167348

STOCHASTIC VOLATILITY, A NEW APPROACH FOR VASICEK MODEL WITH STOCHASTIC VOLATILITY

A THESIS SUBMITTED TO
THE INSTITUTE OF APPLIED MATHEMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

SERKAN ZEYTUN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF FINANCIAL MATHEMATICS

SEPTEMBER 2005

Approval of the Institute of Applied Mathematics

Prof. Dr. Ersan AKYILDIZ
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

Prof. Dr. Hayri KÖREZLİOĞLU
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Assoc. Prof. Dr. Azize HAYFAVİ
Supervisor

Examining Committee Members

Prof. Dr. Hayri KÖREZLİOĞLU

Prof. Dr. Ömer GEBİZLİOĞLU

Assoc. Prof. Dr. Azize HAYFAVİ

Assist. Prof. Dr. Kasırga YILDIRAK

Assist. Prof. Dr. Esma GAYGISIZ

Abstract

STOCHASTIC VOLATILITY, A NEW APPROACH FOR VASICEK MODEL WITH STOCHASTIC VOLATILITY

Zeytun, Serkan
M.Sc., Department of Financial Mathematics
Supervisor: Assoc. Prof. Dr. Azize Hayfavi

September 2005, 49 pages

In the original Vasicek model interest rates are calculated assuming that volatility remains constant over the period of analysis. In this study, we constructed a stochastic volatility model for interest rates. In our model we assumed not only that interest rate process but also the volatility process for interest rates follows the mean-reverting Vasicek model. We derived the density function for the stochastic element of the interest rate process and reduced this density function to a series form. The parameters of our model were estimated by using the method of moments. Finally, we tested the performance of our model using the data of interest rates in Turkey.

Keywords: Stochastic volatility, ARCH processes, Vasicek model, distribution function, parameter estimation.

Öz

STOKASTİK VOLATİLİTE, STOKASTİK VOLATİLİTELİ VASICEK MODELİ İÇİN YENİ BİR YAKLAŞIM

Zeytun, Serkan
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi: Doç. Dr. Azize Hayfavi

Eylül 2005, 49 sayfa

Vasicek modelinde volatilitenin analiz dönemi boyunca sabit kaldığı düşünülerek faiz haddi hesaplanmaktadır. Bu çalışmada faiz haddi için bir stokastik volatilite modeli kurduk. Modelimizde sadece faiz haddinin değilde, volatilitenin de Vasicek modelini takip ettiğini düşündük. Faiz haddi sürecinin stokastik kısmı için yoğunluk fonksiyonunu çıkardık ve daha sonra bu yoğunluk fonksiyonunu seri formuna indirgedik. Momentler metodunu kullanarak modelimizim parametrelerini tahmin ettik. En son olarak Türkiye'nin faiz oranları verisini kullanarak modelimizin performansını test ettik.

Anahtar Kelimeler: Stokastik volatilite, ARCH modelleri, Vasicek modeli, yoğunluk fonksiyonu, parametre tahmini.

To my family

ACKNOWLEDGMENTS

I would like to thank Prof. Dr. Hayri Körezlioğlu and Assoc. Prof. Dr. Azize Hayfavi for guiding, motivating and encouraging me throughout this study.

I would also like to thank my parents for their support. Without their encouragement this thesis would not have been possible.

I am grateful to Prof. Dr. İ. Ferit Öktem for his great assistance in the mathematical framework, and Assist. Prof. Dr. Kasırga Yıldırak for his important suggestions throughout this study.

I am indebted to Yeliz Yolcu, and Çetin Dişibüyük for their valuable helps.

I would also like to thank Ayşegül İscanoğlu, Zehra Ekşi, İrem Yıldırım and Oktay Sürücü for their support.

Finally, I would like to thank all my friends and the members of the Institute of Applied Mathematics for their support.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

In the stochastic volatility models volatility changes randomly according to some stochastic differential equations or some discrete processes. These models have been of growing interest in the last two decades and have been used by many authors. In this work, we aimed to improve the Vasicek model using stochastic volatility. In the original Vasicek model interest rates are calculated using constant volatility. We assumed that not only interest rates but also the volatility process follows the mean-reverting Vasicek model.

This work is organized as follows: In the following chapter, we introduced some important studies on nonconstant volatility models especially stochastic volatility models. In Chapter 3, we gave some information about stochastic processes and methods of parameter estimation. Nonconstant volatility models were introduced in Chapter 4. In this chapter, we focused on stochastic volatility models and ARCH processes. Our works on Vasicek model constitute Chapter 5. In this chapter, we constructed a model with stochastic volatility where also the volatility process follows the Vasick model and then we discretisized this model. We derived the density function for the stochastic element of the interest rate process and then we reduced this density function to a series form. In Chapter 6, we estimated the parameters of our model using the method of moments. We found the parameter values of our model using the data of interest rates in Turkey for the period June, 1, 2001 to June, 1, 2004. Parameters of ARCH and GARCH models were found using the same data and then we tested the out of sample performance of our model against ARCH and GARCH models. The conclusions of this study are presented in Chapter 7.

Chapter 2

LITERATURE REVIEW

There has been a growing interest in time series models of nonconstant volatility since the paper "the pricing of options and corporate liabilities" of Black and Scholes (1973) was published. This paper specifies the first successful option pricing formula and describes a general framework for pricing other derivative instruments. Black and Scholes in their pricing formula assume the volatility of the underlying stock price remains constant over the period of analysis. In the subsequent years empirical analysis of stock volatility have shown that the volatility does not seem to be constant. To get more realistic models many authors have used nonconstant volatility in their models. Among these models stochastic volatility models have obtained great popularity especially in the framework of option pricing.

Geske (1979), assuming volatility was not constant, derived the price valuation equation for a call as a compound option in continuous time. He considered the variance of the rate of return on the stock as a function of the level of the stock price. In this study, it was shown that the stock's return variance is monotonically increasing with leverage in the compound option model. Thus, this model corrects some important biases of the Black-Scholes model.

Engle (1982) introduced a class of stochastic processes called autoregressive conditional heteroscedastic (ARCH) processes. These processes have nonconstant variances conditional on the past, but constant unconditional variances. In these processes the variance is described as a linear function of the recent past values of the squared errors. Engle derived the likelihood function of these processes and described the maximum likelihood estimators.

Bollerslev (1986) extended the Engle's ARCH processes allowing for a much more flexible lag structure. Bollerslev's Generalized ARCH (GARCH) processes allow the conditional variance to be dependent upon not only previous squared errors but also previous own lags. Bollerslev derived the conditions for stationarity of this class of processes. He also discussed maximum likelihood estimation of the linear regression model with GARCH errors.

Wiggins (1987) studied the call option valuation problem assuming return volatility follows a continuous stochastic process. In this study, the stochastic volatility valuation problem was described and numerically solved. Using a method of moments approach, statistical estimators for volatility process parameters were then derived. The empirical study of Wiggins showed that the Black-Scholes formula overvalues out-of-the-money calls in relation to in-the-money calls.

Hull and White (1987) in their study concentrated on the problem of pricing a European call option on a stock with a stochastic volatility. They assumed the case of dependence of stochastic volatility and stock price, and the case of independence. For the case in which the stochastic volatility is uncorrelated with the security price the option price was determined in series form using Taylor expansion. They used numerical methods to find the option price for the case in which the volatility and stock price are correlated. They compared the results with the results of Black-Scholes formula and they found that frequently the Black-Scholes price overvalues options.

Johnson and Shanno (1987) in their study used stochastically changing variance and they focused on pricing a call option. To find the prices they used Monte Carlo method. Simulation results showed that call prices are changing with the correlation coefficient between the volatility and the stock price.

Scott(1987) studied on pricing of European call options on stocks by considering a model that allows the variance rate to change randomly according to an independent diffusion process. He developed techniques for estimating parameters of the variance process. Scott could not develop an analytical formula that gives option prices but he derived a model that can produce accurate estimates of option prices via the method of Monte Carlo simulations.

Melino and Turnbull (1990) focused on obtaining a closer correspondence to the empirical distribution of exchange rate and on the subsequent consequences for option pricing. They used a diffusion model for exchange rates with stochastic volatility. The parameters of the model were estimated and then the estimates are used to price foreign currency options. With this study they concluded that the stochastic volatility model dominates the standard model which assumes a log-normal probability distribution for exchange rates and a constant volatility.

Stein and Stein (1991) studied on stock price processes with stochastically varying volatility parameter. They assumed the volatility is driven by an arithmetic Ornstein-Uhlenbeck process, which raises the possibility that σ can be negative. Assuming volatility is uncorrelated with the asset price, an exact closed-form solution for the stock price distribution was derived. They also used analytic techniques to develop an approximation to the distribution. Then, they used their results to develop closed form option prising formulas, and to sketch some links between stochastic volatility and the nature of fat tails in stock price distributions.

Heston (1993) used characteristic functions to derive a closed-form solution for the price of a European call option on an asset with stochastic volatility. He adapted the model to incorporate stochastic interest rates and showed how to apply the model to bond options and currency options. Heston assumed the spot asset's price is correlated with the volatility and concluded that correlation between the spot asset's price and the volatility is important for explaining return skewness and strike-price biases in the Black-Scholes model.

Hobson and Rogers (1998) proposed a class of nonconstant volatility models, which can be extended to include the single-factor stochastic volatility models, but also share many characteristics with the multi-factor models. They defined the instantaneous volatility in terms of exponentially weighted moments of historical log-price. Therefore the instantaneous volatility is driven by the same stochastic factors as the price of the process and an additional source of randomness is not necessary. In this study, a partial differential equation for the price of a European call option was derived and the existence of skews and smiles via numerical solution of this partial differential equation were demonstrated.

CHAPTER 3

PRELIMINARIES

3.1 STOCHASTIC PROCESSES

A **stochastic process** is a collection of random variables $\{X_t, t \in T\}$ on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The points of the *index* or *parameter* set T are thought of representing time. If T is countable, especially if $T = 0, 1, 2, 3, ... \equiv \mathbb{N}$, i.e. the set of non-negative integers, then the process is called **discrete parameter process.** If $T = \mathbb{R}$ or $T = [a, b]$ for a and b real numbers or $T = [0, \infty)$, i.e. if T is uncountable, then we have a **continuous parameter process.**

In a stochastic process $\{X_t, t \in T\}$ the relationship among the random variables of the process, say $X_{t_1}, ..., X_{t_n}$ for $t_1, ..., t_n \in T$, is specified by the joint distribution function of these variables given by

$$F_{t_1,....,t_n}(x_1, ..., x_n) = \mathbb{P}(X_{t_1} \leq x_1, ..., X_{t_n} \leq x_n).$$

The collection of all disjoint distributions $F_{t_1,....,t_n}(x_1, ..., x_n)$ of the random variables $X_{t_1}, ..., X_{t_n}$, $n \geq 1$, $t_1, ..., t_n \in T$, is called **finite-dimensional distributions** of the process.

We say that a stochastic process $\{X_t\}$ is **stationary** if for any $n \geq 1$ and $t_1, ..., t_n \in T$, its finite-dimensional distributions $F_{s+t_1,....,s+t_n}$ are independent of s (we assume all $s + t_k \in T$). This means that the distributional (or statistical) properties of the process remain unchanged as time elapses.

Definition 3.1. Brownian Motion or Wiener Process:
Let $\{W_t, t \geq 0\}$ be a stochastic process defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

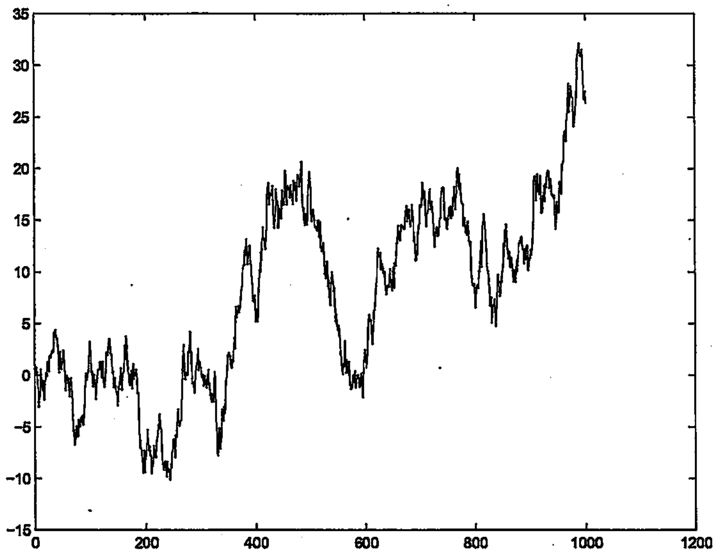

Figure 3.1: *A sample path of standard Brownian motion*

The process $\{W_t\}$ is called a **standard Brownian motion** if

- $W_0(x) = 0$ almost surely (a.s.) (or with probability 1), i.e. we assume the process starts at 0.
- For each $x \in \Omega$, $W_t(x)$ is continuous in t, for $t \geq 0$.
- For all $0 \leq s < t$, the increment $W_t - W_s$ is normally distributed with mean 0 and variance $\sigma^2 = (t - s)$, independent of time t, i.e. it has stationary (normally distributed) increments.
- The increments of the process $W_{t_i} - W_{s_i}$ over intervals $(s_i, t_i]$ are independent.

3.2 DISCRETIZATION

To draw inference about a continuous time model we have to rely on N discrete realizations. Thus, we divide the sample period $[0, T]$ into N intervals corresponding to the discrete-time data (generally equally spaced observations are used but this can be relaxed). Then the continuous time process is replaced with a piecewise-constant process and in the each interval $[t_i, t_{i+1})$, $i = 1, 2, ..., N-1$, it is assumed that the process is constant but from one interval to the next it is changing.

In this part we will present two discretization methods for an Ito process: Euler scheme and Milstein scheme (our main reference for these models is [25]).

3.2.1 Euler Scheme

Consider the process $X = \{X_t, t_0 \leq t \leq T\}$ satisfying the following differential equation:

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t$$

on $t_0 \leq t \leq T$ with the initial value

$$X_{t_0} = X_0.$$

Then, if we divide the time interval $[t_0, T]$ into N intervals where $t_0 < t_1 < ..... < t_n = T$, the Euler scheme is

$$Y_{n+1} = Y_n + a(t_n, Y_n)(t_{n+1} - t_n) + b(t_n, Y_n)(W_{t_{n+1}} - W_{t_n}) \tag{3.2.1}$$

for $n = 0, 1, 2, ..., N-1$ with initial value

$$Y_0 = X_0.$$

If the intervals are equally spaced then

$$\Delta = t_n - t_{n-1} = \frac{T - t_0}{N}$$

and, (3.2.1) can be written in the form

$$Y_{n+1} = Y_n + a\Delta + b\Delta W_n$$

where $\Delta W_n = W_{t_{n+1}} - W_{t_n}$.

3.2.2 Milstein Scheme

Consider the process $X = \{X_t, t_0 \leq t \leq T\}$ satisfying

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t.$$

Then, the Milstein scheme is

$$\begin{aligned} Y_{n+1} &= Y_n + a(t_n, Y_n)(t_{n+1} - t_n) + b(t_n, Y_n)(W_{t_{n+1}} - W_{t_n}) \\ &\quad + \frac{1}{2} b(t_n, Y_n) b'(t_n, Y_n)\{(W_{t_{n+1}} - W_{t_n})^2 - (t_{n+1} - t_n)\} \end{aligned}$$

and in the case of equally spaced intervals

$$Y_{n+1} = Y_n + a\Delta + b\Delta W_n + \frac{1}{2} bb'\{(\Delta W_n)^2 - \Delta\}.$$

The Minstein scheme is equal to the Euler Scheme with additional term

$$\frac{1}{2} bb'\{(\Delta W_n)^2 - \Delta\}.$$

3.3 MOMENTS OF RANDOM VARIABLES

3.3.1 Moments

Definition 3.2. Let X be a random variable with probability density function $f(x)$ (or with probability distribution p(x)). The r**th moment about the origin** of X, denoted by μ'_r, is the expected value of X^r; symbolically

$$\mu'_r = E(X^r) = \sum_x x^r p(x)$$

for $r = 0, 1, 2, ...$ provided $\sum_x |x|^r p(x) < \infty$ when X is discrete, and

$$\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx$$

provided $\int_{-\infty}^{\infty} |x|^r f(x) dx < \infty$ when X is continuous.

Observe that the value of any moment depends only on the probability density function (or probability distribution) of the random variable.
Note that first moment of a random variable about the origin is the mean (expected value) of the random variable.

Definition 3.3. The **rth moment about the mean** of a random variable X, denoted by μ_r, is the expected value of $(X - \mu)^r$; symbolically

$$\mu_r = E[(X - \mu)^r] = \sum_x (x - \mu)^r p(x)$$

for $r = 0, 1, 2, ...$ when X is discrete, and

$$\mu_r = E[(X - \mu)^r] = \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx$$

when X is continuous, provided the finiteness conditions.

Note that the second moment of a random variable about the mean is the variance of the random variable.
Note also that if the kth moment of a random variable exists, all moments of order less than k exist.

Comment: We can express μ_r in terms of μ'_j, $j = 1, 2, ..., r$, by simply writing out the binomial expansion of $(X - \mu)^r$:

$$\mu_r = E[(X - \mu)^r] = \sum_{j=0}^{r} \binom{r}{j} E(X^j)(-\mu)^{r-j}.$$

3.3.2 Moment-Generating Functions

Finding moments of random variables directly, especially for the higher moments, can be quite problematic. Moment-generating functions provide an easier way to calculate moments of random variables.

Definition 3.4. Let X be a random variable. The **moment-generating function (mgf)** for X is denoted $M_X(t)$ and given by

$$M_X(t) = E(e^{tx}) = \begin{cases} \sum_{allx} e^{tx}p(x) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} e^{tx}f(x)dx & \text{if } X \text{ is continuous} \end{cases}$$

at all values of t for which the expected value exists.

Note that the moment generating function is not defined for all random variables and, when it is, $M_X(t)$ is not necessarily finite for all t.

Finding Moments Using Moment-Generating Functions

Theorem 3.1. Let X be a random variable with probability density function $f_X(x)$. [If X is continuous, $f_X(x)$ must be sufficiently smooth to allow the order of differentiation and integration to be interchanged.] Let $M_X(t)$ be the moment-generating function for X. Then, provided the rth moment exists,

$$M_X^{(r)}(0) = E(X^r) = \mu_r'$$

where $M_X^{(r)}(t)$ denotes the rth derivative of M_X at t.

3.4 PARAMETER ESTIMATION METHODS

Almost all econometric models contain unknown parameters. To use an econometric model first we need to estimate these parameters. Typically, this is done by taking a random sample of observations and using these observations to estimate the unknown parameters. In the estimation the idea is that the sample represents the population from which it has been drawn.

For the same parameter of a population we can apply different methods of estimation, so there can be many different estimators of the same parameter. In this part we will present two methods of parameter estimation, the method of maximum likelihood and the method of moments.

3.4.1 The Method of Maximum Likelihood

This estimation method assumes that the distribution of an observed phenomenon is known, except for a finite number of unknown parameters. Then, the unknown parameters will be estimated by looking at the sample values and then choosing our estimates of the unknown parameters the values for which the probability of getting the sample values is a maximum.

Let the probability density function (pdf) for a random variable y, conditioned on a set of parameters, θ, be denoted by $f(y|\theta)$. Assume the observed sample values are $y_1, y_2, ..., y_n$, then the probability of getting them is $f(y_1, y_2, ..., y_n|\theta)$ which is the joint probability density of the entire sample. If the observations are independent and identically distributed (i.i.d.) then the joint density is the product of the individual densities;

$$f(y_1, y_2, ..., y_n|\theta) = \prod_{i=1}^{n} f(y_i|\theta).$$

The likelihood function for the sample data is then given by

$$L(\theta|y) = \prod_{i=1}^{n} L_i(\theta|y_i) = \prod_{i=1}^{n} f(y_i|\theta)$$

which is defined as a function of the unknown parameters, θ, where y is the collection of sample data and $L_i(\theta|y_i)$ are the individual likelihood contributions to the likelihood function $L(\theta|y)$. Then the logarithm of the likelihood function is

$$\log L(\theta|y) = \sum_{i=1}^{n} \log L_i(\theta|y_i) = \sum_{i=1}^{n} \log f(y_i|\theta).$$

Since logarithm is a monotonic function then whatever values of θ maximizes the log-likelihood function must also maximizes the likelihood function. Working with the log-likelihood function is usually simpler and it is used instead of likelihood function.

Since maximum likelihood estimation consist in maximizing the likelihood function (or log-likelihood function) with respect to θ then the necessary condition is

$$\frac{\partial(\log L(\theta|y))}{\partial\theta} = \sum_{i=1}^{n} \frac{\partial(\log L_i(\theta|y_i))}{\partial\theta} = 0. \tag{3.4.2}$$

This is called the likelihood equation and the maximum likelihood estimation is a root of this equation. Since we are searching for a maximum, we also have to satisfy the following condition:

$$\frac{\partial^2(\log L(\theta|y))}{\partial\theta^2} = \sum_{i=1}^{n} \frac{\partial^2(\log L_i(\theta|y_i))}{\partial\theta^2} < 0. \tag{3.4.3}$$

Then the parameters satisfying (3.4.2) and (3.4.3) are the maximum likelihood estimations.

3.4.2 The Method of Moments

Method of moments is historically one of the oldest methods. In this method population means are replaced by sample means. When the underlying model has multiple parameters the method of moments is often more tractable than the method of maximum likelihood.

In the method of moments the first few moments of a population are equated to the corresponding moments of a sample, then these equations are solved for the unknown parameters of the population. The number of moments which are needed depends on the number of unknown parameters.

Suppose X is a random variable and its probability density function is a function of k unknown parameters, $\theta_1, \theta_2, ..., \theta_k$. Then the first k moments of X, if they exist, are given by

$$\mu'_r = E(X^r) \qquad r = 1, 2, ..., k.$$

If we have a set of observations $x_1, x_2, ..., x_n$ then first k moments of X are equated to the corresponding sample moments; symbolically,

$$\begin{aligned} E(X) &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ E(X^2) &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 \\ .... &= .... \\ .... &= .... \\ E(X^k) &= \frac{1}{n}\sum_{i=1}^{n} x_i^k. \end{aligned}$$

Solving these k equations for the unknown parameters $\theta_1, \theta_2, ..., \theta_k$, we can get a set of estimates, $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$.

CHAPTER 4

VOLATILITY

4.1 NONCONSTANT VOLATILITY MODELS

One of the most striking development in financial economics is the work on option pricing of Black and Scholes (1973). Black and Scholes in their pricing formula for stock options assume that the price $(S_t)_{t\leq T}$ of a stock is the solution to a stochastic differential equation (SDE)

$$dS_t = S_t(\mu dt + \sigma dB_t)$$

where σ is a known and constant volatility parameter and B is a Brownian motion.

In the subsequent years, empirical analysis of stock volatility have shown that the volatility is not constant. The volatility cannot be observed directly since it is not traded. However from the empirical studies of the stock price we can drive the stock price return by dS/S, and from this estimate the volatility. In these observations, in general, the volatility is low for several days, then high for a period and so on, that is it changes in clusters. For this reason a number of authors have constructed models of changing volatility. Among these models the most widespread models are stochastic volatility models which define the volatility as an autonomous diffusion driven by a second Brownian motion.

4.1.1 STOCHASTIC VOLATILITY MODELS

Stochastic volatility models have been widely used in the mathematical option pricing literature, where such models provide a natural relaxtion of the Black-Scholes assumption that volatility is fixed, throughout the life of the option. The assumption of the constant volatility is the most disputatious assumption of the Black-Scholes model. Excess kurtosis in financial time series, leverage effects, and the smiles and skew patterns in implied volatilities all contradict the assumption of constant volatility [20]. The study of Belledin and Schlag (1999) tests various stochastic volatility models empirically, and finds these models superior in terms of pricing performance to the Black-Scholes model. However, estimation of these models' parameters has some complexities. Therefore, these models have not been popular in empirical discrete-time financial applications.

In a stochastic volatility model the volatility is changes randomly according to some stochastic differential equations or some discrete processes. Stochastic volatility models are divide into two broad classes: "single-factor" and "multi-factor" models. In the single-factor models the original Brownian motion B_t is the only source of randomness, and in the multi-factor models further Brownian motions or other random elements are introduced.

4.1.1.A SINGLE-FACTOR MODELS

In the single-factor models the volatility is a deterministic function of present or past values of the underlying price. In these models, the volatility and asset price are perfectly correlated, therefore we still have only one source of randomness, and this is an apparent advantage of single on multi-factor models. The dependence of the volatility to the past or present value of the underlying price makes the arithmetic more challenging and commonly precludes the existence of a closed-form solution. However, the arbitrage argument based on portfolio replication and a complete market still goes through unchanged [7].

The simplest single-factor model is the so-called **level-dependent volatility**. In the level-dependent volatility the underlying price S_t satisfies

$$dS_t = \mu S_t dt + \sigma(S_t, t) S_t dB_t$$

where B is a Brownian motion and $\sigma(S_t, t)$ is a deterministic function of the current value S_t and time. For example, in their Constant Elasticity of Variance (CEV) model, Cox and Ross (1976) take $\sigma(S_t) = (S_t)^{(\alpha-1)}$, $\alpha \in (0,1)$.

Empirical analysis have shown that level dependent volatility still fails to price derivative securities better than does the usual Black-Scholes model. This is the conclusion of the study by Dumas, Fleming and Whaley (1998) who suggest a volatility related to past changes of the underlying prices, as the last candidate model before resorting the fully stochastic volatility [31].

4.1.1.B MULTI-FACTOR MODELS

Multi-factor models define the volatility as an autonomous diffusion driven by a second Brownian motion (the asset price process is driven by the first Brownian motions). In the multi-factor models the stock price S and volatility σ satisfy the following stochastic differential equations:

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma_t S_t dW_{1t} \\ d\sigma_t &= a(S_t, \sigma_t)dt + b(S_t, \sigma_t)dW_{2t} \end{aligned}$$

where a, b define the volatility model and a govern the drift of the volatility process. W_{1t}, W_{2t} are standard Brownian motions. The Brownian motions have constant correlation $EdW_{1t}dW_{2t} = \rho dt$, so movements of volatility are possibly correlated with movements of underlying asset price. The Brownian motions have the relation

$$W_{2t} = \rho W_{1t} + \rho' W'_{1t}$$

where W'_{1t} is a Brownian motion independent of W_{1t} (that is W_{1t} and W'_{1t} have no correlation) and $\rho' = \sqrt{1-\rho^2}$.

In this part we will introduce some of the multi-factor models.

Hull-White Model:

In their model, Hull and White considered a derivative asset f with a price that depends upon some security price, S, and its instantaneous variance, $V_t = \sigma_t^2$, which are assumed to obey the following stochastic processes:

$$\begin{aligned} dS_t &= \phi S_t dt + \sigma_t S_t dW_t \\ dV_t &= \mu V_t dt + \xi V_t dZ_t. \end{aligned}$$

In this model, the variable ϕ is a parameter that may depend on S, σ and t. It is assumed that the variables μ and ξ may depend on σ and t but they do not depend on S. The Wienner processes (Brownian motions) Z_t and W_t have correlation ρ. Hull and White in their studies analyzed the model for the cases $\rho = 0$ and $\rho \neq 0$. The actual process that a stochastic variance follows is probably fairly complex. It cannot take on negative values, so the instantaneous standard deviation must approach zero as σ^2 tends to zero. In this model, since S and σ^2 are the only state variables affecting the price of the derivative security, f, the risk-free rate must be constant or at least deterministic [23].

Using this model Hull and White have priced a European call on an asset that has a stochastic volatility. The option price is determined in series form for the case in which the stochastic volatility is independent of the stock price, i.e. for the case $\rho = 0$. Also, numerical solutions are produced for the case in which the volatility is correlated with the stock price, i.e. for $\rho \neq 0$. In their empirical analysis, Hull and White found that Black-Scholes price frequently overprices options and that time to maturity affects the degree of overpricing with the same direction.

Cox-Ingersoll-Ross (CIR) Model:

Cox-Ingersoll-Ross model is a mean-reverting model. The term "mean-reverting" refers to the characteristic (typical) time it takes for a process to get back to the mean level of its invariant distribution (the long-run distribution of the process).

CIR model considers a stock, whose price S_t, as a function of time t, and its volatility σ_t satisfying the following stochastic differential equations:

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma_t S_t dW_{1t} \\ dV_t &= (a + bV_t)dt + c\sqrt{V_t}dW_{2t} \end{aligned}$$

where $V_t = \sigma^2$ is the variance. Here σ_t is the time dependent volatility and c is a parameter that we call the variance noise. In the above stochastic differential equations α is the drift parameter and a, b and c are constants. W_{1t} and W_{2t} are standard Wiener processes. The process W_{1t} and W_{2t} are correlated according to

$$dW_{2t} = \rho dW_{1t} + \sqrt{1-\rho^2}dW'_{1t}$$

where W'_{1t} is a Wiener process independent of W_{1t}, and $\rho \in [-1, 1]$ is the correlation coefficient (a negative correlation is known as the leverage effect).

The V_t process is mean-reverting if $a > 0$ and $b < 0$. In the mean-reverting case V_t tends to revert around a level $-a/b$ with a reversion rate $-b$. In the CIR model V_t has a non-central chi-squared distribution and the expectation and variance are given by

$$\begin{aligned} E(V(t)|V(0) = y) &= -\frac{a}{b} + (y + \frac{a}{b})e^{-|b|t} \\ Var(V(t)|V(0) = y) &= \frac{ac^2}{2b^2} - \frac{c^2}{b}(y + \frac{a}{b})e^{-|b|t} + \frac{c^2}{b}(y + \frac{a}{2b})e^{-|b|t} \end{aligned}$$

and, from the above equations one can find that the limiting distribution of V_t is a gamma distribution with expectation $-a/b$ and variance $ac^2/2b^2$ [2].

Log Ornstein-Uhlenbeck Model:

In the log Ournstein-Uhlenbeck model the stock price process and the volatility are modelled as

$$\begin{aligned} dS_t &= \alpha S_t dt + \sigma_t S_t dW_{1t} \\ dY_t &= (a + bY_t)dt + cdW_{2t} \end{aligned}$$

where a, b, c and α are constants. The volatility is given by $\sigma_t = e^{Y_t}$ which implies $Y_t = \log \sigma_t$. The Wiener processes W_1 and W_2 are correlated according to $dW_{2t} = \rho dW_{1t} + \sqrt{1-\rho^2} dW_{3t}$ where W_{3t} is a Wiener process independent of W_{1t}.

In this model Y_t has normal distribution with expectation

$$E[Y_t | Y_0 = y] = -\frac{a}{b} + (y + \frac{a}{b}) e^{-|b|t}$$

and variance

$$Var[Y_t | Y_0 = y] = \frac{c^2}{2|b|} + (1 - e^{-2|b|t}).$$

Therefore, the limiting distribution for Y_t is a normal distribution with mean $-a/b$ and variance $c^2/2|b|$ [2].

Johnson and Shanno's Model:

In this model, stock price, S, and variance, σ^2, satisfy the following stochastic processes:

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma_t S_t^{\alpha} dW_{1t} \\ d\sigma_t &= \xi \sigma_t dt + \phi \sigma_t^{\beta} dW_{2t} \end{aligned}$$

where μ, ξ, ϕ are constants, $\alpha \geq 0$ and $\beta \geq 0$. W_{1t} and W_{2t} are Wiener processes with correlation coefficient ρ, i.e., they satisfy

$$dW_{2t} = \rho dW_{1t} + \sqrt{1-\rho^2} dW'_{1t}$$

where W'_{1t} is a Wiener process independent of W_{1t}, and $\rho \in [-1, 1]$.

Using this model, Johnson and Shanno generated call prices by using the Monte Carlo method.

Heston Model:

In the Heston model, the stock price return process and the the variance process are modelled as

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{V_t} S_t dW_{1t} \\ dV_t &= \kappa(\theta - V_t)dt + \varepsilon \sqrt{V_t} dW_{2t} \end{aligned}$$

where μ, κ, ϕ, ε are constants and W_{1t}, W_{2t} are Wiener processes with correlation ρ. Here $\sigma_t = \sqrt{V_t}$ is the volatility of the stock price return process. The variance process is a mean-reverting process with long-run mean of θ and with mean-reversion speed determined by κ.

Using this model, Heston used a new technique, based on characteristic functions, to derive a closed-form solution for the price of a European call option on an asset.

4.1.2 ARCH PROCESSES

Autoregressive Conditional Heteroscedasticity (ARCH) Models

ARCH models were first introduced by Engle (1982) and this models are used to model the conditional variance. The essence of the method picks up on the often observed characteristic that large shocks to the unpredictable component of returns tends to occur in cluster (not necessarily of the same sign), and the histogram of shocks has fatter tails than would be expected if they had been generated from a normal distribution. The key feature is that there seems to be an autoregressive nature to the shocks so Engle's ARCH model allow the conditional variance to vary over time driven by past shocks.

Consider the process

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \qquad |\phi_1| < 1$$

with the error term

$$\varepsilon_t = u_t \sigma_t$$

where u_t has a standard normal distribution. Then the conditional variance of ε_t given ε_{t-1} is

$$Var(\varepsilon_t | \varepsilon_{t-1}) = Var(u_t \sigma_t) = \sigma_t^2 Var(u_t) = \sigma_t^2.$$

Under the ARCH model the "autocorrelation in volatility" is modelled by allowing the conditional variance of the error term, σ_t^2, to depend on the immediately previous value of the squared error

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2.$$

The above model is known as an ARCH(1), since the conditional variance depends on only one lagged squared error.

The generalization of this model can be gotten by including more lags of ε_t. Thus, an ARCH(q) model is:

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots\dots\dots + \alpha_q \varepsilon_{t-q}^2 \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2. \end{aligned}$$

The unconditional variance (which is also the long-run variance here) is denoted

σ^2 and defined as

$$\begin{aligned} E(\sigma_t^2) &= E(\alpha_0) + \alpha_1 E(\varepsilon_{t-1}^2) + ... + \alpha_q E(\varepsilon_{t-q}^2) \\ \text{using } E(\varepsilon_{t-i}^2) = \sigma_{t-i}^2 & \\ &= \alpha_0 + \alpha_1 \sigma_{t-1}^2 + ... + \alpha_q \sigma_{t-q}^2 \\ \Longrightarrow & \\ \sigma^2 &= \alpha_0 + \alpha_1 \sigma^2 + ... + \alpha_q \sigma^2 \\ \sigma^2 &= \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i}. \end{aligned}$$

In the above calculations it is assumed that in the long run the conditional variances are constant and equal to the long-run variance σ^2.

Stationarity of the ARCH(q) model imposes conditions on the α_i coefficients. For σ_t^2 to be nonnegative (because the negative variance is not sensible), whatever the values of ε_{t-i}^2 and σ^2 to be finite and nonnegative, we must have $\alpha_0 \geq 0$, $\alpha_i \geq 0$,for $i = 1, 2, ..., q$, and $0 \leq \sum_{i=1}^{q} \alpha_i < 1$.

In the ARCH models the unconditional distribution of returns has fat tails giving a relatively large probability of outliers relative to the normal distribution. This property was showed by Engle (1982) using kurtosis coefficient, κ. ARCH(1) model has stationary moments of order 2 and 4, if $3a^2 < 1$ and these moments are

$$\begin{aligned} E(\varepsilon_t^2) &= \frac{\alpha_0}{1 - \alpha_1} \\ E(\varepsilon_t^4) &= \frac{3\alpha_0^2}{(1 - \alpha_1)^2} \frac{1 - \alpha_1^2}{1 - 3\alpha_1^2}. \end{aligned}$$

Then

$$\kappa = \frac{E(\varepsilon_t^4)}{(E(\varepsilon_t^2))^2} = \frac{3(1 - \alpha_1^2)}{1 - 3\alpha_1^2} > 3$$

that is κ is the ratio of the fourth moment to the squared second moment, and $\kappa = 3$ for the normal distribution. In the ARCH(1) case $\kappa > 3$ indicating leptokurtosis and, hence, fat tails.

Generalized ARCH (GARCH) Models

In the ARCH models the conditional variance σ_t^2 is taken positive. To provide this, all α_is must be non-negative. But when q large unconstrained estimation will often lead to the violation of the non-negativity constraints on the α_is. Because of this fact, in early applications of the ARCH models, many authors prefered small lag structures. To obtain more flexibility, GARCH models were developed independently by Bollerslev (1986) and Taylor (1986) as an extension of ARCH models.

The GARCH models allow the conditional variance to be dependent upon previous own lags, so that the conditional variance equation in the simplest case is

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2.$$

This model is known as GARCH(1,1) and can be extended to a GARCH(p,q) formulation, where the current conditional variance is parametrised to depend upon q lags of the squared error and p lags of the conditional variance

$$\begin{aligned}\sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + ... + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + ... + \beta_p \sigma_{t-p}^2 \\ &= \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2.\end{aligned}$$

GARCH models better therefore a far more widely used models than ARCH models. Because GARCH models are less likely to breach non-negativity constraints.

The unconditional variance for the GARCH models can be motivated in the same way as for the ARCH models (for the formal proof see Bollerslev (1986, theorem 1)). For the GARCH(p,q) model the unconditional variance is

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^{q} \alpha_i - \sum_{j=1}^{p} \beta_j}. \tag{4.1.1}$$

Stationarity of GARCH(p,q) model imposes conditions on the α_i and β_j coefficients. If $\alpha_0 \geq 0$, $\alpha_i \geq 0$ for $i = 1, 2, ..., q$, $\beta_j \geq 0$ for $j = 1, 2, ..., p$ and $0 \leq \sum_{i=1}^{q} \alpha_i + \sum_{j=1}^{p} \beta_j < 1$ then σ^2 is nonnegative and finite and given by (4.1.1) (due to Bollerslev (1986)).

Parameter Estimation of ARCH-GARCH Models Using Maximum Likelihood Method

As we mentioned in the previous chapter, in the maximum likelihood method the parameters are estimated by taking those values for them that give the observed values the highest probability. For this, first, we need to cunstruct the likelihood function, then we need to find the parameters that maximize this likelihood function.

As an example, we will work on the process

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \sim N(0, \sigma_t^2)$$

with stochastic volatility satisfying GARCH(1,1) model:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2.$$

For simplicity we are taking GARCH(1,1) model, but the parameters of all ARCH-GARCH models can be estimated using the same way.

Since $\varepsilon_t \sim N(0, \sigma_t^2)$, then $y_t \sim N(\phi_0 + \phi_1 y_{t-1}, \sigma_t^2)$. So, the probability density function for the process y_t is

$$f(y_t|\phi_0+\phi_1 y_{t-1},\sigma_t^2) = \frac{1}{\sigma_t\sqrt{2\pi}}\exp(-\frac{1}{2}\frac{(y_t-\phi_0-\phi_1 y_{t-1})^2}{\sigma_t^2}) = \frac{1}{\sigma_t\sqrt{2\pi}}\exp(-\frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2}).$$

Suppose we have observations $y_1, ..., y_T$. Then, the joint probability density function for all ys is the product of the individual density functions, i.e.,

$$f(y_1,....,y_T|\phi_0+\phi_1 y_{t-1},\sigma_t^2) = \prod_{t=1}^{T} f(y_t|\phi_0+\phi_1 y_{t-1},\sigma_t^2) = \prod_{t=1}^{T}\frac{1}{\sigma_t\sqrt{2\pi}}\exp(-\frac{1}{2}\frac{\varepsilon_t^2}{\sigma_t^2}).$$

This is the likelihood function. If we take the logarithm of this function then it will be turned to an additive function of the sample data, and working on the new function will be easier. The logarithm of the likelihood function, i.e. the log-likelihood function, is

$$LLF = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T} log(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{\sigma_t^2}.$$

Maximizing the log-likelihood function for a model with time-varying variance is not easy. Derivatives of log-likelihood function with respect to the parameters are complicated, so a numerical procedure is often used instead to maximize the log-likelihood function.

All methods work by searching over the parameter-space until the values of the parameters that maximize the log-likelihood function are found. These optimization methods exist in some of the packaged softwares such as EViews and RATS. These methods are based on the determination of the first and second derivatives of the log-likelihood function with respect to the parameter values at each iteration (i.e. the gradient and Hessian matrices, respectively). A famous algorithm for optimization is BHHH (Berndt, Hall, Hall and Hausman (1974)), and this algorithm employs only first derivatives and approximations for the second derivatives are calculated [6].

4.2 VASICEK MODEL

Vasicek model is one of the earliest stochastic models of the short term interest rate. Vasicek assumes that the instantaneous interest rate r_t follows the so-called Ornstein-Uhlenbeck process,

$$dr_t = \beta(\alpha - r_t)dt + \sigma dW_t \tag{4.2.2}$$

where α, β, σ are non-negative constants. Here, the parameter α is the long-run normal interest rate. The Vasicek model exhibits mean-reversion. The instantaneous drift $\beta(\alpha - r_t)$ represents a force that keeps pulling the process towards its long-run mean α with magnitude proportional to the deviation of the process from the mean. The coefficient β is the speed of adjustment of the interest rate towards its long-run level. The stochastic element σdW, which has a constant instantaneous variance σ^2 (that is, a variance per unit of time dt), causes the process to fluctuate around the level α in an erratic, but continuous, fashion.

This model has many advantages, but it has also shortcomings. The main advantage of this model is that it has an explicit solution. Since the distribution of r_t is normal then negative interest rates are possible and this is a major shortcoming of this model.

The solution of the stochastic differential equation (4.2.2) can be found as (see [38])

$$r_t = e^{-\beta(t-u)}r_u + \alpha(1 - e^{-\beta(t-u)}) + \sigma \int_u^t e^{-\beta(t-s)}dW_s$$

where $u \leq t$.

The conditional mean and variance of r_t are:

$$
\begin{aligned}
E[r_t|r_u] &= e^{-\beta(t-u)} r_u + \alpha(1 - e^{-\beta(t-u)}) && (4.2.3) \\
&= \alpha + (r_u - \alpha) e^{-\beta(t-u)} && (4.2.4)
\end{aligned}
$$

$$
\begin{aligned}
Var[r_t|r_u] &= E\{(\sigma \int_u^t e^{-\beta(t-u)} dW_u)^2 | r_u\} \\
&= \frac{\sigma^2}{2\beta}(1 - e^{-2\beta(t-u)}).
\end{aligned}
$$

for all $u \leq t$. From the equation (4.2.3) we can say that according to Vasicek model, the conditional expectation of the short rate is a weighted average of the long-run mean and the last period short rate. It is apparent from equation (4.2.4), when the current short rate is above the long-run interest rate then for the future a decrease in short rate is expected, and when it is bellow the long-run interest rate then an increase in the short rate expected. This situation is illustrated in Figure 1. At the Point 1, it is expected that the future short rate will decrease. When the current short rate is equal to long-run mean then it is expected that future short rates remains at this level and such a case is illustrated at the Point 2 [16].

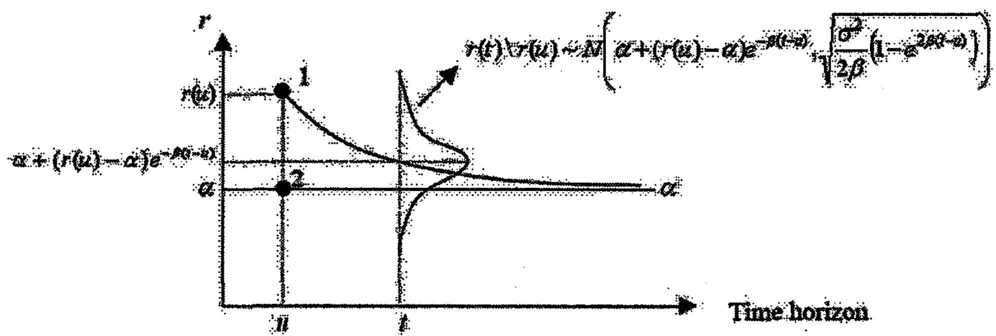

Figure 4.1: *Ornstein-Uhlenbeck process for very short rate.*

4.2.1 STOCHASTIC VOLATILITY VASICEK MODELS

In the stochastic volatility Vasicek models, short term interest rate process follows Vasicek model with stochastic volatility. For the volatility process, different stochastic differential equations can be taken. For example;

Cotton, Fouque, Papanicolau and Sircar [8] in their study used the following model:

$$\begin{aligned} dr_t &= a(\mu - r_t)dt + f(Y(t))dW_t \\ dY_t &= \alpha(m - Y_t) + dB_t \end{aligned}$$

where the stochastic volatility σ_t given by a non-negative function $f(Y_t)$. In this model, the Brownian motions satisfy

$$dB_t = \rho dW_t + \rho' dZ_t$$

where W_t and Z_t are independent standard Brownian motions, i.e., the interest rate and volatility are correlated, and $\rho' = \sqrt{1-\rho^2}$. Using this model, Cotton, Fouque, Papanicolau and Sircar priced bond options.

CHAPTER 5

STOCHASTIC VOLATILITY VASICEK MODEL: A NEW APPROACH

5.1 DERIVATION OF THE MODEL

In this section we constructed a stochastic volatility model using Vasicek model.

Suppose that the short rate interest rate satisfies the mean-reverting Vasicek model with stochastic volatility where volatility also satisfies the mean-reverting Vasicek model, that is

$$dr_t = \phi(b - r_t)dt + \sigma_t dW_{1t} \quad (5.1.1)$$

$$d\sigma_t = \gamma(k - \sigma_t)dt + \varepsilon dW_{2t} \quad (5.1.2)$$

where W_1 and W_2 are two independent Wiener processes and ϕ, b, γ and k are constants. Here ε is the volatility of the volatility process.

Our aim is to estimate the parameters of the above model. In the application part since we will use discrete data, then we need to use a discrete model. So, first we need to discretisize this model. For discretization we will use Euler method.

Applying the Euler method to (5.1.1) and (5.1.2) we get

$$\begin{aligned} r_{t+1} - r_t &= \phi(b - r_t)\Delta t + \sigma_t(\Delta t)^{1/2} Z_{1t} \\ \sigma_{t+1} - \sigma_t &= \gamma(k - \sigma_t)\Delta t + \varepsilon(\Delta t)^{1/2} Z_{2t} \end{aligned}$$

where Z_{1t} and Z_{2t} are independent and identically distributed (i.i.d.) standard normal distributions, i.e. $Z_1, Z_2 \sim N(0,1)$, and ε is constant. If we take the observations daily, i.e. $\Delta t = 1$, and rearrange the equations we get

$$\begin{aligned} r_{t+1} &= \phi b + (1-\phi) r_t + \sigma_t Z_{1t} \\ \sigma_t &= \gamma k + (1-\gamma)\sigma_{t-1} + \varepsilon Z_{2t}. \end{aligned}$$

For simplicity, take $\alpha = \phi b$, $\beta = (1-\phi)$, $a = (1-\gamma)$, $x = \gamma k$ and $\xi_t = \varepsilon Z_{2t}$ where ξ_t is the innovation term. Then

$$r_{t+1} = \alpha + \beta r_t + \sigma_t Z_{1t} \qquad (5.1.3)$$

$$\sigma_t = x + a\sigma_{t-1} + \xi_t \qquad (5.1.4)$$

and we will work on these equations.

The volatility process (5.1.4) can be written in the form

$$\sigma_t = a^t \sum_{k=0}^{t} a^{-k}(\xi_k + x). \qquad (5.1.5)$$

We can show the equality of (5.1.4) and (5.1.5) as follows:

$$\begin{aligned}
\sigma_t &= a^t \sum_{k=0}^{t} a^{-k}(\xi_k + x) \\
&= a^t \sum_{k=0}^{t-1} a^{-k}(\xi_k + x) + \xi_k + x \\
&= \underbrace{a\, a^{t-1} \sum_{k=0}^{t-1} a^{-k}(\xi_k + x)}_{\sigma_{t-1}} + \xi_k + x \\
&= x + a\sigma_{t-1} + \xi_t .
\end{aligned}$$

Since Z_{2t} has the standard normal distribution then $\xi_t = \varepsilon Z_{2t}$ has a normal distribution with zero mean and variance equal to ε^2. Then the expectation of σ_t is

$$\begin{aligned}
E(\sigma_t) &= E(a^t \sum_{k=0}^{t} a^{-k}(\xi_k + x)) \\
&= a^t \sum_{k=0}^{t} a^{-k} E(\xi_k + x) \\
&= a^t \sum_{k=0}^{t} a^{-k}(E(\xi_k) + E(x))
\end{aligned}$$

since ξ_k has zero mean and x is constant then

$$\begin{aligned}
E(\sigma_t) &= a^t \sum_{k=0}^{t} a^{-k} x \\
&= a^t(1 + a^{-1} + a^{-2} + \ldots\ldots + a^{-t})x \\
&= (1 + a^1 + a^2 + \ldots\ldots + a^t)x \\
&= \frac{1 - a^{t+1}}{1 - a} x
\end{aligned}$$

and, the variance of σ_t is

$$\begin{aligned}
Var(\sigma_t) &= E[(\sigma_t - E(\sigma_t))^2] \\
&= E[(a^t \sum_{k=0}^{t} a^{-k}(\xi_k + x) - a^t \sum_{k=0}^{t} a^{-k}x)^2] \\
&= E[(a^t\xi_0 + a^{t-1}\xi_1 + a^{t-2}\xi_2 + ... + a^0\xi_t)^2]
\end{aligned}$$

since Z_ts are i.i.d. then ξ_ts are i.i.d., so $Cov(\xi_i\xi_j) = 0$ for $i \neq j$. Therefore

$$\begin{aligned}
Var(\sigma_t) &= E[(a^t\xi_0)^2] + E[(a^{t-1}\xi_1)^2] + E[(a^{t-2}\xi_2)^2] + ... + E[(a^0\xi_t)^2] \\
&= Var(a^t\xi_0) + Var(a^{t-1}\xi_1) + Var(a^{t-2}\xi_2) + ... + Var(a^0\xi_t) \\
&= a^{2t}Var(\xi_0) + a^{2(t-1)}Var(\xi_1) + a^{2(t-2)}Var(\xi_2) + ... + a^0 Var(\xi_t) \\
&= a^{2t}\varepsilon^2 + a^{2(t-1)}\varepsilon^2 + a^{2(t-2)}\varepsilon^2 + ... + a^0\varepsilon^2 \\
&= [(a^2)^0 + (a^2)^1 + (a^2)^2 + ... + (a^2)^t]\varepsilon^2 \\
&= \frac{1-(a^2)^{t+1}}{1-a^2}\varepsilon^2.
\end{aligned}$$

If σ_t satisfies the equation (5.1.4) then $|a|$ must be less than 1. Because if $|a| > 1$ then the volatility will increase in time and it will tend to infinity. So, the process of σ_t is stationary iff $|a| < 1$. Since $|a| < 1$ then the expectation and variance of σ_t will converge to the asymptotic mean and asymptotic variance, that is

$$\frac{1-a^{t+1}}{1-a}x \longrightarrow \frac{x}{1-a} \qquad \text{as } t \longrightarrow \infty$$

and

$$\frac{1-(a^2)^{t+1}}{1-a^2}\varepsilon^2 \longrightarrow \frac{\varepsilon^2}{1-a^2} \qquad \text{as } t \longrightarrow \infty.$$

To deal with the solution of this model we will use the asymptotic mean and the asymptotic variance as the mean and the variance of the volatility process.

Let

$$\sigma Z = Y.$$

Then for a given $\sigma = s$ the distribution of Y is normal with mean 0 and variance

s^2, i.e. $Y \sim N(0, s^2)$. So

$$f(Y|\sigma = s) = \frac{1}{\sqrt{2\pi}s} \exp(-\frac{y^2}{2s^2})$$

where $f(Y|\sigma = s)$ is the conditional density of Y given $\sigma = s$. Since the conditional distribution of Y and the distribution of σ are Gaussian, and using the conditional density formula

$$f(Y, \sigma) = f(Y|\sigma)f(\sigma)$$

and since

$$f(Y) = \int_{-\infty}^{\infty} f(Y|\sigma)f(\sigma)d\sigma$$

then the density of Y is

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}s} \exp(-\frac{y^2}{2s^2}) \frac{1}{\sqrt{2\pi}\Sigma} \exp(-\frac{(s-M)^2}{2\Sigma^2})ds \qquad (5.1.6)$$

or

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}s} \exp(-\frac{y^2}{2s^2}) \frac{1}{\sqrt{2\pi}\sqrt{\frac{\varepsilon^2}{1-a^2}}} \exp(-\frac{(s-\frac{x}{1-a})^2}{\frac{2\varepsilon^2}{1-a^2}})ds. \qquad (5.1.7)$$

5.2 REPRESENTATION OF THE DENSITY IN A SERIES FORM

Let the integral (5.1.7) be denoted by I. Setting

$$\frac{1}{2}y^2 = m^2, \quad \frac{2\varepsilon^2}{1-a^2} = c^2, \quad \frac{x}{1-a} = b \quad (m \geq 0, \quad c > 0)$$

we have

$$I = \frac{\sqrt{2}}{2\pi c} \int_{-\infty}^{\infty} \frac{1}{s} \exp(-\frac{m^2}{s^2}) \exp[-(\frac{s-b}{c})^2]ds.$$

If we introduce the variable $t = \frac{s}{c}$ and put $\frac{m}{c} = \lambda$, $\frac{b}{c} = \mu$ we get

$$I = \frac{\sqrt{2}}{2\pi c} \int_{-\infty}^{\infty} \frac{1}{t} \exp(-\frac{\lambda^2}{t^2}) \exp[-(t-\mu)^2] dt.$$

As

$$\int_{-\infty}^{\infty} f(t) dt = \int_{0}^{\infty} [f(t) + f(-t)] dt$$

the integral takes the form

$$I = \frac{\sqrt{2}}{\pi c} \exp(-\mu^2) \int_{0}^{\infty} \frac{1}{t} \exp(-t^2 - \frac{\lambda^2}{t^2}) \sinh(2\mu t) dt.$$

Let

$$A(\lambda, \mu) = \int_{0}^{\infty} \frac{1}{t} \exp(-t^2 - \frac{\lambda^2}{t^2}) \sinh(2\mu t) dt.$$

In view of the expansion

$$\frac{1}{t} \sinh(2\mu t) = \sum_{n=0}^{\infty} \frac{(2\mu)^{2n+1}}{(2n+1)!} t^{2n}$$

we have

$$A(\lambda, \mu) = \int_{0}^{\infty} \exp(-t^2 - \frac{\lambda^2}{t^2}) \{ \sum_{n=0}^{\infty} \frac{(2\mu)^{2n+1}}{(2n+1)!} t^{2n} \} dt.$$

Consider the series with general term

$$U_n = \exp(-t^2 - \frac{\lambda^2}{t^2}) \frac{(2\mu)^{2n+1}}{(2n+1)!} t^{2n}.$$

As $U_n \geq 0$, the series can be integrated term by term, giving

$$A(\lambda, \mu) = \sum_{n=0}^{\infty} \frac{(2\mu)^{2n+1}}{(2n+1)!} A_n(\lambda)$$

with

$$A_n(\lambda) = \int_{0}^{\infty} \exp(-t^2 - \frac{\lambda^2}{t^2}) t^{2n} dt \qquad (n = 0, 1, 2, ...).$$

The integral I then becomes

$$I = \frac{\sqrt{2}}{\pi c} \exp(-\mu^2) \sum_{n=0}^{\infty} \frac{(2\mu)^{2n+1}}{(2n+1)!} A_n(\lambda). \qquad (5.2.8)$$

It can be shown, by differentiation under the integral sign, that

$$A_n'(\lambda) = -2\lambda A_{n-1}(\lambda) \qquad (n = 1, 2, 3, ...)$$

with

$$A_0(\lambda) = c \exp(-2\lambda) \qquad (\lambda \geq 0).$$

We thus obtain

$$A_n(\lambda) = \int_0^{\lambda} -2t A_{n-1}(t) dt + C_n$$

with

$$C_n = A_n(0) = \frac{1}{2}\Gamma(n + \frac{1}{2}) = \frac{(2n-1)!!}{2^n} \frac{\sqrt{\pi}}{2} \qquad (-1!! = 1).$$

Knowing $A_0(\lambda)$ and letting succesively $n = 1, 2, 3, ...$ in the above recurence formula all the $A_n(\lambda)$s can be determined. Then using these $A_n(\lambda)$s we can calculate the expression (5.2.8), so we can get the density of $Y = \sigma Z$.

CHAPTER 6

PARAMETER ESTIMATION AND APPLICATIONS

In this chapter, the estimates of our model will be derived using method of moments. These estimates will then be used to establish term structure of interest rates in Turkey.

6.1 PARAMETER ESTIMATION USING METHOD OF MOMENTS

Consider the model we derived in the previous chapter:

$$\begin{aligned} r_t &= \alpha + \beta r_{t-1} + \sigma_{t-1} Z_{1t} \\ \sigma_{t-1} &= x + a\sigma_{t-2} + \xi_{t-1}. \end{aligned}$$

From the first equation we get

$$\sigma_{t-1} Z_{1t} = r_t - \alpha - \beta r_{t-1}. \tag{6.1.1}$$

Taking the expectation of both sides we have

$$E(\sigma_{t-1} Z_{1t}) = E(r_t - \alpha - \beta r_{t-1}).$$

Since any expectation can be written in the form of expectation of conditional expectation and since σ_{t-1} is $\mathcal{F}_{t-1}$ measurable then

$$\begin{aligned} E(\sigma_{t-1}Z_{1t}) &= E[E(\sigma_{t-1}Z_{1t}|\mathcal{F}_{t-1})] \\ &= E[\sigma_{t-1}E(Z_{1t}|\mathcal{F}_{t-1})] \\ &= 0. \end{aligned}$$

Thus,

$$E(r_t - \alpha - \beta r_{t-1}) = 0. \tag{6.1.2}$$

If we take the square of both sides in equation (6.1.1) and then take the expectation of both sides we get

$$E(\sigma_{t-1}^2 Z_{1t}^2) = E[(r_t - \alpha - \beta r_{t-1})^2].$$

Since

$$\begin{aligned} E(\sigma_{t-1}^2 Z_{1t}^2) &= E[E(\sigma_{t-1}^2 Z_{1t}^2|\mathcal{F}_{t-1})] \\ &= E[\sigma_{t-1}^2 E(Z_{1t}^2|\mathcal{F}_{t-1})] \\ &= E(\sigma_{t-1}^2) \end{aligned}$$

then

$$E[(r_t - \alpha - \beta r_{t-1})^2] = E(\sigma_{t-1}^2). \tag{6.1.3}$$

Remark 6.1.1. Let Z be a random variable with standard normal distribution, i.e. $Z \sim N(0,1)$, Then

$$E(Z^t) = \begin{cases} (t-1).(t-3).....5.3.1 & \text{for t even} \\ 0 & \text{for t odd} \end{cases}$$

Taking the third and fourth powers and then the expectations of both sides in (6.1.1) and using the above remark we get

$$E[(r_t - \alpha - \beta r_{t-1})^3] = 0 \quad (6.1.4)$$

$$E[(r_t - \alpha - \beta r_{t-1})^4] = 3E(\sigma_{t-1}^4) \quad (6.1.5)$$

$E(\sigma^2)$ and $E(\sigma^4)$ are the second and fourth moments of σ, respectively. Thus, we can calculate these moments using moment-generating function for σ. From the previous chapter we know that σ has the normal distribution with mean $M = \frac{x}{1-a}$ and variance $\Sigma^2 = \frac{\varepsilon^2}{1-a^2}$. Therefore, the moment-generating function for σ is

$$v(t) = \exp\{Mt + \frac{\Sigma^2 t^2}{2}\}.$$

Using this function we can find that

$$E(\sigma^2) = v^{(2)}(0) = M^2 + \Sigma^2$$

and

$$E(\sigma^4) = v^{(4)}(0) = M^4 + 3\Sigma^4 + 6\Sigma^2 M^2.$$

Thus, we have the following system of equations:

$$(*) \begin{cases} E[(r_t - \alpha - \beta r_{t-1})] = 0 \\ E[(r_t - \alpha - \beta r_{t-1})^2] = M^2 + \Sigma^2 \\ E[(r_t - \alpha - \beta r_{t-1})^3] = 0 \\ E[(r_t - \alpha - \beta r_{t-1})^4] = 3(M^4 + 3\Sigma^4 + 6\Sigma^2 M^2) \end{cases}$$

The solutions of this system of equations are the estimates of our model's parameters.

6.2 APPLICATIONS

Data: The short rate interest rate (1-day) in Turkey for the period June, 1, 2001 to June, 1, 2004. Nominal interest rates were used. Interest rates were derived from the rates of treasury bills using Nelson-Siegel method.

First of all we checked whether the data have autocorrelation or not. For this we applied Ljung-Box Q-statistic and the results for this test are as follows:

	H	**pValue**	**Qstat**	**CriticalValue**
lbqtest	1	0	1.4400e+004	31.4104

Table 6.1: *Ljung-Box Q-Statistic*

In this test the null hypothesis is "H_0=there is no autocorrelation". Since the results give H=1 then we reject the null hypothesis.
Also, we applied the Arch Test to the data and we gave the results of this test in the following table:

	H	**pValue**	**ARCHstat**	**CriticalValue**
archtest	1	0	775.1894	3.8415

Table 6.2: *Arch Test*

Since the results give $H = 1$ then we reject the null hypothesis (H_0 =There is no ARCH effect) of the Arch Test.
From the above tests we conclude that the data have autocorrelation, so it is convenient to work on.

We estimated our model's parameters using the equation system $(*)$. We made constrained optimization using Matlab function "fmincon". We constructed our objective function by getting 4 equations which equal to 0 from the $(*)$ and then summing up the absolute values of these expressions. Our constraints are $|b| < 1$, $M > 0$ and $\Sigma > 0$. We take $M > 0$ because using the series form of the density function of σZ, i.e., from the equation (5.2.8), we can conclude that $x > 0$ which indicates $M > 0$. Results of the constrained optimization are given in the following table:

α	β	M	Σ
8.1299	0.7901	2.7124	1.3879

Table 6.3: *Parameter Values*

The linear regression model which takes volatility constant gives $\alpha = 0.0401$ and $\beta = 0.9976$. Our stochastic volatility model showed that the parameter β is not very close to 1 in contrast to the linear regression model with constant volatility. After filtering the volatility it is obtained that the β parameter is about 0.79 for our data.

Since $M = \frac{x}{1-a}$ and $\Sigma = \frac{\varepsilon}{\sqrt{1-a^2}}$ then

$$0.7901 = \frac{x}{1-a} \Longrightarrow x = 0.79(1-a)$$

and

$$1.3879 = \frac{\varepsilon}{\sqrt{1-a^2}} \Longrightarrow \varepsilon = 1.3879(1-a^2)^{1/2}.$$

So, the volatility process of our model is

$$\sigma_t = 0.79(1-a) + a\sigma_{t-1} + 1.3879(1-a^2)^{1/2} Z_t.$$

By searching for the value of a which minimizes

$$\sum_{t=1}^{T} (\sigma_{observed} - \sigma_{model})^2$$

with constraint $|a| < 1$ we can find $a = 0.9965$. Thus

$$\begin{aligned} x &= 0.79(1-a) = 0.0095 \\ \varepsilon &= 1.3879(1-a^2)^{1/2} = 0.1160 \end{aligned}$$

Therefore, our model for the data of interest rates in Turkey is

$$
\begin{aligned}
r_{t+1} &= 8.1299 + 0.7901 r_t + \sigma_t Z_{1t} \\
\sigma_t &= 0.0095 + 0.9965 \sigma_{t-1} + 0.1160 Z_{2t}.
\end{aligned}
$$

6.2.1 RESULTS FOR ARCH AND GARCH MODELS

We applied ARCH and GARCH models to our data and we gave the results in this section. In the following table we give the Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) values for some ARCH processes.

	ARCH(1)	**ARCH(2)**	**ARCH(3)**	**ARCH(4)**
AIC	5187.8	5188.6	5190.5	5192.5
BIC	5201.7	5207.2	5213.8	5223.2

Table 6.4: *AIC and BIC for ARCH models*

AIC and BIC are used for model order selection. When using either AIC or BIC, models that minimize the criteria are preferred. AIC and BIC support ARCH(1) among ARCH models for our data.
The values of AIC and BIC for some GARCH models are given in the following table.

	G(1,1)	**G(2,1)**	**G(1,2)**	**G(2,2)**	**G(1,3)**	**G(2,3)**	**G(3,3)**
AIC	5188.5	5190.5	5190.5	5192.2	5192.5	5194.4	5196.5
BIC	5207.2	5213.9	5213.9	5220.5	5220.5	5227.2	5233.8

Table 6.5: *AIC and BIC for GARCH models*

AIC and BIC support GARCH(1,1) among GARCH models for our data.
Comparing AIC and BIC values for ARCH and GARCH models we can conclude that ARCH(1) is the most appropriate model for our data.

Using ARCH(1) and GARCH(1,1) models we can find the following parameter values for our data.

	α_0	α_1	β_1
ARCH(1)	0.2657	1	
GARCH(1,1)	0.2107	0.8948	0.1052

Table 6.6: *Parameter values for ARCH(1) and GARCH(1,1)*

So, the ARCH(1) model is

$$\sigma_t^2 = 0.2657 + \varepsilon_{t-1}^2$$

and the GARCH(1,1) model is

$$\sigma_t^2 = 0.2107 + 0.8948\varepsilon_{t-1}^2 + 0.1052\sigma_{t-1}^2.$$

6.2.2 TESTING THE PERFORMANCE OF OUR MODEL

In this section we tested the out of sample performance of our stochastic volatility model against ARCH and GARCH models. We used 10-days volatility forecasts of the models. The forecasts of the models and the observed volatility (we took innovations as the volatility) are given in the following table:

	t=0	**t=1**	**t=2**	**t=3**	**t=4**	**t=5**
SV	18.9445	18.8877	18.8311	18.7747	18.7185	18.6625
ARCH(1)	18.9445	18.9515	18.9585	18.9655	18.9725	18.9795
GARCH(1,1)	18.9445	18.9753	18.9808	18.9863	18.9918	18.9973
OBSERVED	18.9445	18.7956	18.1177	18.2731	18.4042	18.5750
	t=6	**t=7**	**t=8**	**t=9**	**t=10**	
SV	18.6066	18.5510	18.4956	18.4403	18.3853	
ARCH(1)	18.9865	18.9935	19.0005	19.0075	19.0145	
GARCH(1,1)	19.0028	19.0083	19.0138	19.0193	19.0248	
OBSERVED	18.7002	18.8284	18.8387	18.8492	18.6372	

Table 6.7: *Volatility forecasts*

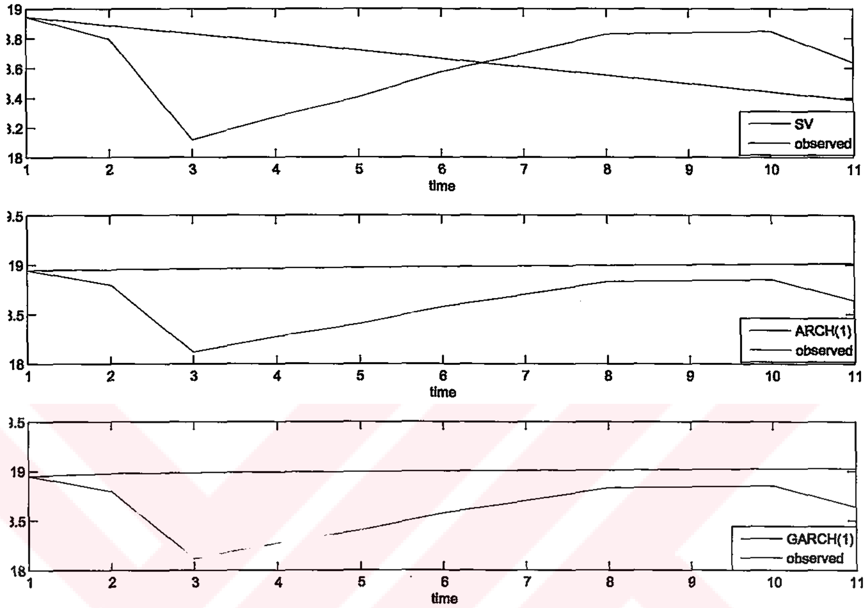

Figure 6.1: *Volatility forecasts and observed data*

The sum of squares of the deviations from the observed volatility for each model's forecasts are as follows:

	SV	ARCH(1)	GARCH(1,1)
$\sum_{t=1}^{10}(\sigma_{observed} - \sigma_{model})^2$	1.3095	2.0001	2.1433

Table 6.8: *Sum of squared deviations*

So, our stochastic volatility model gives closer predictions to the observed volatility than ARCH and GARCH models for our data in the studied time interval.

CHAPTER 7

CONCLUSIONS

In this study we focused on nonconstant volatility models, especially on stochastic volatility models. We introduced stochastic volatility and ARCH processes and gave some parameter estimation methods. Using Vasicek model we constructed a stochastic volatility model and then we worked on this model. We derived the density function for the stochastic element of the interest rate process. To make the calculations easier we reduced the density function from the integral form to a series form. For the estimation of parameters we used the method of moments. In the application part we used the interest rate data of Turkey. We calculated the parameters of our stochastic volatility model and the parameters of ARCH and GARCH models. Then, we tested the out of sample performance of our model against the ARCH and GARCH models. The results showed that in the studied time interval our stochastic volatility model gives closer predictions to the observed volatility than ARCH and GARCH models for our data.

REFERENCES

[1] Alizadeh, S., Brandt, M., W., and Diebold, F., X. (2002). *Range-based estimation of stochastic volatility models.* Journal of Finance, 57 : 1047-1091.

[2] Andersson, K. (2003). *Stochastic volatility* . U.U.D.M. Project Report 2003 : 18.

[3] Belledin, M., and Schlag, C. (1999). *An empirical comparison of alternative stochastic volatility models.* Goethe University Frankfurt am Main, Working Papar Series: Finance and Accounting, 38.

[4] Bollerslev, T. (1986). *Generalized Autoregressive Conditional Heteroscedasticity.* Journal of Econometrics, 31 : 307-327.

[5] Borovkov, K. (2003). *Elements of stochastic modelling.* World Scientific Publishing Co. Pte. Ltd., Singapore.

[6] Brooks, C. (2002). *Introductary econometrics for finance.* Cambridge University Press, Cambridge.

[7] Buff, R. (2002). *Uncertain volatility models-Theory and applications.* Springer-Verlag Berlin Heidelberg.

[8] Cotton, P., Fouque, J.-P., Papanicolaou, G., and Sircar, R. (2004). *Stochastic volatility corrections for interest rate derivatives* . Mathematical Finance 14(2) : 173-200.

[9] Cox, J., C., Ingersoll, J., E., and Ross, S., A. (1985). *A theory of the term structure of interest rates* . Econometrica 53(2) : 385-408.

[10] Davidson, R., and MacKinnon, J., G. (2004). *Econometric theory and methods.* Oxford University Press, New York.

[11] Davis, M., H. (2004). *Complete-market models of stochastic volatility.* Royal Society of London Proceedings, Serie A, 460(2041) : 11-26.

[12] Dragulescu, A., A., and Yakovenko, V., M. (2002). *Probability distribution of returns in the Heston model with stochastic volatility.* Quantitative Finance, 2 : 443-453.

[13] Dumas, B., Fleming, J., and Whaley, R., E. (1998). *Implied volatility functions: Empirical tests.* National Bureau of Economic Research, Inc., NBER Working Papers, 5500.

[14] Engle, R., F. (1982). *Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.* Econometrica, 50 : 987-1008.

[15] Fouque, J.-P., Papanicolaou, G., and Sircar, K., R. (2000). *Derivatives in financial markets with stochastic volatility.* Cambridge University Press, Cambridge.

[16] Georges, P. (2003). *The Vasicek and CIR models and the expectation hypothesis of the interest rate term structure.* Depertment of Finance Canada, Working Papers, 2003-17.

[17] Geske, R. (1979). *The valuation of compaund options.* Journal of Financial Economics, 7 : 63-81.

[18] Gourieroux, C. (1997). *ARCH models and financial applications.* Springer-Verlag New York, Inc.

[19] Greene, W., H. (2003). *Econometric analysis.* Pearson Education, Inc., Upper Saddle River, New Jersay.

[20] Henderson, V. (2002). *Analiytical comperations of option prices in stochastic volatility models.* Oxford Financial Research Center, OFRC Working Paper Series, 2000mf03.

[21] Heston, S., L. (1993). *A closed-form solution for options with stochastic volatility with application to bond and currency options.* The Review of Financial Studies, 6(2) : 327-343.

[22] Hobson, D., G., and Rogers, L., C., G. (1998). *Complete models with stochastic volatility.* Mathematical Finance 8(1) : 27-48.

[23] Hull, J., and White, A. (1987). *The pricing of options on assets with stochastic volatilities* . The Journal of Finance 42(2) : 281-300.

[24] Johnson, H., and Shanno, D. (1987). *Option pricing when the variance is changing.* The Journal of Financial and Quantitative Analysis, 22(2) : 143-151.

[25] Kloeden, P., E., and Platen, E. (1992). *Numerical solution of stochastic differential equations.* Springer-Verlag Berlin Heidelberg.

[26] Larsen, R., J., and Marx, M., L. (2001). *An introduction to mathematical statistics and its applications.* Prentece-Hall, Inc., Upper Saddle River, New Jersey.

[27] Malliaris, A., G., and Brock, W., A. (1982). *Stochastic mothods in economics and finance.* Elsevier Science Publishers B. V., Amsterdam, Netherlands.

[28] Melino, A., and Turnbull, S., M. (1990). *Pricing foreign currency options with stochastic volatility.* Journal of Econometrics, 45 : 239-265.

[29] Miller, I., and Miller, M. (1999). *John E. Freund's mathematical statistics.* Prentice Hall, Upper Saddle River, New Jersey.

[30] Mills, T., C. (1993). *The econometric modelling of financial time series.* Cambridge University Press, Cambridge.

[31] Pascucci, A. (2004). *Analysis of an uncertain volatility model.* 5. Workshop di Franza Quantitativa, Siena.

[32] Patterson, K., D. (2000). *An introduction to applied econometrics: A time series approach.* Macmillan, Basingstoke, England.

[33] Scott, L., O. (1987). *Option pricing when the variance changes randomly: Theory, esimation and application.* Journal of Financial and Quantitative Analysis, 22(4) : 419-438.

[34] Stein, E., M., and Stein, J., C. (1991). *Stock price distributions with stochastic volatility: An analytic approach* . Rewiev of Financial Studies 4(4) : 727-752.

[35] Vasicek, O. (1977). *An equilibrium characterization of the term structure.* Journal of Financial Economics, 5 : 177-188.

[36] Verbeek, M. (2000). *A guide to modern econometrics.* John Wiley and Sons, Ltd., Chichester, England.

[37] Wiggins, J., B. (1987). *Option values under stochastic volatility: Theory and empirical estimates.* Journal of Financial Economics, 19 : 351-372.

[38] Yolcu, Y. (2005). *One-factor interest rate models: Analytic solutions and approximations.* Master Thesis.