of the coefficient matrix are functions of time t. We will compute the period loss function for 500 periods and compute the loss function for different value of k and θ. We can approximate $c = -\frac{v'_+(\theta,\theta)}{v'_-(\theta^-,\theta)}$ by $c = -\frac{v_+(\theta+1,\theta)-v_+(\theta,\theta)}{v_-(\theta-2,\theta)-v_-(\theta-1,\theta)}$, where $\theta = 2, 3....$ The result of this section is to compute c in (2.13).

Given Cúrdia and Woodford (2009)'s model, the condition in Proposition 2.2.4 is satisfied, so

$$c = -\frac{v_+ (\theta + 1, \theta) - v_+ (\theta, \theta)}{v_- (\theta - 2, \theta) - v_- (\theta - 1, \theta)}$$

is a constant.

The proof of this proposition is straightforward, hence is omitted.

2.3.2 Optimal Stopping Problem

Once c is computed, we can start solving the optimal stopping problem from Proposition 2.2.4

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

with the process of credit spread

$$\omega_t = \chi_\gamma + \sigma_\gamma \xi_t,$$

where

$$\gamma = \begin{cases} 0 & t < \theta \\ 1 & t \geq \theta \end{cases}.$$

We assume a geometric prior distribution

$$P(t = k) = \begin{cases} \pi, & if\, k = 0 \\ (1 - \pi)\, \rho\, (1 - \rho)^{k-1}, & if\, k = 1, 2, ... \end{cases}$$

Define $\pi_k = \tilde{P}(t \leq k | \mathcal{F}_k)$, $k = 0, 1, ...$ Then the sequence $\{\pi_k\}$ evolves according to the recursion[5]

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

[5] Please see equation (5.26) of Poor and Hadjiliadis (2009).

π_k is the posterior probability that the structure break happens before k.

2.3.3 The Alternatives

Intuitively, the alternative is what the policy maker expects after the structure break. How to specify the alternative is an open question. In this section, we will try different specifications of the alternative.

The specification of the first alternative is estimated from data ranging from August 8, 2007 to July 31, 2008. To do so, it is assumed that the policy maker knows the specification of the economy after the structural break.

Table 2.1: Specifications of Different Alternatives

First Alternative		Second Alternative		Third Alternative	
χ_1	0.7558	χ_1	0.4	χ_1	1.2
σ_1	0.1547	σ_1	0.1547	σ_1	0.1547

To use the result stated in Section 2.2.4 to compute the value function and the optimality condition. Figure 2.3 shows the optimality condition with the first specification of the alternative.

Figure 2.3: Optimal Stopping Time with the First Specification of the Alternative

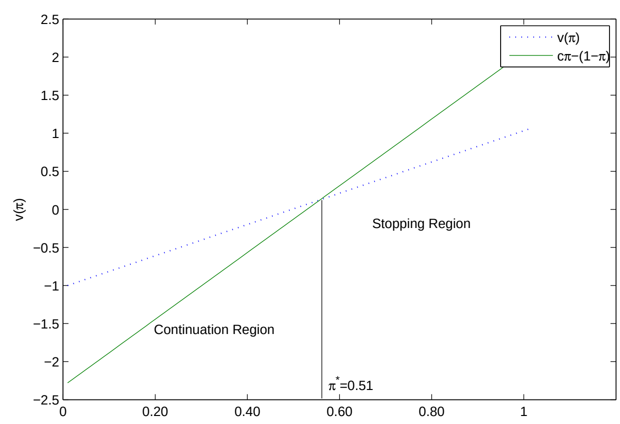


Figure 2.4 shows the evolution of π_t and the optimal stopping time with the first specification of the alternative.

Figure 2.4: Optimal Stopping Rule with the First Specification of the Alternative

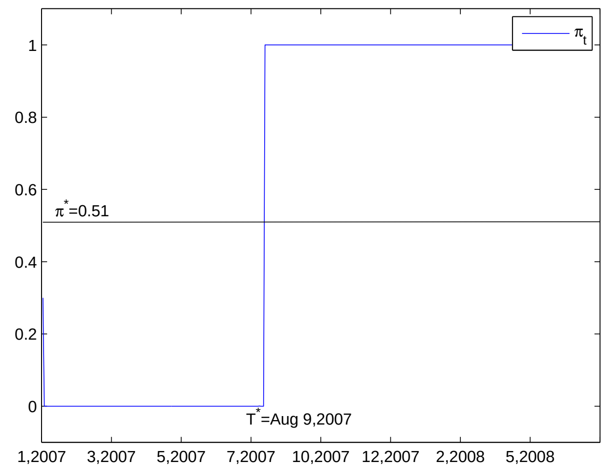
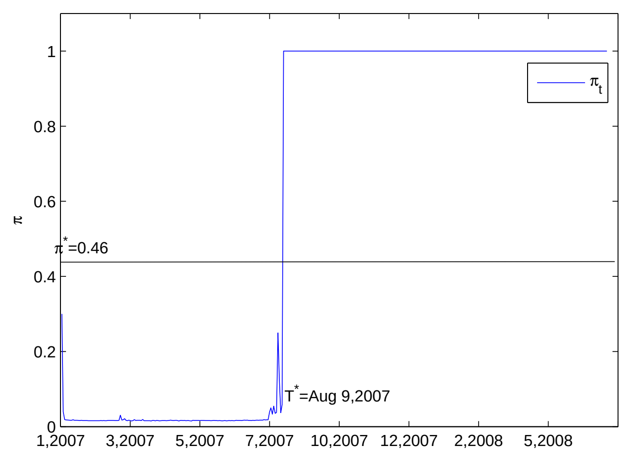


Figure 2.5: Optimal Stopping Rule with the Second Specification of the Alternative

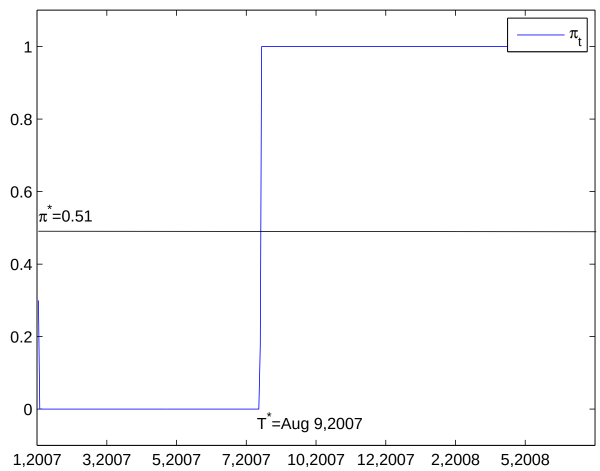


The next figure shows the evolution of π_t and the optimal stopping time with the third specification of the alternative.

Table 2.3: Optimal Stopping Time with Different Priors

$\rho = 0.05$	$\pi^* = 0.94$	$T^* = Aug\,9$
$\rho = 0.1$	$\pi^* = 0.87$	$T^* = Aug\,9$
$\rho = 0.3$	$\pi^* = 0.51$	$T^* = Aug\,9$

Figure 2.6: Optimal Stopping Rule with the Third Specification of the Alternative



The optimal stopping times are presented in the following table.

Table 2.2: Optimal Stopping Rules with Different Specifications of Alternatives

	First Alternative	Second Alternative	Third Alternative
T^*	Aug 9,2007	Aug 9, 2007	Aug 9,2007
π^*	0.51	0.46	0.51
c	1.1392	1.1392	1.1392

Based on the calculations shown above, it is optimal to switch policy rule on Aug 9, 2007, the Fed reduced the primary rate on Aug 17, 2007; it seems hard to justify the Federal Reserve's response time even with different specifications of the alternative. We have considered the trade-off between false alarm and delay detection, this model still can not justify the timing decision of the Federal Reserve in the recent financial crisis. What happens on Aug 9, 2007 in the money market is sure something hard to miss.

This result is derived under the assumption that the Federal Reserve knows the alternative, detecting the change is the only task, which might be too strong an assumption. In the next section, this assumption will be relaxed by assuming unknown alternatives.

2.4 Unknown Alternatives

2.4.1 Introduction

The previous section shows the case when the alternative is known, in this section this assumption will be relaxed. Assuming that there are a set of potential alternatives indexed by $\mu \in \mathcal{M}, \mathcal{M} = \{1, 2, ..., M\}$. We define an sequential decision rule $\{T, d\}$, where T is the stopping time and d determines the alternative to accept.

The mathematical tool adopted is developed by Dayanik et al. (2008) known as the *Bayesian sequential change diagnosis*. Bayesian sequential change diagnosis solves a combined problem of detection and identification of an abrupt change in the distribution of a random sequence. In this problem, there is a sequence of i.i.d random variables $X_1, X_2, ...$, with values in a measurable space $(E, \mathcal{E})$. Initially, there is some known probability measure $\mathbb{P}_0$ on $(E, \mathcal{E})$, then at some unknown and unobservable time θ, the common probability measure changes to another probability measure $\mathbb{P}_\mu$, where $\mu \in \mathcal{M}$, $\mathcal{M} \equiv \{1, 2, ..., M\}$ is unknown and unobservable. The objective is to detect the change in probability measure as soon as possible, at the mean time, identify the probability distribution after the change as accurately as possible. In order to reduce the risk of identifying the wrong alternative probability distribution, it might be optimal for the decision maker to delay in order to collect more information about the probability distribution after the change. This introduces another trade-off into the decision making process that is not taken into account in the case of a known alternative. In the case with a known alternative, to sound the alarm as soon as the change occurs is the only concern of the decision maker. Raising alarm as soon as the change occurs may be advantageous for the task of change detection, but it is not desirable to identify the right alternative, because the longer one waits, the more accurately one can identify the right alternative, hence the risk of accepting the wrong alternative is reduced by waiting longer. The welfare gain of increased accuracy of adopting the right policy response by waiting longer, is defined as the option value of waiting.

On Aug 9, 2007, it might be clear that the structural break has happened, but it is not clear that the policy maker is certain how serious the credit crunch might be. In particular, the mean of the credit spread after the structure break is unknown to the policy maker. The policy maker may be contemplating several levels of the mean of the credit spread, indicating different intensity of the credit crunch.

Assume that the Federal Reserve is contemplating two alternative specifications of the economy after the structural break; each alternative is a version of Curdia and Woodford (2009a)'s model with different sets of parameter values (See Table B.3). The alternatives can be specified by the following

$$\begin{aligned} H_1 : \chi_1 &= 0.7558, \\ H_2 : \chi_2 &= 0.45; \end{aligned}$$

against the null

$$H_0 : \chi_0 = 0.0914.$$

H_2 represents a milder version of the financial crisis. Besides χ_μ, π_b will also take on different values in different regimes of the economy, π_b is the probability of a household to be a borrower. With different π_b corresponding to different alternatives, it is necessary for monetary policy to have different response in ϕ_ω^d. Hence, corresponding to each versions of the financial crisis that is contemplated by the policy maker, there is a monetary policy rule with different intensity to respond to the evolution of the credit spread.

$$i = \phi_y Y_t + \phi_\pi \pi_t - 1_{\{t \geq k\}} \phi_\omega^d \omega_t, \; where\, d \in \{1, 2\}.$$

We assume $\phi_\omega^1 > \phi_\omega^2$, this suggests the Federal Reserve will respond to the milder version of financial crisis with less intensive policy rule. With multiple alternatives, the cost of a sequential decision rule $\{T, d\}$ comes form different sources, first, stopping at the wrong time; second, accepting the wrong alternative.

2.4.2 The Process and Probability Measure

The period loss function is given as $L(\xi_t, u_{Tt})$, the D-dimensional controlled stochastic process with u_t as the control is given by

$$d\xi_t = k_\gamma(\xi_t, u_{Tt})\, dt + \sigma dW_t, \tag{2.22}$$

alternatively

$$\xi_t = \xi_0 + \int_0^t k_\gamma(\xi_s, u_{Ts})\, ds + \sigma \int_0^t dW_s$$

with

$$u_{Tt} = \begin{cases} u_{0t}, & t < T \\ u_{dt}, & t \geq T \end{cases},$$

$$\begin{aligned} \gamma &= \begin{cases} 0, & t < \theta \\ \mu, & t \geq \theta \end{cases} \\ \mu &\in \{1, ..., M\} \\ d &\in \{1, ..., M\}. \end{aligned}$$

θ and μ are random variables with probability measure ϕ_θ and ϕ_μ:

$$\begin{aligned} \phi_\theta (\theta = 0) &= \pi, \\ \phi_\theta (\theta \geq s) &= (1 - \pi) e^{-\lambda s}, \; \lambda > 0, \\ \phi_\mu (\mu = i) &= \nu_i, \\ \sum_{i \in \mathcal{M}} \nu_i &= 1. \end{aligned} \tag{2.23}$$

The prior for θ and μ is given by

$$\begin{aligned} \phi (\theta = 0, \mu = i) &= \pi \nu_i, s \\ \phi (\theta \geq s, \mu = i) &= (1 - \pi) \nu_i e^{-\lambda s}. \end{aligned}$$

The process ξ_t is D-dimensional, and there is one signal process among ξ_t. This signal process is assumed to contain all the information available about θ and μ. The decision maker will monitor this signal process and determine whether the structural break has happened yet, and upon stopping, determine the alternative to accept. In this paper, this signal process is the credit spread, an indicator of distress in the financial sector, the credit spread differs in mean and standard deviation indicating a different version of the financial crisis. The Federal Reserve will monitor this process to infer whether a financial crisis has happened and which version of the financial crisis has happened, and determine when to switch and which policy rule to adopt correspondingly. In this paper, there are two potential states of the economy: mild financial crisis and severe financial crisis.

We propose a sequential decision rule (T, d), where T is a stopping time and $d : \mathcal{G} \rightarrow \mathcal{M}$ is the terminal decision rule, the control rule determined by this sequential decision rule is given by

$$u_{Tt}=\begin{cases} u_{0t}, & when\, t<T \\ u_{dt}, & d\in\mathcal{M},\, when\, t\geq T \end{cases}, \tag{2.24}$$

u_{0t} will be the Taylor rule 2.18, and u_{dt} will be the credit spread adjusted Taylor rule 2.19 for each alternative $d \in \mathcal{M}$. The sequential decision rule consist of announcing when to switch policy rule and which policy rule will be adopted upon switching.

We consider the following measurable spaces

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

where $C[0,\infty)$ is the space of continuous functions. We consider the following filtration $\mathcal{G}_t = \sigma\{T \leq t, d, \xi_s, s \leq t\}$ and $\mathcal{J}_t = \sigma\{k,i,j,l,\xi_s, s\leq t\}$. $\mathcal{J}_t$ represents full information, we know the structural break time, the stopping time, we also know which alternative is true, and which alternative we will accept. $\mathcal{G}_t$ is the information set of the policy maker at t.

For any given $k \in \mathcal{R}^+$and $i \in \mathcal{R}^+$, here k represents the realization of stopping time and i represents the realization of structure break time θ, j is the true alternative, l is the alternative accepted.

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 $P_0\left(W_0 = 0\right) = 1$.

When $k < i$, define function

$$K_-^{k,i,j,l}(s) = k_0\left(\xi_s, u_{0,t}\right) 1_{\{s<k\}} + k_0\left(\xi_s, u_{l,t}\right) 1_{\{k\leq s<i\}} + k_j\left(\xi_s, u_{l,t}\right) 1_{\{i\leq s\}};$$

when $k \geq i$, define function

$$K_+^{k,i,j,l}(s) = k_0\left(\xi_s, u_{0,t}\right) 1_{\{s<i\}} + k_j\left(\xi_s, u_{0,t}\right) 1_{\{i\leq s<k\}} + k_j\left(\xi_s, u_{l,t}\right) 1_{\{k\leq s\}}.$$

When $k < i$,

$$\frac{dP_-^{k,i,j,l}}{dP_0} \equiv \exp\left\{-\int_0^t \sum_{d=1}^{D} K_-^{k,i,j,l(d)}(s)\, dW_s^{(d)} - \frac{1}{2}\int_0^t \left\|K_-^{k,i,j,l}(s)\right\|^2 ds\right\}; \tag{2.25}$$

when $k \geq i$

$$\frac{dP_{+}^{k,i,j,l}}{dP_0} \equiv \exp\left\{ -\int_0^t \sum_{d=1}^{D} K_{+}^{k,i,j,l(d)}(s)\, dW_s^{(d)} - \frac{1}{2}\int_0^t \left\| K_{-}^{k,i,j,l}(s) \right\|^2 ds \right\}. \tag{2.26}$$

Define

$$\frac{dP^{k,i,j,l}}{dP_0} \equiv \frac{dP_{-}^{k,i,j,l}}{dP_0} 1_{\{k<i\}} + \frac{dP_{+}^{k,i,j,l}}{dP_0} 1_{\{k\geq i\}}.$$

The next thing we need to do is to average out k, l and i, j, with probability measure $P_{\theta\mu}$ and ϕ.

For a given stopping time T, define

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

Define probability measure

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

given a stopping time T, and correspondingly defined T_n, $P_{\theta,n}$ tells us the probability of $T_n = k$ and $d = l$, when structural break happens at θ and the true alternative is μ.

For any event $A \in \Lambda$, stopping time T and terminal decision rule d,

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

$P^{k,i,j,l}(A)$ is the probability of observing event A, given that the structural break happens at i and policy switches at k and alternative l is accepted and the true alternative is j. $E_{\theta\mu}\left(1_{\{A\}}\right)$ tells us, given stopping time T and terminal decision rule d, and the structural break happens at i and the true alternative is μ, the probability of observing event A.

$$P_{\theta\mu}(A) = E_{\theta\mu}\left\{1_{\{A\}}\right\} \tag{2.27}$$

$P_{\theta\mu}(A)$ is the probability of A conditional on θ and μ.

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

$$P(A) = E_{\pi}\left(E_{\theta\mu}\left(1_{\{A\}}\right)\right) \tag{2.28}$$

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 the Novikov condition

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

The expectation is defined under probability measure P.

2.4.3 The Objective of Unknown Alternatives

Similar to the case of unknown alternatives,

For $T < \theta$,

$$\begin{aligned} V_-(T,d) &= E_{\theta\mu}\left\{\int_0^T e^{-\rho t}L(\xi_t,u_{0t})\,dt + \int_T^\theta e^{-\rho t}L(\xi_t,u_{dt})\,dt + \int_\theta^\infty e^{-\rho t}L(\xi_t,u_{dt})\,dt\right\} \\ d\xi_t &= k_0(\xi_t,u_{0t})\,dt + \sigma dW_t,\ t<T \\ d\xi_t &= k_0(\xi_t,u_{dt})\,dt + \sigma dW_t,\ T\le t<\theta \\ d\xi_t &= k_\mu(\xi_t,u_{dt})\,dt + \sigma dW_t,\ \theta\le t. \end{aligned}$$

For $T \ge \theta$,

$$\begin{aligned} V_+(T,d) &= E_{\theta\mu}\left\{\int_0^\theta e^{-\rho t}L(\xi_t,u_{0t})\,dt + \int_\theta^T e^{-\rho t}L(\xi_t,u_{0t})\,dt + \int_T^\infty e^{-\rho t}L(\xi_t,u_{dt})\,dt\right\} \\ d\xi_t &= k_0(\xi_t,u_{0t})\,dt + \sigma dW_t,\ t<\theta \\ d\xi_t &= k_\mu(\xi_t,u_{0t})\,dt + \sigma dW_t,\ \theta\le t<T \\ d\xi_t &= k_\mu(\xi_t,u_{dt})\,dt + \sigma dW_t,\ T\le t. \end{aligned}$$

$$V(T,d) = E_\pi\left\{V_-(T,d)\,1_{\{T<\theta\}} + V_+(T,d)\,1_{\{T\ge\theta\}}\right\}$$

$E_{\theta\mu}$ is the conditional expectation given θ and μ, while E_π is expectation with probability distribution ϕ over θ and μ.

Normalizing $V(T,d)$ by subtracting $V_+(\theta,\mu)$, $V_+(\theta,\mu)$ is the value of stopping perfectly and accepting the right alternative upon stopping

$$\begin{aligned} R(T,d) &\equiv V(T,d) - V_+(\theta,\mu) \\ &= E_\pi\left\{[V_-(T,d) - V_+(\theta,\mu)]\,1_{\{T<\theta\}} + [V_+(T,d) - V_+(\theta,\mu)]\,1_{\{T\geq\theta\}}\right\}. \end{aligned}$$

Our problem is to find a sequential decision strategy $(T,d) \in \Delta$ with minimum risk

$$R^* = \inf_{(T,d)\in\Delta} R(T,d). \tag{2.29}$$

Given $T = k$, $\theta = i$, $\mu = j$ and $d = l$, and $k < i$, define

$$\begin{aligned} v_-(k,i,j,l) &= E_-\left\{\int_0^k e^{-\rho t} L(\xi_t, u_{0t})\,dt + \int_k^i e^{-\rho t} L(\xi_t, u_{lt})\,dt + \int_i^\infty e^{-\rho t} L(\xi_t, u_{lt})\,dt\right\} \\ d\xi_t &= k_0(\xi_t, u_{0t})\,dt + \sigma dW_t,\ t \leq k \\ d\xi_t &= k_0(\xi_t, u_{lt})\,dt + \sigma dW_t,\ k < t < i \\ d\xi_t &= k_j(\xi_t, u_{lt})\,dt + \sigma dW_t,\ i \leq t. \end{aligned}$$

When $k \geq i$, define

$$\begin{aligned} v_+(k,i,j,l) &= E_+\left\{\int_0^i e^{-\rho t} L(\xi_t, u_{0t})\,dt + \int_i^k e^{-\rho t} L(\xi_t, u_{0t})\,dt + \int_k^\infty e^{-\rho t} L(\xi_t, u_{lt})\,dt\right\} \\ d\xi_t &= k_0(\xi_t, u_{0t})\,dt + \sigma dW_t,\ t < i \\ d\xi_t &= k_j(\xi_t, u_{0t})\,dt + \sigma dW_t,\ i \leq t < k \\ d\xi_t &= k_j(\xi_t, u_{lt})\,dt + \sigma dW_t,\ k \leq t. \end{aligned}$$

Given $\mu = j$ and $d = l$, the value function for any given $T = k, \theta = i$, when $k < i$, let $i^- = i - \delta$, for any $\delta > 0$

$$\begin{aligned} v_-(k,i,j,l) - v_+(i,i,j,j) &= \underbrace{v_-(k,i,j,l) - v_-\left(i^-,i,j,l\right)}_{1} + \underbrace{v_-\left(i^-,i,j,l\right) - v_+(i,i,j,l)}_{2} + \\ &\quad \underbrace{v_+(i,i,j,l) - v_+(i,i,j,j)}_{3}. \end{aligned}$$

In the above equation, part 1 represents the loss of stopping at the wrong time, 3 represents the cost of stopping at the right time but accepting the wrong alternative. Similar to the case of known alternative, we will have the following approximations

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

since by construction $\lim_{\delta\to 0} v_-\left(i^-,i,j,l\right) = v_+(i,i,j,l)$.

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

When $k \geq i$,

$$\begin{aligned} v_+(k,i,j,l) - v_+(i,i,j,j) &= v_+(k,i,j,l) - v_+(i,i,j,l) + v_+(i,i,j,l) - v_+(i,i,j,j) \\ &\approx v'_+(j,l)(k-i)^+ + v_+(i,i,j,l) - v_+(i,i,j,j). \end{aligned}$$

For any $T = k$, $\theta = i$, $\mu = j$ and $d = l$, we have

$$v(k,i,j,l) - v_+(i,i,j,j) = 1_{\{k<i\}}\left[v_-(k,i,j,l) - v_+(i,i,j,j)\right] + 1_{\{k\geq i\}}\left[v_+(k,i,j,l) - v_+(i,i,j,j)\right]$$

We have

$$R(T,d) \approx E\left\{v(T,\theta,\mu,d)\right\}.$$

The objective $R(T,d)$ in 2.29 can be approximated by

$$\tilde{E}\left\{\left(\theta^- - T\right)^+ + c_{\mu d}(T-\theta)^+ + c'_{\mu d}1_{\{T<\theta\}} + c''_{\mu d}1_{\{T\geq\theta\}}\right\},$$

where1

$$\begin{aligned} c_{\mu d} &= -\frac{v'_+(\theta,\theta,\mu,d)}{v'_-\left(\theta^-,\theta,\mu,d\right)}, \\ c'_{\mu d} &= -\frac{v_-\left(\theta^-,\theta,\mu,d\right) - v_+(\theta,\theta,\mu,d)}{v'_-\left(\theta^-,\theta,\mu,d\right)} + \delta \\ c''_{\mu d} &= -\frac{v_+(\theta,\theta,\mu,d) - v_+(\theta,\theta,\mu,\mu)}{v'_-\left(\theta^-,\theta,\mu,d\right)}. \end{aligned}$$

It is easy to verify the condition in Proposition 2.2.3 is satisfied, $c_{\mu d}$, $c''_{\mu d}$ and $c'_{\mu d}$ are constants with respect to θ.

c_{jl} is the weight assigned to delay detection given alternative $\mu = l$ is true and policy maker accept alternative $d = j$; c'_{jl} is the loss associated with accepting alternative $d = j$ when alternative $\mu = l$ is true in the event of false alarm; the interpretation of c''_{jl} is similar to that of c'_{jl}, but in the event of delay detection. In general, waiting for longer time will increase the probability of event $\{0 \leq T < \infty, d = \mu\}$, hence reduce expected loss of identifying the wrong alternative $\tilde{E}\left\{c''_{\mu d} 1_{\{T \geq \theta\}}\right\}$, this will give rise to the definition of option value of waiting.

For any two stopping times T and S, satisfying $T \geq S > \theta$, given a decision rule d, the option value of waiting between sequential decision rule (T, d) and (S, d) is defined as

$$OV(S, T, d) \equiv \tilde{E}\left\{c_{\mu d}(T - \theta)^{+} + c''_{\mu d} 1_{\{T \geq \theta\}}\right\} - \tilde{E}\left\{c_{\mu d}(S - \theta)^{+} + c''_{\mu d} 1_{\{S \geq \theta\}}\right\}.$$

2.4.4 Discrete Time

In discrete time, the problem became

$$\inf_{(T,d)\in\Delta} \tilde{R}(T, d) = \inf_{(T,d)\in\Delta} \tilde{E}\left\{(\theta - 1 - T)^{+} + c_{\mu d}(T - \theta)^{+} + c'_{\mu d} 1_{\{T < \theta\}} + c''_{\mu d} 1_{\{T \geq \theta\}}\right\}. \tag{2.30}$$

Suppose $\tilde{E}(\theta) < \infty$, $\tilde{E}(T) < \infty$ and define sequence $\left\{\pi_k = \left(\pi_k^0, \pi_k^1, ..., \pi_k^M\right)\right\}$ by

$$\begin{aligned} \pi_m^0 &= \tilde{P}(\theta > n | \mathcal{G}_m) \\ \pi_m^i &= \tilde{P}(\theta \leq n, \mu = i | \mathcal{G}_m), \end{aligned}$$

then the objective of

$$\tilde{R}(T, d) = \tilde{E}\left\{\theta - \sum_{n=0}^{T} \pi_n^0 + \sum_{j=1}^{M} 1_{\{d=j\}} \left[\sum_{n=0}^{T-1} \sum_{i=1}^{M} c_{ij}\left(1 - \pi_n^0\right) + \sum_{i=1}^{M} c'_{ij} \pi_T^0 + \sum_{i=1}^{M} c''_{ij} \pi_T^i\right]\right\}$$

Please see Appendix B.9.

The evolution of π_k is defined by Dayanik et al. (2008), equation (3.1)-(3.3).

Define

$$\begin{aligned} h(\pi) &= \min_{j\in\mathcal{M}} h_j\left(1-\pi^0\right) \\ h_j(\pi) &= \sum_{i=1}^{M} c_{ij}\left(1-\pi^0\right) \\ h'(\pi) &= \min_{j\in\mathcal{M}} h'_j(\pi)-\pi^0 \\ h'_j(\pi) &= \sum_{i=1}^{M}\left[c''_{ij}\pi^i+c'_{ij}\pi^0\right] \end{aligned}$$

The problem

$$\inf_{(T,d)\in\Delta} \tilde{R}(T,d)$$

can be solved with the solution provided by Dayanik et al. (2008).

Define operator

$$(\mathbb{M}f)(\pi) \equiv \min\left\{h'(\pi), h(\pi)-\pi^0+(\mathbb{M}f)\left(\pi'\right)\right\},$$

then

$$V(\pi) \equiv \lim_{N\to\infty}\left(\mathbb{M}^N h\right)(\pi).$$

The optimality condition is defined as

$$V(\pi) = (\mathbb{M}V)(\pi),$$

which is the Proposition 4.1 of Dayanik et al. (2008).

Given the value function $V(\pi)$, we can present an optimal sequential decision strategy for this problem. First define

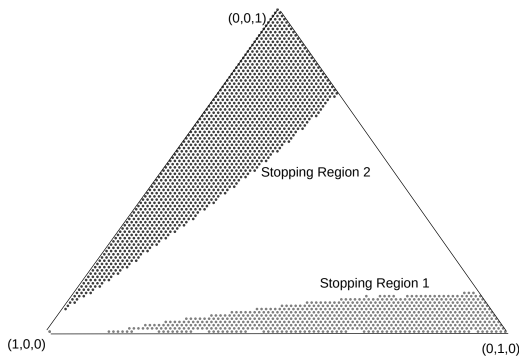
$$\begin{aligned} \Gamma &\equiv \left\{\pi\in S^M | V(\pi)=h'(\pi)\right\}, \\ \Gamma^j &\equiv \Gamma\cap\left\{\pi\in S^M | h'(\pi)=h'_j(\pi)\right\}. \end{aligned} \tag{2.31}$$

Based on Theorem 4.1 of Dayanik et al. (2008), the optimal sequential decision strategy (T^*, d^*) is defined as

$$\begin{aligned} T^* &= \inf\left\{n\geq 0|\pi_n\in\Gamma\right\}, \\ d^* &= j \,such\, that\, T^*=n\, and\, \pi_n\in\Gamma^j. \end{aligned}$$

The next figure shows the stopping region determined by Equation 2.31.

Figure 2.7: Stopping Regions for the Second Set of Specification of the Alternatives



The following tables shows the specifications of two sets of alternatives and their corresponding stopping times.

Table 2.4: Specifications of First Set of Null and Alternatives

Specification of Null	Specification of Alternative 1	Specification of Alternative 2
$\chi_0 = 0.0914$	$\chi_1 = 0.7558$	$\chi_2 = 1$
$\sigma_0 = 0.047$	$\sigma_1 = 0.1547$	$\chi_2 = 0.1547$
Stopping time: Aug 10, 2007, Alternative 1 is accepted.		

Table 2.5: Specifications of Second Set of Null and Alternatives

Specification of Null	Specification of Alternative 1	Specification of Alternative 2
$\chi_0 = 0.0914$	$\chi_1 = 2$	$\chi_2 = 2.2$
$\sigma_0 = 0.047$	$\sigma_1 = 0.1547$	$\chi_2 = 0.1547$
Stopping time: Aug 15, 2007, Alternative 1 is accepted.		

2.5 Conclusion and Discussion

Did the Federal Reserve respond too late to the recent financial crisis? How to evaluate the merits of such criticism? The appropriate way to evaluate this claim is to compute the optimal timing conditional on the information available when the decision was made. We adopt the mathematical tools developed in quickest detection literature to develop a framework to compute the optimal timing of the policy response to the 2007 financial crisis. The optimal timing is studied in two cases: a known alternative and unknown alternatives. In the case of a known alternative, we find that the optimal timing is Aug 9, 2007. In the case of unknown alternatives, we find that the optimal timing of the policy response is Aug 15, 2007 with certain specifications of the alternatives, it is optimal for the Fed to delay a while to accumulate more information in order to reduce the risk of identifying the wrong alternative. The contrast between the two cases highlights the importance of uncertainty for the timing decision of the policy maker. In other words, without knowing the uncertainty faced by the policy maker at the time, there is barely any merits to criticize the delay of Federal Reserve's policy response.

Chapter 3

Monetary Policy Coordination Revisited With Robust Control

3.1 Introduction

Booming international trade and financial market integration have caused members of participating countries to be unprecedentedly interdependent; and as prosperity and economic growth rate of participating countries reaching a new high level, this interdependency also brings new challenges and constrains on national policy making.

Governments often have a variety of macroeconomic targets, such as low unemployment rate, low inflation, high economic growth and current account balance surplus; equipped with fewer macroeconomic policy instruments, governments often find it impossible to achieve all targets simultaneously; policy makers will have to strike a balance among all targets. More so, the objectives of governments may be inconsistent with each other at the mean time due to heterogeneous economic structure and short run macroeconomic shocks and even the belief about the underlying macroeconomic models they employ. As has been argued by Ghosh and Masson (1994), parallel to the argument of comparative advantage in international trade literature, gain of efficiency from policy coordination arises from comparative advantage of national policy: national policy may be more efficient in achieving one or more macroeconomic targets of domestic or abroad. This comparative advantage of national policy can lead governments to exploit trade-offs among macroeconomic targets more efficiently through policy coordination.

Linked with exchange rate and interest rate, one country's macroeconomic policy may have different impact abroad, beneficial or detrimental. Due to the inward looking nature of governments, these spill-over effects are often ignored by their initiators. As a result, naturally, this leads to inefficiency of national policies. These are the two rationales often being quoted by proponents of policy coordination.

In practice, since World War II, from Bretton Woods to the European Monetary System, international monetary policy cooperation has taken place in different forms, the G8 summit, for instance, driven by the belief that independent policy making will lead to unfavorable outcomes, but its justification and magnitude of gain is till being debated among scholars.

Early contributions to international policy coordination literatures are credited to Meade (1951), **?** and Mundell (1963), they model the transmission effects of monetary and fiscal policies between two countries. Hamada (1976) studies the strategic interplay of monetary authorities in a multicounty setup with fixed exchange rate. He shows that monetary cooperation if possible can improve welfare.

Oudiz and Sachs (1984) are pioneers in exploring macroeconomic economic policy coordination; within Mundell-Fleming framework, they identify the externality of monetary policy and this leads to inefficiency of Nash Equilibrium of macroeconomic policy competition, hence policy coordination can lead to Pareto welfare improvement. In their analysis, Oudiz and Sachs (1984) assume that all participating countries know the "true model" and also agree on what the "true model" should be. They also show that since Nash Equilibrium and coordinated equilibrium don't coincide, the participating countries have incentive to deviate, hence it is necessary to design some mechanism to insure coordinated outcome. Oudiz and Sachs (1984) also provide the first estimate of welfare gain from coordination, their finding is, given a set of plausible assumptions, the welfare gain from policy coordination is no more than one percent of GDP.

It should be noted that not all scholars agree that coordination is always welfare improving, Rogoff (1985) shows that due to the time inconsistency of optimal policies, policy coordination could lead to welfare deteriorating. Other scholars disagree with this conclusion and point out that in Rogoff's paper, he assumes the structure model faced by policy makers are invariant with monetary policy regimes, i.e. coordination vs independent policy making; others challenge that assuming monetary authority has different objective to agent's might be inappropriate. Kehoe (1989) presents an alternative analysis on monetary policy cooperation, he argues that cooperation may not be desirable due to that tax policy competition may benefit residents of both countries. In his model, a benevolent government chooses tax rate on saving and labor supply, under non-cooperation, governments compete for foreign investment by lowering saving tax rate, at equilibrium, both country will not tax saving; when under cooperation, in order to reduce the distortion on labor supply caused by labor tax, both governments have the incentive to reduce labor tax, hence increase tax on saving, since saving decision has already been made before government choose saving tax rate. In equilibrium, this will result zero saving, which has lower welfare rank than non-cooperation outcome.

Obstfeld and Rogoff (2002) design a two country model featured with preset wage and monopolistic competition, they show that under either conditions, non-coordinated Nash Equilibrium coincide with coordinated equilibrium, the conditions are unitary intertemporal elasticity of consumption and shocks are global. They also calibrate the model and conclude that even gain from policy coordination exist, quantitatively it is not significant; in order to generate significant gain, unreasonably high value of intertemporal elasticity of consumption has to be assumed. Obstfeld and Rogoff (2002) and **?** and others introduce a new generation of models studying policy coordination. These new generation of models featured with monopolistic competition, sticky nominal wage and price setting. Recent contribution including Pappa (2004), Pappa and Liu (2005), and **?**. In their studies, equipped with a metric measuring social welfare with a quadratic approximation of agent's utility function, they conclude that even coordination gain exists in general, it is quantitatively small.

Another strand of literature concerns about the roll of model uncertainty in the application of international monetary coordination, Ghosh and Masson (1994) make a comprehensive survey on this topic. They conclude, in general, model uncertainty will enhance the welfare gain from policy coordination. But they assume that even though agents can not forecast the outcome, they have perfect knowledge about the distribution of the model. The question about the legitimacy of the above assumption, gives rise to another school of thoughts.

This approach presumes that policy makers are aware that their models are only approximations to the real data generating process and the distribution of potentially correct models are unknown to policy makers, then policy maker takes this possible model misspecification into account. It assumes that the true model lies in a neighborhood of the approximation model, and there is a fictitious malevolent agent who will choose the worst model in the possible model space, then decision maker choose a policy to maximize his objective, hence, his policy setting will be robust to the worst case scenario, i.e. the misspecified model that leads to the worst outcome. This school of thoughts is referred as robust control approach, which is originated from physics and engineering, introduction of robust control into economics is credited to Hansen and Sargent (2007). Robust control approach makes particular sense when applied to model policy maker's behavior, since in an world infested with stochastic shocks, the macroeconomic performance of a policy is often judged by its ex post result, even though ex ante, the policy may be justified by maximizing expected social welfare.

Few researchers have devoted to robust control approach of monetary policy design in open economies, efforts have been made by Dennis et al. (2006), Dennis et al. (2007), and Leitemo and Soderstrom (2005c).

In this paper I will study international monetary policy coordination with robust control approach. In a two country setup, both central banks minimize a social welfare loss function subject to domestic and foreign IS and Phillips curve, this is because whenever the central bank implements a policy, they know changes of foreign variables will be induced by changes of domestic variables, changes of foreign variables are governed by a set of foreign models, and this will in turn feedback to its own system; but central banks are more concerned about model misspecification coming from foreign models; when under coordination, with information swapped and action coordinated, this concern about foreign model misspecification will be gone, that is to say home central bank will be as conficent about domestic models as foreign models. I propose this can give rise to another source of gain in monetary policy coordination.

Then why would home central bank be less conficent about foreign models? As Ghosh and Masson (1994) assert that basically there two types of uncertainty coming from foreign sources, first is technological, i.e. real shocks, foreign authorities are more readily to observe and forecast these real shocks happened in their own country than home country authorities; second is strategical, it could be that foreign authorities strategically send noisy signals to mislead other parts of the world to exploit benefit. This will cause home monetary authority particularly concern about foreign model misspecification.

As has been mentioned above, current literature have concluded that the source of gain from policy coordination is mainly through endogenizing spill-over effect of independent policy making, and this gain is quantitatively small, attempts have been made to explore other sources of gain in order to enhance the gain quantitatively, for example Pappa and Liu (2005). The contribution of this paper can be summarized as twofold: first, another source of gain is identified; second, I calibrate the model and find that quantitatively, the gain from this source is significant.

This paper is organized into the following parts: the first part is introduction; in second part, I lay out the underlying model; I solved the uncoordinated problem in third part and investigate the implication of robust preference on monetary policy; in the forth part, I solve the coordinated problem; Calibration of the models can be found in fifth part; a concluding remark will be made in sixth part.

3.2 Model

Following Pappa (2004) in a two country model, economies are characterized by home and foreign countri's IS and Phillips curves[1]:

$$\begin{aligned}\pi_t &= \beta E_t\{\pi_{t+1}\} + \left(k_c + \frac{\sigma k_s}{1-\alpha}\right)c_t - \left(\frac{k_s\sigma}{1-2\alpha}\right)c_t^* &(3.1)\\ c_t &= E_t\{c_{t+1}\} - \left(\frac{1-2\alpha}{(1-\alpha)\sigma}\right)[i_t - E_t\{\pi_{t+1}\}] - \left(\frac{\alpha}{1-\alpha}\right)E_t\{c_{t+1}^* - c_t^*\} &(3.2)\end{aligned}$$

$$\begin{aligned}\pi_t^* &= \beta E_t\{\pi_{t+1}^*\} + \left(k_c + \frac{(1-\alpha)\sigma k_s}{\alpha(1-2\alpha)}\right)c_t^* - \left(\frac{(1-\alpha)\sigma k_s}{\alpha(2\alpha-1)}\right)c_t &(3.3)\\ c_t^* &= E_t\{c_{t+1}^*\} - \left(\frac{2\alpha-1}{\sigma\alpha}\right)[i_t^* - E_t\{\pi_{t+1}^*\}] - \left(\frac{1-\alpha}{\alpha}\right)E_t\{c_{t+1} - c_t\}, &(3.4)\end{aligned}$$

where $k = \frac{1-\gamma}{\gamma}(1-\gamma\beta)$, $k_c = \frac{k(\sigma+\omega)}{1+\theta\omega}$, $k_s = \frac{k\alpha(1+\psi_s)}{1+\theta\omega}$, $\psi_s = 2\eta(1-\alpha) - \frac{(1-2\alpha)}{\sigma}$. Lower case letters represent log deviation from deterministic steady state level.

Let e_t be the nominal exchange rate, from Uncovered Interest Parity condition

$$e_t = E_t e_{t+1} - (i_t - i_t^*) \tag{3.5}$$

With complete pass-through, law of one price holds

$$P_{F,t} = e_t P_{F,t}^* \tag{3.6}$$

The term of trade can be represented \notagin log form as

$$s_t = e_t + p_{F,t}^* - p_{H,t}.$$

From UIP, we'll have

$$i_t - E_t\pi_{t+1} - \left(i_t^* - E_t\pi_{t+1}^*\right) = E_t s_{t+1} - s_t, \tag{3.7}$$

we also have

$$s_t = \frac{\sigma}{1-2\alpha}(c_t - c_t^*), \tag{3.8}$$

then equations (3.1), (3.2), (3.3), (3.4), (3.7) and (3.8) consist a set of constrains that monetary policy authority has to face when solving its problem.

[1]Details of derivation can be found in Pappa (2004).

3.3 Non-Coordination

I borrowed the objective function form Pappa (2004), Under non-coordination, the home country monetary authority's problem is

$$\min_{\{i_t\}}\max_{\{v_{j,t}\}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{1}{2} \left(\pi_t^2 + \phi_c c_t^2 + \phi_s s_t^2 + \phi_{sc} s_t c_t\right), \tag{3.1}$$

$$\begin{aligned}
s.t.\pi_t &= \beta E_t \{\pi_{t+1}\} + \left(k_c + \frac{\sigma k_s}{1-\alpha}\right) c_t - \left(\frac{k_s \sigma}{1-2\alpha}\right) c_t^*, \\
c_t &= E_t \{c_{t+1}\} - \left[\frac{1-2\alpha}{(1-\alpha)\sigma}\right] [i_t - E_t \{\pi_{t+1}\}] - \left[\frac{\alpha}{1-\alpha}\right] E_t \{c_{t+1}^* - c_t^*\}, && (3.2) \\
\pi_t^* &= \beta E_t \{\pi_{t+1}^*\} + \left(k_c + \frac{(1-\alpha)\sigma k_s}{\alpha(1-2\alpha)}\right) c_t^* - \left[\frac{(1-\alpha)\sigma k_s}{\alpha(2\alpha-1)}\right] c_t + \Sigma_\pi (v_{\pi,t} + \varepsilon_{\pi,t}), \\
c_t^* &= E_t \{c_{t+1}^*\} - \left(\frac{2\alpha-1}{\sigma\alpha}\right) [i_t^* - E_t \{\pi_{t+1}^*\}] - \left(\frac{1-\alpha}{\alpha}\right) E_t \{c_{t+1} - c_t\} + \Sigma_c (v_{c,t} + \varepsilon_{c,t}), && (3.3) \\
i_t &= E_t \pi_{t+1} + \left(i_t^* - E_t \pi_{t+1}^*\right) + E_t s_{t+1} - s_t + \Sigma_e (v_{e,t} + \varepsilon_{e,t}), && (3.4) \\
s_t &= \frac{\sigma}{1-2\alpha} (c_t - c_t^*), && (3.5) \\
\eta &\geq E_0 \sum_{t=0}^{\infty} \beta^t \left[(v_{\pi,t})^2 + (v_{c,t})^2 + (v_{e,t})^2\right], && (3.6)
\end{aligned}$$

Where $\varepsilon_{\pi,t}, \varepsilon_{c,t}$ and $\varepsilon_{e,t}$ are white noises, and Σ_π, Σ_c and Σ_e are the standard deviations of respective shocks. In order to simplify computation, only shocks to foreign IS and Phillips curve and UIP are introduced. In this problem, home authority takes foreign policy as given, and ignores the spill-over effect of domestic policy on foreign welfare. Following Hansen and Sargent (2007), $v_{j,t}$, where $j \in \{\pi, c, e\}$ are processes that can feedback to the histories of π_t^*, c_t^* and i_t, which are controlled by a malevolent agent. In order to simplify computation, I assume monetary authority doesn't concern about model misspecification stemming from domestic models, they only concern about model misspecification coming from foreign models. $v_{j,t}$, where $j \in \{\pi, c, e\}$ are designed to capture this extra concern about model misspecification coming from foreign models. Equation (3.6) is the constrain for malevolent agent, the size of η indicates how confident is the monetary authority about their models; when $\eta = 0$, then there is no model misspecification.

I only consider the case where monetary authority has no commitment technology to implement their policy, so monetary authority will take all the expectations as given. I will solve this problem as a multiplier problem described in Hansen and Sargent (2007), the malevolent agent and monetary authority move simultaneously. The Lagrangian of this problem is given by

$$L = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} \frac{1}{2}\left[\pi_t^2 + \phi_c c_t^2 + \phi_s s_t^2 + \phi_{sc} s_t c_t - \theta_E \left((v_{\pi,t})^2 + (v_{c,t})^2 + (v_{e,t})^2\right)\right] + \\ \mu_{\pi,t}\left[\beta E_t\{\pi_{t+1}\} - \pi_t + \left(k_c + \frac{\sigma k_s}{1-\alpha}\right) c_t - \left(\frac{k_s \sigma}{1-2\alpha}\right) c_t^*\right] + \\ \mu_{c,t}\left[E_t\{c_{t+1}\} - c_t - \left[\frac{1-2\alpha}{(1-\alpha)\sigma}\right][i_t - E_t\{\pi_{t+1}\}] - \left[\frac{\alpha}{1-\alpha}\right] E_t\left\{c_{t+1}^* - c_t^*\right\}\right] + \\ \mu_{\pi,t}^*\left[\beta E_t\left\{\pi_{t+1}^*\right\} - \pi_t^* + \left(k_c + \frac{(1-\alpha)\sigma k_s}{\alpha(1-2\alpha)}\right) c_t^* - \left[\frac{(1-\alpha)\sigma k_s}{\alpha(2\alpha-1)}\right] c_t + \Sigma_\pi\left(v_{\pi,t} + \varepsilon_{\pi,t}\right)\right] + \\ \mu_{c,t}^*\left[E_t\left\{c_{t+1}^*\right\} - c_t^* - \left(\frac{2\alpha-1}{\sigma\alpha}\right)\left[i_t^* - E_t\left\{\pi_{t+1}^*\right\}\right] - \left(\frac{1-\alpha}{\alpha}\right) E_t\{c_{t+1} - c_t\} + \Sigma_c\left(v_{c,t} + \varepsilon_{c,t}\right)\right] + \\ \mu_{e,t}\left[E_t \pi_{t+1} - i_t + \left(i_t^* - E_t \pi_{t+1}^*\right) + E_t s_{t+1} - s_t + \Sigma_e\left(v_{e,t} + \varepsilon_{e,t}\right)\right] + \\ \mu_{s,t}\left[\frac{\sigma}{1-2\alpha}\left(c_t - c_t^*\right) - s_t\right]. \end{array} \right\} \tag{3.7}$$

From first order conditions, we can find that

$$\pi_t = -\left[k_c + \left(\frac{\delta - 1 + 2\alpha}{1-2\alpha}\right) k_s\right]^{-1} \left(\phi_c c_t + \frac{\phi_{sc}}{2} s_t\right) \tag{3.8}$$

$$v_{\pi,t} = \frac{1}{2\theta_E} \pi_t \Sigma_\pi \tag{3.9}$$

$$v_{c,t} = -\frac{1}{2\theta_E} \frac{\alpha\sigma}{1-2\alpha}\left(k_s \pi_t + \phi_s s_t + \frac{\phi_{sc}}{2} c_t\right) \Sigma_c \tag{3.10}$$

$$v_{e,t} = \frac{1}{2\theta_E}\left(k_s \pi_t + \phi_s s_t + \frac{\phi_{sc}}{2} c_t\right) \Sigma_e \tag{3.11}$$

Given the parameter values in Section 3.6, we have $A \equiv k_c + \left(\frac{\delta-1+2\alpha}{1-2\alpha}\right) k_s > 0$, $\phi_c > 0$ and $\phi_{sc} > 0$.

$$v_{\pi,t} = \frac{1}{2\theta_E} \pi_t \Sigma_\pi \tag{3.12}$$

$$v_{c,t} = -\frac{1}{2\theta_E} \frac{\alpha\sigma}{1-2\alpha}\left[\left(k_s - \frac{\phi_{sc} A}{2\phi_c}\right) \pi_t + \left(\phi_s - \frac{\phi_{sc}^2}{4\phi_c}\right) s_t\right] \Sigma_c \tag{3.13}$$

$$v_{e,t} = \frac{1}{2\theta_E}\left[\left(k_s - \frac{\phi_{sc} A}{2\phi_c}\right) \pi_t + \left(\phi_s - \frac{\phi_{sc}^2}{4\phi_c}\right) s_t\right] \Sigma_e \tag{3.14}$$

Some results established in Leitemo and Soderstrom (2005a) in a small open economy still hold:

(Optimal output-inflation trade-off) The optimal trade-off between inflation and output gap is not affected by central bank's preference for robustness.

The optimal trade-off relationship between output and inflation does not depend on central bank's preference for robustness, which is obvious from equation (3.8) Walsh (2004) derives similar result, he compare robustly optimal instrument rules Giannoni and Woodford (2003a) and Giannoni and Woodford (2003b) with robust control approach of optimal instrument rule Hansen and Sargent (2001), and conclude that the two approaches yield the same result which doesn't depend on the central bank's robustness preference. The intuition is robustness preference doesn't change the channel through which central bank counter fight inflation pressure caused by cost push shocks Leitemo and Soderstrom (2005b), i.e. changing output gap in the opposite direction.

(Misspecification and shocks) Given the preference for robustness, the absolute level of misspecification increases with variance the corresponding model.

This is evident from equation (3.9) to (3.11) this is because as the information regarding a specific model getting noisy, it is more difficult for the central bank to approximate the true model; hence they will be less confident about their model.

(Misspecification and inflation) The absolute level of misspecification increase with deviation of inflation from steady state level.

This is also evident from equation (3.12) to (3.14) the reason, as has been stated by Leitemo and Soderstrom (2005a), is that central bank fears all inflationary shocks, which will cause central bank to further contract output to counter fight inflation; hence central bank wants to design a policy that is robust to potential model misspecification concerning inflation.

There are there shocks in the system $\varepsilon_{\pi,t}, \varepsilon_{c,t}, \varepsilon_{e,t}$, each has mean zero and unitary standard deviation. I will solve for endogenous variables $\pi_t, c_t, \pi_t^*, c_t^*$ and control i_t and i_t^*. I guess that the solution should be of form

$$\begin{bmatrix} \pi_t \\ y_t \\ \pi_t^* \\ y_t^* \\ i_t \\ i_t^* \end{bmatrix} = \begin{bmatrix} a_\pi & a_c & a_e \\ b_\pi & b_c & b_e \\ c_\pi & c_c & c_e \\ d_\pi & d_c & d_e \\ e_\pi & e_c & e_e \\ f_\pi & f_c & f_e \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{c,t} \\ \varepsilon_{e,t} \end{bmatrix}. \tag{3.15}$$

The malevolent agent's strategy can be solved as of the form

$$\begin{bmatrix} v_{\pi,t} \\ v_{c,t} \\ v_{e,t} \end{bmatrix} = \begin{bmatrix} \hat{a}_\pi & \hat{a}_c & \hat{a}_e \\ \hat{b}_\pi & \hat{b}_c & \hat{b}_e \\ \hat{c}_\pi & \hat{c}_c & \hat{c}_e \\ \hat{d}_\pi & \hat{d}_c & \hat{d}_e \\ \hat{e}_\pi & \hat{e}_c & \hat{e}_e \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{c,t} \\ \varepsilon_{e,t} \end{bmatrix} \tag{3.16}$$

Given the first order condition (3.8)–(3.11) and the budget constrains, I can solve coefficients of (3.15) and (3.16) With proper parameter values, I can calibrate the objective function; by setting $\theta_E = \infty$, I can also study the case when monetary authority does not concern about model misspecification with other features of the model unchanged. This exercise is necessary because there are two sources of gain for monetary policy coordination in this model, one is the concern about foreign model misspecification, the other is the externality of monetary policy **?**. I need to identify welfare gain from each source. By comparing welfare loss when $\theta_E = \infty$ and $\theta_E = 0.0875$, I can find the welfare improvement arising from the second source of gain.

I solved the non-coordination problem is two ways: the worst case model and the approximating model Giordani and Soderlind (2004). In this paper, I only study the case when it is difficult to distinguish between the approximating model and the alternative models, that is when the distortion of alternative model is not too big relative to approximating model. I choose $\theta_E = -0.085$, which corresponds to 10% of detection error probability Hansen and Sargent (2007).

3.4 Coordination

Under coordination, I assume there is a supranational monetary authority whose objective is a weighted sum of both countries' objective function. The objective function is borrowed from Pappa (2004), which the supranational monetary authority assigns equal weight to home and foreign country. The monetary authority's problem is

$$\min_{\{i_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{1}{2} \left[\phi_c \left(c_t^2 + (c_t^*)^2 \right) + \left(2\phi_s + \phi_{sc} \left(\frac{1-2\alpha}{\sigma} \right) \right) s_t^2 + \pi_t^2 + (\pi_t^*)^2 \right] \tag{3.1}$$

$$\begin{aligned}
s.t.\pi_t &= \beta E_t \{\pi_{t+1}\} + \left(k_c + \frac{\sigma k_s}{1-\alpha} \right) c_t - \left(\frac{k_s \sigma}{1-2\alpha} \right) c_t^*, && (3.2)\\
c_t &= E_t \{c_{t+1}\} - \left[\frac{1-2\alpha}{(1-\alpha)\sigma} \right] [i_t - E_t \{\pi_{t+1}\}] - \left[\frac{\alpha}{1-\alpha} \right] E_t \{c_{t+1}^* - c_t^*\}, && (3.3)\\
\pi_t^* &= \beta E_t \{\pi_{t+1}^*\} + \left(k_c + \frac{(1-\alpha)\sigma k_s}{\alpha(1-2\alpha)} \right) c_t^* - \left[\frac{(1-\alpha)\sigma k_s}{\alpha(2\alpha-1)} \right] c_t + \Sigma_\pi \varepsilon_{\pi,t}, \\
c_t^* &= E_t \{c_{t+1}^*\} - \left(\frac{2\alpha-1}{\sigma\alpha} \right) [i_t^* - E_t \{\pi_{t+1}^*\}] - \left(\frac{1-\alpha}{\alpha} \right) E_t \{c_{t+1} - c_t\} + \Sigma_c \varepsilon_{c,t}, && (3.4)\\
i_t &= E_t \pi_{t+1} + \left(i_t^* - E_t \pi_{t+1}^* \right) + E_t s_{t+1} - s_t + \Sigma_e \varepsilon_{e,t}, && (3.5)\\
s_t &= \frac{\sigma}{1-2\alpha} (c_t - c_t^*), && (3.6)
\end{aligned}$$

Solving the first order conditions, we can get

$$\pi_t + \pi_t^* = -\frac{\phi_c}{k_c}\left(c_t + c_t^*\right) \tag{3.7}$$

We can solve the system similar to the method used in solving the non-coordination problem in section 3.3.

$$\begin{bmatrix} \pi_t \\ y_t \\ \pi_t^* \\ y_t^* \\ i_t \\ i_t^* \end{bmatrix} = \begin{bmatrix} \tilde{a}_\pi & \tilde{a}_c & \tilde{a}_e \\ \tilde{b}_\pi & \tilde{b}_c & \tilde{b}_e \\ \tilde{c}_\pi & \tilde{c}_c & \tilde{c}_e \\ \tilde{d}_\pi & \tilde{d}_c & \tilde{d}_e \\ \tilde{e}_\pi & \tilde{e}_c & \tilde{e}_e \\ \tilde{f}_\pi & \tilde{f}_c & \tilde{f}_e \end{bmatrix} \begin{bmatrix} \varepsilon_{\pi,t} \\ \varepsilon_{c,t} \\ \varepsilon_{e,t} \end{bmatrix}.$$

Results of calibration can be found in Table C.2 at the end of paper.

3.5 Incentive for Deviation

In this section, I will exam whether partners of a monetary policy coordination have incentive to deviation. I will solve home country monetary authority's problem given foreign authority is following the agreed upon path of consumption, inflation and interest rate under coordination. The only shock in this system is Uncovered Interest Parity shock.

$$\min_{\{i_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{1}{2}\left(\pi_t^2 + \phi_c c_t^2 + \phi_s s_t^2 + \phi_{sc} s_t c_t\right), \tag{3.1}$$

$$\begin{aligned} s.t.\pi_t &= \beta E_t\left\{\pi_{t+1}\right\} + \left(k_c + \frac{\sigma k_s}{1-\alpha}\right) c_t - \left(\frac{k_s \sigma}{1-2\alpha}\right) c_t^*, & (3.2)\\ c_t &= E_t\left\{c_{t+1}\right\} - \left[\frac{1-2\alpha}{(1-\alpha)\sigma}\right]\left[i_t - E_t\left\{\pi_{t+1}\right\}\right] - \left[\frac{\alpha}{1-\alpha}\right] E_t\left\{c_{t+1}^* - c_t^*\right\}, & (3.3)\\ i_t &= E_t \pi_{t+1} + \left(i_t^* - E_t \pi_{t+1}^*\right) + E_t s_{t+1} - s_t + \Sigma_e \varepsilon_{e,t}, & (3.4) \end{aligned}$$

$$L = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \begin{array}{l} \frac{1}{2}\left[\pi_t^2 + \phi_c c_t^2 + \phi_s s_t^2 + \phi_{sc} s_t c_t\right] + \\ \mu_{\pi,t}\left[\beta E_t\left\{\pi_{t+1}\right\} - \pi_t + \left(k_c + \frac{\sigma k_s}{1-\alpha}\right) c_t - \left(\frac{k_s \sigma}{1-2\alpha}\right) c_t^*\right] + \\ \mu_{c,t}\left[E_t\left\{c_{t+1}\right\} - c_t - \left[\frac{1-2\alpha}{(1-\alpha)\sigma}\right]\left[i_t - E_t\left\{\pi_{t+1}\right\}\right] - \left[\frac{\alpha}{1-\alpha}\right] E_t\left\{c_{t+1}^* - c_t^*\right\}\right] + \\ \mu_{e,t}\left[E_t \pi_{t+1} - i_t + \left(i_t^* - E_t \pi_{t+1}^*\right) + E_t s_{t+1} - s_t + \Sigma_e \varepsilon_{e,t}\right]. \end{array} \right\} \tag{3.5}$$

I calibrate this model and find that by deviation, welfare loss can be minimized to almost zero. Readers can find the results in Table C.3 at the end of paper.

3.6 Calibration

I calibrated the results of the two problems with a set of parameter value from Pappa (2004), parameter values can be found in Table C.1 at the end of paper. As has been argued, there are two sources of welfare gain from monetary coordination: 1. monetary policy externality, 2. concern about foreign model misspecification. The proportion of welfare improvement from monetary coordination arise from concern about foreign model misspecification consists of 10% total welfare improvement from monetary policy coordination.

3.7 Concluding Remarks

In a world with insurmountable complexity, no models can capture all aspects of the economy; together with stochastic shocks, decision makers often do not trust the underlying models they employ, but only consider them as approximations to the true data generating process. Given this concern about model misspecification, decision makers want their policy to be robust to the worst case scenario. In the context of international monetary policy design, policy makers are more concern about model misspecification coming from foreign sources due to technical and strategical reasons. When central banks coordinate with each other, with information swapped and a common objective, there is no need to worry about foreign model misspecification, and this in turn, can result in welfare gain. As the result of calibration reveals, welfare gain from concern about foreign model misspecification consists of 10% of total welfare improvement from policy coordination; hence, concern about foreign model misspecification is a significant source of gain. It is evident from Table C.3, once coordination is established, both participating members have incentive to deviate; hence in order to guarantee the coordinated outcome, some punishment mechanism must be implemented.

References

Bibliography

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.

Blanchard, Olivier Jean and Charles M Kahn, "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, July 1980, *48* (5), 1305–11.

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.

Cecchetti, Stephen G., "Crisis and Responses: the Federal Reserve and the Financial Crisis of 2007-2008," Working Paper 14134, National Bureau of Economic Research June 2008.

Curdia, Vasco and Michael Woodford, "Credit frictions and optimal monetary policy," BIS Working Papers 278, Bank for International Settlements March 2009.

_ **and** _ , "Credit Spreads and Monetary Policy," NBER Working Papers 15289, National Bureau of Economic Research, Inc August 2009.

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

Dayanik, Savas, Christian Goulding, and H. Vincent Poor, "Bayesian sequential change diagnosis," *Math. Oper. Res.*, 2008, *33* (2), 475–496.

Dennis, Richard, Kai Leitemo, and Ulf Soderstrom, "Monetary Policy and Model Uncertainty in a Small Open Economy," Computing in Economics and Finance 2006 14, Society for Computational Economics July 2006.

_ **,** _ **, and** _ , "Monetary Policy in a Small Open Economy with a Preference for Robustness," CEPR Discussion Papers 6067, C.E.P.R. Discussion Papers January 2007.

Freedman, Charles, "General Discussion: Monetary Policy and Real Stabilization," in "Rethinking Stabilization Policy, A Symposium Sponsored by the Federal Reserve Bank of Kansas City" Federal Reserve Bank of Kansas City 2002, pp. 319–331.

Ghosh, Atich R and Paul R Masson, *Economic Cooperation in an Uncertain World*, Blackwells, Oxford, 1994.

Giannoni, Marc P. and Michael Woodford, "Opitmal Interest-Rate Rules: I. General Theory," NBER Working Papers 9419, National Bureau of Economic Research, Inc January 2003.

_ **and** _ , "Optimal Interest-Rate Rules: II. Applications," NBER Working Papers 9420, National Bureau of Economic Research, Inc January 2003.

Giordani, Paolo and Paul Soderlind, "Solution of macromodels with Hansen-Sargent robust policies: some extensions," *Journal of Economic Dynamics and Control*, December 2004, *28* (12), 2367–2397.

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.

Hamada, Koichi, "A Strategic Analysis of Monetary Interdependence," *The Journal of Political Economy*, 1976, *84* (4), 677–700.

Hansen, Lars Peter and Thomas J. Sargent, "Robust Control and Model Uncertainty," *The American Economic Review*, 2001, *91* (2), 60–66.

_ **and** _ , *Robustness*, Princeton University Press, 2007.

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

John, Authers, "Bernanke Acts Too Late to Help Nikkei," *Financial Times*, Aug 18 2007.

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.

Kehoe, Patrick J., "Policy Cooperation Among Benevolent Governments May Be Undesirable," *The Review of Economic Studies*, 1989, *56* (2), 289–296.

Leitemo, Kai and Ulf Soderstrom, "Robust Monetary Policy in a Small Open Economy," Working Papers 290, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University 2005.

_ **and** _ , "Robust monetary policy in the New-Keynesian framework," Macroeconomics 0508032, EconWPA August 2005.

_ **and** _ , "Simple monetary policy rules and exchange rate uncertainty," *Journal of International Money and Finance*, April 2005, *24* (3), 481–507.

Liptser, Robert S. and Albert N. Shiryaev, *Statistics of random processes. I*, expanded ed., Vol. 5 of *Applications of Mathematics (New York)*, Berlin: Springer-Verlag, 2001. General theory, Translated from the 1974 Russian original by A. B. Aries, Stochastic Modelling and Applied Probability.

Meade, James, *The Balance of Payments*, London: Oxford University Press, 1951.

Mishkin, Frederic S., "Monetary policy flexibility, risk management, and financial disruptions," *Journal of Asian Economics*, June 2010, *21* (3), 242–246.

Mizen, Paul, "The credit crunch of 2007-2008: a discussion of the background, market reactions, and policy responses," *Review*, 2008, (Sep), 531–568.

Mundell, Robert A., "Capital Mobility and Stabilization Policy under Fixed and Flexible Exchange Rates," *The Canadian Journal of Economics and Political Science / Revue canadienne d'Economique et de Science politique*, 1963, *29* (4), 475–485.

Nickell, Stephen, "Monetary Policy Issues: Past, Present, Future," in "Speech on June 19, 2002" Bank of England 2002.

Obstfeld, Maurice and Kenneth Rogoff, "Global Implications Of Self-Oriented National Monetary Rules," *The Quarterly Journal of Economics*, May 2002, *117* (2), 503–535.

Oudiz, Gilles and Jeffrey Sachs, "Macroeconomic Policy Coordination among the Industrial Economies," *Brookings Papers on Economic Activity*, 1984, *15* (1984-1), 1–76.

Pappa, Evi, "Do the ECB and the fed really need to cooperate? Optimal monetary policy in a two-country world," *Journal of Monetary Economics*, May 2004, *51* (4), 753–779.

_ **and Zheng Liu**, "Gains from International Monetary Policy Coordination: Does It Pay to Be Different?," Computing in Economics and Finance 2005 457, Society for Computational Economics November 2005.

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

Poole, William, "Milton and Money Stock Control," Federal Reserve Bank of St. Louise July 2007.

Poor, Vincent and Olympia Hadjiliadis, *Quickest Detection*, Cambridge University Press, 2009.

Rogoff, Kenneth, "Can international monetary policy cooperation be counterproductive?," *Journal of International Economics*, May 1985, *18* (3-4), 199–217.

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.

Svensson, Lars E O, "Optimal Policy with Low-Probability Extreme Events," CEPR Discussion Papers 4218, C.E.P.R. Discussion Papers February 2004.

Svensson, Lars E.O, "Monetary policy and real stabilization," *Proceedings*, 2002, pp. 261–312.

Taylor, John B., "Monetary Policy and the State of the Economy," testimony before the Committee on Financial Services, U.S. House of Representatives February 2008.

_ **and John C. Williams**, "A Black Swan in the Money Market," Working Paper 13943, National Bureau of Economic Research April 2008.

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

Walsh, Carl E., "Robustly Optimal Instrument Rules and Robust Control: An Equivalence Result," *Journal of Money, Credit and Banking*, 2004, *36* (6), 1105–1113.

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.

Appendix A

Appendix for Chapter 1

A.1 Some Figures

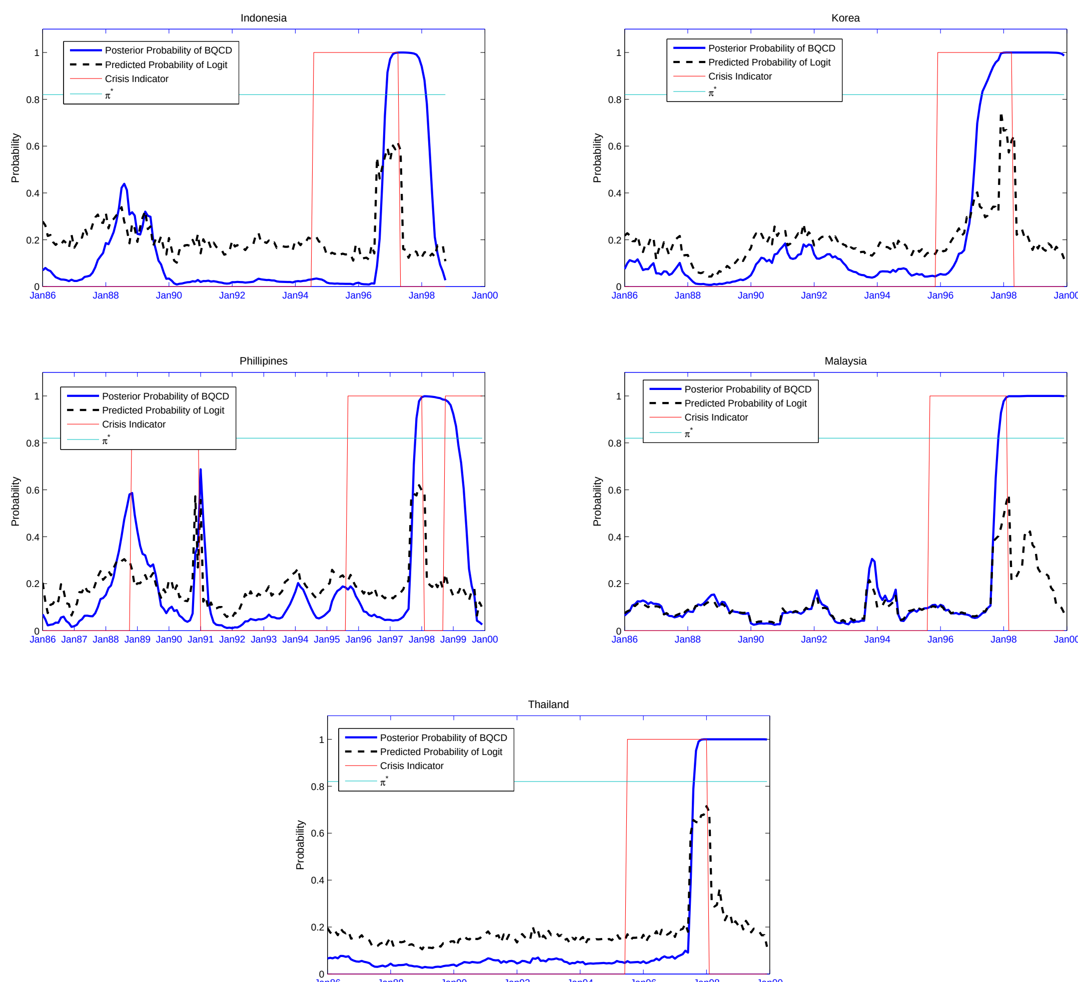


Figure A.1: Out-of-Sample Predicted Probability of 1997 Asian Financial Crises with $\pi^* = 0.82$

A.2 Proof of Proposition 1.2.4

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 $k \leq l$

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

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

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

Where E_- and E_+ are defined with respect to probability measure constructed through Radon-Nikodým derivative (1.3) and (1.4) respectively. Next I 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. I 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 $k \leq i^-$.

Given $\theta = l$ and $\delta > 0$ for $l+h \geq k \geq l^+$, and let $l^+ = l+\delta$, I have[1]

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

[1] $x^- = -\inf\{x,0\}$ and $x^+ = \sup\{x,0\}$.

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

hence I will have

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

for $k \leq l$, I have

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

Hence,

$$\begin{aligned} v(k,l) - v_-(l,l) &= [v_-(k,l) - v_-(l,l)]\,1_{\{k\leq l\}} + [v_+(k,l) - v_-(l,l)]\,1_{\{k>l\}} && \text{(A.1)} \\ &\approx -v'_-(l,l)(l-k)^+ + v'_+(l^+,l)(k-l^+)^+ + v'_+(l^+,l)\,\delta 1_{\{k>l\}} && \text{(A.2)} \\ &= -v'_-(l,l)(l-k)^+ + v'_+(l^+,l)(k-l)^+ && \text{(A.3)} \end{aligned}$$

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

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

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

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

T_n is a stopping time and $\lim_{n\to\infty} T_n = T$[2].

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

[2]Please see Problem 2.24 in Karatzas and Shreve (1991)

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

where E_θ is expectation conditional on θ, and P_θ is the probability distribution given θ, defined by (1.5).

$$\begin{aligned} E_\pi\left\{E_\theta\left[v(T,l)-v(l,l)\right]\right\} &\approx E_\pi\left\{E_\theta\left[-v'_-(l,l)(l-T)^+ + v'_+\left(l^+,l\right)1_{\{T>l\}}\right]\right\} \\ E\left\{v(T,\theta)-v(\theta,\theta)\right\} &\approx E\left\{-v'_-(\theta,\theta)(\theta-T)^+ + v'_+\left(\theta^+,\theta\right)1_{\{T>\theta\}}\right\} \end{aligned} \tag{A.4}$$

here E is the expectation under probability measure P, defined by (1.6). 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.

A.3 Proof of Proposition 1.2.5

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

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

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

I 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 I have

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

Let

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

The rest is easy to show that

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

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

A.4 Proof of Proposition 1.2.4

Given the condition in Lemma 1.2.4 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 $c=-\frac{v'_-\left(\theta,\theta\right)}{v'_+\left(\theta^+,\theta\right)}$ is a constant.

A.5 Proof of Proposition 1.2.5

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, I 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.

A.6 Logit Regression

Table A.1: Logit Maximum Likelihood Estimates

Variables	Coefficient	t-statistics	t-probability
Constant	-1.218744	-6.424941	0.000000
Growth of International Reserves	-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		

Appendix B

Appendix for Chapter 2

B.1 Appendix

Given the risk function (2.9)

$$R(T) = E_{\pi}\left\{\left[V_{-}(T,\theta) - V_{+}(\theta,\theta)\right]1_{\{T<\theta\}} + \left[V_{+}(T,\theta) - V_{+}(\theta,\theta)\right]1_{\{T\geq\theta\}}\right\},$$

Given $T = k$ and $\theta = i$, define

$$v(k,i) = v_{-}(k,i)\,1_{\{k<i\}} + v_{+}(k,i)\,1_{\{k\geq i\}},$$

where, when $k < i$

$$\begin{aligned}
v_{-}(k,i) &= E_{-}\left\{\int_0^k e^{-\rho t}L(\xi_t,u_{0,t})\,dt + \int_k^i e^{-\rho t}L(\xi_t,u_{1,t})\,dt + \int_i^\infty e^{-\rho t}L(\xi_t,u_{1,t})\,dt\right\} \\
d\xi_t &= k_0(\xi_t,u_{0,t})\,dt + \sigma dW_t,\ t \leq k \\
d\xi_t &= k_0(\xi_t,u_{1,t})\,dt + \sigma dW_t,\ k < t < i \\
d\xi_t &= k_1(\xi_t,u_{1,t})\,dt + \sigma dW_t,\ i \leq t < \infty;
\end{aligned}$$

and when $k \geq i$

$$\begin{aligned}
v_{+}(k,i) &= E_{+}\left\{\int_0^i e^{-\rho t}L(\xi_t,u_{0,t})\,dt + \int_i^k e^{-\rho t}L(\xi_t,u_{0,t})\,dt + \int_k^\infty e^{-\rho t}L(\xi_t,u_{1,t})\,dt\right\} \\
d\xi_t &= k_0(\xi_t,u_{0,t})\,dt + \sigma dW_t,\ t < i \\
d\xi_t &= k_1(\xi_t,u_{0,t})\,dt + \sigma dW_t,\ i \leq t < k \\
d\xi_t &= k_1(\xi_t,u_{1,t})\,dt + \sigma dW_t,\ k \leq t < \infty.
\end{aligned}$$

Where E_- and E_+ are defined with respect to probability measure constructed through Radon-Nikodým derivative (2.4) and (2.5) respectively. Next we need to linearize $v_-(k,i)$ around $k = i^-$, since $v_-(k,i)$ is defined on $k < i$, $v_-(i,i)$ is not defined. We need to choose an approximation, let $\delta > 0$ be an arbitrarily small positive number, then $v_-(k,i)$ can be linearized on $i^- = i - \delta$, for $k \leq i^-$.

Given $\theta = i$ and $\delta > 0$ for $k \leq i^-$, and let $i^- = i - \delta$, we have[1]

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

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

hence we will have

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

for $k \geq i$, we have

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

Hence,

$$\begin{aligned} v(k,i) - v_+(i,i) &= [v_-(k,i) - v_+(i,i)]\,1_{\{k<i\}} + [v_+(k,i) - v_+(i,i)]\,1_{\{k\geq i\}} && \text{(B.1)} \\ &\approx -v'_-\left(i^-,i\right)\left(i^- - k\right)^+ + v'_+(i,i)(k-i)^+ - v'_-\left(i^-,i\right)\delta 1_{\{k<i\}} && \text{(B.2)} \end{aligned}$$

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

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

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

T_n is a stopping time and $\lim_{n\to\infty} T_n = T$[2].

Conditional on $\theta = i$ and given $\delta > 0$, we have the following:

[1] $x^- = -\inf\{x,0\}$ and $x^+ = \sup\{x,0\}$.

[2] Please see Problem 2.24 in Karatzas and Shreve (1991)

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

where E_θ is expectation conditional on θ, and P_θ is the probability distribution given θ, defined by (2.27).

$$\begin{aligned}
R\left(T\right) &\approx E_\pi\left[E_\theta\left(-v'_-\left(i^-,i\right)\left(i-T\right)^+ + v'_+\left(i,i\right)\left(T-i\right)^+ - v'_-\left(i^-,i\right)\delta 1_{\{T<i\}}\right)\right] \\
&= E\left[-v'_-\left(\theta^-,\theta\right)\left(\theta-T\right)^+ + v'_+\left(\theta,\theta\right)\left(T-\theta\right)^+ - v'_-\left(\theta^-,\theta\right)\delta 1_{\{T<\theta\}}\right] \qquad \text{(B.3)}
\end{aligned}$$

here E is the expectation under probability measure P, defined by (2.28).

B.2 Appendix

Given the objective of Equation (2.12)

$$\tilde{E}\left\{\left(\theta^- - T\right)^+ + c\left(T-\theta\right)^+ + \delta 1_{\{T<\theta\}}\right\}$$

Define $\pi_k=\tilde{P}\left(\theta\le k|\mathcal{F}_k\right)$. Let's first consider $\tilde{E}\left[c\left(T-\theta\right)^+\right]$, let

$$\begin{aligned}
D_k &= c\tilde{E}\left\{\left(k-\theta\right)^+|\mathcal{F}_k\right\} \\
&= c\int_0^k\left(k-s\right)d\tilde{P}\left(\theta\le s|\mathcal{F}_k\right) \\
&= \left(k-s\right)\tilde{P}\left(\theta\le s|\mathcal{F}_k\right)|_0^k-\int_0^k\tilde{P}\left(\theta\le s|\mathcal{F}_k\right)d\left(k-s\right) \\
&= \int_0^k\tilde{P}\left(\theta\le s|\mathcal{F}_k\right)ds \\
&= \int_0^k\tilde{P}\left(\theta\le s|\mathcal{F}_s\right)ds+\int_0^k\left[\tilde{P}\left(\theta\le s|\mathcal{F}_k\right)-\tilde{P}\left(\theta\le s|\mathcal{F}_s\right)\right]ds \\
&= \int_0^k\pi\left(s\right)ds+\int_0^k\left[P\left(\theta\le s|\mathcal{F}_k\right)-\pi\left(s\right)\right]ds \\
&= \int_0^k\pi\left(s\right)ds+M_k,
\end{aligned}$$

where

$$M_k = \int_0^k \left[\tilde{P}(\theta \le s|\mathcal{F}_k) - \pi(s)\right] ds.$$

Next, we need to show $E(M_T) = 0$, we do so by employing Optional Sampling theorem(Poor and Hadjiliadis, 2009, page 24). To employ Optional sampling theorem, we need to show

1. $\{M_k, \mathcal{F}_k\}$ is a martingale.

2. There is an integrable M such that $M_k = E\{M|\mathcal{F}_k\}$, so M_k is regular[3].

then by Optional Sampling Theorem $E(M_T) = E(M_0)$ for any stopping time T.

B.2.0.1 M_k is a Martingale

$$M_k = \int_0^k \left[\tilde{P}(\theta \le s|\mathcal{F}_k) - \pi(s)\right] ds$$

for $m < k$,

$$\begin{aligned}
\tilde{E}(M_k|\mathcal{F}_m) &= E\left\{\int_0^k \left[\tilde{P}(\theta \le s|\mathcal{F}_k) - \pi(s)\right] ds|\mathcal{F}_m\right\} \\
&= \int_0^k E\left\{\left[\tilde{P}(\theta \le s|\mathcal{F}_k) - \tilde{P}(\theta \le s|\mathcal{F}_s)\right]|\mathcal{F}_m\right\} ds \\
&= \int_0^m \left[\tilde{P}(\theta \le s|\mathcal{F}_m) - \tilde{P}(\theta \le s|\mathcal{F}_s)\right] ds + \int_m^k \left[\tilde{P}(\theta \le s|\mathcal{F}_m) - \tilde{P}(\theta \le s|\mathcal{F}_m)\right] ds \\
&= M_m
\end{aligned}$$

B.2.0.2 M is Integrable

let

$$M = \int_0^\infty \left[1 - \pi_m - 1_{\{\theta > m\}}\right] dm$$

[3]Definition of regular Martingale can be found in Poor and Hadjiliadis (2009), page 23.

First, consider

$$\begin{aligned}\tilde{E}\left\{\int_0^k (1-\pi_m)\,dm\right\} &= \int_0^k \tilde{P}(\theta > m)\,dm \\ &= m\left(1-\tilde{P}(\theta \le m)\right)|_0^k - \int_0^k md\left(1-\tilde{P}(\theta \le m)\right) \\ &= k\left(1-\tilde{P}(\theta \le k)\right) + \int_0^k md\tilde{P}(\theta \le m) \\ &\to E(\theta) < \infty\end{aligned}$$

Similarly

$$\begin{aligned}\tilde{E}\left\{\int_0^k 1_{\{\theta > m\}}dm\right\} &= \int_0^k \tilde{P}(\theta > m)\,dm \\ &\to \tilde{E}(\theta) < \infty\end{aligned}$$

Then we have

$$\begin{aligned}\tilde{E}\{|M|\} &\le \tilde{E}\left\{\int_0^k |1-\pi_m|\,dm + \int_0^k \left|1_{\{\theta > m\}}\right| dm\right\} \\ &\le 2\tilde{E}\{\theta\} < \infty\end{aligned}$$

B.2.0.3 M is regular

$$\begin{aligned}E(M|\mathcal{F}_k) &= \int_0^\infty \left\{E\left[\tilde{P}(\theta > m|\mathcal{F}_m)\,|\mathcal{F}_k\right] - \tilde{P}(\theta > m|\mathcal{F}_k)\right\} dm \\ &= \int_0^k \left\{\tilde{P}(\theta > m|\mathcal{F}_m) - \tilde{P}(\theta > m|\mathcal{F}_k)\right\} dm \\ &= \int_0^k \left\{\tilde{P}(\theta \le m|\mathcal{F}_k) - \tilde{P}(\theta \le m|\mathcal{F}_m)\right\} dm \\ &= M_k\end{aligned}$$

So M_k is regular.

By Optimal Sampling Theorem, we have showed

$$E\{M_T\} = E(M_0) = 0,$$

hence, we have

$$\tilde{E}\left\{(T-\theta)^{+}\right\}=\tilde{E}\left\{\int_{0}^{T}\pi\left(s\right)ds\right\}.$$

B.2.0.4 $\tilde{E}\left\{\left(\theta^{-}-T\right)^{+}\right\}$

Now, let's see second part of the proof. We have

$$\tilde{E}\left\{(\theta-\delta-T)^{+}\right\}=\tilde{E}\left\{(\theta-(T+\delta))^{+}\right\}$$

let $T'=T+\delta$,

$$\tilde{E}\left\{\left(\theta-T'\right)^{+}\right\}=\tilde{E}\left\{D_{T'}\right\}$$

$$\begin{aligned}
D_k &= \tilde{E}\left\{(m-k)^{+}|\mathcal{F}_k\right\}\\
&= \int_k^{\infty}(m-k)\,d\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\\
&= -\int_k^{\infty}(m-k)\,d\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)\\
&= -(m-k)\left(1-\tilde{p}\left(\theta\le m|\mathcal{F}_k\right)\right)|_k^{\infty}+\int_k^{\infty}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)d\left(m-k\right)\\
&= \int_k^{\infty}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)dm\\
&= \int_0^{\infty}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)dm-\int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)dm\\
&= \tilde{E}\left\{\theta|\mathcal{F}_k\right\}-\int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)dm\\
&= \tilde{E}\left\{\theta|\mathcal{F}_k\right\}-\int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_m\right)\right)dm+\int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_m\right)\right)dm-\int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)dm\\
&= \tilde{E}\left\{\theta|\mathcal{F}_k\right\}-\int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_m\right)\right)dm+N_k
\end{aligned}$$

where

$$\begin{aligned}
N_k &= \int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_m\right)\right)-\int_0^{k}\left(1-\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)\right)dm\\
&= \int_0^{k}\left(\tilde{P}\left(\theta\le m|\mathcal{F}_k\right)-\tilde{P}\left(\theta\le m|\mathcal{F}_m\right)\right)dm
\end{aligned}$$

we already have $\tilde{E}\left\{N_T\right\}=\tilde{E}\left\{N_0\right\}=0$ from previous result.

Hence we have prove the following theorem

$$\tilde{E}\left\{\left(\theta^{-}-T\right)^{+}+c\left(T-\theta\right)^{+}+1_{\{T<\theta\}}\right\} = \tilde{E}\left\{\theta+\left(1-\pi_{T}\right)+\int_{0}^{T}\left[c\pi_{m}-\left(1-\pi_{m}\right)\right]dm\right\}$$

B.3 Appendix

In the context of Liptser and Shiryaev (2001, chapter 9.4), $E=\{0,1\}$, the prior $P\left(\theta\geq t|\theta>0\right)=e^{-\lambda t}, P\left(\theta=0\right)=p.$

$$\lambda_{00}=-\lambda, \lambda_{01}=\lambda, \lambda_{10}=0, \lambda_{11}=0,$$

$$\pi\left(t\right)=P\left(\theta_{t}=1|\mathcal{F}_{t}^{\xi}\right)\ or, =P\left(\theta\leq t|\mathcal{F}_{t}^{\xi}\right)$$

here $\pi_{\beta}\left(t\right)=\pi\left(t\right)$, where $\beta=1.$ we solve the following elements in theorem 9.1

$$\begin{aligned}\Omega^{*}\pi\left(u\right) &= \sum_{\gamma\in\{0,1\}}\lambda_{\gamma1}\pi_{\gamma}\left(u\right)\\ &= \lambda_{01}\pi_{0}\left(u\right)+\lambda_{11}\pi_{1}\left(u\right)\\ &= \lambda\left(1-\pi\left(u\right)\right)\end{aligned}$$

$$\begin{aligned}\bar{A}_{u}\left(\xi\right) &= \sum_{\gamma\in\{0,1\}}A_{u}\left(\gamma,\xi\right)\pi_{\gamma}\left(u\right)\\ &= k_{0}\left(\xi_{t},u_{0,t}\right)\pi_{0}\left(u\right)+k_{1}\left(\xi_{t},u_{0,t}\right)\pi\left(u\right)\\ &= k_{0}\left(\xi_{t},u_{0,t}\right)\left(1-\pi\left(u\right)\right)+k_{1}\left(\xi_{t},u_{0,t}\right)\pi\left(u\right)\end{aligned}$$

then the third part of theorem 9.1 can be presented as

$$\begin{aligned}&\int_{0}^{t}\pi\left(u\right)\frac{k_{1}\left(\xi_{t},u_{0,t}\right)-\left[k_{0}\left(\xi_{t},u_{0,t}\right)\left(1-\pi\left(u\right)\right)+k_{1}\left(\xi_{t},u_{0,t}\right)\pi\left(u\right)\right]}{\sigma}d\bar{W}_{u}\\ &=\int_{0}^{t}\pi\left(u\right)\left(1-\pi\left(u\right)\right)\frac{k_{1}\left(\xi_{t},u_{0,t}\right)-k_{0}\left(\xi_{t},u_{0,t}\right)}{\sigma}d\bar{W}_{u}\end{aligned}$$

where

$$d\bar{W}_{t}=\frac{1}{\sigma}\left\{d\xi_{t}-\left[k_{0}\left(\xi_{t},u_{0,t}\right)\left(1-\pi\left(t\right)\right)+k_{1}\left(\xi_{t},u_{0,t}\right)\pi\left(t\right)\right]dt\right\}$$

B.4 Appendix

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

For the part $E\left[c\,(T-\theta)^{+}\right]$, the proof is similar to Poor and Hadjiliadis (2009), hence we have

$$\begin{aligned} E\left[c\,(T-\theta)^{+}\right] &= E\left\{c\sum_{m=0}^{T-1}\pi_m\right\} \\ E\left[1_{\{T<\theta\}}\right] &= 1-\pi_T. \end{aligned}$$

For $E\left[(\theta-1-T)^{+}\right]$, we need to show, again let $E\left[(\theta-1-T)^{+}\right]=E\left[C_T\right]$, where

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

then following the similar proof before, and we can show

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

hence we have

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

B.5 Appendix

Rewrite the model as

$$g_0 \begin{bmatrix} E_t\Omega_{t+1} \\ E_tY_{t+1} \\ E_t\pi_{t+1} \\ u_{t+1} \end{bmatrix} = c + g_1 \begin{bmatrix} \Omega_t \\ Y_t \\ \pi_t \\ u_t \end{bmatrix} + g_2 \begin{bmatrix} \epsilon_t \\ \varepsilon_{t+1} \end{bmatrix},$$

where

$$g_0 = \begin{bmatrix} \delta & 0 & 0 & 0 \\ \bar{\sigma}\left(s_\Omega + \psi_\Omega\right) & 1 & \bar{\sigma} & 0 \\ 0 & 0 & -\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$c = \begin{bmatrix} -\chi_0 - \left(\chi_1 - \chi_0\right) 1_{t\geq i} \\ \bar{\sigma}\left(\chi_0 + \left(\chi_1 - \chi_0\right) 1_{\{t\geq i\}}\right)\left(\pi_b - \phi_\omega 1_{t\geq k}\right) 1_{t\geq i} \\ 0 \\ 0 \end{bmatrix},$$

$$g_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \bar{\sigma}s_\Omega & \left(1 + \bar{\sigma}\phi_y\right) & \bar{\sigma}\phi_\pi & 0 \\ \xi\left(s_\Omega + \pi_b - \gamma_b\right) & \kappa & -1 & 1 \\ 0 & 0 & 0 & \rho \end{bmatrix},$$

$$g_2 = \begin{bmatrix} -\left(\sigma_0 + \left(\sigma_1 - \sigma_0\right) 1_{\{t\geq i\}}\right) & 0 \\ \bar{\sigma}\left(\pi_b 1_{t\geq i} - \phi_\omega 1_{t\geq k}\right)\left(\sigma_0 + \left(\sigma_1 - \sigma_0\right) 1_{\{t\geq i\}}\right) & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

The model can further rewritten as

$$\begin{bmatrix} E_t\Omega_{t+1} \\ E_tY_{t+1} \\ E_t\pi_{t+1} \\ u_t \end{bmatrix} = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} + A \begin{bmatrix} \Omega_t \\ Y_t \\ \pi_t \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \begin{bmatrix} \epsilon_t \\ \varepsilon_t \end{bmatrix}.$$

Now proceed with Blanchard and Kahn (1980)'s method, decompose $A = \Lambda^{-1} J \Lambda$. At any t, C_1, D_1 will only depend on the deep parameters including k, i. Let $x_t = [\Omega_t, Y_t, \pi_t]'$, then we can write the system as

$$\begin{bmatrix} E_t \acute{x}_{t+1} \\ \acute{u}_t \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} + \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix} \begin{bmatrix} \acute{x}_t \\ \acute{u}_{t-1} \end{bmatrix} + \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} \begin{bmatrix} \epsilon_t \\ \varepsilon_t \end{bmatrix},$$

where

$$\begin{bmatrix} \acute{x}_t \\ \acute{u}_t \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix};$$
$$\begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix};$$
$$\begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}.$$

Now we have the explosive component

$$E_t \acute{x}_{t+1} = C_1 + J_1 \acute{x}_t + D_1 f_t,$$
$$\acute{x}_t = J_1^{-1} E_t \acute{x}_{t+1} - J_1^{-1} C_1 - J_1^{-1} D_1 f_t,$$

where $f_t = [\xi_t, \varepsilon_t]'$, solve this forward

$$\acute{x}_t = J_1^{-2} E_t \acute{x}_{t+2} - J_1^{-2} C_1 - J_1^{-1} C_1 + J_1^{-2} D_1 E_t f_{t+1} - J_1^{-1} D_1 f_t,$$

since ξ_t are i.i.d. shocks

$$\begin{aligned} \acute{x}_t &= J_1^{-2} E_t \acute{x}_{t+2} - J_1^{-2} C_1 - J_1^{-1} C_1 - J_1^{-1} D_1 f_t \\ &= J_1^{-\infty} E_t \acute{x}_\infty - \sum_{i=0}^{\infty} J_1^{-(i+1)} C_1 - J_1^{-1} D_1 f_t \end{aligned}$$

we have

$$\begin{aligned}
\acute{x}_t &= \Lambda_{11}x_t + \Lambda_{12}u_t \\
x_t &= \Lambda_{11}^{-1}\acute{x}_t - \Lambda_{11}^{-1}\Lambda_{12}u_t \\
&= -\Lambda_{11}^{-1}\left(\Lambda_{12}u_t + \sum_{i=0}^{\infty} J_1^{-(i+1)}C_1 + J_1^{-1}D_1 f_t\right) \\
&= -\Lambda_{11}^{-1}\left(\Lambda_{12}u_t + J_1^{-1}D_1 f_t + \sum_{i=0}^{\infty} J_1^{-(i+1)}C_1\right).
\end{aligned}$$

The last equation is true, because J_1 only contains explosive eigenvalues. Then the variance of x_t at t can be expressed as

$$var(x_t) = \left(\Lambda_{11}^{-1}\Lambda_{12}\right) var(u_t) \left(\Lambda_{11}^{-1}\Lambda_{12}\right)' + \left(\Lambda_{11}^{-1}J_1^{-1}D_1\right) var(f_t) \left(\Lambda_{11}^{-1}J_1^{-1}D_1\right)',$$

B.6 Appendix

Given the condition in Lemma 2.2.3 is satisfied, the period loss function can then be expressed as

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

For $k \geq i$

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

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

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

B.7 Appendix

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

$$\begin{aligned} v_{-}\left(i-2,i\right)-v_{-}\left(i-1,i\right) &= \sum_{t=0}^{\infty}\beta^{t}\left\{L\left(1_{\{t\geq i-2\}},1_{\{t\geq i\}}\right)-L\left(1_{\{t\geq i-1\}},1_{\{t\geq i\}}\right)\right\} \\ &= \beta^{i-2}\left\{\underbrace{L(1,0)}_{t=i-2}-\underbrace{L(0,0)}_{t=i-2}\right\} \end{aligned}$$

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

Finally, we have

$$\begin{aligned} c &= -\frac{v_{+}\left(i-1,i\right)-v_{+}\left(i,i\right)}{v_{-}\left(i-2,i\right)-v_{-}\left(i-1,i\right)} \\ &= -\beta^{-2}\frac{L\left(1,0\right)-L\left(0,0\right)}{L\left(0,1\right)-L\left(1,1\right)}. \end{aligned}$$

Clearly, c is not a function of i.

B.8 Appendix

Given the objective, we will solve this problem as an optimal stopping problem. First define

$$\pi_{k}=\tilde{P}\left(\theta\leq k|\mathcal{F}_{k}^{\xi}\right). \tag{B.4}$$

We can derive $d\pi_k$ with Theorem 9.1 of Liptser and Shiryaev (2001)[4].

$$d\pi_{t}=\lambda\left(1-\pi_{t}\right)dt+\frac{1}{\sigma}\pi_{t}\left(1-\pi_{t}\right)\left(k_{1}^{s}\left(\xi_{t}^{s},u_{0,t}^{*}\right)-k_{0}^{s}\left(\xi_{t}^{s},u_{0,t}^{*}\right)\right)d\bar{W}_{t}, \tag{B.5}$$

Superscript s denotes π_t is derived based on the signal process.

[4]Please see Appendix B.3 for detailed derivation

Assume $\tilde{E}(\theta)$ and $\tilde{E}(T)$ are finite, then

$$\tilde{E}\left\{\left(\theta^{-}-T\right)^{+}+c\left(T-\theta\right)^{+}+1_{\{T<\theta\}}\right\}=\tilde{E}\left\{\theta+\left(1-\pi_{T}\right)+\int_{0}^{T}\left[c\pi_{m}-\left(1-\pi_{m}\right)\right]dm\right\}$$

Please see Appendix B.2.

B.8.1 Solving the Optimal Stopping Problem

Given the objective, we are finally ready to solve this problem subject to (B.5)

$$\inf_{T\in\mathcal{T}}\tilde{E}\left\{\left(1-\pi_{T}\right)+\int_{0}^{T}\left[c\pi_{m}-\left(1-\pi_{m}\right)\right]dm\right\}=\inf_{T\in\mathcal{T}}\tilde{E}\left\{\left(1-\pi_{T}\right)+\int_{0}^{T}c'\pi_{m}dm-T\right\},$$

where $c'=c-1$.

Let $h(x)=f(x)+g(x)$,

$$\begin{aligned} f(x) &= -\frac{2}{\mu^{2}}\int_{\pi}^{x}e^{-\Lambda H(y)}\int_{\pi}^{y}e^{\Lambda H(z)}\frac{1}{\left[z\left(1-z\right)\right]^{2}}dzdy \\ g(x) &= \frac{2c'}{\mu^{2}}\int_{\pi}^{x}e^{-\Lambda H(y)}\int_{\pi}^{y}e^{\Lambda H(z)}\frac{1}{z\left(1-z\right)^{2}}dzdy \end{aligned}$$

where $\Lambda=2\lambda/\mu^{2}$ and $H(y)=\log\left(\frac{y}{1-y}\right)-\frac{1}{y}$, $\mu=\frac{k_{1}(\xi)-k_{0}(\xi)}{\sigma}$, then

$$dh(x)=(-1+x)\,dt+\Gamma dI_{t},$$

where I_t is some standard Brownian motion, hence we have

$$\tilde{E}\left\{\int_{0}^{T}c'\pi_{m}dm-T\right\}=\tilde{E}\left\{h\left(\pi_{T}\right)\right\}.$$

Following this proposition, the objective can be written in the above way where $l(x)=1-x+f(x)+g(x)$, $l(x)$ is convex since $f(x)$ and $g(x)$ are convex, and have global minimum at

$$l'(x_{0})=0. \tag{B.6}$$

Notice x_0 satisfies the above equation is indexed ξ the realization of ξ_t.

We have the following theorem:

The problem (2.13) with prior (2.3) and (2.23) is solved by the stopping time

$$T^{*}=\inf\left\{t\geq 0|\pi_{t}\geq x_{0}\right\},$$

Known Alternative					
$c = 1.1392$					
Unknown Alternatives					
$c_{11} = 8.28$	$c_{12} = 5.43$	$c'_{11} = 0$	$c'_{12} = 24.77$	$c''_{11} = 0$	$c''_{12} = 25.14$
$c_{21} = 4.28$	$c_{22} = 2.13$	$c'_{21} = 27.50$	$c'_{22} = 0$	$c''_{21} = 28.58$	$c''_{22} = 0$

Table B.1: Value of c

where x_0 is given by (B.6).

B.9 Appendix

The part $E\left\{(\theta - 1 - T)^+\right\} = E\left\{\theta - \sum_{k=0}^{T}\left(\pi_k^0\right)\right\}$ is similar to previous proof.

For any $\mu = i \in \mathcal{M}$,

$$\begin{aligned} E\left\{c_{\mu d}\left(T-\theta\right)^+\right\} &= E\left\{\sum_{n=0}^{\infty} 1_{\{\theta \le n < T\}} 1_{\{d=j\}} c_{ij}\right\} \\ &= E\left\{\sum_{n=0}^{\infty} 1_{\{n<T, d=j\}} c_{ij} P\left(\theta \le n | \mathcal{F}_n\right)\right\} \\ &= E\left\{\sum_{n=0}^{T-1}\sum_{j=1}^{M} 1_{\{d=j\}} c_{ij}\left(1-\pi_n^0\right)\right\}. \end{aligned}$$

$$E\left\{c'_{\mu d} 1_{\{T<\theta\}}\right\} = E\left\{\sum_{j=1}^{M} 1_{\{d=j\}} c'_{ij} \pi_T^0\right\},$$

$$E\left\{c''_{\mu d} 1_{\{T\ge\theta\}}\right\} = E\left\{\sum_{j=1}^{M} 1_{\{d=j\}} c''_{ij} \pi_T^i\right\}.$$

Plug these expression back into (2.30), we get

$$E\left\{\theta - \sum_{n=0}^{T} \pi_n^0 + \sum_{j=1}^{M} 1_{\{d=j\}}\left[\sum_{n=0}^{T-1}\sum_{i=1}^{M} c_{ij}\left(1-\pi_n^0\right) + \sum_{i=1}^{M} c'_{ij}\pi_T^0 + \sum_{i=1}^{M} c''_{ij}\pi_T^i\right]\right\}$$

B.10 Appendix

Values of parameters are taken from Curdia and Woodford (2009a)

Alternative 1	$\pi^* = 0.58$	Stopping Time: Aug 9 2007
Unknown Alternative		Stopping Time: August 16 2007, Alternative 1 accepted.

Table B.2: Optimal Stopping Time

Name[a]	Value	Name	Value
s_Ω	-0.0526	ψ_Ω	0
$\bar{\sigma}$	6.25	ϕ_y	0.125[b]
γ_b	0.5	ϕ_π	1.1
κ	0.0725	ϕ_ω	0.8
ξ	0.0388	π_b^1	0.5
β	0.9560[c]	π_b^2	0.6
π_b	0.5	ϕ_ω^1	0.45
$\bar{\sigma}$	6.25	ϕ_ω^2	0.55
$\hat{\delta}$	0.9696		

Table B.3: Parameter Values

[a] Superscript or subspript with 0,1 or 2 indicate the value with corresponding alternatives.

[b] This is quarterly value, needs to be converted into daily value.

[c] This is quarterly value, the corresponding daily value is 0.9995

B.11 Appendix: The Federal Reserve's Aug 17, 2007 Press Release

Release Date: August 17, 2007

For immediate release

To promote the restoration of orderly conditions in financial markets, the Federal Reserve Board approved temporary changes to its primary credit discount window facility. The Board approved a 50 basis point reduction in the primary credit rate to 5-3/4 percent, to narrow the spread between the primary credit rate and the Federal Open Market Committee's target federal funds rate to 50 basis points. The Board is also announcing a change to the Reserve Banks' usual practices to allow the provision of term financing for as long as 30 days, renewable by the borrower. These changes will remain in place until the Federal Reserve determines that market liquidity has improved materially. These changes are designed to provide depositories with greater assurance about the cost and availability of funding. The Federal Reserve will continue to accept a broad range of collateral for discount window loans, including home mortgages and related assets. Existing collateral margins will be maintained. In taking this action, the Board approved the requests submitted by the Boards of Directors of the Federal Reserve Banks of New York and San Francisco.

Appendix C

Appendix for Chapter 3

Table C.1: Benchmark Parameter Values

Modle	Description	Value
β	Discount factor	0.99
σ	Relative risk: aversion coefficient	2
η	Elasticity of substitution between home and foreign goods	1.5
θ	Elasticity of substitution between domestic goods	8.14
$1-\alpha$	Home bias in consumption	0.8
$1/\omega$	Elasticity of labor supply	0.3
$1-\gamma$	probability of a firm is choosen to change price	0.25

Table C.2: Welfare Loss

Parameter	Welfare Loss
Approximation(Non-Coordination)	10.72
Distorted(Non-Coordination)	11.90
Coordination	0.33
Deviation	0

Table C.3: Payoff Matrix

	Coordinate	Deviate
Coordinate	0.33[a]	0
Deviate	0	11.90

[a] Payoff is in terms of welfare loss.