An Option Pricing Model with Regime-Switching Economic Indicators

by

Zongming Ma

Submitted in partial fulfilment of the requirements
for the degree of Master of Arts

at

Dalhousie University
Halifax, Nova Scotia
August 2013

TABLE OF CONTENTS

List of Tables

List of Figures

Abstract

Although the Black-Scholes (BS) model and its alternatives have been widely applied in finance, their flaws have drawn the attention of many investors and risk managers. The Black-Scholes (BS) model fails to explain the volatility smile. Its alternatives, such as the BS model with a Poisson jump process, fail to explain the volatility clustering. Based on the literature, a novel dynamic regime-switching option-pricing model is developed in this thesis, to overcome the flaws of the traditional option pricing models. Five macroeconomic indicators are identified as the drivers of economic states over time. Two regimes are selected among all likely numbers of regimes under the Bayes Information Criterion (BIC). Both in-sample and out-of-sample tests are constructed to examine the prediction of the model. Empirical results show that the two-state regime-switching option-pricing model exhibits significant prediction power.

Keywords: Regime-switching, Option Pricing, Macroeconomic Indicators, Underlying Asset Return, Parameter Estimation, Out-of-sample Test.

List of Abbreviations and Symbols Used

Abbreviations:

BS: Black-Scholes

CAPM: Capital Asset Pricing Model

FF: Fama-French Factor Model

CPI: Consumer Price Index

PPI: Producer Price Index

ETFs: Exchange Traded Funds

STK: S&P 500 Price Index

UCS: U.S. Credit Spread

TYS: Treasury Yield Spread

CCI: Consumer Confidence Index

LEI: Leading Economic Index

BIC: Bayes Information Criterion

MLL: Maximum Likelihood

EM: Expectation and Maximization Algorithm

JB: Jarque-Bera

SR: Single Regime

RS: Regime-switching

VIX: Volatility Index

RSMR: Regime-switching Model Rate

RSLR: Regime-switching Libor Rate

RMSE: Root Mean Standard Error

VAR: Vector Autoregression

DITM: Deeply In the Money

ITM: In the Money

OTM: Out of the Money

DOTM: Deeply Out of the Money

ATM: At the Money

K-S: Kolmogrov-Smirnov

Symbols:

K: Strike Price

S_T: Underlying Stock Price

F: Factors

γ_{M_t}: Regime-dependent Variance-covariance Matrix

α_{M_t}: Regime-dependent Intercepts

β_{M_t}: Regime-dependent Coefficients

p_{ij}: Transition Matrix

R_t: Regime-dependent Asset Returns

σ_{M_t}: Regime-dependent Variance-covariance Matrix for Asset Returns

P_t: Priority Probabilities

r: Risk Free Rate

q^m: Risk Neutral Probabilities

D: Dividend Yield

Γ: Unconditional Expected Variance-covariance Matrix

$f_m(r)$: Regime-dependent Density Function of Underlying Asset Returns

Chapter 1

Introduction

Options are financial securities. The owner of a call option has the right, not an obligation to buy the underlying asset (such as stock and bond) at the strike price (K) on or before expiration date (T). The strike price or exercise price is the price that is paid for the asset if the option is exercised. Maturity or expiration date is the day by which the option is exercised. There are two types of options, European and American. European options have to be exercised on the expiration date, while American options can be exercised at any time on or before expiration. In this thesis, we will discuss the pricing of the European Vanilla call.

When the stock price is higher than the strike price,call option holders are more likely to exercise or sell the call option. On the other hand, if the stock price is lower than the strike price, the option is not being exercised at maturity date. The value of the option is almost zero in this case. Therefore, the higher the underlying stock price, the more valuable the call option is. Assuming the underlying stock price at

maturity date is S_T, the current value of option (C) is equal to the current stock price (S_T) minus the strike price of the option (K), if S_T is greater than K by a big margin. Otherwise, the value of the option at expiration date (T) is zero. This can be expressed as the following fomula:

$$C = max(S_T - K, 0)$$

Thus, the value of an option at expiration date is determined by the price of the underlying asset. It's crucial for investors and risk managers to estimate an accurate price of the underlying asset and standard linear models are widely used.

The traditional asset pricing models, such as the CAPM model, can characterize the asset return successfully. However, the economic episodes, such as the Internet bubble (1999-2001) and the recent economic recession (2008-2009), shows the limitation of the traditional model: it fails to capture the dynamics of asset returns in the financial market. In addition, the standard linear asset pricing models, such as the CAPM model and Fama-French (FF) model, assume the expected values of the error terms jointly follow an independent identical normal distribution (homoscedasticity). Although the normal distribution is widely used in the traditional asset pricing model (see Duffie (1995), Ingersoll (1988), Karatzas and Shreve (1998), Musiela and Rutkowski (1997), and Boyle, Broadie, and Glasserman (1997)), it has been very inconsistent with empirical data. Financial data shows that the distribution of asset returns is fat-tailed and highly skewed. The standard linear model with an independent identical distribution

2

is insufficient for characterizing the asset return.

Although option pricing has been an interesting topic in academics and practical finance since the nineteenth century, the existing options pricing models do not perform well (Bakshi, Cao and Chen, 1997). The log prices of underlying assets are assumed to follow a random walk, which is normally distributed innovation. Most of the options pricing models are based on the Black-Scholes model, which was first published in 1973. However, the Black-Scholes model is designed to capture the price dynamics of the underlying assets following the geometric Brownian motion model:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dw_t$$

where μ is the annualized expected return on the asset, σ is the volatility of the asset return, and w_t is the Brownian motion. The expected rate of return and the volatility are deterministic in the model. The Black-Scholes model cannot capture the dynamics of the parameters. In addition, the Black-Scholes model fails to capture the volatility smile or smirk. Specifically, the implied volatility should be constant according to the Black-Scholes model. However, the empirical results showed that the implied volatility curve with respect to the maturity of the option resembles a curve like "smile". Many researchers have been working to overcome the bias of the prediction based on the Black-Scholes model. In order to relax the assumption of constant volatility, Merton (1976), and Hull and White (1990) add a jump diffusion process to the geometric Brownian motion model. The volatility smile is explained in

Merton's model in terms of short-term options. However, it fails to explain volatility clustering due to the limitation of constant jump over time in Merton's model. The GARCH (general autoregressive and conditional heteroscedasticity) model, which can address the issue of volatility clustering, is widely used to model the volatility (e.g. Engle, 1982, Bollerslev, 1987, Duan,1995). However, none of these papers characterize the source of the volatility.

As a result of emerging interests in the aforementioned challenges, regime-switching models are employed by allowing the parameters of stochastic distribution to change over regimes. While the ultimate goal of this thesis is to develop a novel dynamic option-pricing model, there are two steps in building up the option-pricing model. The first step is to develop an asset pricing model with a broad list of macroeconomic indicators embedded in a regime-switching auto-regressive mechanism. The second step is to establish an option-pricing model based on the risk neutral probability approach.

Many authors have been building for better models for volatility as mentioned above (e.g. the GARCH model), but these models failed to clarify where the volatility comes from. For this reason, we select the macroeconomic indicators to characterize the risk sources of the asset return process. This research contributes to the literature by linking the asset returns with the volatility sources, embedded in the dynamic macroeconomic indicators. The main objective of the fund managers is to maximize the asset returns using the public information available to them, such as

4

the historical aggregate stock market returns and bond market returns. In terms of the pricing process, all useful information will be incorporated into a pricing mechanism. However, the market regimes generated by the macroeconomic indicators are unobservable. There is no single way to tell whether the market is bullish or bearish by observing asset returns directly. For example, market rally can happen both in bear and bull markets, which are known as bear market rally and bull market rally, respectively. From historical data, temporary rally can be observed in a bear market while temporary correction can be observed in a bull market as well. Different future asset returns in different market regimes can be implicated by similar economic indicator values. More specifically, a downward market with high market volatility may not be a sure sign of a future bear market. The market regimes, such as bull, bear and transition market, which are hidden behind the observed regime-dependent process reflected in the macroeconomic indicators, are stochastic and uncertain. Because of these observations, we use the regime-switching model to characterize unobserved economic regimes.

In this research, we use macroeconomic indicators to identify the economic states over time. There are some good reasons to do so. The key reason is that macroeconomic indicators can influence both firms' cash flows and risk-adjusted discount rates. Chen, Ross and Roll (1986) study the relationship between macroeconomic variables and stock returns and find that many macroeconomic indicators have significant influence in predicting security returns, such as unexpected inflation, expected inflation, the spread between long and short interest rates, credit spread, and growth

rate of industrial production. Afterwards, many researchers study the impact of different macroeconomic indicators on asset returns. Mark and Aris (2002) study the daily equity returns by establishing a GARCH model and find that stock returns are negatively related to inflation and money growth. They find several macro factor candidates playing a significant role in asset returns, including the consumer price index (CPI) and producer price index (PPI). However, none of these studies address the sensitivity of regime-dependent asset returns to macroeconomic indicators. In much research that has been done on the Markov-chain regime-switching model in asset returns, macroeconomic indicators are studied to characterize the changing patterns of asset returns over time, thus reflecting the different regimes of the economic market. Liu, Zhao, and Xu (2011) added three additional macroeconomic indicators to the original Fama and French (FF) three factor model to study the selection of Exchange Traded Funds (ETFs). The additional macroeconomic indicators are the implied market volatility index, yield spread, and credit spread. They find the sensitivity of ETFs risk premiums is highly regime varying. Mulvey and Zhao (2011) establish an investment model under the regime-switching framework and employ eight common macroeconomic indicators to infer the economic market regimes. The macroeconomic indicators are the S&P 500 price index, Treasury bond yield, U.S. dollar index, implied volatility index, treasury yield spread, credit spread, aggregate dividend yield and short-term interest rate. They find that the macroeconomic indicators play a significant role in driving the stock market from one regime to another and then the asset returns are highly regulated by the macroeconomic indicators in different market regimes.

6

In the literature, many researchers have documented that the asset returns are highly regime-dependent. Unobservable regimes in the financial market do exist and asset pricing processes change following significant events such as financial shocks (e.g. Jeanne and Masson, 2000; Cerra, 2005; Hamilton, 2005), changes in consumer preference (e.g. Veronesi, 1999) and abrupt changes in government policies (Hamilton, 1989; Sims and Zha, 2004; Davig, 2004). In addition, consumer and investors' behaviors, such as decisions, consumption, savings and investment, may differ during a bull market compare to during a bear market, and may differ for positive news events compare to negative news events. Norden and Schaller (1993) examine the US stock market by testing a single regime in the stock market against three different alternatives and find strong evidence that there are regimes in the US stock market. Therefore, a model that can capture the hidden regimes in the stock market is in high demand and should be utilized.

A regime-switching model is a non-linear time series model, first introduced into economics by Hamilton (1989). It is designed to model the switching behaviors between upward shifts and downward shifts of an underlying economic process. Now the regime-switching model has been widely used in many areas, such as option pricing, the term structure of interest rate, and return volatility. For example, Naik (1993) incorporates a regime-switching model into option pricing mechanisms. There are two regimes in his model, so the parameters of the asset return process will take two sets of different values depending on the change of market regimes. Several studies have

been published under the discrete time framework. For instance, the two states model allows the mean and variance of asset returns to take two sets of values determined by regimes over time (e.g. Chourdakis and Tzavalis, 2000; Campbell and Li, 2002). The discrete shifts in the behaviors of financial time series can be characterized by the switching of the multiple structures of the regime-switching model. The parameters of the data generating process can take different values depending on the switching of the regimes over time.

The literature describes many other methods used to capture market break points. For example, a Kalman filter model with changing alpha and beta is proposed by Mamaysky et al. (2004). This paper uses a dummy variable to clarify the binary economic states and the state probabilities cannot be obtained in continuous time. However, Qiu, Faff, and Benson (2011) use the regime-switching model to obtain the continuous state probabilities to overcome the flaws of Mamaysky et al (2004).

There is no doubt that the regime-switching model is a much better choice than the standard linear model to characterize the dynamics of asset returns. Actually, the standard linear pricing model can be taken as a special case of the regime-switching model with only single regime in the pricing model, which contains only one set of parameters for the data generating process. In this case, the regime-switching model is reduced to the regular mean regression model. However, the mean regression model approach has been challenged, since it does not always provide a good fit for actual financial data. As mentioned above, most financial data have shown that asset returns

are fat-tailed and highly skewed.

The rest of the thesis is organized as follows. Chapter 2 discusses the methodology that the thesis adopts. Chapter 3 gives a description of the candidate's data. Then chapter 4 shows the comprehensive empirical analysis. Finally, the findings and future studies are concluded in chapter 5.

Chapter 2

Methodology

2.1 A Vector Autoregressive Regime-switching Model

In this section, we discuss the vector autoregressive regime-switching model for macroe-conomic indicators and a regime-switching regression model for asset returns. In the first model, all the macroeconomic indicators are embedded into the regime-switching model to characterize the dynamics of macroeconomic indicators in different regimes. The second model includes the selected set of macro indicators, which is used to predict the regime-dependent asset returns. Based on the existing research, we identify five macroeconomic indicators that drive the dynamics of market regimes. These indicators are: the S&P 500 Price Index (STK), the U.S. credit spread (UCS), the Treasury yield spread (TYS), the Consumer Confidence Index (CCI) and the Leading Economic Index (LEI). The data are collected from Datastream and Federal Reserve Bank of St.Louis.

Let $F_t = (F_{t_1}, F_{t_2}, \ldots, F_{t_J})$ be a set of J macroeconomic indicators. We use F_{t_j} to denote the jth economic indicator at time t. The indicators follow a vector auto-regressive process (VAR); the coefficients in this process are changing with the switching of the regimes.

$$F_t = \alpha_{M_t} + \beta_{M_t} F_{t-1} + \gamma_{M_t} \epsilon_t \tag{2.1}$$

We assume that there are M regimes in the market. M_t is a discrete, first order Markov chain with M regimes. $\alpha_{M_t}, \beta_{M_t}$, and γ_{M_t} are set of model parameters determined by the regimes at time t. $\alpha'(\alpha_{1_{M_t}}, \alpha_{2_{M_t}}, \ldots, \alpha_{j_{M_t}})$ is a vector of regime-dependent intercepts of the linear factor model while β is a regime-dependent matrix of the sensitivities of the macroeconomic indicators at time $t-1$ to the macroeconomic indicators on time t at regime M_t.

$$\beta_{M_t} = \begin{bmatrix} \beta_{11M_t} & \beta_{12M_t} & \cdots & \beta_{1JM_t} \\ \beta_{21M_t} & \beta_{22M_t} & \cdots & \beta_{2JM_t} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{J1M_t} & \beta_{J2M_t} & \cdots & \beta_{JJM_t} \end{bmatrix}$$

ϵ is the multivariate independently normally distributed vector of error, with zero means and unit standard deviations.

M_t is inferred from the macroeconomic indicators over time. M_t can only take discrete

values such as $M_t = 1, 2, 3, ..., M$. The transition matrix is given by the probability of M_t depends on the probability of M_{t-1}; that is,

$$
P = \begin{bmatrix}
p_{11} & p_{12} & \cdots & p_{1j} \\
p_{21} & p_{22} & \cdots & p_{2j} \\
\vdots & \vdots & \ddots & \vdots \\
p_{i1} & p_{i2} & \cdots & p_{ij}
\end{bmatrix}
$$

where $Pr(M_t = j | M_{t-1} = i, M_{t-2} = k, ...) = Pr(M_t = j | M_{t-1} = i) = p_{ij}$ is the transition matrix probability from state i to state j. The joint distribution of the macroeconomic indicators at time $t + 1$, depends both on the regime-switching model parameters and the macroeconomic indicators at time t. The distribution of macroeconomic indicators at time t follows a mixture normal distribution with a set of mixing model parameters. The distribution of regimes at time t can be updated by the Bayesian analysis with new information inflows from $t - 1$ to t under a multivariate regime-switching model.

2.2 A Regime-switching Framework for Asset Returns

In order to characterize the dynamics of the asset returns, we specify the following

regime-switching model with M regimes for the asset returns:

$$R_t = a_{M_t} + b_{M_t}F_{t-1} + \sigma_{M_t}\epsilon_t \qquad (2.2)$$

where R_t is a vector of logarithmic returns from time $t - 1$ to t. The returns at every time point will be characterized by one of the market regimes $(1, ..., M)$. We denote the asset price at time t as P_t, the R_t equals $\ln P_t - \ln P_{t-1}$ with continuous compounding. F_t is a set of the macroeconomic indicators at time t. The parameters $(a_{M_t}, b_{M_t}, \sigma_{M_t})$ are regime-dependent. ϵ_{M_t} is the regime-dependent vector of errors that have a identical independent bivariate normal distribution with zero means, and unit standard deviations.There are M distinct regimes and all the regimes follow the first order Markov chain with an initial regime distribution q_0 and a constant transition matrix $P = p_{ij}$, where p_{ij} indicates that the transition probability of the market transfer from regime i at time $t - 1$ to regime j at time t.

In this research, we combine the economic indicator linear factor model (2.1) with the regime-switching model for asset returns (2.2). That is, we include R_t (asset return) as an economic indicator in the linear factor model, which makes the pricing process more concise.

The regime-dependent conditional expected asset returns, given regimes at time t (M_t),

13

are normally distributed as

$$E(R_t|M_t) = a_{M_t} + b_{M_t}F_{t-1}$$

with a variance-covariance matrix σ_{M_t}.

The established regime-switching parameters and the established vector of regime indicators are required for the forecasting of asset returns. The conditional expected asset returns in regime m follow a normal distribution with density $f_m(t)$, that is $P(R_T|M_t) \sim f_{m(t)}$. The joint probability of unconditional expected asset return at time t follows a multivariate mixture normal distribution: $f(t) \sim \sum_{k=1}^{K}(p_{ij}) \cdot f_m(t)$. We denote $p_t(m)$ as the prior probability of regime m at time t. The prior probability $p_t(m)$ is:

$$p_t(m) = \sum_{i=1}^{M}(p_{im} \cdot q_{t-1}(i))$$

where $p_{im} = Pr(M_t = m|M_{t-1} = i)$ is the transition probability from regime i to m, and $q_{t-1}(i)$ is the posterior probability of regime i at time $t-1$.

The unconditional expected asset return from $t-1$ to t is the expectation of regime-dependent expected conditional asset return, with the prior probability $p_t(m)$ at regime m and time t:

$$E(R_t) = E(E(R_t|M_t)) = \sum_{m=1}^{M}(a_{M_t} + F_{t-1}b_{M_t})p_t(m)$$

14

and the variance-covariance matrix is:

$$V(R_t) = \sum_{m=1}^{M}[(E(R_t) - E(R_t|M_t = m))^2 + \sigma_{M_t}]p_t(m)$$

Some studies, such as Mulvey and Zhao (2011),use the predicted values of macroeconomic indicators ($\bar{F}_t$) in their asset pricing models. The reason for this is straightforward: using the predicted values of the economic indicator can reinforce the linkage between the regime-dependent asset returns and macroeconomic indicators. However, in this research, we will use actual observations to construct our in-sample and out-of-sample tests. The underlying reason is if the result based on this method has a strong predictability, then the other method using predicted values, which has less disturbing noises, is expected to be even stronger.

2.3 Optimal Number of Regimes

Having selected the regime-switching framework, we should consider the optimal number of regimes. Since the regimes are unobservable, a proper criterion must be adopted to choose the optimal regimes. Intuitively, when we increase the regimes, the number of parameters in the regime-switching framework will increase correspondingly. Thus, the likelihood of the data will increase when more parameters are added. The larger the likelihood, the better fit the model has. However, this intuition is only valid for our in-sample data. We have to consider the out-of-sample mediations as well. If too

many regimes are embedded in the model, this could creat problems of over-fitting or model specification error. In terms of choosing a most parsimonious, accurate, but no elaborate model, we have to balance between the number of regimes and the predictability of the model. The Bayes information criterion (BIC)[Schwarz (1978)], is applied to select the optimal number of regimes. In this research, we employ the following BIC specification to determine the optimal number of regimes:

$$BIC(M) = -2\ln(L|M, \phi(M)) + f(M, \phi(m))\ln(T)$$

where M is the number of regimes and L is the likelihood function given the number of regimes. T is the number of observed data points. $\phi(m) = \{\alpha_{M_t}, \beta_{M_t}, \gamma_{M_t}, p_{ij}\}$ is a set of parameters while $f(M, \phi(m))$ refers to the number of parameters. By trying different numbers of regimes M, we select the number which can minimize the value of $BIC(M)$ as the optimal number of regimes.

2.4   Expectation and Maximization Algorithm

Since the regimes (M_t) are unobservable, the expectation and maximization algorithm (EM algorithm) will be used to estimate the model parameters. The EM algorithm is an iterative process between E-step (expectation) and M-step (maximization), which was first introduced by Arthur Dempster, Nan Laird, and Donald Rubin(1977). The EM algorithm is efficient for estimating models that have missing data or unobservable

latent variables. The E-step is used to estimate the missing data on regimes based on observed data and current estimates by calculating the expected log likelihood with the updating missing data. The M-step is to maximize the log likelihood function based on the missing data on regimes found in the E-step.

We denote the model parameters as $\phi = \{\alpha_{M_t}, \beta_{M_t}, \gamma_{M_t}\}$, the unobserved regimes over time as M and the observed indicators as X. The EM algorithm can be summarized as the following two steps:

E-step: Set an initial parameter value ϕ_0 for the true parameter set ϕ, calculate the conditional distribution on regimes, $Q(m) = Pr(M|X; \phi_0)$, and determine the expected log likelihood, $E_Q[\ln Pr(X, M; \phi)]$.

M-step: Maximize the expected log likelihood of joint data of X and M with respect to ϕ, to obtain an improved estimate for parameter ϕ. The improved estimate is:

$$\phi_1 = \arg\max_{\phi}\{E_Q[\ln Pr(X, M; \phi)]\}$$

where ϕ_1 is the new initial value for the true parameter ϕ. The algorithm returns to the E-step after a new estimate is obtained. As the aforementioned processes are going on, the parameters are estimated when the log likelihood is maximized.

2.5   Option Pricing with Simulation

An option pricing model can be proposed under the regime-switching framework and

the risk neutral mechanism.

2.5.1 A Risk Netrual Probabilities and the Black-Scholes Formula

In terms of option pricing, most existing models depend either on stochastic discount factor or risk neutral probability. The option pricing model in this thesis is based on the concept of risk neutral valuation. Since the underlying asset price mainly depends on the level of risk that the investors would like to take, it is crucial to know the investor preference. Unfortunately, it is hard to quantify the investor preference. Then risk neutral probability is introduced to address this issue. The risk neutral probability is the probability of future expected returns adjusted for risk, which can be used to calculate the asset price by simply taking the expectation of future payoff under the risk neutral probability. The key assumption for calculating the unique risk neutral probability is that the market is arbitrage free. Mathematically, Cox, Ross and Rubinstein (1979) define the upward risk neutral probability as

$$P = \frac{e^{(r \times \Delta t)} - d}{u - d}$$

where u is the ratio of moving upward and d is the ratio of moving downward. u and d depend on the volatility of the asset returns (σ) and the length of time interval. The upward ratio and downward ratio can be expressed mathematically as

$$u = e^{(\sigma \times \sqrt{\Delta t})} \qquad d = e^{-(\sigma \times \sqrt{\Delta t})} = \frac{1}{u}$$

In the risk neutral world, the expected return is the risk free rate r, while the expected return of stock is μ and the standard error of stock return is σ under the real probability measurement. The stock price follows the geometric Brownian motion:

$$dS_t = rS_t dt + \sigma S_t d\omega_t^*$$

or in forms of integral:

$$S_T = S_0 e^{(r - \frac{1}{2}\sigma^2)T + \sigma \omega_T^*}$$

where $\omega_t^* = \omega_t + \frac{\mu - r}{\sigma} dt$ where ω_t^* is the stochastic variable under risk neutral measure and ω_t is the Brownian motion.

The European call option with strike price K under the risk neutral probability measure is priced as:

$$C(S_0) = e^{-rT} \cdot E_t^Q [S_T - K]^+$$

To be specific,

$$C(S_0) = e^{-rT} \int_{-\infty}^{\infty} (S_T - K)^+ \cdot f(S_T, T | S_0) dS_T$$

and the density function f is

$$f(S_T, T | S_0) = \frac{1}{S_T} \frac{1}{\sqrt{2\pi\sigma^2 T}} e^{-\frac{[ln\frac{S_T}{S_0} - (r - \frac{1}{2}\sigma^2)T]^2}{2\sigma^2 T}}$$

Then, the Black-Scholes formula for the European option is given by:

$$C(S_0) = S_t N(d_1) - Ke^{-rT} N(d_2)$$

where

$$d_1 = \frac{ln\frac{S_t}{K} + (r + \frac{1}{2}\sigma^2)}{\sigma\sqrt{T}}$$

and

$$d_2 = \frac{ln\frac{S_t}{K} + (r - \frac{1}{2}\sigma^2)}{\sigma\sqrt{T}}$$

2.5.2 Regime-switching Option Pricing Model

As we have done in the previous section, the asset return is taken as one of the macroeconomic indicators in the autoregressive regime-switching model.

$$F_t = \alpha_{M_t} + \beta_{M_t} F_{t-1} + \gamma_{M_t} \epsilon_t$$

F_t is the vector of macroeconomic indicators. By taking asset return out of the process, we have the asset return simulation process,

$$R_t = a_{M_t} + b_{M_t} F_{t-1} + \sigma_{M_t} \epsilon_t$$

where $R_t = \log(S_t/S_{t-1})$. S_t is the price of the stock at time t. Both processes, a_{M_t}, b_{M_t}, σ_{M_t} are a set of estimated model parameters, depending on the regimes at time t.

The aim of this section is option-pricing simulation. In the other word, assuming the initial stock price at time $t = 1$ (S_1) is known, we simulate a realization for stock price at time $t = 2$ (S_2). The simulation process is one period. Based on the actual stock price (S_2), we repeat the same process to simulate the realization of stock price at time $t = 3$ (S_3).

The crucial part of option pricing is to find out the risk neutral probability. We will calculate every set of risk neutral probability of each point in time. For each note at $t - 1$, the risk neutral probability ($q^{mt}, m = 1, 2, ...M$) at time t is

$$\min Var^Q[R] = \sum [(\bar{R}_t^m - r_t)^2 \times q^{mt}]$$

subject to

1. $q^{mt} \geqslant 0$

2. $q^{1t} + q^{2t} + ... + q^{Mt} = 1$

3. $\bar{R}_{1t} \times q^{1t} + \bar{R}_{2t} \times q^{2t} + ... + \bar{R}_{Mt} \times q^{Mt} = r_t$

where m is the number of regimes $(1, ..., M)$, q^{mt} is the risk neutral probability of regime m in time t. $\bar{R}_{mt}$ is the expected conditional asset return in regime m at time t. Here comes a problem: what is r_t ? In this case, r_t can be LIBOR rate and taken as an exogenous variable. Also, r_t can be taken as an endogenous variable and derived from the above quadratic programming minimization process. In this research, both approaches will be discussed.

In the existing research, the LIBOR rate is widely used for option pricing. If the LIBOR rate is used as the risk free rate r_t, the objective function becomes:

$$\min_{q^{mt}} Var^Q[R] = \sum(\bar{R}_{mt}^2 \times q^{mt}) - r_t^2.$$

where r_t is only depend on time t and not varying with regimes. The conditional expected asset returns at time t $(\bar{R}_{mt})$ can be obtained from the expected asset return simulation process

$$E(R|M_t, F_{t-1}) = \bar{R}_{M_t} = \alpha_{M_t} + \beta_{M_t} \times F_{t-1}$$

The extreme cases may exist in some points in time: the LIBOR rate may smaller or larger than all the regime-dependent expected asset returns $(\bar{R}_{mt})$. Thus, the last constraint cannot be satisfied and there is no optimal solution in this linear programming problem. A proper way must be adopted to address this issue. In this

thesis, the LIBOR rate is adjusted as the average of the conditional expected returns over regimes in the extreme cases, to make sure the solutions at very time point can be obtained.

Instead of using the LIBOR rate, r_t can be infered from the minimization process. The objective function becomes

$$\min_{q^{mt}, r_t} Var^Q[R] = \sum((\bar{R}_{mt})^2 \times q^{mt}) - r_t^2.$$

Since r_t is established from the quadratic programming method, we refer r_t is as model rate.

Intuitively, the model rate is more suitable for option pricing simulation process, since the model rate has already been determined optimally in the quadratic programming method. The applicability of both methods will be compared in the empirical part. After addressing the problem of risk free rate, we further overcome the risk neutral probabilities q^{mt}, $m = 1, ..., M$. The density function of unconditional asset returns at time t can be expressed as

$$\widetilde{R}_t \sim \sum_{m=1}^{M} (q^{mt} \times f_{mt}(r))$$

where $f_{mt}(r)$ is the density function of asset return $(\widetilde{R}_t)$ conditional on the regimes at time $t(1, \ldots, T)$, $f_{m_t}(r) = f(R_t|m_t)$. The risk neutral probability is used to weight the density fucntions of conditional expecetd returns across regimes. Hence, the

uncontional expected return from time $t-1$ to t is

$$E(R_t) = E(E(R_t|M_{t-1}, F_{t-1})) = \sum_{m=1}^{M}((\alpha_{M_t} + F_{t-1}\beta_{M_t})q^{mt})$$

with an expeceted variance-covariance matrix

$$\Gamma_R = Var^Q[R] = E^Q[Var^Q(R|M)] + Var^Q[E^Q(R|M)]$$

To be specific,

$$\Gamma_R = \sum_{m=1}^{M}(q^{mt}\Sigma_m) + \sum_{m=1}^{M}((E[R_t|M_T = m] - E(R_t)^2)q^{mt})$$

where q^{mt} is risk neutral probability at time t. Assume R is a vector, Σ_m is variance-covariance matrix for conditional expected VAR process at regime m.

The realized stock price at time t is $\widetilde{S}_t = S_{t-1} \cdot e^{\widetilde{R}_t}$. A set of realizations $\widetilde{S}_t$ from $t = 2$ to T can be obtained by repeating the same one period process, assuming that the initial asset return (S_1) is known.

We have discussed the boundary condition for the European call option previously:

$$V_T = \max\{S_T - K, 0\}$$

24

Therefore, the option pricing simulation process is:

$$C = e^{-rT} \times E^Q(\widetilde{S}_{t+1} - K)^+$$

where K is the strike price at the expire date T. Five different strike prices are chosen at each point in time. The option pricing simulation process can be expressed as

$$C_i = e^{-rT} \int_{-\infty}^{+\infty} [(S_0 e^r - K_i)^+ \cdot \sum_{m=1}^{M} q^m \cdot f_m(r)]$$

By substitute the density function $f_m(r)$ into the above equation, the process is simplified into:

$$C_i = \sum_{m=1}^{M} q^m (S_0 e^{(\mu_m - r - D)T} N(d_m) - K_i e^{-rT} N(d_m - \sigma\sqrt{T}))$$

where
$$d_m = \frac{ln\frac{S_0}{K_i} + (\mu_m - D + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$$

μ_m is the conditional expected return on regime m, and D is the dividend yiled.

C is a set of prices of European call option in our option pricing simulation. We will compare the simulation option prices with the actual prices to see whether actual option price is undervalued or not. The BS model will be taken as the benchmark and compared with the new option pricing model. In addition, the out-of-sample test will be constructed to test the effectiveness of the new option-pricing model.

Chapter 3

Data

This chapter reports a comprehensive analysis of the dataset including summary statistics, the data resources and measurement of all the dynamic macroeconomic indicators.

3.1 Data Description

In this thesis, five macroeconomic indicators are used for identify market regimes. These macroeconomic indicators include the S&P 500 Price Index (STK), Treasury yield spread (TYS), U.S. credit spread (UCS), U.S. Consumer Confidence Index (CCI), and U.S Leading Economic Index (LEI). All of the indicators have been studied in the literature and utilized to characterize the dynamics of economic states over time. TYS is the difference between the yields on 10 year Treasury bond and 2 year Treasury bond. UCS defines the spread between the 3-month LIBOR rate and

3-month treasury rate. A different measurement for UCS is adopted in this thesis, which is the difference between the yields on A rated corporate bond and B rated corporate bond. Some evidence shows that higher yield spread and higher credit spread are more likely to be associated with financial turmoil. Investors are motived to find potential investment opportunities based on the movement of the historical yield and credit spreads. The full sample period is from 1973/02 to 2013/06, while the in-sample period is from 1973/02 to 2010/12 and the out-of-sample data is from 2011/01 to 2013/06. The good economic reason to choose this in-sample period is that this in-sample period covers the dot-com bubble (1997-2000) and the recent financial crisis (2008-2009). These significant financial difficulties are expected to be captured in the empirical analysis, which shows that the model has good predictability of selected macroeconomic indicators.

The data are monthly. Table 3.1 gives the detailed description of the dataset, including variable name and their abbreviation, data sources and data description.

Table 3.1: Macro Economic Factors: Names, Abbreviations, Description and Sources

Abbr.	Variables	Measurements	Sources
STK	Stock Return	$100 * (\ln(P_t) - \ln(P_{t-1}))$	S&P 500 Index, *Datastream*.
UCS	U.S. Credit Spread	$Baa_t - Aaa_t$	Moody's Seasoned Aaa Corporate Bond Yield(Aaa)-Moody's Seasoned Baa Corporate Bond Yield(Baa), *Federal Reserve Bank of ST. Louis*
TYS	Treasury Yiled Spread	TYS_t	US Interest Rate Spread-10 Year Treasury Bonds Less Federal Bond
CCI	U.S. Consumer Confidence Index	$100 * (\ln(CCI_t) - \ln(CCI_{t-1}))$	US Consumer Confidence Index, *Datastream*
LEI	changes in U.S. LeadingEconomic Index	$LEI_t - LEI_{t-1}$	The US Conference Borad Leading Economic Indicators Index, *Datastream*

Table 3.1 gives the variable names,abbreviations, descriptions and sources. All the data are collected from Datastream except for U.S credit spread

3.2   Summary Statistics

The summary statistics of the full sample data from 1978/02 to 2013/06 are shown in Table 3.2. The data are measured in term of percentage; the summary statistics are also in terms of percentage. The mean value of stock return is 0.5426 while the standard deviation is 4.6197, indicating the high volatility in stock returns. All macroeconomic indicators exhibit some excess kurtosis and skewness. Kurtosis measures the heaviness of the tail of a distribution. All kurtosis values of the indicators are greater than the value (3) to the normal distribution. The distributions of the data are more likely to have extreme values than the normal distribution has, with the evidence of fat tails. Skewness measures the level of asymmetry of a distribution. All skewness values are negative except for that of UCS. Thus, the distribution of UCS is right skewed and has more values concentrating on the left of the mean while the distributions of other values are all left skewed. Therefore, the financial data are highly skewed. In addition, the JB (Jarque-Bera) test shows that the data are not normally distributed. Therefore, standard linear pricing models cannot capture the stock return well enough.

What we aim to establish is an option pricing model taht can take into account of the properties of our data. In order to better visualize the distribution of stock returns, both stock returns and their histograms are plotted in Figure 3.1. The lower figure shows that the amplitude of stock returns changes over time, indicating the risk level are time varying. Around year 2000 (internet bubble) and year 2008 (financial crisis),
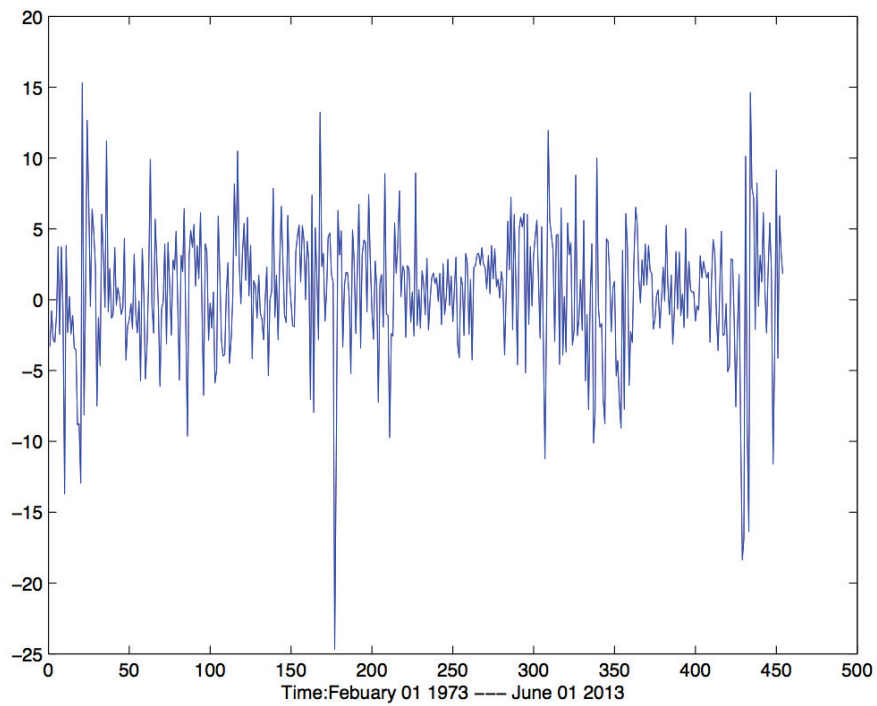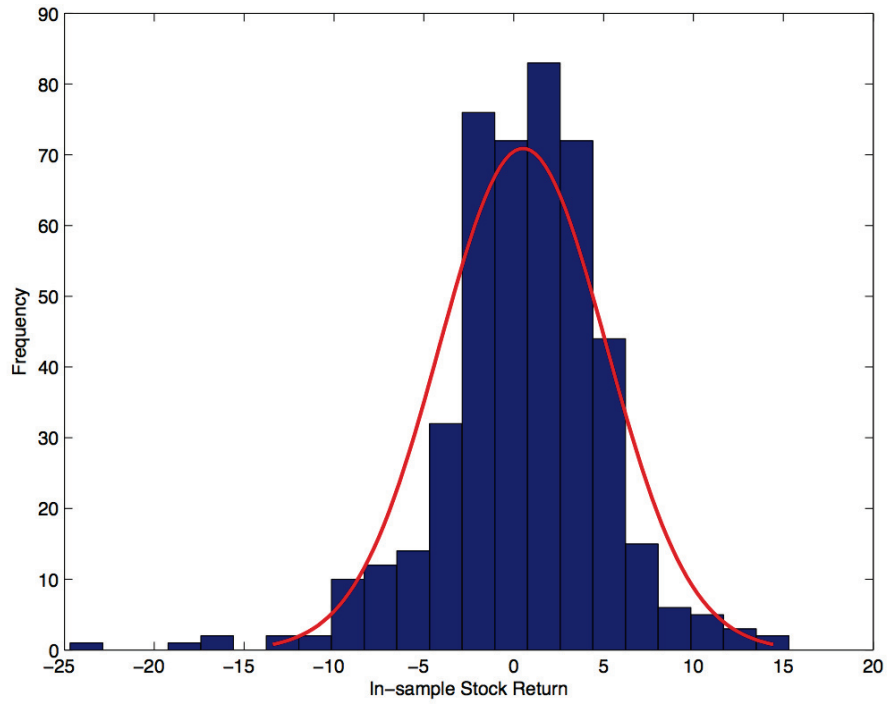
Figure 3.1: Histogram and Magnitude of Stock Returns

Figure 3.1:The upper graph shows the distribution of stock return with estimated in-sample mean and standard deviation. The lower graph plots the magnitude of stock return of in-sample data.

the return volatility is much greater than other time periods, inferring the risk level is high during the big financial events. In addition, low returns tend to be followed by low returns while high returns are likely followed by high returns. Volatility clustering appears clearly in stock returns. The full sample historical performance of five macroeconomic indicators is displayed in Figure 3.2.

In Figure 3.2, the upper figure shows the changing patterns of STK, CCI, and UCS over time. Wider credit spread is always associated with higher stock aptitude and downward market states. The lower figure shows the changing patterns of TYS and LEI. It is obvious that the two indictors moves in the opposite direction and are negatively correlated over time. There is always an increase in TYS right after a decrease in LEI. Large TYS is highly correlated with economic recession. This is consistent with some previous evidence in the literature, such as Chen et al. (1986). The macroeconomic indicators reflect the overall financial market reasonably well. However, the stock market performance cannot be simply determined by observing the dynamic movements of macroeconomic indicators. The market performance, such as the market volatility and stock return, is likely to be regime-dependent over time. If so, it is critical to characterize the financial market by regimes. Otherwise, some information will be averaged across market regimes.
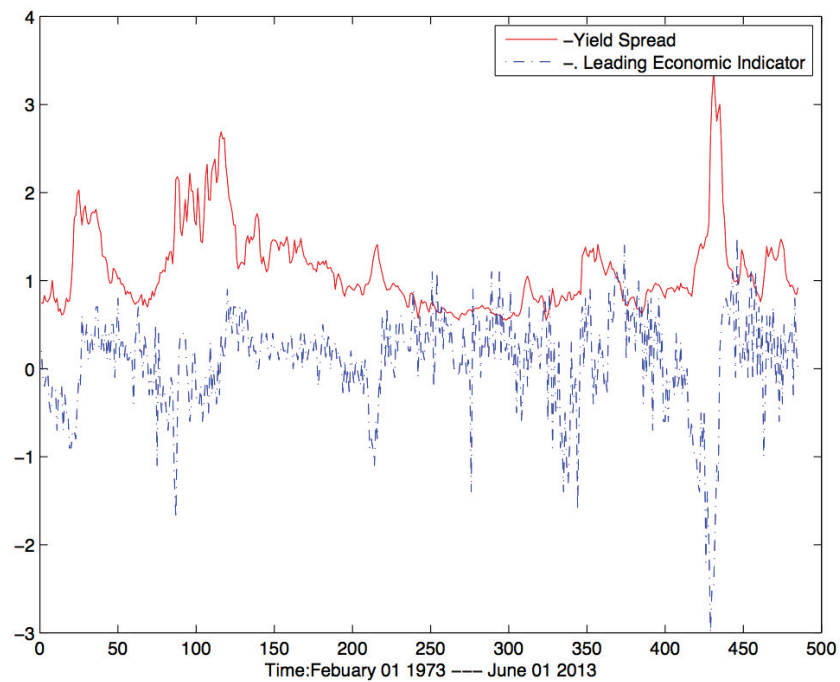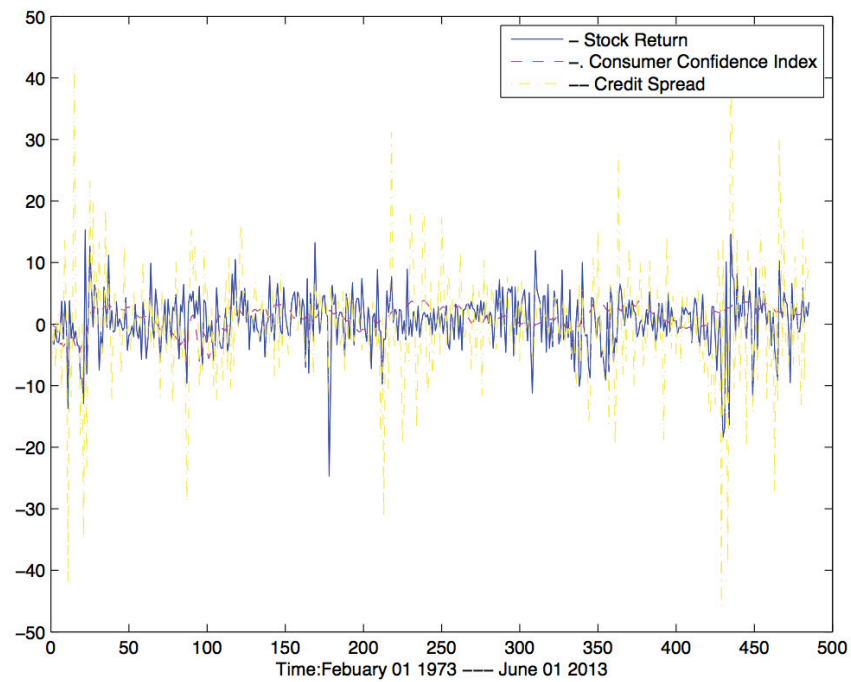
Figure 3.2: Historical Performance of Economic Indicators

Figure 3.2: The upper graph shows the dynamic changes of stock return, consumer condident index and U.S. creadit spread over the full sample period. The lower graph shows the patterns of treasury yield spread and leading economic indicator over the full sample. The full sample period is from 1973/01 to 2013/06

Table 3.2: Summary Statistics of the Macroconomic Indicators

	STK	TYS	UCS	CCI	LEI
Panel A					
mean	0.5426	1.1172	1.1207	-0.0714	0.0816
standard deviation	4.6197	0.4722	1.841	8.9372	0.5718
median	0.9514	0.97	1.42	0	0.2
skewness	-0.7395	1.6623	-1.1956	-0.2674	-1.3277
kurtosis	6.0083	6.4045	4.8369	8.7324	6.8158
JB-test	0.001	0.001	0.001	0.001	0.001
Panel B					
Correlation					
STK	1				
TYS	-0.0395	1			
UCS	0.0738	0.0226	1		
CCI	0.3901	-0.0126	0.1506	1	
LEI	0.3743	-0.26	0.4146	0.3593	1

Table3.2 shows the summary statistics of five macroeconomic indicators including mean, standard deviation, median, skewness, kurtosis, Jarque-Bera test and correlation. The null hypothesis of the JB-test is that the macroeconomic indicators indicators follows the normal distribution while the alternative hypothesis is that the indicator does not come from the normal distribution. The significance level is the 5 %. We can reject the null hypothesis at 5% significance level for all macroeconomic indicators.

Chapter 4

Empirical Analysis

We report empirical analysis in this chapter. The chapter has five sections: the optimal number of regimes, interpretation of the regimes, regime-switching model estimation, out-of-sample performance and the performance of option pricing model.

4.1 Optimal Number of Regimes

In order to choose the optimal number of regimes, we set a range for K, i.e $K = 1, 2, 3, 4, 5, 6$. An information criterion is utilized to determine the optimal number of regimes. A detailed application to choose the regimes in joint distribution of bond yield and stock returns is provided in a research paper by Guidolin and Timmermann (2005). In their research, a four-state model is chosen based on the Bayes information criterion (BIC). The maximized log likelihood function and the values of BIC for different numbers of regimes are displayed in Table 4.1.

Table 4.1: MLL and BIC for Different Regimes

Regime	1	2	3	4	5	6
MLL	-3189	-2717.2	-2589.7	-2493.7	-2419.0	-2336.2
BIC	6653.4	6003.4	6054.3	6180.4	6361.4	6538.5

Although the maximized log likelihood function appear to increase monotonically as the number of regimes increases, the most parsimonious model appears to be the two-regime model based on the lowest value of BIC. Therefore, the number of two regimes is selected under the selected macroeconomic indicators.

4.2 Interpretation of the Regimes

After obtaining the optimal number of regimes, we relate each regime to market sentiments. Since the number of regimes is two, the estimated transition matrix is a 2-by-2 matrix.

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} 0.8625 & 0.1375 \\ 0.0509 & 0.9491 \end{pmatrix}$$

The transition matrix gives the probabilities that the market moves from one regime to the other conditional on the existing regime. In this case, the unobserved market states are transiting between those two regimes over time. From the transition matrix, the transition probability from regime 1 to regime 2 is 0.1375 and from regime 2 to regime 1 is 0.0509, indicating that the probabilities of transfer between regimes are

pretty low. Among them, the probability from regime 1 to regime 2 is relatively higher than the probability from regime 2 to regime 1, which means the market is more likely to transfer from regime 1 to 2 than the other direction. Moreover, Both regimes are highly persistent, as they have a very high retaining probability, 0.8625 for regime1 and 0.9491 for regime2. Table 4.2 shows the summary statistics of macroeconomic indicators and correlation matrix by regime. The different characteristics of these indicators across regimes are the key to define these regimes. The mean return of STK is negative and the standard deviation is relatively high (6.887) in regime 1. While regime 2 has a positive mean return and a low standard deviation (3.4156). We can infer that the regime 1 is a bear state and regime 2 is bull state from the performance of STK. In addition, the value of UCS is larger in regime 1 than in regime 2, which is a reinforcement of what we would expect. The CCI and LEI in regime 1 are both negative while they are positive in regime 2, indicating that regime 1 is bear and regime 2 is bull. As expected, the TYS is smaller in regime 1 than in regime 2.The yield curve is likely to be upward sharply during a bull market while it is likely to be downward sharply during a bear market. Therefore, regime 1 can be labeled as the bearish state, and it should coincide with all financial turmoil and crises from 1973 to 2013, including the oil crisis in the 1970s, the black Monday on 1987, the dot-com bubble during 1997 to 2000 and the financial crisis in 2008. Since all the characteristics of regimes 2 are reversed from those in regime 1, regime 2 can be labeled as bullish state. The correlation between indicators also is changed significantly by regimes. In the full sample period, the correlation between STK and UCS is negative (-0.0395) without regimes. After addressing the market regimes,

36

the correlation in the in-sample period becomes positive in both regimes, at 0.058 (regime1) and 0.0131 (regime2). The correlations between STK and CCI increased in regime 1 while the correlations of them decreased in regime 2, compared to the correlations without characterizing regimes.

The bull market is more stable than the bear market, and has a relatively high retaining probability. If the market regime shifts, it is more likely to shift from the bear to bull market. The average steady probability reflects the mean time of the market staying at each regime in the long run and is found to be 0.2703 (regime1/bear) and 0.7297 (regime2/bull), respectively. Therefore, the market stays more often in the bull regime rather than in the bear regime in the long run. In order to visualize the regimes over time, Figure 4.2 shows the frequency of the market regimes over time. The probability of bear regimes is high from year 1973 to 1975, showing the two oil crises in the early 1970s. After that, the market stepped into the bull regime and it lasted until the end of the 1970s. Generally, the market stayed at the bull regime, consistent with the transition probability. In 1998, there was turmoil in the stock market, so the probability of the bear regime started to rise, capturing the Internet bubble. Afterwards, the bear state stayed from 2004 to 2006. The probability of the bear regime rose sharply in 2008, corresponding to the 2008 financial crisis. Thereafter, the market had been recovering until the end of 2010, shown in the gradually falling probability of bear regime. Figure 4.1 plots the state probability for regime-switching model to better visualize the dynamic changing patterns of unobserved market regimes.

Table 4.2: Factor Statistics by Regimes

	STK	UCS	TYS	CCI	LEI
Regime 1					
Mean	-0.4771	1.5702	0.011	-1.1511	-0.3393
Standard Deviation	6.887232516	0.606896976	2.519231291	13.59867078	0.700006774
Correlation	1				
	0.058	1			
	0.0948	0.2996	1		
	0.4534	-0.0056	0.1622	1	
	0.4947	-0.1369	0.1529	0.3774	1
Regime 2					
Mean	0.8838	0.9539	1.4425	0.2548	0.2259
Standard Deviation	3.415558659	0.290416347	1.396726616	6.094246516	0.438678422
Correlation	1				
	0.0131	1			
	-0.0594	0.2632	1		
	0.226	0.0803	0.1179	1	
	0.2012	0.1107	0.4994	0.3724	1

Table 4.2 shows the statistics for five macroeconomic indicators by regimes.The upper pannel is for regime 1, inluding the mean, standard deviation and correlation. The lower pannel is for regime 2, including the mean, standard deviation and correlation. The caculation are for the in-sample period, from 1973/2 to 2010/12
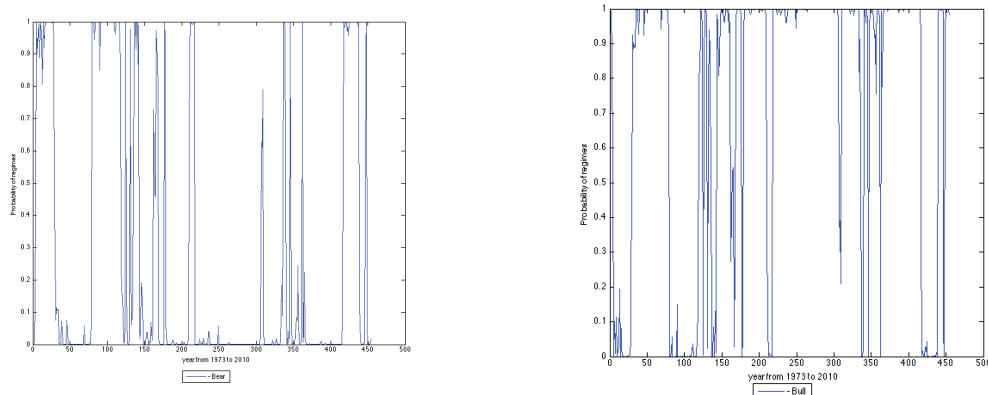
Figure 4.1: Posterior Probabilities of Regimes

Figure 4.1 shows the in-sample implied posterior probability of two regimes. The left figure is the implied posterior probability for the bear market while the right graph is the implied posterior probability for the bull market.

4.3 Model Estimation

The performance of the regime-switching model needs to be estimated. Table 4.3 documents the estimated parameters of the RS and SR models, while Table 4.4 documents the standards errors of corresponding estimated parameters of the RS and SR models. In terms of the SR model, only TYS and LEI have some predictive power for stock returns. None of other macroeconomic indicators has any significant predictive power for stock returns. While in the RS model, most macroeconomic indicators exhibit some predictive power in a high significant level. It is interesting to note that some indicators have some regime-dependent predictive power. While the variable is significant in one regime, it would be insignificant in the other. For example, UCS shows such characteristic. UCS is statistically significant at the 1% level in regime 2, while it is insignificant in regime 1. In addition, it is surprising to see that CCI is insignificant in neither the RS nor SR model for stock returns. Also, CCI is likely
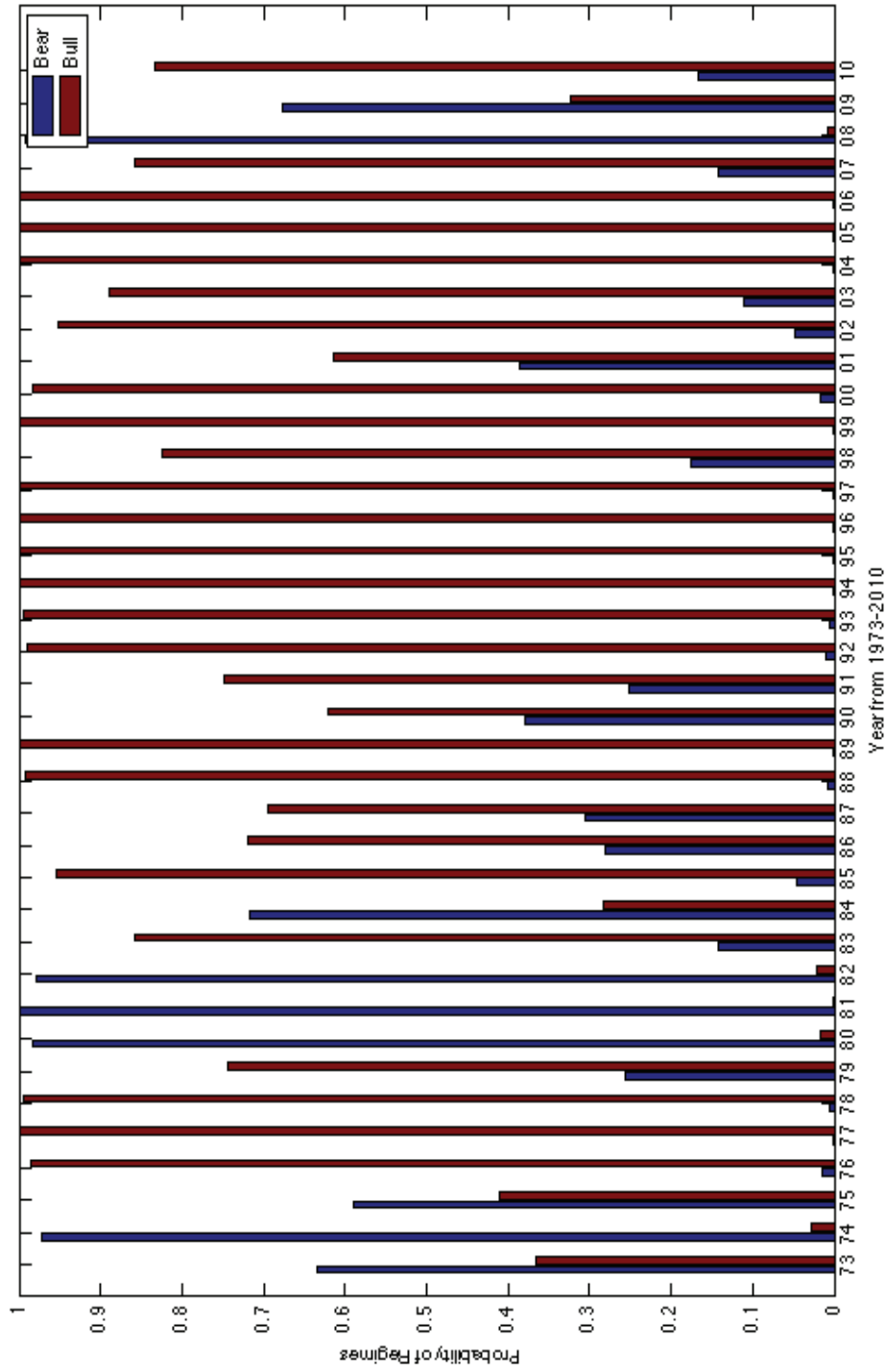
Figure 4.2: Frenquency of Regimes over time

Figure 4.2 shows the frequency of two regimes from 1973 to 2010 (the in-sample period). For each year; The left bar represents the bear market while the right bar represents the bull market. Internet bubble and recent financial crisis are relected clearly by the frequency of regimes.

to be insignificant in other processes most of the time as well. For example, CCI is insignificant in its own autoregressive process both in the RS and SR models. Therefore, CCI shows relatively weak predictability overall. In general, the significance of parameters in the RS model is substantially different from that in the SR model. The RS model has more significant parameters in predictability than the SR model has.

Figure 4.3 compares five macroeconomic indicators under the SR and RS models with actual observations in in-sample period. As we can see from this figure, the predicted values from both the SR and RS models fit the actual data well. Among them, the predicted values from the RS model are much closer to actual observations than that from the SR model. The first graph of Figure 4.3 shows the fitted stock returns in the RS and SR models, and actual stock returns. Fitted stock returns in the RS model are weighted by the posterior probability of each regime over time. It gives a visual comparison of the fitness of the RS and SR models. There is no doubt that the cumulative predicted stock returns of the RS model are closer to the actual stock returns than that of the SR model. In particular, the RS model captures the dynamic movement of cumulative stock returns well at the end of 1990s. Visually, the predictability of the RS model is better than that of the SR model. Predictive accurate rate is used to provide statistical support for the above observation. The prediction accurate rate, which measures the uniformity of the prediction and actual stock returns movement directions (upwards and downwards) one step forward under

Table 4.3: Estimated Parameters for the RS and SR models

		Intercept	STK	TYS	UCS	CCI	LEI
STK :							
RS Model	*Regime 1*	-3.0481***	-0.1249**	2.4129***	-0.0823	0.0301	3.407***
	Regime 2	0.3762	-0.1153**	0.956*	-0.5232***	0.0252	2.0075***
SR Model		-0.6124	-0.0838	1.0489**	-0.174	0.0289	2.5949***
TYS :							
RS Model	*Regime 1*	0.1647***	-0.0033*	0.8905***	0.0069	-0.0015*	-0.0848***
	Regime 2	0.0517***	0.0004	0.9412***	-0.0015	0.0003	-0.0149*
SR Model		0.0704***	-0.0022	0.9426***	-0.0002	-0.0008	-0.0599***
UCS :							
RS Model	*Regime 1*	-0.3542***	0.0023	0.2613***	0.9213***	-0.0123***	-0.1452*
	Regime 2	-0.1286***	0.007**	0.0908**	0.9969***	-0.0025	0.0139
SR Model		-0.1514**	0.0014	0.1785***	0.9604***	-0.0088***	-0.0289
CCI :							
RS Model	*Regime 1*	-2.5823	-0.3301***	2.233**	0.8445***	-0.0386	5.8812***
	Regime 2	-1.659*	0.0382	1.2906	0.0798	0.0271	1.9955**
SR Model		-2.6795**	-0.0853	1.7156**	0.4267*	-0.0096	3.2008***
LEI :							
RS Model	*Regime 1*	-0.2884***	-0.0136***	0.1724***	0.0057	-0.0004	0.8895***
	Regime 2	-0.0014	0.0095*	-0.0219	0.1569***	0.0059**	0.0176
SR Model		0.021	0.0021	-0.0495	0.0674***	0.0021	0.5118***

Table 4.3 displays the estimated parameters for the regime swiching model across regimes and the SR model.***represents that the coefficient is statistically significant at the 1% significance level. ** represents that the coefficient is statistically significant at the 5% significance level.* represents that the coefficient is statistically significant at the 10% significance level

Table 4.4: Standard Errors for the RS and SR Models

		Intercept	STK	TYS	UCS	CCI	LEI
STK :							
RS Model	*Regime 1*	0.841	0.0571	0.5228	0.1285	0.0267	0.5215
	Regime 2	0.537	0.0455	0.5493	0.1387	0.026	0.4528
SR Model		0.5565	0.0512	0.458	0.1258	0.0268	0.4693
TYS :							
RS Model	*Regime 1*	0.0276	0.0019	0.0172	0.0042	0.0009	0.0171
	Regime 2	0.0093	0.0008	0.0095	0.0024	0.0005	0.0079
SR Model		0.0155	0.0014	0.0127	0.0035	0.0007	0.013
UCS :							
RS Model	*Regime 1*	0.1317	0.0089	0.0818	0.0201	0.0042	0.0816
	Regime 2	0.0398	0.0034	0.0407	0.0103	0.0019	0.0336
SR Model		0.071	0.0065	0.0584	0.016	0.0034	0.0599
CCI :							
RS Model	*Regime 1*	1.705	0.1157	1.0598	0.2604	0.0542	1.0572
	Regime 2	0.9516	0.0807	0.9733	0.2458	0.0461	0.8024
SR Model		1.0631	0.0977	0.8749	0.2402	0.0512	0.8966
LEI :							
RS Model	*Regime 1*	0.0541	0.0037	0.0336	0.0083	0.0017	0.0336
	Regime 2	0.0596	0.0051	0.0609	0.0154	0.0029	0.0502
SR Model		0.0543	0.005	0.0447	0.0123	0.0026	0.0458

Table 4.4 shows the standard errors for the regime swiching (RS) model and single regime (SR) model. The regime swiching model is defined as

$$F_t = \alpha_{M_t} + \beta_{M_t} F_{t-1} + \gamma_{M_t} \epsilon$$

The SR model is estimated for comparison, the estimated parameters are presented in Table 4.3. Standard VAR (1) is defined as

$$F_t = \alpha_t + \beta_t F_{t-1} + \epsilon_t$$

a model, to further evaluate the findings. For example, if stock returns are predicted to move up one step forward under a model, meanwhile, the actual stock returns move as predicted. The prediction accurate of the model is thought to be as high. For our in-sample data, the prediction accurate of the RS and SR models are calculated. Table 4.5 demonstrates the prediction accurate rate of five indicators.

Table 4.5: Accurate Rate of the RS Model

Model	STK	UCS	TYS	CCI	LEI
RS	0.6233	1	0.9537	0.5308	0.7819
SR	0.5925	1	0.9537	0.5176	0.7577

None of the prediction accurate rates from the RS model are less than those rates from the SR model model. The predictive accurate rate for the RS model is 0.6233 against 0.5925 under the SR model. The predictive accurate rate of a model is important for investors. Investors can make decisions to rebalance their portfolios to minimize the risk and maximize the returns based on the prediction accurate rate. The root mean standard errors (RMSE) measures the difference between predicted values under a model and actual values. It is a good measurement of accuracy across different models. RMSE is defined mathematically as $RMSE = \sqrt{\frac{\sum_1^N (x_i - \bar{x}_i)^2}{N}}$, where x_i is the actual value and $\bar{x}_i$ is model predict value. N is the number of observations. Table 4.6 shows the RMSE in prediction for five macroeconomic indicators.

Table 4.6: RMSE for Risk Factors

Model	STK	UCS	TYS	CCI	LEI
SR	0.044	0.0012	0.0057	0.849	0.0043
RS	0.0019	0.000001399	0.000030837	0.007	0.00001437

The values of RMSE are incomparable among macroeconomic indicators under a same model. However, RMSE for all macroeconomic indicators under the RS model are lower than RMSE under the SR model and the differences are substantial for model prediction errors. In summary, the performance of the RS model is superior than that of the SR model in the in-sample data as expected.

Figure 4.3: In-sample Predictions of and Actual Observations of the Economic Indicators
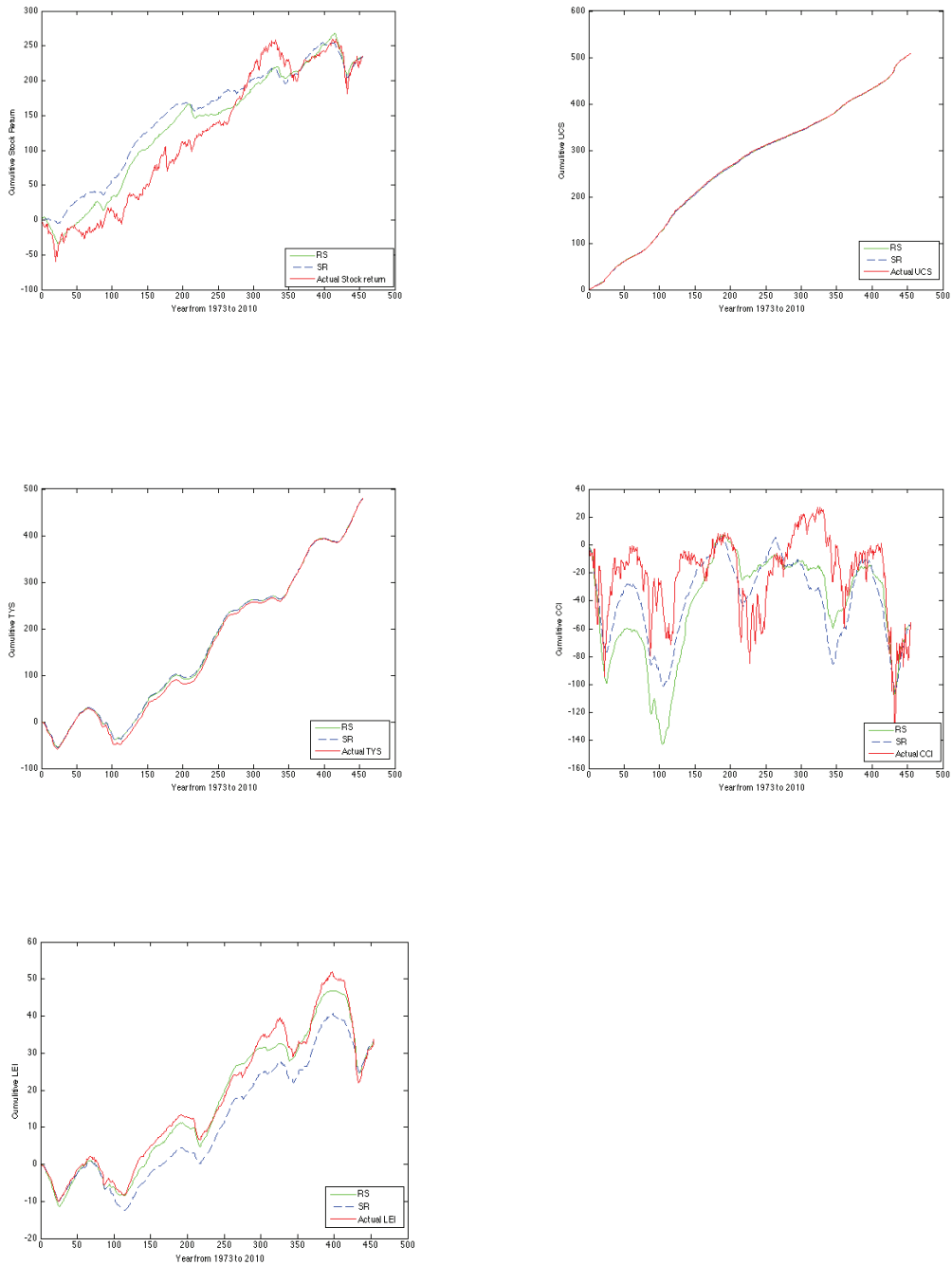

Figure 4.3 shows the in-sample comparison of predicted macroeconomic indicators by the RS and SR model, and actual obeservations. From left to right and up to down, the graphs represent SKT,UCS,TYS,CCI and LEI,respectively.

4.4 Out-of-sample Performance

As the in-sample evidence has shown that the RS model perform better in predictions than the SR model does, the out-of-sample test is constructed to further investigate the RS model performance. The out-of-sample period is from2011/01 to 2013/06.

Figure 4.4 provides the out-of-sample comparison among actual observations, predicted values under the RS and the SR models for the five macroeconomic indicators. The first graph shows the cumulative return. Most of the time, the predicted returns form the RS models are below actual returns, but the RS model can still capture the dynamic movements of returns over time. The predicted stock returns under the SR model are relatively consistent with the trend of the actual observations. However, the RS model shows better prediction performance when the stock market goes down. Out-of-sample predicted values of UCS and TYS are not as close to the actual observations as the predicted values in the in-sample period, but they are still much better than predicted values of other macroeconomic indicators in the out-of-sample data. The fitness of predicted CCI is the worst one among five macroeconomic indicators, which is consistent with the findings in Table 4.3. Predicted values of CCI from both the RS and the SR models do not fit the actual values of CCI. With respect to LEI, the RS model has better prediction power than the SR model before the middle date of the out-of-sample data (2012/03), while the SR model fit actual data better afterwards. Overall, the out-of-sample performance of RS model is not as good as expected. However, to some extent, the RS model still shows strong prediction power.

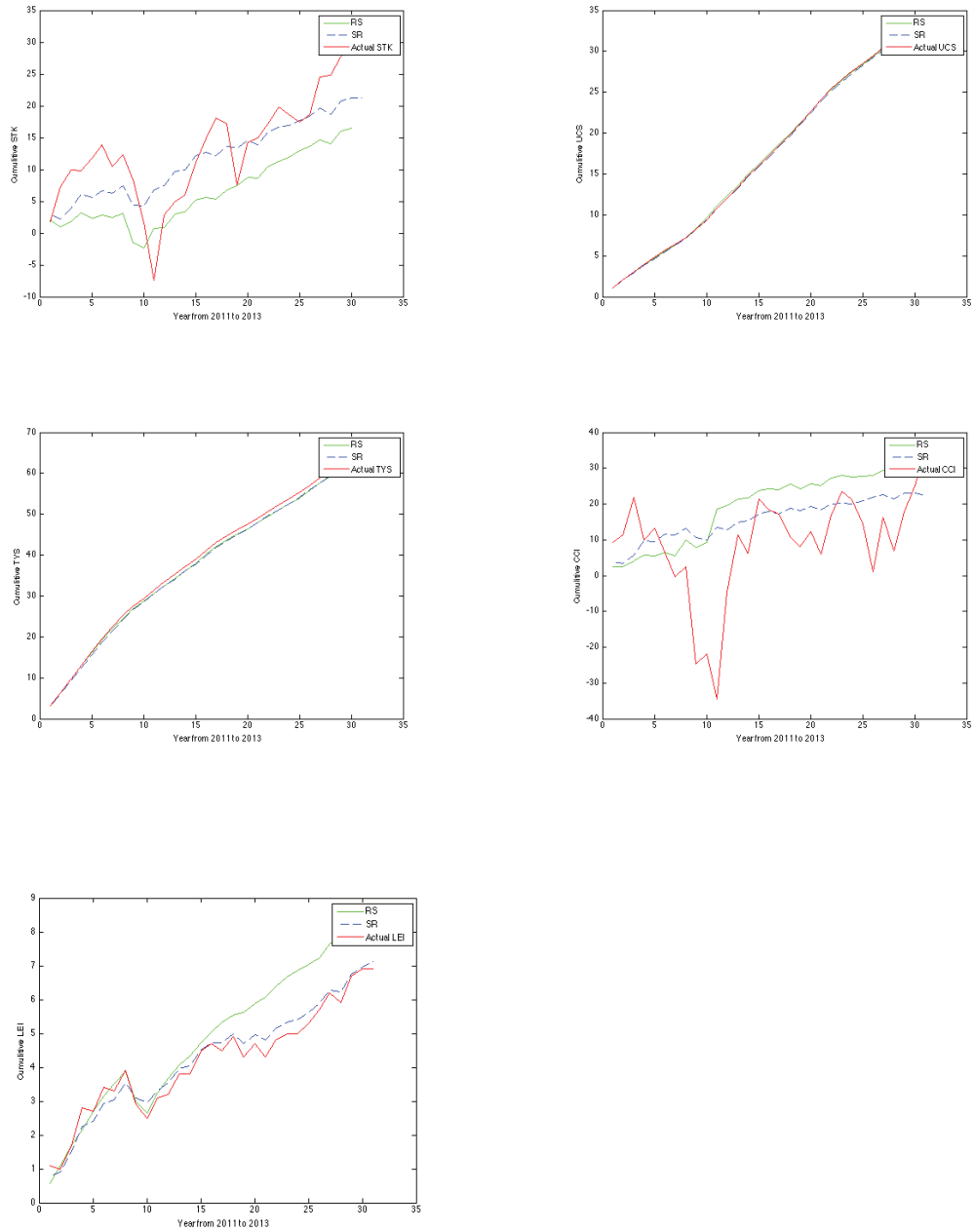Figure 4.4: Out-of-sample Predictions and Observed Factors

Figure 4.4 shows the out-of-sample comparison of predicted macroeconomic indicators by the RS and SR models, and actual obeservations. From left to right and up to down, the graphs represent SKT, UCS, TYS, CCI and LEI.

4.5 The Option Pricing Model

In this section, results on the option pricing model will be discussed. The data of option prices of S&P 500 are collected from Datastream. Option Index is from 2006 to 2013. The in-sample period is from 2006 to the end of 2010 due to the availability of option data. The dividend yield on S&P 500 Index is usually modeled as a constant. We assume that the dividend yield is identical across regimes. The out-of-sample period is from 2010/12 to 2012/10. The regime-switching option pricing model with model rate (RSMR) and regime-switching option pricing model with LIBOR rate (RSLR) are compared with the standard Black-Scholes model (BS) option-pricing model, with both volatility index (VIX) and historical volatility.

The selection of strike prices is not random. As mentioned in Chapter 2, five sets of strike prices $(k_1, k_2, k_3, k_4, k_5)$ are selected around the price of underlying assets at each time period, with a price interval of 25. From the highest strike price to the lowest one, we label them as deeply out of the money, out of the money, at the money, in the money and deeply in the money. Since what we study is the call option, the definition of "in the money" is fairly intuitive. For a call option, if the strike price is lower than the market price of the underlying asset, it is called "in the money" (ITM). "Out of the money" (OTM) is the opposite of in the money. "At the money" (ATM) is when the strike price is the same as the spot price of the underlying asset. Five situations are considered and compared with actual option prices individually. Both

in-sample and out-of-sample tests are constructed to test the validity of the option pricing model. Since there are not many out-of-sample empirical comparison of option pricing models, the out-of-sample test here is itself a contribution to the literature. In addition, the regime-switching option pricing model with two different discount rates (RSMR and RSLR) are discussed in Chapter 2. The LIBOR rate is considered as a common discount rate while the model rate is an endogenous variable derived from the model.

We use figures to visualize the comparison of option prices across the strike prices for the option pricing models. Figure 4.5 shows the in-sample comparison of actual option prices and predicted prices under RSMR and RSLR over five different strike prices. Figure 4.6 demonstrates the patterns of actual option price, predicted prices under the BS model with the VIX index, and predicted prices under the BS model with historical VIX data. Figure 4.7 and Figure 4.8 give the corresponding out-of-sample performances to Figure 4.5 and Figure 4.6, respectively. To start up with in-sample data, it is easy to notice that the BS model with VIX index gives higher predicted option prices than actual ones across five strike prices. No matter how actual prices move, the predicated option prices under the BS model with VIX index always move above actual ones. Therefore, the BS model with the VIX index overvalues the option price. Generally, the predicted prices under BS model with historical VIX data have an upward trend from the beginning of in-sample data (2006/05) to the end of in-sample data (2010/12). While actual option prices move up and down over time and peak at the middle of in-sample period (2008). By and large, the BS model with

historical VIX data undervalues option prices in the first half of the in-sample period and overvalues option price at the last half period. The regime-switching option pricing model has a different perspective when compared with the BS model. The unobservable market state is divided into two regimes (bullish and bearish) under the RS framework. Instead of simply averaging the information of regimes over time, the RS model can capture the shift of market regimes. Options are separately priced under different regimes and weighed with the risk neutral probabilities. The two option pricing models, which are nested in the RS model, exhibit similar prediction price patterns in the first half of the in-sample period and showed some differences in the later part of period. By capturing the market shifts[1] over time, the model predicted prices move up and down around actual prices, except for the middle of the in-sample period.

Out-of-sample performance is considered as the true measurement of model predictability. Parameter estimated from time t is used to predict the option price at time $t + 1$. Basically, the BS model with the VIX index is consistent with its performance in the in-the-sample data. It overvalues the option pricing model over time and across different strike prices. The BS model with historical VIX data overvalues option prices as well, except for the middle of out-of-sample data. RS nested models show different characteristics across different strike prices.

To determine which model has better performance than another, MSRE is used for

---

[1]Market shift from bullish to bearish market or from bearish to bullish market.

comparison.

Table 4.7: RMSD for Various Option Pricing Models

	K	RSMR	RSLR	BS (VIX index)	BS (historical VIX)
	K_1	7.6182	7.7638	9.9112	5.7489
	K_2	11.8209	12.0618	12.0610	7.4308
In Sample	K_3	15.8262	16.0865	13.0110	8.2406
	K_4	19.3792	19.6248	13.3442	8.0693
	K_5	22.8040	23.4516	12.2422	7.4768
	K_1	9.6490	9.7333	9.1721	7.7518
	K_2	13.3229	13.7473	11.8554	11.1610
Out of Sample	K_3	16.1555	16.8050	13.1474	13.5203
	K_4	18.0298	18.8555	12.8741	14.5677
	K_5	18.6462	19.7365	11.9288	14.7888

Table 4.7 gives the MSRE of the BS nested models and RS nested models varying different strike prices.

With respect to the in-sample period, the BS model with historical VIX data is the winner of the four models. It has the smallest MSRE across five strike prices for the in-sample data. For DOTM and OTM, the RS models perform better than the BS model with the VIX Index. While the BS model with the VIX index has smaller MSRE than the RS models when option is ATM, ITM and DITM.

In the out-of-sample evaluation, the BS model with historical VIX data has the smallest MSRE when option is DITM and ITM. Instead, the BS model with the VIX index is the winner when option is DOTM, OTM and ATM. The RSMR model has smaller MSRE than RSLR model both in-sample and out of sample. So among all RS nested models, RSMR has better performance. Overall, different models have their own advantages and disadvantages when pricing options with different strike prices. The BS

nested model performs better out of sample. One explanation is that the BS model is widely used to price options by traders, and, hence, option pricing model emerge to those of the BS model. In addition, Kolmogorov Smirnov tests (K-S test) are constructed to evaluate the overall performance of option pricing models. The K-S test can be used to measure the distance between two empirical distributions. The null hypothesis for the K-S test is the difference between the two distributions is zero.

Table 4.8: K-S Tests for Option Pricing Models

		RSMR	RSLR	BS(VIX)	BS(h)
In Sample:	h	0	0	1	0
	p-value	0.1695	0.0599	0.0001	0.0929
Out of Sample:	h	0	1	1	1
	p-value	0.1312	0.0034	0.0001	0.0001

Table 4.8 shows the results of K-S tests over BS and RS option pricing models. $h = 0$ means the null hypothesis is not rejected at the 1 % significance level while $h = 1$ means the null hypothesis is rejected at the 1 % significance level

We note that only RSMR fails to reject the null hypothesis statistically in sample and out of sample at the 1% significance level. The RSLR and BS model with historical VIX data fail to reject the null hypothesis at the 1% significance level in sample, but reject the null out of sample. The BS with VIX index model performs even worse. The null can be rejected at the 1% significance level both in sample and out of sample. The K-S test shows that the RSMR model has the best performance overall, which is as expected.

Figure 4.5: In-sample Comparison of RSMR,and RSLR predictions and Actual Option
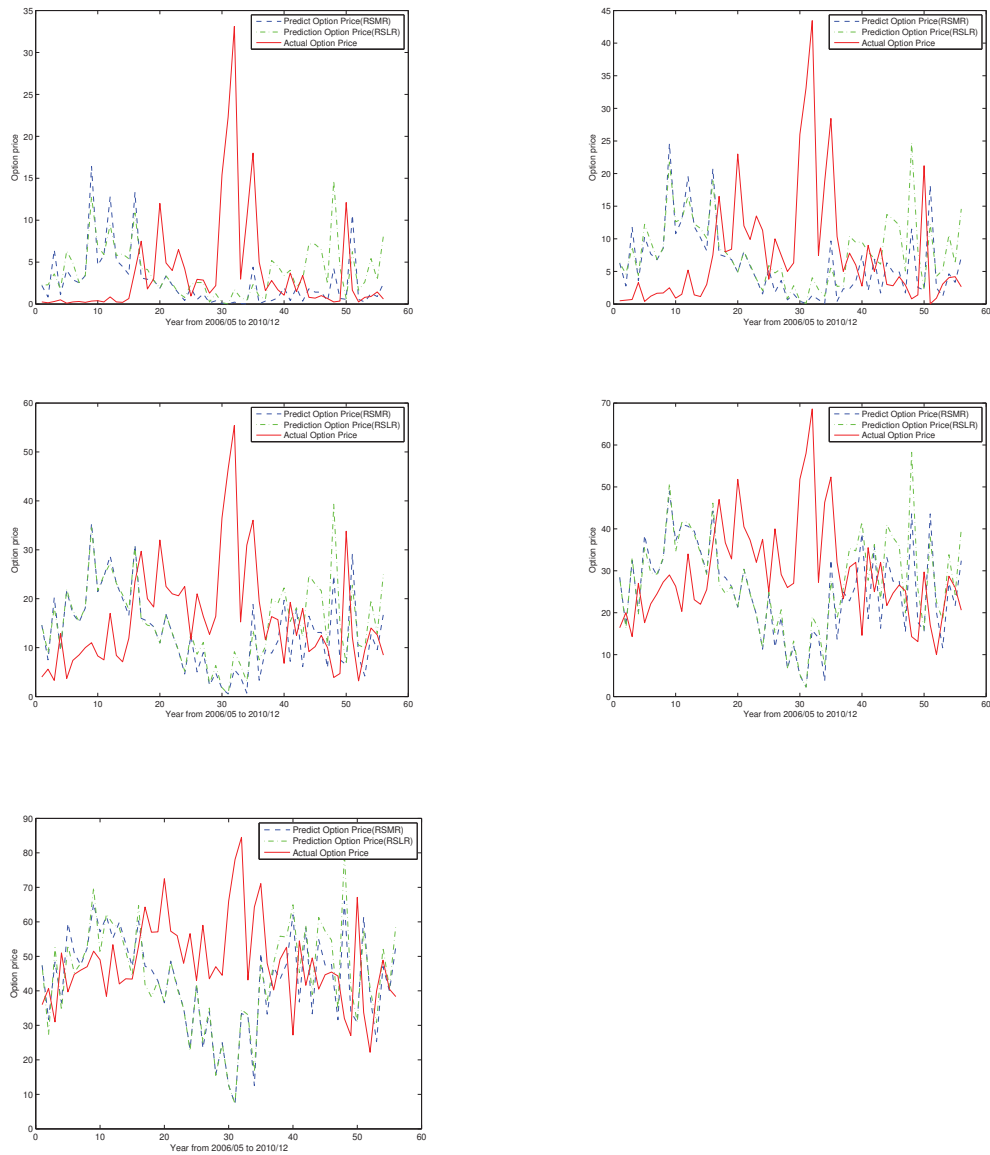Prices

Figure 4.5 shows the in-sample comparison of RSMR, RSLR and actual option prices.
From left to right and up to down, the graph represents DOTM, OTM, ATM, ITM and
DITM

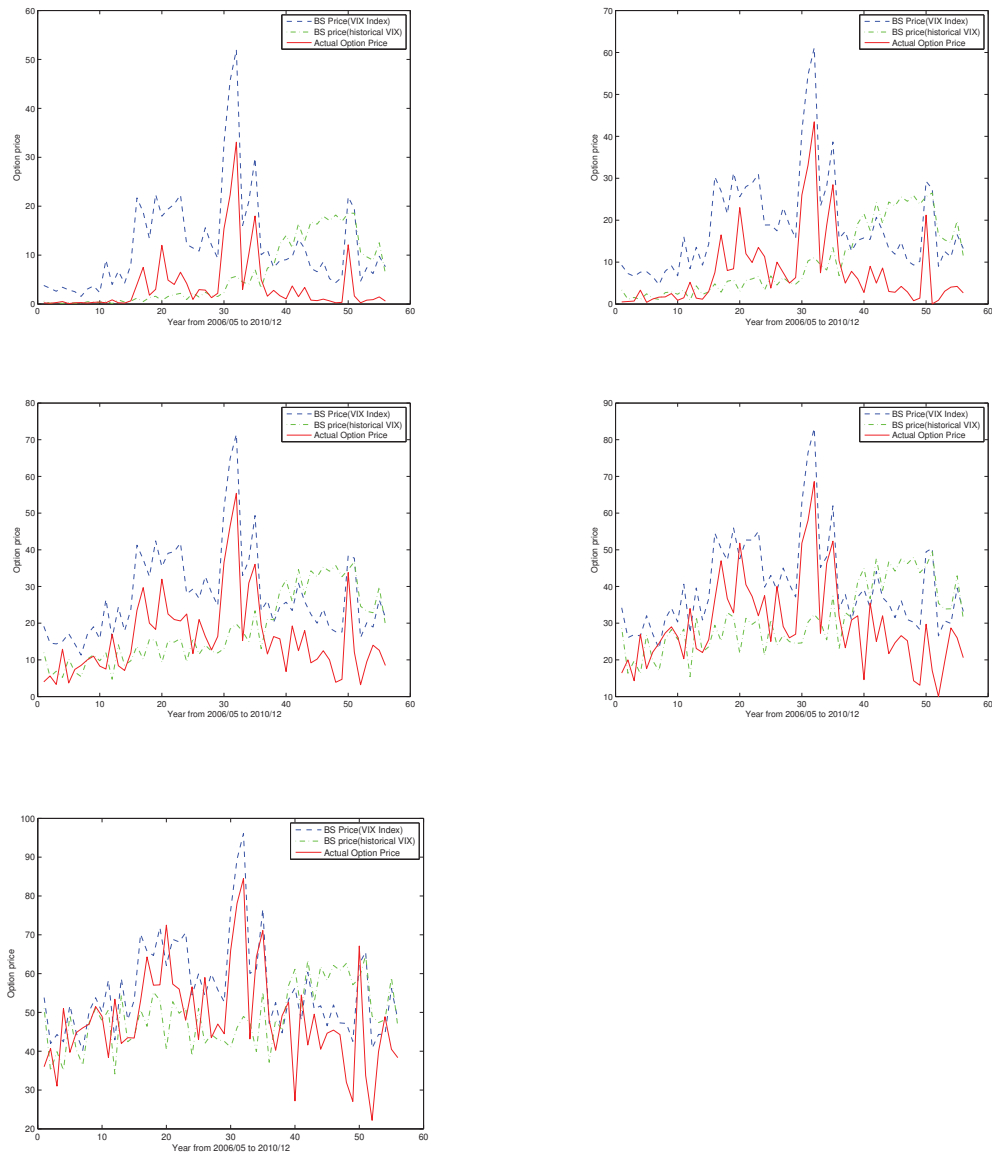Figure 4.6: In-sample Comparison of the BS Model Predictions and Actual Option Prices

Figure 4.6 shows the in-sample comparison of the BS model predictions and actual option prices. From left to right and up to down, the graph represent DOTM, OTM, ATM, ITM and DITM

Figure 4.7: Out-of-sample Comparison of RSMR,and RSLR Predictions and Actual
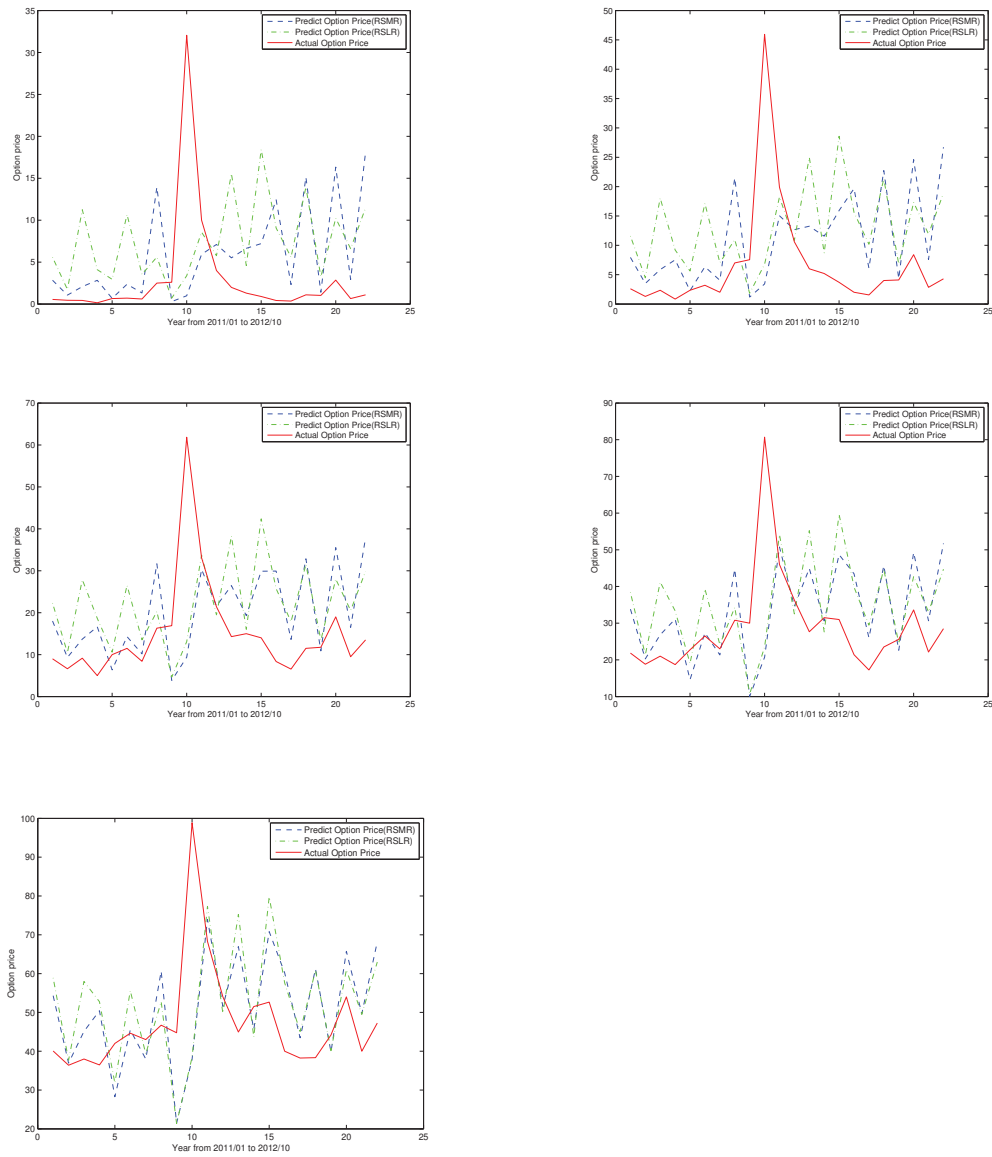Option Prices

Figure 4.7 shows the out-of-sample comparison of RSMR, RSLR and actual option prices.
From left to right and up to down, the graph represents DOTM, OTM, ATM, ITM and
DITM

Figure 4.8: Out-of-sample Comparison of the BS Model Prediction and Actual Option Prices
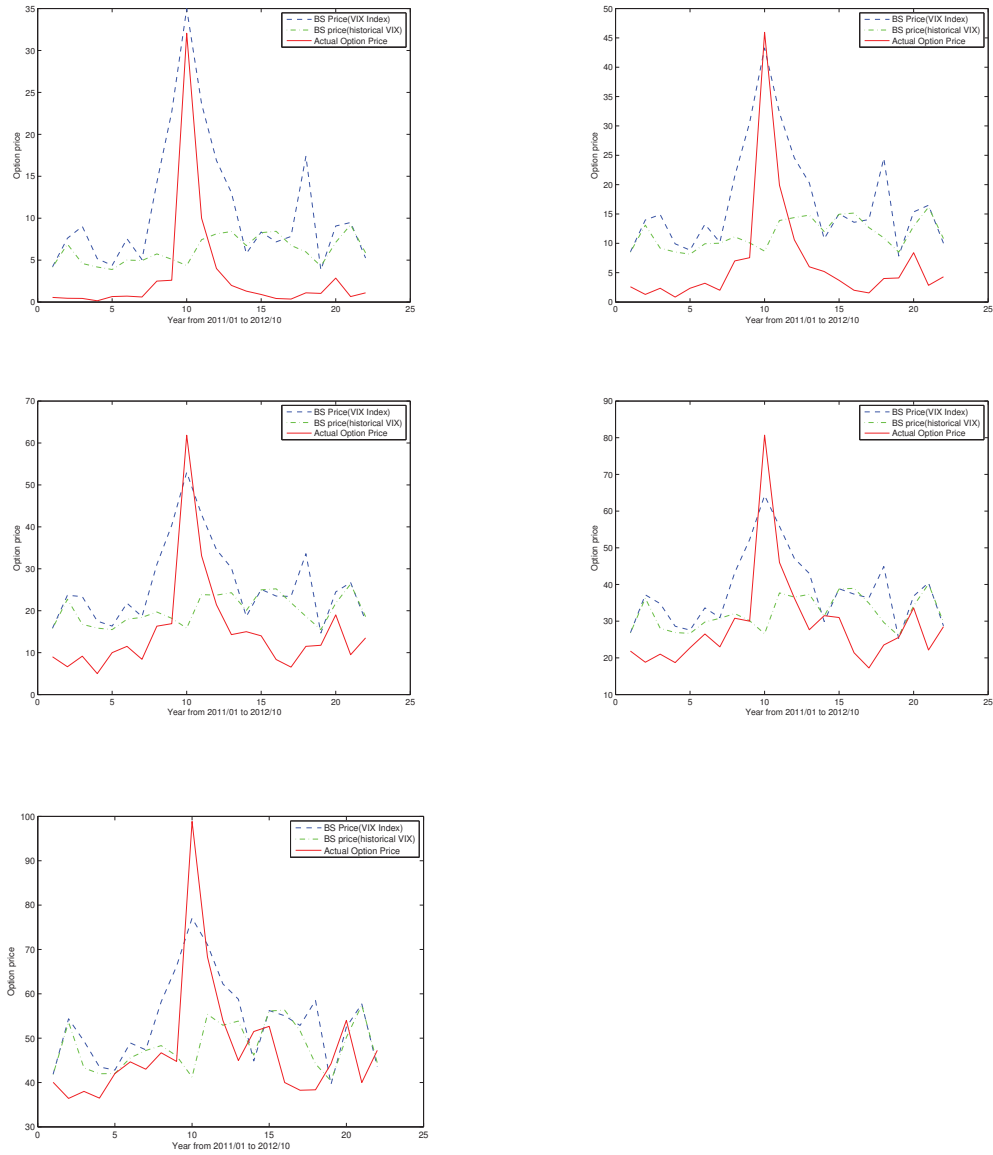

Figure 4.8 shows the out-of-sample comparison of BS model predictions and actual option prices. From left to right and up to down, the graph represents DOTM, OTM, ATM, ITM and DITM

Chapter 5

Conclusion

The standard linear asset pricing model is found to have many flaws, especially after the financial turmoils, such as the oil crisis in 1970s and recent financial crisis in 2008. As shown in many pervious studies, asset returns are highly regime-dependent. By ignoring the probability of changing regimes, the information among various states will be simply averaged. The Markov regime-switching model is employed to address the issue by characterizing the dynamic nature of data generating process over time.

In this thesis, in-sample and out-of-sample tests are both constructed to evaluate the performance of the dynamic regime-switching option-pricing model. The in-sample period spans from 1973/01 to 2010/12 while the out-of-sample period is from 2011/01 until 2013/06. The in-sample data shows clearly that the predictabilities of regime-switching VAR model is much better than the standard linear VAR model. This result is consistent with the findings by Gray (1996). He shows that using a regime-switching model to forecast is more sensible than using a constant-variance model.

With respect to the out-of-sample performance, the regime-switching model does not perform as well as relation to the in-sample model. The regime-switching model has less pricing errors than the traditional linear factor model does in general. Economic indicators are the key information source that the regime-switching model based on. The changing patterns of these indicators drive the market switching from one regime to the other, from the bearish state to the bullish state and vice versa. All macroeconomic indicators are selected based on previous studies. Many papers have shown the strong evidence that these indicators have significant correlation with the market, either positive or negative. However, there are still some biases in the selection of macroeconomic indicators. A further research can be constructed in selecting macroeconomic indicators with respect to the regime-switching option pricing model.

Risk neutral valuation is employed to price options under the fundamental theory for asset pricing. After adjusted by the risk neutral probability, the expected payoff in the future will be discounted as zero. This concept is heavily used in financial security pricing. In this research, a result worth of noting is the selection of risk free rate. Both the LIBOR rate and model rate are tested in pricing options. LIBOR rate is collected from Datastream and it is available from 1986/02 to 2013/06. Therefore, the in-sample data is rearranged from 1986/02 to 2010/12. The model rate is inferred from the quadratic programming process together with risk neutral probabilities. The empirical analysis shows the expectation that the model rate is more consistent than the LIBOR rate. In addition, comparison between the RS and BS models has been done. Different models exhibit a difference prediction power across different strike

prices. But in general, the distribution of the predicted values under RSMR option pricing model is the same as that of actual option prices. The limitations of the regime-switching model in this thesis are (1) the regime dependent constant variance-covariance matrix and (2) the constant transition matrix over time. These limitations will be addressed in the future studies.

Bibliography

[1] Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance, 52(5)*, 2003-2049.

[2] Black, F., & Myron Scholes. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81(3)*, 637-654.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31(3)*, 307-327.

[4] Boyle, P., Broadie, M., & Glasserman, P. (1997). Monte carlo methods for security pricing. *Journal of Economic Dynamics and Control, 21(8)*, 1267-1321.

[5] Campbell, S. D., & Li, C. (2002). *Option Pricing in Regime-Switching Environments* Working paper.

[6] Cerra, V., & Saxena, S. C. (2005). Did output recover from the asian crisis? *IMF Staff Papers, 52(1)*, 1-23.

[7] Chen, N., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *The Journal of Business, 59(3)*, 383-403.

[8] Chourdakis, K. M., Tzavalis, E., & Mary, Q. (2000). *Option Pricing under Discrete Shifts in Stock Returns*, University of London Working paper.

[9] Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics, 7(3)*, 229-263.

[10] Davig, T. (2004). Regime-switching debt and taxation. *Journal of Monetary Economics, 51(4)*, 837-859.

[11] Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society.Series B (Methodological), 39(1)*, 1-38.

[12] Duan, J. (1995). The GARCH option pricing model. *Mathematical Finance, 5(1)*, 13-32.

[13] Duffie, D. (2010). *Dynamic Asset Pricing Theory*, Princeton University Press.

[14] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 987-1007.

[15] Flannery, M. J., & Protopapadakis, A. A. (2002). Macroeconomic factors do influence aggregate stock returns. *Review of Financial Studies, 15(3)*, 751-782.

[16] Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics, 42(1)*, 27-62.

[17] Guidolin, M., & Timmermann, A. (2005). *Strategic asset allocation and consumption decisions under regime-switching*, Federal Reserve Bank of St. Louis, Research Division.

[18] Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica,57(2)*, 357-384.

[19] Hamilton, J. D. (2005). *What's Real about the Business Cycle?* NBER Working Paper.

[20] Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies, 3(4)*, 573-592.

[21] Ingersoll, J. E. (1987). *Theory of financial decision making* Rowman & Littlefield.

[22] Jeanne, O., & Masson, P. (2000). Currency crises, sunspots and markov-switching regimes. *Journal of International Economics, 50(2)*, 327-350.

[23] Karatzas, I., & Shreve, S. E. (1998). *Methods of Mathematical Finance* Springer.

[24] Liu, P. P., Xu, K., & Zhao, Y. (2011). Market regimes, sectorial investments, and time-varying risk premiums. *International Journal of Managerial Finance, 7(2)*, 107-133.

[25] Mamaysky, H., Spiegel, M., & Zhang, H. (2008). Estimating the dynamics of mutual fund alphas and betas. *Review of Financial Studies, 21(1)*, 233-264.

[26] Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics, 3(1)*, 125-144.

[27] Mulvey, J. M., & Zhao, Y. (2011). *An Investment Model Via Regime-Switching Economic Indicators*. Working paper.

[28] Musiela, M., & Rutkowski, M. (2005). *Martingale methods in financial modelling*, Springer.

[29] Naik, V. (1993). Option valuation and hedging strategies with jumps in the volatility of asset returns. *The Journal of Finance, 48(5)*, 1969-1984.

[30] Qiu, J., Faff, R. W., & Benson, K. (2011). Business cycles and mutual fund timing performance: An application of regime-switching and GARCH modeling. Paper presented at the *24th Australasian Finance and Banking Conference.*

[31] Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6(2)*, 461-464.

[32] Sims, C. A., & Zha, T. (2006). Were there regime switches in US monetary policy? *The American Economic Review,96(1)*, 54-81.

[33] Veronesi, P. (1999). Stock market overreactions to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies, 12(5)*, 975-1007.