

# Stochastic Modeling and Estimation of Market Volatilities with Applications in Financial Forecasting

Jin Zheng (Correspondence author)

Department of Mathematical Sciences, University of Liverpool

Liverpool L69 3BX, UK

E-mail: Jin.Zheng@liverpool.ac.uk

Dejun Xie

Department of Mathematical Sciences, Xian Jiaotong Liverpool University

Suzhou 215123, China

E-mail: Dejun.Xie@xjtlu.edu.cn

Received: February 15, 2012 Accepted: February 27, 2012 Published: May 1, 2012

doi:10.5539/ijsp.v1n1p7 URL: <http://dx.doi.org/10.5539/ijsp.v1n1p7>

## Abstract

This paper aims to provide a framework for modeling and estimating the implied volatilities of stock prices using stochastic processes. The CIR (Cox-Ingersoll-Ross) model is proposed to capture the mean reversion characteristic as shown in the movement of the daily implied volatilities of the S&P 500 Index and Nikkei 225 Index. The maximum likelihood procedure is applied to estimate the parameters appearing in the model, where both analytical and approximation techniques are sought to handle the difficulties arising from the corresponding optimization problem. The procedure is validated with varying sampling methods by setting different time horizons and observation intervals. Results from numerical experiments show that the stochastic volatility model proposed in this paper functions well in both the US and Japan markets. As one of the applications, our approach is tested to be effective in detecting jumps in likelihood ratios, hence useful for forecasting stock market shocks and crisis.

**Keywords:** Stock volatility, CIR model, Maximum likelihood ratio, Financial forecasting

## 1. Introduction

Implied volatility plays a key role in studies of financial options and derivatives. A classic example is the Black-Schole option pricing method (Black and Scholes, 1973), where the implied volatilities of the same underlying asset with the same expiration date are assumed as constant, even when the strike prices are different. Here one of the most important assumptions of BS model is the constant volatility, or local volatility. However, abundant studies have questioned the validity of the constant volatility assumption. Dumas et al. (1998) analyzes the prices of S&P 500 index options for the period from June 1988 to December 1992, and shows that the volatility is a function of asset price and time to maturity. Derman et al. (1994) argues that the implied volatility is a function of share price, exercise price, time to maturity and a drift function. In the experiment of Peña et al. (1999), moneyness is considered as a determining factor of the implied volatility for options on the Spanish IBEX-35 exchange market. In Cont and Fontseca (2002), a contrast study focusing on out of money options is reported using S&P500 index and FTSE100 index. Cassese and Guidolin (2003) studies whether the implied volatility is a function of time to maturity, moneyness, or the interaction between the two with data of the DOTM options of MIB30 index on Italian market.

Understanding implied volatility is important to both theoretical and practical finance as it offers a reliable barometer to forecast the trend of financial market and largely affects the values of option contracts. Most of the previous literatures in this topic are of empirical in nature. In this work, we intend to address the problem by structural modeling approach where the CIR mean-reverting process is adopted to describe the movement of implied volatilities over time. The exponential average method is used to generate the daily implied volatilities for S&P 500 and Nikkei 225, for the purpose of which Chaikin's volatility function is implemented in finding the spread between the stocks' high price and low price and in determining the percent of changes in moving average of high price against low price for a specified time interval. The maximum likelihood procedure is carried out to estimate the parameters of the CIR model based on discretization of the model and approximation of the corresponding transitional probability density function. Matlab algorithms are proved to be helpful in solving the pertaining optimization problems arising from the maximum likelihood procedure. Likelihood

ratio analysis of historical movement of volatilities is carried out for predicting stock market shocks.

The rest of the paper is organized as follows. Section 2 discusses the methodology and procedures of our study, including data description and model derivation. Section 3 contains data analysis and empirical results of our study. Section 4 is to show the applications of our modeling and estimation procedure for financial crisis prediction. Concluding remarks and possible future directions are provided in Section 5.

## 2. Methodology and Procedures

### 2.1 Data for the Study

The data sets used in this paper are S&P 500 index and Nikkei 225 index. The S&P 500 index is chosen for the following considerations. First, since its inception in 1957, S&P 500 index has been a barometer for the American economy, which draws interests from many researchers. The market trend predicted by the many financial analysts is mostly based on observing the movement of S&P 500 index, although such a trend analysis may not always quickly foresee the potential bubbles existed in the market. Second, since there is a large trading volume of S&P 500 index every day, the data set is abundant and representative for the experiment in this model. Third, the data are maintained continuous and consistent in S&P 500 index, which is critically important for model testing and selection. Nikkei 225 is a stock market index for the Tokyo Stock Exchange. As the most widely quoted average of Japanese equities, Nikkei 225 is a benchmark to value the Japanese economy. Previous literatures such as Kunitomo and Sato (2009) have emphasized the importance of Nikkei 225 for Japanese market. The financial systems are relatively mature in the United States and Japan, which is also one reason to choose the indexes of these two countries. The data entry is composed of trading date, high price and low price. The data sets are openly available at yahoofinance. All public holidays are deleted from the raw data entries for considerations of consistency and continuity. The implied volatilities can be computed by the method of exponential average, for instance. The parameters for CIR model can be estimated through maximum likelihood procedure, although technical points such as the singularities arising from the optimization iterations involving Bessel type of functions have to be addressed separately.

### 2.2 Model Description

The adoption of a mean-reverting stochastic process expressed in terms of, say, the Ornstein Uhlenbeck equation

$$d\sqrt{v(t)} = -\beta\sqrt{v(t)}dt + \delta dZ(t) \quad (1)$$

where  $v(t)$  is the variance of the stock, and  $Z(t)$  is a Brownian Motion, and  $\beta$  and  $\delta$  are positive constants, for modeling volatilities can be traced back to Stein and Stein (1991), for instance, where several theoretical favorable consequences ensuing from such a gratification have been discussed, including the mathematical tractability in the case of simple coupon bond pricing. On the other hand, some undesirable implications, mainly the allowing of negative variance, have also been noticed and explained in Stein and Stein (1991). Therefore, it is more theoretically plausible and practically useful to adopt a model with only non-negative trajectories.

To transform equation (1), let  $x(t) = \sqrt{v(t)}$ , one gets

$$dx = -\beta x dt + \delta dZ(t)$$

Applying Ito's lemma with  $df = dv(t) = dx^2(t)$  (Heston, 1993) gives

$$\begin{aligned} dx^2(t) &= \frac{\partial f}{\partial x(t)} dx(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x(t)^2} dx^2(t) \\ &= 2x(t)(-\beta x(t)dt + \delta dZ(t)) + \delta^2 dt \\ &= (\delta^2 - 2\beta x^2(t))dt + 2x(t)\delta dZ(t) \end{aligned}$$

Replacing  $x(t)$  by  $\sqrt{v(t)}$  yields

$$dv(t) = (\delta^2 - 2\beta v(t))dt + 2\delta\sqrt{v(t)}dZ(t), \quad (2)$$

which is a special case of the usual CIR model:

$$dv(t) = k(\theta - v(t))dt + \sigma\sqrt{v(t)}dZ(t) \quad (3)$$

In this way, one can capture the variance or the volatility of the stock prices described by the CIR model. While Heston (1992) mainly concerns the theoretical benefits of adopting such models in terms of the option pricing, the work contained herein focuses on the validity and robustness of such modelings from statistical analysis perspective. In particular, we

are more interested in the consistency and convergence of the parameter estimation via maximum likelihood procedure, for instance. We are also interested in whether the model is applicable to predicting stock plunges and economic cycles, which are more practical and reasonable concerns from practitioner and policy maker's point of view.

### 2.3 ML Estimation for CIR Process

Under the Euler-Maruyama approximation, we can rewrite equation (3) as

$$\Delta v(t) = k(\theta - v(t))\Delta t + \sigma \sqrt{v(t)}\Delta Z(t)$$

or equivalently, by the definition of  $\Delta v(t) = v(t + \Delta t) - v(t)$  and  $\varepsilon_t = \Delta Z(t)$

$$v(t + \Delta t) - v(t) = k(\theta - v(t))\Delta t + \sigma \sqrt{v(t)}\varepsilon_t \quad (4)$$

where  $\varepsilon_t$  is to be minimized in the following maximum likelihood procedure.

To obtain maximum likelihood estimation of the CIR process, we first notice that the transitional probability density for the process is explicitly given as (Cox et al., 1985)

$$p(\Delta, x, y) = c * \exp(-u - v) * \left(\frac{u}{v}\right)^{\frac{q}{2}} * I_q(2\sqrt{uv})$$

where

$$\begin{aligned} c &= \frac{2k}{\sigma^2(1 - \exp(-k\Delta))} \\ q &= \frac{2k\theta}{\sigma^2} - 1 \\ u &= cx * \exp(-k\Delta) \\ v &= cy \end{aligned}$$

from which one can derive the log-likelihood function of CIR process :

$$\begin{aligned} \ln L(m) &= (N - 1) \ln c + \sum_{i=1}^{N-1} (-u_{t(i)} - v_{t(i+1)}) \\ &\quad + 0.5q \ln \frac{v_{t(i+1)}}{u_{t(i)}} + \ln I_q[2\sqrt{u_{t(i)}v_{t(i+1)}}] \end{aligned}$$

where  $m = (k, \theta, \sigma)$  denotes the estimation of the parameters for the model, and N is the total number of observations. To obtain the estimated parameters  $m = (k, \theta, \sigma)$ , one needs to maximize the log-likelihood function  $\ln L(m)$ . The maximum of  $\ln L(m)$  can be solved by the system of equations (5) :

$$\begin{cases} \frac{\partial \ln L(m)}{\partial k} = 0 \\ \frac{\partial \ln L(m)}{\partial \theta} = 0 \\ \frac{\partial \ln L(m)}{\partial \sigma} = 0 \end{cases} \quad (5)$$

However, the estimation of  $m = (k, \theta, \sigma)$  is not trivial with a modified Bessel function of the first kind  $I_q(2\sqrt{uv})$  contained in the expression, which is related one of the two linearly independent solutions to the following modified Bessel's differential equation:

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} - (x^2 + n^2)y = 0$$

Consider the definition of the Modified Bessel's function of the first kind of order n:

$$\begin{aligned} I_n(x) &= \frac{x^n}{2^n \Gamma(n+1)} \left[ 1 + \frac{x^2}{2(2n+2)} + \frac{x^4}{2 * 4(2n+2) + (2n+4)} + \dots \right] \\ &= \sum_{k=0}^{\infty} \frac{x^n}{2^n \Gamma(n+1)} * \frac{(\frac{x}{2})^{2k}}{\frac{(n+k)!}{n!} k!} \\ &= \sum_{k=0}^{\infty} \frac{(\frac{x}{2})^{2k+n}}{n! \frac{(n+k)!}{n!} k!} \\ &= \sum_{k=0}^{\infty} \frac{(\frac{x}{2})^{2k+n}}{k! \Gamma(n+k+1)} \end{aligned}$$

Substituting  $x = 2\sqrt{uv}$ ,  $n = q$  into the equation  $I_x(n)$  yields

$$I_q(2\sqrt{uv}) = \sum_{k=0}^{\infty} \frac{(-1)^k \left(\frac{2\sqrt{uv}}{2}\right)^{q+2k}}{k! \Gamma(q+k+1)}$$

It is costly to solve the system of the equations (5) for obtaining the estimated parameters  $m = (k, \theta, \sigma)$  without leveraging on computing packages provided by the `fminsearch` function in Matlab, for instance. However, the Matlab routine produces different answers for the estimated parameters with different initial values. In this paper, the ordinary least square procedure proposed by Kladykova (2011) is used to obtain the initial points of optimization.

Plug in the observed sample values into equation (4) and rearrange it, we have

$$\sum_{i=1}^{N-1} \varepsilon_i^2 = \sum_{i=1}^{N-1} \left[ \frac{v(t_{i+1}) - v(t_i)}{\sigma \sqrt{v(t_i)}} - \frac{k(\theta - v(t_i))\Delta t}{\sigma \sqrt{v(t_i)}} \right]^2 \quad (6)$$

where the estimation of initial value of  $\sigma$  is the standard deviation of residuals, after which the initial values of  $k$  and  $\theta$  can be obtained using the following procedure of ordinary least squares:

$$\begin{cases} \frac{\partial \sum_{i=1}^{N-1} \varepsilon_i^2}{\partial k} = 0 \\ \frac{\partial \sum_{i=1}^{N-1} \varepsilon_i^2}{\partial \theta} = 0 \end{cases} \quad (7)$$

#### 2.4 Procuring Daily Volatilities

As the main objective is to model the implied volatilities of S&P 500 index and Nikkei 225 index as well as to estimate the parameters of the governing model, we need a proper scheme to measure the daily volatility of a specific stock given the historical data, where the model under consideration in this work is given in the equation (3).

Since  $v(t)$ , the daily variance is unknown, the daily volatility has to be retrieved with implicit method. In this paper the spread between the stock's high price and low price in a specified interval such as a trading day is computed in order to determine the Chaikin's volatility using percentage change in the moving averages for a given period. There are two major steps for the calculation.

First, we calculate an exponential moving average of the difference between the daily high and low prices. Compared to simple average method, the exponential moving average method assumes the latest price contributes more weight than the prices in more distant periods. Generally, a 10-day moving average is recommended (Achelis, 2011). The general form of a series  $Y$  is defined as

$$S_t = \alpha * Y_{t-1} + (1 - \alpha) * S_{t-1}$$

where  $S_t$  is the exponential average value of stock price at time  $t$ ,  $Y_t$  is the observation at time  $t$ ,  $\alpha$  represents the degree of weighting decrease. According to Hunter (1986), we can rewrite as a weighted sums of the following series:

$$S_t = \alpha * [Y_{t-1} + (1 - \alpha) * Y_{t-2} + (1 - \alpha)^2 * Y_{t-3} + \dots + (1 - \alpha)^9 * Y_{t-10}] + (1 - \alpha)^{10} * S_{t-10}$$

where  $Y_t$  represents difference between the high price and the low price of a specific stock at time  $t$ .

Second, we calculate the percentage of changes over a specified time period according to the formula

$$\sigma = \frac{S_t - S_{t-10}}{S_{t-10}} * 100\%.$$

### 3. Data Analysis and Numerical Results

#### 3.1 Model Calibration Using Historical Data

We present the results for estimated parameters,  $k$ ,  $\theta$ , and  $\sigma$  using daily observations of S&P 500 Index and Nikkei 225 index. Since the observations are daily recorded, the time interval,  $\Delta$ , is set as  $1/365$  in the unit of number of years. The Figure 1 and 2 below show the historical behavior of S&P 500 Index and Nikkei 225. The daily volatility against time is plotted in the graph. The data of S&P 500 Index and Nikkei 225 Index both have a duration of 12 years from 1995 to 2006. The results of parameter estimations from the implementation of the maximum likelihood procedure are presented in the following Table 1 and Table 2.

Table 1 shows the results for the estimated parameters  $k$ ,  $\theta$  and  $\sigma$  for S&P 500 index. In this experiment, 5 years', 8 years', 10 years' and 12 years' data are used for comparative analysis. As shown in Table 1, the speed that the process comes back to the mean,  $k$ , decreases smoothly from 5 years data to 12 years data. The estimated long-term means,  $\theta$ , are close to the means of historical data for different durations, which means the value for  $\theta$  approaches to the corresponding value of the mean of  $v(t)$ . In addition, the values for estimated long-term means,  $\theta$ , converge to 0.03 in different sub-periods. The estimated values for volatility,  $\sigma$ , for the periods of 5 years, 8 years, 10 years and 12 years are all located in the interval [0.05 0.08]. The estimated value for  $\theta$  has shown slight increment with the increase of time duration, although the increment is not significant.

Table 2 shows the results for the estimated parameters  $k$ ,  $\theta$  and  $\sigma$  for Nikkei 225 index. In this experiment, again, we use 5 years', 8 years', 10 years' and 12 years' data to make comparisons. In Table 2, the drift parameter,  $k$ , varies in a wide range for different periods. The sudden change of  $k$  is a symbol of the dramatic changes in the implied volatility, which has to do with the economic shocks. For instance, the Asian market saw financial and currency crisis in 1997. Russia experienced debt default and long-term capital management crisis in 1998. Both of the two crises had a significant impact on major Asian markets like Japan's Nikkei 225, which can explain why  $k$  jumped from 4 to 26.8. Moreover, the estimated long-term means approach to the means of historical data, the trend of which is the same as S&P 500 index. And the estimated volatility,  $\sigma$ , with the time period of 5 years, 8 years, 10 years and 12 years are approaching to 0.07. Similarly, the values for  $\theta$  are not exactly the same for different periods, but the difference is insignificant.

Overall, two stock markets have profiled similar trends of volatility. Compared to those for Nikkei 225 Index, the estimated parameters of the volatilities of S&P 500 Index are more consistent. As shown in Table 1, the estimated parameters,  $k$ ,  $\theta$  and  $\sigma$  show small difference when using different sampling periods. However, for Nikkei 225, the estimates to the speed of adjustment for the process,  $k$ , has dramatic changes from one sampling period to another, which is due to the economic crisis in Asia. The volatility model in this paper, or CIR model, functions well in a stable market.

To make the results more comprehensive, we extend our experiment to include the cases of using short-term data, i.e., 1 year and 2 years of S&P 500 Index and Nikkei 225 Index to obtain the estimates of the model parameters. However, the results are not as consistent as expected. In particular,  $k$ , the speed that the process will come back to the mean, decreases dramatically from 1 year's data to 2 years' data for S&P 500 Index. By observing the daily recorded data in 2006, the value for  $k$  is 144.2, while when we use 2 years' data in 2005 and 2006,  $k$  jumps to 45.8. As shown in Table 1, the estimates of  $k$  using medium to long term observation durations approach to 6, implying that the model is more suitable for medium to longterm forecast. In addition, we use 1 year's data to 2 years' data for Nikkei 225 Index to estimate the parameters. The estimations of  $k$ ,  $\theta$  and  $\sigma$  are consistent in different samplings, but the estimation of  $k$  flutters due to economic shocks.

### 3.2 Influence of Observation Frequency

#### 3.2.1 Volatility Indicator Based on EMA

Chainkin's volatility is used in this paper to obtain the daily implied volatilities for S&P 500 Index and Nikkei 225 Index. Practically a 10-day exponential moving average is generally used for computing the volatility. However, there is not a generally agreed horizon for risk management. The horizons differ from different derivatives, or financial instruments. The equity and foreign exchange have horizons of 7 to 10 days, while the interest rate instruments have horizons for about 30 days (Diebold et al., 1998). Traders will choose a time period fitting to their investment time frame or preference. Generally, risk is assessed at a short time horizon (Christoffersen et al., 1998). To make the results more conclusive, we repeat the estimation process and compare the volatilities obtained by 5-day and 15-day EMA methods, in comparison with the previous results using 10-day EMA method.

Table 3 and 4 show the results of estimated parameters for S&P 500 Index and Nikkei 225 Index for 5-day moving average, 10-day moving average and 15-day moving average from 1995 to 2006. As shown in Table 3, for S&P 500 Index,  $k$  fluctuates wildly when using different time periods for averaging. In addition,  $k$  achieves the smallest value when 10 day moving average is used. The values for  $\sigma$  and  $\theta$  vary insignificantly with different time periods. Table 4 shows the similar results for the values of estimated parameters for Nikkei 225 index.

#### 3.2.2 Influence of Observation Frequency

In this paper, abundant observations are recorded with daily recorded data. It is desirable to investigate the influence of observation frequency on the parameter estimation. For the selected CIR model in this paper, we carry out the estimation process using not only daily observations, but also weekly observations and monthly observations. Table 5 and 6 present the results of estimated parameters for S&P 500 Index and Nikkei 225 Index for daily, weekly and monthly data from 1995 to 2006. The values of the estimated parameters using the data set of S&P 500 Index do not change significantly except for  $k$ . The value for  $k$  achieves the maximum for monthly recorded data, and the minimum for daily recorded data.

The results of the estimated parameters using Nikkei 225 index are consistent with S&P 500 Index, implying that time increment in sampling or increase of observation frequencies result in different parameter values for stochastic volatility models.

In this study, the parameters of the stochastic volatility model are estimated through the maximum likelihood procedure. We would like to remark that there exist alternative approaches to estimate the parameters for stochastic models, such as the Bayesian estimation method. Compared to maximum likelihood method, Bayesian estimation requires the specification of a prior distribution for unknown parameters. And also, Bayesian method is often realized through sampling simulations with augmented latent points, giving rises of the issue of inconsistency. Thus maximum likelihood estimation is still more preferred for the current work. The results of the maximum likelihood procedure through this work show that the parameters  $\theta$  and  $\sigma$  are more consistent for different sampling time periods. However, the estimated values for  $k$  are not quite consistent for different sampling periods, even for different time horizons of averaging. Similar findings, particularly the difficulty of achieving convergent values for  $k$ , have been discussed in Feng and Xie (2011, 2012), for instance.

#### 4. Application in Financial Crisis Prediction

##### 4.1 Economic Cycle

Stock volatility, which represents massiveness of the change of stock prices (Olowe & Ayodeji, 2009), is a major indicator of future economic activity (Romer, 1990). It is generally believed that the stock market fluctuation is closely related to the macroeconomic development. The increased stock volatility usually signals higher uncertainty for future economic activity. The investors or the stock holders will tend to spend less money or wealth on consumption and investment, which will lead to decrease in aggregate demand, even economic slack. On the other hand, the decrease in stock volatility encourages the investors to make consumption or investment, which is often the first move in a cycle towards the economic prosperity.

##### 4.2 Financial Crisis Prediction

Financial crisis, a symbol of economic downturn, or recession, happens when financial institutions or financial markets lose large part of their value. The stock volatility is a powerful indicator of financial crisis because it is closely related to business or economic cycle. By observing the U. S. stock market's implied volatility and quarterly percentage growth of real GDP, Raunig and Scharler (2011) carried out an experiment on the relationship between stock volatility and stock market shocks and found a negative relationship between volatility and GDP growth.

From Figure 7, one can see that the implied volatility fluctuates during the period from 1984 to 2006. As mentioned, stock market shocks are often a strong indicator of economic recessions. From the historical record of the implied volatility, it can be seen that the implied volatility fluctuates intensely in the period from 1986 to 1994. During these years, Japan suffered several major financial crises, including the burst of Japanese real estate bubble. From 1986 to 1991 the property market of Japan had undergone a great inflation, which explains the fluctuant movement of volatilities during the same period. In the following years, the world saw the Japan's ERM crisis in 1992, Mexican peso crisis in 1994, the Asian financial and currency crisis in 1997, Russian debt default and long-term capital crisis in 1998. All these crises had severally and collectively impacted the Japanese economy and its stock market.

##### 4.3 Case Study: Stock Market Crisis in 1987

The U. S. stock market suffered one of its largest daily percentage decline on October 19, 1987, on which day the volatility saw a huge jump. Black (1988) largely attributed the stock market plunges on October 19 to the errors in the perception of mean-reversion expectations of investors, or equivalently, the so called mean reversion illusion (Hillebrand, 2003). Here the mean reversion speed is the speed that the process will come back to the mean, which corresponds to the parameter,  $k$ , in the stochastic volatility model in this paper.

According to Hillebrand (2003), the mean reversion in returns is phenomenal, which is transient but recurring. In Table 7 and Figure 8, one can observe the fact that the value of  $k$ , the mean reversion speed, is higher after the stock plunge than before the plunge, which supports the concept of mean-reversion disillusion. In addition,  $k$  is higher at about 9 months before the plunge, which is consistent with the experiment carried out by Hillebrand (2003). The underestimation of the mean reversion is the leading reason for the stock plump.

In order to provide a further proof of the mean-reversion disillusion, which assumes the mean-reversion rate is higher after stock fluttering period, we carry out the statistical t-test with the following hypotheses:

$$H_0 : k_1 < k_2$$

$$H_1 : k_1 \geq k_2$$

where  $k_1$  refers to the mean of  $k$  before the stock plump and  $k_2$  is the mean of  $k$  after the plump. We use two sampled t-test to evaluate the null hypothesis and alternative hypothesis:

$$t = \frac{k_2 - k_1}{S_{(k_2-k_1)}}$$

$$S_{(k_2-k_1)} = \sqrt{\frac{S_{k_1}^2}{n_1} + \frac{S_{k_2}^2}{n_2}}$$

where  $S_i$  is the unbiased estimator of the variance of the two samples,  $n_1$  = number of participants before the market plunge,  $n_2$  = number of participants after the plunge. The degree of freedom, denoted as d. f., of our approximation is

$$d.f. = \frac{\left(\frac{S_{k_1}^2}{n_1} + \frac{S_{k_2}^2}{n_2}\right)^2}{\left(\frac{S_{k_1}^2}{n_1}\right)^2/(n_1 - 1) + \left(\frac{S_{k_2}^2}{n_2}\right)^2/(n_2 - 1)}$$

The calculation shows that the p-value of the t-statistics 0.1695, which means we can not reject  $H_0$  even at 10% significant level.

We apply Hillebrand (2003)'s approach of using likelihood ratio to test whether the mean-reversion speed is higher after the stock market plump. For this purpose, the following two models are compared.

$$\text{Model1} : d \ln S_t = \left(u - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t$$

$$\text{Model2} : d \ln S_t = \left(u - \frac{1}{2}\sigma^2\right)dt + \lambda(\ln V_t - \ln S_t)dt + \sigma dW_t$$

The unconditional distribution of the log-price processes above are given by:

$$\text{Model1} : (\ln S_t - \ln S_{t-1}) \sim N\left(u - \frac{1}{2}\sigma^2, \sigma^2\right)$$

$$\text{Model2} : (\ln S_t - \ln S_{t-1}) \sim N\left(u - \frac{1}{2}\sigma^2 + \lambda(\ln V_t - \ln S_t), \sigma^2\right)$$

where  $S_t$  is the stock price at time  $t$ ,  $V_t = S_0 \exp\left((u - \frac{1}{2}\sigma^2)t\right)$ ,  $u$  is the long-term mean growth of the market,  $\lambda$  is the mean-reversion speed,  $\sigma$  is the volatility of the stock price, and  $W_t$  is the standard Brownian Motion. The hypotheses are

$$H_0 : \lambda = 0$$

$$H_1 : \lambda \neq 0$$

The likelihood-ratio statistics is used to evaluate the null hypothesis and alternative hypothesis, where the likelihood ratio statistics is defined as

$$L = \frac{\sup_{u \geq 0, \sigma \geq 0} \left[ \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{((y_i - x_i) - (u - \frac{1}{2}\sigma^2))^2}{2\sigma^2}\right) \right]}{\sup_{u \geq 0, \sigma \geq 0, \lambda \geq 0} \left[ \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{((y_i - x_i) - (u - \frac{1}{2}\sigma^2 + \lambda(\ln V_i - \ln S_i)))^2}{2\sigma^2}\right) \right]}$$

where  $y_i$  and  $x_i$  represent the logarithms of stock prices on  $i$ th and  $(i - 1)$ th day.

We apply the transition probability functions of Model (1) and Model (2) to the likelihood ratio statistics using four different sampling periods. The parameters  $(u, \sigma)$  in Model (1) and  $(u, \sigma, \lambda)$  in Model (2), are estimated by MLE using daily stock index values. The results are presented in Figure 9.

#### 4.4 Further Applications

The implied volatility model is widely accepted and used in financial market. In risk management and financial mathematics, value-at-risk is a widely used risk measure technique which estimates the risk of loss on a specific portfolio of financial assets. To be more concrete, value-at-risk of a given portfolio demonstrates the expected maximum loss over an aimed horizon within a given confidence interval (Larsen, 2001). Since the implied volatility provides an unbiased and efficient forecast of future market, capturing the volatility process is a main approach for value-at-risk estimation. Hence, the implied volatility models in Section 2 can be used for computing value-at-risk portfolios, which include assets whose payoffs/returns are functions of the S&P 500 index or Nikkei 225 Index (Cassese & Guidolin, 2003), although it is not

often to have closed form solution for value-at-risk. The chosen portfolio can be combination of equity shares, assets and other financial derivatives.

## 5. Concluding Remarks

To estimate whether the implied volatilities are constant during the whole duration of options as Black-Schole model assumed, we use CIR model to describe the implied volatilities and maximum likelihood method to estimate the parameters. The data set in this paper are S&P 500 Indexes and Nikkei Index from 1995 to 2006. According to the statistical analysis results, the selected stochastic volatility model fits well the S&P 500 Index and Nikkei Index. The robustness of our model is calibrated with historical data with varying sampling methods. The estimated parameters are convergent in the stable market.

The implied volatilities can be used as an unbiased and efficient forecast of the performance of financial market. The model of implied volatility can be applied in many fields of financial market, including for calculation of the value-at-risk for portfolios and for stock crisis prediction. The results support the idea that the implied volatility tends to be higher in the recession period. In the meantime, the mean-reverting disillusion, which explains the mean-reverting speed is higher after the stock fluttering period than before, is proved by both the t-statistics and likelihood ratio test in the study. The approaches of financial modeling and the relating parameter estimation procedures contained in this study also provide useful supports to neighboring studies such as characterization of mortgage loans and American put options (Xie et al., 2007, 2011).

One limitation of the current study is the relatively narrow scope of the observations, where only 12 years of data of two mature markets are considered. For future research, the data set period should be chosen to allow more frequent regime switches. Also an improved modeling should be able to achieve more stable and convergent estimation of the mean-reverting speed,  $k$ . In addition, the current work only concerns the stochastic modeling for S&P 500 Index and Nikkei Index. A further study is recommended to be carried on for other types of index such as KOSPI 200 in Korea and the Hushen 300 Index in China.

## Acknowledgment

The research is partially supported by the Natural Science Foundation of Jiangsu Province of China. The authors would like to thank two anonymous referees for their helpful comments that have improved this manuscript.

## References

- Achelis, S. B. (1999). Trading Solutions Function Library. [Online] Available: <http://www.tradingsolutions.com/functions/ChaikinsVolatility.html> (November 7, 2011).
- Black, F., & Scholes, M. (1973). The pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81, 637-654. <http://dx.doi.org/10.1086/260062>
- Black, F. (1988). An Equilibrium Model of the Crash. *NBER Macroeconomics Annual*, 3, 269-275. <http://dx.doi.org/10.2307/3584955>
- Cassese, G., & Guidolin, M. (2003). Modeling the Implied Volatility Surface: Does Market Efficiency Matter? *University of Virginia Working Paper*.
- Christoffersen, P., Diebold, F., & Schuermann, T. (1998). Horizon Problems and Extreme Events in Financial Risk Management. *Economic Policy Review, Federal Reserve Bank of New York*, 109-118.
- Cont, R., & Fonseca, J. D. (2002). Dynamics of Implied Volatility Surfaces. *Quantitative Finance*, 2, 40-56. <http://dx.doi.org/10.1088/1469-7688/2/1/304>
- Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53, 385-408. <http://dx.doi.org/10.2307/1911242>
- Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied Volatility Functions: Empirical Tests. *The Journal of Finance*, 6, 2059-2106. <http://dx.doi.org/10.1111/0022-1082.00083>
- Derman, E., & Iraj, K. (1994). The Volatility Smile and Its Implied Tree. *Quantitative Strategies Research Notes*, Goldman Sachs.
- Diebold, F., Hickman, A., Inoue, A., & Schuermann, T. (1998). Scale Models. *Risk*, 11, 104-107.
- Feng, X. X., & Xie, D. J. (2011). Application of MCMC Algorithm in Interest Rate Modeling. *Proceeding of the International MultiConference of Engineers and Scientists*, 16-18.

- Feng, X. X., & Xie, D. J. (2012). Bayesian Estimation of CIR model. *Journal of Data Science*, 10, 271-280.
- Hillebrand, E. (2003). A Mean-Reversion Theory of Stock-Market Crashes. Center for Complex Systems and Visualization, Universitat Bremen, Department of Mathematics, Stanford University.
- Hunter, J. (1986). The Exponentially Weighted Moving Average. *Journal of Quality Technology*, 18, 203-209.
- Heston, S. L. (1992). Invisible Parameters in Option Prices. *The Journal of Finance*, 48, 933-947. <http://dx.doi.org/10.2307/2329021>
- Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6, 327-343. <http://dx.doi.org/10.1093/rfs/6.2.327>
- Kunitomi, N., & Sato, S. (2009). Realized Volatility, Covariance and Hedging Coefficient of the Nikkei-225 Futures with Micro-Market Noise. *18th World IMACS/MODSIM Congress*, 17-19.
- Kladivko, K. (2011). Maximum Likelihood Estimation of The Cox-Ingersoll-Ross Process: The MATLAB Implementation. Department of Statistics and Probability Calculus, University of Economics, Prague and Debt Management Department, Ministry of Finance of the Czech Republic.
- Larsen, N., Mausser, H., & Ursvsev, S. (2001). Algorithms for Optimization of Value-at-Risk. *Financial Engineering, E-Commerce and Supply Chain*, 129-157.
- Olowe, R., & Ayodeji, R. (2009). Stock Return, Volatility And The Global Financial Crisis in An Emerging Market: The Nigerian Case. *International Review of Business Research Papers*, 5, 426-447.
- Peña, I., Serna, G., & Rubio, G. (1999). Why do We Smile? On the Determinants of The Implied Volatility Function? *Journal of Banking and Finance*, 23, 1151-1179.
- Romer, C. (2011). The Great Crash and the Onset of the Great Depression. *The Quarterly Journal of Economics*, 105, 597-1179.
- Raunig, B., & Scharler, J. (2011). Stock Market Volatility and the Business Cycle. *Monetary Policy & Economy*, Q2, 54-63.
- Stein, E. M., & Stein, J. C. (1991). Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Review of Financial Studies*, 4, 727-752.
- Xie, D. D., Chen, X. F., & Chadam, J. (2007). Optimal Payment of Mortgages. *European Journal of Applied Mathematics*, 18, 363-388.
- Xie, D. D., Edwards, D. A., Schleiniger, G., & Zhu, Q. H. (2011). Characterization of the American Put Option Using Convexity. *Applied Mathematical Finance*, 18, 353-365.

Table 1. Maximum Likelihood estimates of CIR model using historical daily data of S&P 500 Index. The starting dates for the data are January 3rd, 1995, January 2nd, 1997, January 4th, 1999 and January 2nd, 2002, respectively. The closing date is December 29th, 2006

SP 500	$k$	$\theta$	$\sigma^2$	$v(\bar{t})$
1995–2006	5.751066	0.042041	0.087743	0.041898
1997–2006	6.071555	0.039969	0.082144	0.038925
1999–2006	7.057089	0.034299	0.059359	0.034474
2002–2006	7.611455	0.029819	0.050125	0.029942

Table 2. Maximum Likelihood estimates of CIR model using historical daily data of Nikkei 225 Index. The starting dates for the data are January 4th, 1995, January 6th, 1997, January 4th, 1999 and January 4th, 2002, respectively. The closing date is December 29th, 2006

Nikkei 225	$k$	$\theta$	$\sigma^2$	$v(\bar{t})$
1995–2006	4.008421	0.051751	0.086224	0.045755
1997–2006	26.79940	0.044146	0.083093	0.044349
1999–2006	627.2639	0.041893	0.064212	0.042393
2002–2006	85.44733	0.041163	0.064792	0.041298

Table 3. Maximum Likelihood estimates of CIR model using historical daily data of S&P 500 Index from January 3rd, 1995 to December 29th, 2006. The table shows 5-day moving average, 10-day moving average and 15-day moving average for Chaikins Volatility

SP 500	$k$	$\theta$	$\sigma^2$
$n=5$	226.3401	0.057505	0.132672
$n=10$	5.751065	0.042041	0.087743
$n=15$	95.10144	0.038993	0.078978

Table 4. Maximum Likelihood estimates of CIR model using historical daily data of Nikkei 225 Index from January 3rd, 1995 to December 29th, 2006. The table shows 5-day moving average, 10-day moving average and 15-day moving average for Chaikins Volatility

Nikkei 225	$k$	$\theta$	$\sigma^2$
$n=5$	174.3905	0.060898	0.121118
$n=10$	4.008421	0.051751	0.086224
$n=15$	79.60699	0.040406	0.060591

Table 5. Maximum Likelihood estimates of CIR model using historical daily, weekly and monthly data of S&P 500 Index from January 3rd, 1995 to December 29th, 2006

SP 500	$k$	$\theta$	$\sigma^2$
Daily	5.751066	0.042041	0.087743
Weekly	17.58786	0.059539	0.080261
Monthly	299.2399	0.085253	0.060909

Table 6: Maximum Likelihood estimates of CIR model using historical daily, weekly and monthly data of Nikkei 225 Index from January 3rd, 1995 to December 29th, 2006

Nikkei 225	$k$	$\theta$	$\sigma^2$
Daily	4.008421	0.051751	0.086224
Weekly	13.63029	0.053029	0.081287
Monthly	164.2547	0.045228	0.053284

Table 7. The estimated results  $k$  of S&P 500 index on sample periods before and after the day of stock plunge on October 16, 1987. The data from October 16, 1987 to October 26, 1987 are not included so that the fluttering itself does not affect the sampling process

quarters before Oct.16	$k$	quarters after Oct.26	$k$
1	133.61	1	41.823
2	63.106	2	92.779
3	107.88	3	76.940
4	85.051	4	196.08
5	50.758	5	324.93
6	24.601	6	35.162
7	42.492	7	129.60
8	9.5577	8	11.100
9	60.435	9	72.700
10	77.535	10	4.8500

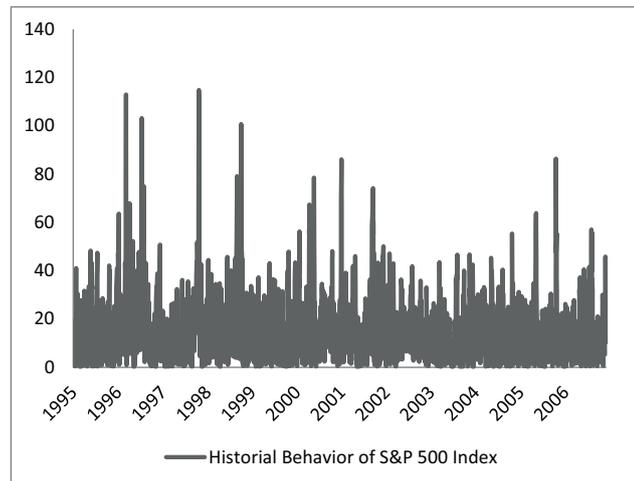


Figure 1. The daily implied volatility of S&P 500 Index over a duration of 12 years

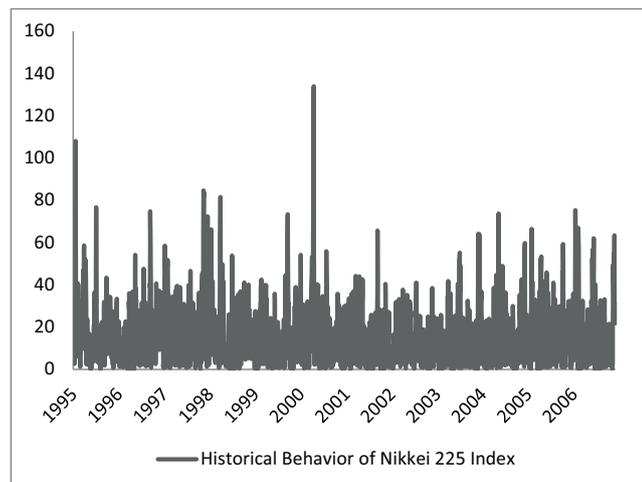


Figure 2. The daily implied volatility of Nikkei 225 Index over a duration of 12 years

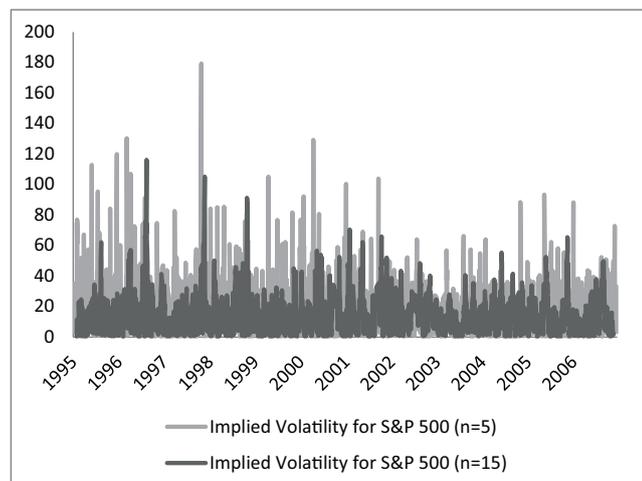


Figure 3. The daily implied volatility of S&P 500 Index over a duration of 12 years, for  $n=5$  and  $n=15$

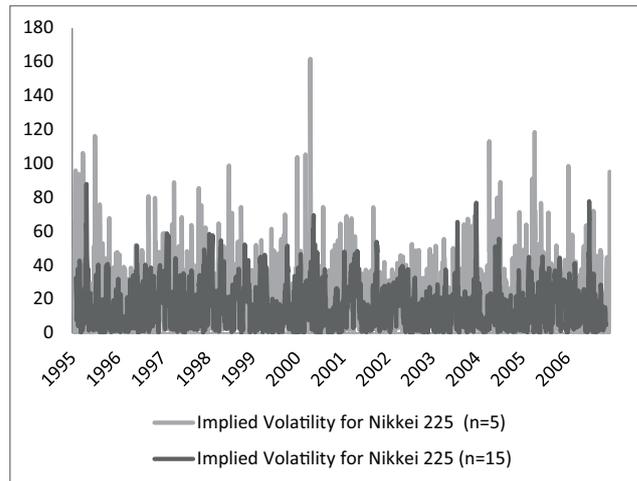


Figure 4. The daily implied volatility of Nikkei 225 Index over a duration of 12 years, for  $n=5$  and  $n=15$

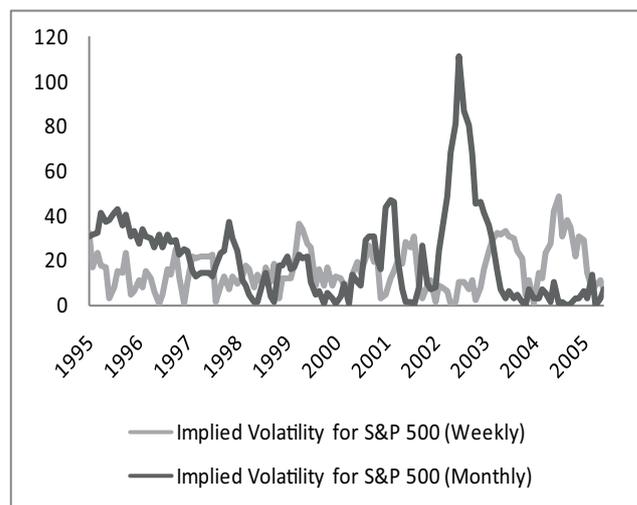


Figure 5. The weekly and monthly implied volatility of S&P 500 Index over a duration of 12 years

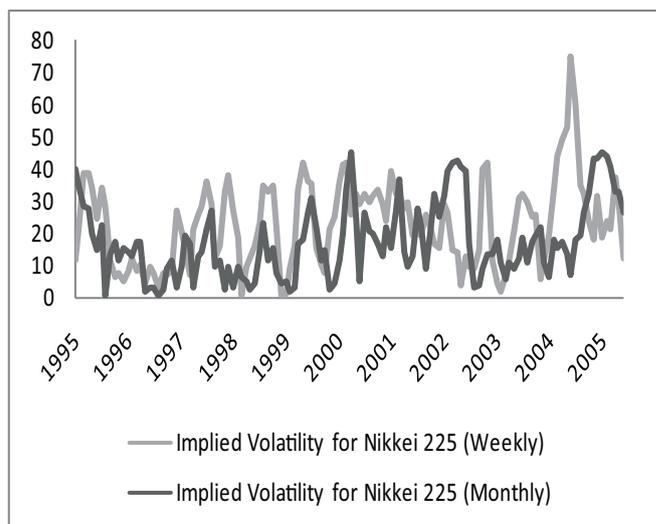


Figure 6. The Weekly and Monthly implied volatility of Nikkei 225 Index over a duration of 12 years

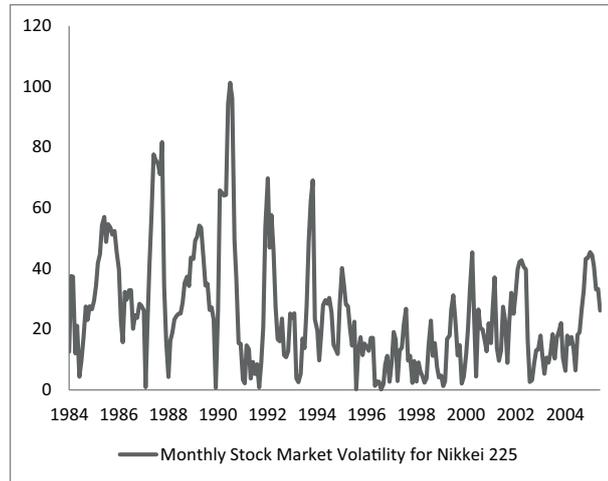


Figure 7. The monthly implied volatility of Nikkei 225 Index over a duration of 2 decades from 1984 to 2006



Figure 8. The estimated value for  $k$  of S&P 500 Index over a duration of 20 quarters before and after the stock crash on October 16, 1987. Specifically, the data set is from April 16, 1985 to April 26, 1990. The x-axis refers to  $n$ th quarter before or after the stock crisis

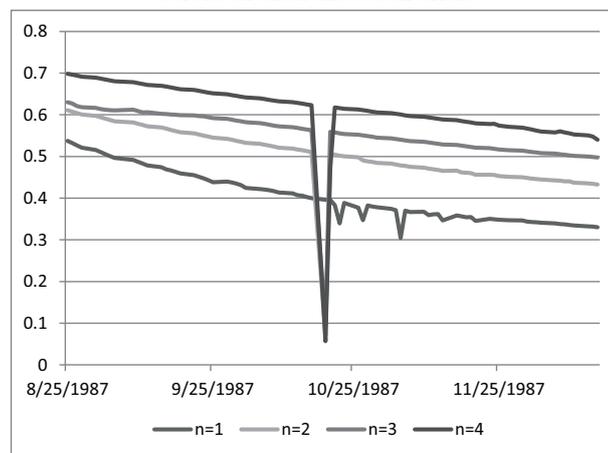


Figure 9. Likelihood-ratio test for Model 1 against Model 2 in Section 4.3. The estimation calculated for  $n = 1, 2, 3, 4$ , where  $n$  represents the quarter before or after the crash. The likelihood ratio jumps dramatically near to zero when the crisis happens, which means Model 2 is better than the null model during the period of the crisis. The graph shows the speed of mean-reversion time is higher after the stock crisis