

แบบประเมินราคาอปชันสำหรับกระบวนการแพร่อย่างกระโดด
กับความผันผวนสโตนสโตกและดอกเบี้ยสโตนสโตก



วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรดุษฎีบัณฑิต

สาขาวิชาคณิตศาสตร์ประยุกต์

มหาวิทยาลัยเทคโนโลยีสุรนารี

ปีการศึกษา 2556

**A JUMP-DIFFUSION OPTION PRICING MODEL WITH
STOCHASTIC VOLATILITY AND STOCHASTIC
INTEREST RATE**

Paiboon Peeraparp



**A Thesis Submitted in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Applied Mathematics**

Suranaree University of Technology

Academic Year 2013

**A JUMP-DIFFUSION OPTION PRICING MODEL
WITH STOCHASTIC VOLATILITY AND
STOCHASTIC INTEREST RATE**

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

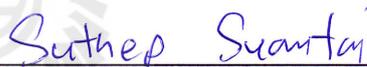
Thesis Examining Committee


(Assoc. Prof. Dr. Prapasri Asawakun)

Chairperson


(Prof. Dr. Pairote Sattayatham)

Member (Thesis Advisor)


(Prof. Dr. Suthep Suantai)

Member


(Asst. Prof. Dr. Eckart Schulz)

Member


(Asst. Prof. Dr. Arjuna Chaiyasena)

Member


(Prof. Dr. Sukit Limpijumnong)

Vice Rector for Academic Affairs


(Assoc. Prof. Dr. Prapun Manyum)

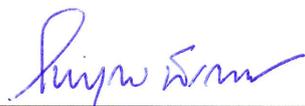
Dean of Institute of Science

ไพบูรณ์ พิรภาพ : แบบประเมินราคาอปชันสำหรับกระบวนการแพร่อย่างกระโดดกับความผันผวนสโตแคสติกและดอกเบี้ยสโตแคสติก (A JUMP-DIFFUSION PROCESS OPTION PRICING MODEL WITH STOCHASTIC VOLATILITY AND STOCHASTIC INTEREST RATE) อาจารย์ที่ปรึกษาวิทยานิพนธ์ : ศาสตราจารย์ ดร. ไพโรจน์ สัตยธรรม, 115 หน้า.

วิทยานิพนธ์ฉบับนี้เสนอการนำวิธีการเปลี่ยนเวลาของกระบวนการเลวิมาประยุกต์ในกระบวนการแพร่อย่างกระโดดกับความผันผวนแบบสโตแคสติกและดอกเบี้ยแบบสโตแคสติก โดยมีการนำทั้งวิธีการแปลงฟูเรียร์ และวิธีการกำหนดราคาแบบมอดคูลาร์ มาใช้ในการหาค่าสิทธิที่จะซื้อแบบยุโรป วิธีการทั้งสองนี้จะทำให้การสร้างสูตรค่าสิทธิที่จะซื้อแบบยุโรปค่อนข้างรวดเร็วและมีประสิทธิภาพในการคำนวณราคาด้วยวิธีการทางตัวเลข นอกจากนี้ได้มีการประมาณค่าพารามิเตอร์ของตัวแบบกระบวนการแพร่อย่างการกระโดดซึ่งมีความผันผวนและอัตราดอกเบี้ยแบบสโตแคสติก (JDSVSI) และตัวแบบกระบวนการแพร่อย่างกระโดดซึ่งมีความผันผวนแบบสโตแคสติก (JDSV) โดยใช้เทคนิคการหาค่าเหมาะสมแบบวงกว้างและแบบเฉพาะที่

สาขาวิชาคณิตศาสตร์
ปีการศึกษา 2556

ลายมือชื่อนักศึกษา



ลายมือชื่ออาจารย์ที่ปรึกษา



PAIBOON PEERAPARP : A JUMP-DIFFUSION PROCESS OPTION
PRICING MODEL WITH STOCHASTIC VOLATILITY AND
STOCHASTIC INTEREST RATE. THESIS ADVISOR : PROF. PAIROTE
SATTAYATHAM, Ph.D. 115 PP.

TIME CHANGED LEVY PROCESS/ CALIBRATION/ STOCHASTIC INTEREST
RATE/ STOCHASTIC VOLATILITY/ JUMP-DIFFUSION/ BLACK-SCHOLES
EQUATION

In this thesis, we present the application of the time changed Levy method to model a jump-diffusion process with stochastic volatility and stochastic interest rate. We apply the Lewis Fourier transform method as well as the modular pricing method to derive a pricing formula for a European call option price. Combining these methods gives quite a short route to derive a European call option formula and makes it efficient to compute option prices. We also show the calibration of the jump-diffusion process with stochastic volatility (JDSV) and the jump-diffusion process with stochastic volatility and stochastic interest rate (JDSVSI) model to real market prices with global and local optimization algorithms.

School of Mathematics

Academic Year 2013

Student's Signature

Paiboon Peeraparp

Advisor's Signature

P. Sattayatham

ACKNOWLEDGEMENTS

Writing the PhD. thesis required not only a lot of efforts from myself but also help and support from teachers, friends and family members. My special thanks goes to my advisor, Prof. Dr. Pairote Sattayatham who initialed and continuous to improve the great PhD. program in financial mathematics. Without him, I could not have come this far.

I also like to express my gratitude to Asst. Prof. Dr. Eckart Schulz and Asst. Prof. Dr. Arjuna Chaiyasena who helped correcting my Mathematics and English in the proposal and thesis. This helped me to learn how to write Mathematics. Assoc. Prof. Dr. Prapasri Asawakun always gave good support and chaired the thesis defense committee. I also like to thank Prof. Dr. Suthep Suantai for his long-term guidance and for having been available to be an external committee member for this thesis defense.

Lastly, without my wife's understanding, I could not have made my dream to learn more on financial mathematics come true.

Paiboon Peeraparp

CONTENTS

	Page
ABSTRACT IN THAI.....	I
ABSTRACT IN ENGLISH	II
ACKNOWLEDGEMENTS	III
CONTENTS.....	IV
LIST OF TABLES	VII
LIST OF FIGURES	VIII
CHAPTER	
I INTRODUCTION.....	1
1.1 Requirements for a Better Model	1
1.1.1 Specifications of the Model	2
1.1.2 Estimation of the Model Parameters.....	3
1.2 Jump-Diffusion Model with Stochastic Volatility and Stochastic Interest Rate.....	4
1.3 Outline of this Thesis	5
II MATHEMATICAL AND PROBABILITY CONCEPTS	7
2.1 Probability and Stochastic Process	7
2.2 Stochastic Calculus for a Continuous Stochastic Process.....	9
2.3 Feynman Kac and Girsanov Theorems.....	11
2.4 Jump Process.....	13

CONTENTS (Continued)

	Page
2.4.1 Poisson and Compound Poisson Process	13
2.4.2 Stochastic Integral for a Jump Process and its Integral.....	17
2.5 Levy Process	18
2.5.1 Levy Process and its Properties.....	18
2.5.2 Levy Ito Decomposition and Levy Kitchen Theorem.....	19
2.5.3 Subordination	24
III DERIVATIVE PRICING	26
3.1 No-Arbitrage Principle.....	26
3.2 Incomplete Market.....	29
3.3 Derivatives and Their Uses	31
3.4 Dynamic Hedging and Option Trading.....	34
3.5 Empirical Surveys.....	36
3.6 Jump-Diffusion Model.....	40
3.7 Stochastic Volatility Model	45
IV MODEL SPECIFICATIONS.....	48
4.1 Introduction.....	48
4.2 The Typical Model.....	48
4.3 The Lewis Fourier Transform Method	49
4.4 Derivation of the Pricing Formula	52
4.5 Numerical Integration	62

CONTENTS (Continued)

	Page
4.5.1 The Trapezoidal Rule	62
4.5.2 Numerical Calculation of the Pricing Formula	63
V CALIBRATION	68
5.1 Calibration Problem	68
5.2 Calibration as an Inverse Problem	71
5.3 Optimization for Continuous Variables	78
5.4 Simulated Annealing (SA).....	80
5.4.1 General Simulated Annealing (GSA).....	80
5.4.2 Adaptive Simulated Annealing	85
5.4.3 Reannealing	86
5.4.4 Matlab Optimization Toolbox	87
5.5 Calibrating the Model	89
5.5.1 Implementation.....	89
5.5.2 Calibration Results	90
VI CONCLUSION AND RESEARCH POSSIBILITIES.....	101
6.1 Conclusion	101
6.2 Research Possibilities.....	102
REFERENCES	104
APPENDIX	109
CURRICULUM VITAE	115

LIST OF TABLES

Table	Page
4.1 Option prices and execution times.....	67
5.1 Implied volatility surface of DAX index options on Jul. 5, 2002.....	92
5.2 Comparison of the sum of square price differences in each model.	93
5.3 JDSV parameters from calibration.....	96
5.4 JDSVSI parameters from calibration.....	96
5.5 Implied volatility term structure generated from JDSV model.	97
5.6 Implied volatility term structure generated from JDSVSI model.....	97
5.7 Square of errors of implied volatilities of JDSV compared to market prices..	99
5.8 Square of errors of implied volatilities of JDSVSI compared to market prices.....	100

LIST OF FIGURES

Figure	Page
3.1 The S&P index option implied volatility term structure.....	37
3.2 Three months implied volatilities of SPX Options vs. the S&P index price. ...	37
4.1 Illustration of Trapezoid rule.	63
4.2 The integrand function $f(z)$ with $T=0.5$	65
4.3 The integrand function $f(z)$ with $T=1.0$	66
4.4 The integrand function $f(z)$ near the origin.	66
5.1 The objective function graph with varied kappa and sigma	72
5.2 The objective function graph near the minimum point.....	73
5.3 Characteristic of the error function.....	76
5.4 Local and global minimums of the objective function.	78
5.5 Error of each data point for JDSV (left) and JDSVSI (right).	98

CHAPTER I

INTRODUCTION

1.1 Requirements for a Better Model

The requirements for usability of an option pricing model are almost similar to applications and practices of a model in other fields of applied sciences. That is, a model should describe the characteristics of a real phenomenon well, and the derivations of the model and its properties should not be too complicated with available resources. This can be achieved by a proper design of the components of the model and we call this process the *specification of the model*. Therefore, a good model should have many parameters, but not too many, to reflect the phenomena sufficiently. With these considerations, the practitioners in all applied sciences always need to balance between the accuracy of the model and the efficiency of the model.

The distinct characteristic that differentiates the study of an option pricing model from other models in the applied sciences is that the study of an option pricing model is to deal with the attitudes of collections of investors which value an option price based on their risk preferences. The importance of this risk preference has been shown in various empirical studies which agree that investors always need to be compensated more than the risk free rate asset for taking risks in investing in a risky financial market. We use the word “*risk premium*” to represent the additional return that investors need to be compensated for taking more risks. But the derivation of the

risk premium is very complicated and rests on ideal assumptions. The practical alternative is to implicitly embed this risk premium in the parameters of the model. With these requirements for a better model, the study of option price modeling has long been mainly divided into two parts,

- Specifications of the model and
- Estimation of the model parameters.

1.1.1 Specifications of the Model

The study of modern option pricing models started with the work of Black and Scholes (1973) who implemented the geometric Brownian motion for the asset price dynamics. This model of an asset price dynamic is known as the *Black Scholes model* (*BS model*). However, given the strict assumptions of the model, this model could not explain the real asset price dynamics sufficiently well. The work of Merton (1976) added a compound Poisson process whose jump size is log-normally distributed to the original BS model and this model is called the *jump-diffusion model*. By adding this process, the improved model can describe the asymmetry and fat-tailed distribution of the asset return.

The shared properties of both Brownian motion and Poisson processes are the independent and stationary increments of their paths. To combine these two generic processes and even more into a single process, the Levy process is proposed. By including both continuous and discontinuous processes, a Levy process is a versatile model to be used for a model of an asset price dynamic. But the implementation of the pure Levy process model is not a convenient way to describe some observed properties of the variance process of an asset. By many empirical observations, the variance of the asset return is not constant and shows some other properties including

volatility clustering and correlation of the asset return and its volatility. To explain these mentioned properties, Heston (1993) introduced the stochastic volatility model by adopting a mean reverting square root process which has a correlation with the asset return to represent the variance process. This work also employed the characteristic function method to derive an option pricing formula. When the characteristic function of the asset price dynamic is analytical, we refer to this model as the *analytical tractable model*. With this tractable model, an option pricing formula can be recovered with the inverse transform method. Another way to explain the mentioned empirical properties of the variance, the time changed Levy process was brought in to complement a pure Levy process by Carr, Geman, Madan and Yor (2003).

The stochastic volatility model also has its drawback in that it is not able to fit prices of short date options well. Later Bates (1996) included a jump component into the stochastic volatility model to be able to explain the prices of short date options. The work of Bakshi, Cao and Chen (1997) included a stochastic interest rate component in order to correct prices of long date options. Adding more features to the model is not without problems. The long computation time, the difficulties in implementation and the complicated pricing formula have prevented a highly accurate model to be used in the financial industry.

1.1.2 Estimation of the Model Parameters

To identify the risk premium associated with an option model, most models implicitly include this risk premium in the parameters of the models. By observing the time series data of asset and option prices, we may be able to estimate these parameters by a statistical method, but there is no guarantee that the model's

parameters from this method will generate a price that fits with the market price. A process to identify the coefficients of the variables in the model that are consistent with the market prices is called the *calibration process*. However this process normally leads to a nonlinear inverse problem which is usually an ill-posed problem. Also in a real market, prices of options are not available for all maturities and strikes, an optimization algorithm is therefore needed to find the best possible solution to identify the coefficients of parameters. Though there are many efficient optimization algorithms for the solution to this problem, many algorithms succeed only in finding a local solution. Global optimization algorithms can overcome this difficulty but most of the global optimization algorithms are not efficient, and it is difficult to prove for their global convergence.

1.2 Jump-Diffusion Model with Stochastic Volatility and Stochastic Interest Rate

Inspired by the works of Bakshi, Cao and Chen (1997), Pinkham and Sattayatham (2011) and Carr and Wu (2004), our model will combine a time changed Levy process, a compensated compound Poisson process (jump component) and a stochastic interest rate component. The set up of this model specification is close to the model in Pinkham and Sattayatham (2011) but we will not time-change the jump component, the reason for which we will give in the later part of the thesis. This specification of stochastic components will help to explain the empirical characteristics of asset prices and their volatilities which are described in Bakshi, Cao and Chen (1997) and Schoutens (2003). In this specification, the jump component will represent the jump of prices due to the rare events, and the time changed Levy process

will explain the characteristic of asset price volatility including the *volatility clustering property* and the *leverage effect* (The correlation of the volatility process and the asset price process). The derivation will utilize mostly the *risk neutral expectation pricing method* (3.4), instead of the partial differential equation method, which should give a more simple and efficient way to derive and to compute an option price.

All of 3 papers mentioned in the first paragraph of this section on which our work is based use the statistical methods to estimate the parameters of the models. We shall extend our work by having our model to be calibrated to the market prices instead of using a statistical methodology. There are some previous works using a calibration method such as Moodley (2005), Nassar (2010) and Mikhailov and Nogel (2006), but most of them give a crude description of the optimization algorithms and do not show the results on the numerical methods.

1.3 Outline of this Thesis

Here, we provide the outline of this thesis.

Chapter II gives a brief background of the mathematics for continuous time processes, jump processes and Levy processes. The intention in this chapter is to provide readers with the general concepts and the introductions to the notations that we will use in the later chapters.

For Chapter III, we begin with the background of the risk neutral pricing method in both complete and incomplete markets. Then we provide the practicability side of the option market in terms of their uses and trading. The empirical evidences and the definitions of these stylized facts are briefly introduced here. In the last part,

we also include the importance and derivation of the jump-diffusion and stochastic volatility model.

For Chapter IV which is the first main part of this thesis, we combine the Lewis Fourier transform method with the Modular Pricing method to derive a European call option pricing formula of the jump-diffusion model with stochastic volatility and stochastic interest rate. We then present the numerical algorithm applied for the calculation of option prices generated from our pricing formula.

In Chapter V, which is the second main part of this thesis, we start with the background of the nonlinear inverse problem and the regularization method that are applicable to our calibration problem. Then, the concept of global optimization is introduced. In the last section, we implement the calibration method to the option model derived in Chapter IV. Both simulated annealing algorithm and local search algorithm are employed to obtain the solution of this calibration problem.

We then conclude our work in Chapter VI and discuss the possibilities for further research.

CHAPTER II

MATHEMATICAL AND PROBABILITY CONCEPTS

In this chapter, we review the concepts of probability and stochastic calculus for continuous processes, jump processes and Levy processes that will be tools for the following chapters. Rather than giving a thorough mathematical treatment, we present the ideas and general facts that are applicable to our problems.

2.1 Probability and Stochastic Process

In financial modeling, we deal with the characteristics of the prices of underlying assets and its derivatives. The dynamics of these prices are random and governed by the desired probability laws set by the model specification. The theorems of probability and stochastic process will provide the general foundation for the modeling work.

To save a repetition of a general description, we assume there is always a probability triple (Ω, \mathcal{F}, P) or (Ω, \mathcal{F}, Q) associated with the stochastic process we refer to. This probability triple is always defined and can be extended to accommodate our uses. Each component of the probability triple is defined as follows,

- Ω denotes set of all states of the world which relates to the model, where an element $\omega \in \Omega$ is referred to be either a sample point or a sample path.
- \mathcal{F} denotes the σ -algebra which is the set of possible events.

- P denotes the physical or real probability measure and Q denotes the risk neutral probability measure.
- A *filtration* F is a non decreasing family $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ which represents the increasing known information.
- The stochastic process $X = (X_t)_{0 \leq t \leq T}$ is adapted to the filtration F , if X_t is \mathcal{F}_t measurable for each t . That means given the filtration F , at each t , we know the possible values of X_t .

Definition 2.1. (Stochastic Process). Let (Ω, \mathcal{F}, P) be a probability space, T an index set, (E, \mathcal{B}) a measurable space. An (E, \mathcal{B}) -valued *stochastic process* on (Ω, \mathcal{F}, P) is a collection $(X_t)_{t \in T}$ of random variables $X_t : (\Omega, \mathcal{F}, P) \rightarrow (E, \mathcal{B})$ for $t \in T$.

- For each $t \in T$, we have a random variable $X(\cdot, t) : \omega \rightarrow X(\omega, t)$ for all $\omega \in \Omega$ and
- For each $\omega \in \Omega$, we have a path $X(\omega, \cdot) : t \rightarrow X(\omega, t)$ for all $t \in T$.

Definition 2.2. (Martingale). A stochastic process X is a *martingale* with respect to a probability space (Ω, \mathcal{F}, P) with a filtration F if

- X is F adapted,
- $E[|X_t|] < \infty$ for all $t \geq 0$,
- $E[X_t | \mathcal{F}_s] = X_s$ for $0 \leq s \leq t$

Definition 2.3. (Brownian Motion). A real-valued process $(W_t)_{t \geq 0}$ is a *Brownian Motion* (BM.) if it satisfies the following properties,

- $W_0 = 0$ almost surely.
- For every $s \leq t \leq u$, $W_u - W_t$ is independent with $W_t - W_s$.
- For every $0 \leq s \leq t$, $W_t - W_s$ is normally distributed with mean 0 and variance $t - s$. We will denote the normal distribution with mean 0 and variance $t - s$ by the symbol $N(0, t - s)$.

2.2 Stochastic Calculus for a Continuous Stochastic Process

Definition 2.4. (Ito Process). An *Ito process* X_t is a process that can be represented by

$$X_t = X_0 + \int_0^t a(s, X_s) ds + \int_0^t b(s, X_s) dW_s, \quad (2.1)$$

where $a(t, X_t)$ and $b(t, X_t)$ are adapted processes which satisfy some conditions to guarantee the existence and uniqueness of the solution (Refer to these conditions in Øksendal (1998)). The first integral on the right-hand side is defined as the Riemann integral while the second integral is the Ito integral (Definition 2.5) with respect to a Brownian motion W_t .

We often write (2.1) in differential form as

$$dX_t = a(t, X_t) dt + b(t, X_t) dW_t, \quad (2.2)$$

with the initial condition X_0 .

Definition 2.5. (Ito Integral). The *Ito integral* $\int_0^t f(s, X_s) dW_s$ can be defined for a function $f(t, X_t)$ adapted to $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ such that $\int_0^t E(f(s, X_s)^2) ds < \infty$ by

$$\int_0^t f(s, X_s) dW_s = \lim_{n \rightarrow \infty} \sum_{i=1}^n f\left(\left(i-1\right)\frac{t}{n}, W_{\left(i-1\right)\frac{t}{n}}\right) \left(W_{\frac{it}{n}} - W_{\left(i-1\right)\frac{t}{n}}\right).$$

As consequences of the Ito integral, we can have 3 important results of the Ito integral,

- *Martingality*, $\int_0^t f(s, X_s) dW_s$ is a continuous martingale,
- *Zero-Mean Property*, $E\left(\int_0^t f(s, X_s) dW_s\right) = 0$,
- *Ito Isometry*, $E\left(\int_0^t f(s, X_s) dW_s\right)^2 = E\left(\int_0^t f(s, X_s)^2 ds\right)$.

Theorem 2.6. (Ito's Formula). Let $dX_t = a(t, X_t)dt + b(t, X_t)dW_t$ be an Ito process where $a(t, X_t)$ and $b(t, X_t)$ are adapted processes. Define the process Y by $Y_t = f(t, X_t)$, where f is a $C^{1,2}$ function. Then

$$dY(t, X_t) = \frac{\partial f(t, X_t)}{\partial t} dt + \frac{\partial f(t, X_t)}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f(t, X_t)}{\partial x^2} dX_t dX_t, \quad (2.3)$$

or in integral form,

$$Y(T, X_T) - Y(0, X_0) = \int_0^T \frac{\partial f(t, X_t)}{\partial t} dt + \int_0^T \frac{\partial f(t, X_t)}{\partial x} dX_t + \frac{1}{2} \int_0^T \frac{\partial^2 f(t, X_t)}{\partial x^2} d[X, X]_t.$$

The term $dX_t dX_t$ in (2.3) is defined to be the differential form of quadratic variation of process X where the integral form of this term is denoted by $[X, X]_t$. We report the values of the common quadratic variation terms here without detailed calculation. For a time variable t and Brownian motions W^i, W^j with correlation or with no correlation, the values of these terms are

$$\begin{aligned}
dWdt &= 0, \\
dW^i dW^i &= dt, \\
dW^i dW^j &= 0 \text{ if there is no correlation,} \\
dW^i dW^j &= \rho dt \text{ if there is correlation}
\end{aligned}$$

where ρ denotes the correlation coefficient of W^i and W^j .

The Ito formula in (2.3) applies to a function of a one dimensional process. In case, the function Y_t depends on more than one stochastic variable, we can extend the formula to the *Ito formula for multidimensional processes*.

Theorem 2.7. (Ito's Formula for n-dimensional process).

Let $dX_t^i = a_i(t, X_t) + \sum_{j=1}^m b_{ij}(t, X_t)W_t^j$, $i = 1, \dots, n$ be a multidimensional process where

$a_i(t, X_t)$ and $b_{ij}(t, X_t)$ are adapted processes. Define the process Y by $Y_t = f(t, X_t)$

where f is a $C^{1,2}$ function. Then we have

$$dY_t = \frac{\partial f(t, X_t)}{\partial t} dt + \sum_{i=1}^n \frac{\partial f(t, X_t)}{\partial X_t^i} dX_t^i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f(t, X_t)}{\partial X_t^i \partial X_t^j} dX_t^i dX_t^j. \quad (2.4)$$

2.3 Feynman Kac and Girsanov Theorems

The *Feynman Kac theorem* allows to write the solution of a partial differential equation (PDE) in terms of the expectation of the stochastic differential equation (SDE) whose drift and diffusion coefficients are defined by the coefficients of the given PDE. In the option pricing application, this theorem will give the flexibility to write the solution of the PDE generated from a riskless portfolio as the expectation of the discount payoff function whose asset dynamic is on the risk neutral measure and vice versa.

Theorem 2.8. (Feynman Kac Theorem). Given a stochastic differential equation of the form,

$$dX_t = a(t, X_t)dt + b(t, X_t)dW_t \quad (2.5)$$

which is an Ito process starting at time $t = 0$. Suppose, $y(T, x)$ satisfies the following PDE,

$$\frac{\partial y(T, x)}{\partial t} = a(x, T) \frac{\partial y(T, x)}{\partial x} + \frac{1}{2} b^2(T, x) \frac{\partial^2 y(T, x)}{\partial x^2} - q(x)y(T, x), \quad (2.6)$$

with an initial condition

$$y(T = 0, x_0) = f(0, x_0),$$

where $q(x)$ is a continuous function.

Then the solution of the PDE (2.6) can be expressed as the expected value of the discount payoff,

$$y(T, x_0) = E[\exp(-\int_0^T q(X(t))dt) f(X(T))]. \quad (2.7)$$

The *Girsanov theorem* describes the way to change the measure of an Ito process from one measure to another measure. In the application for the derivatives pricing, we need this theorem to change the dynamic of the underlying asset from a physical measure to a risk neutral measure.

Theorem 2.9. (Girsanov Theorem). Given an Ito process X_t with well defined coefficients $a(X_t, t)$ and $b(t, X_t)$,

$$dX_t = a(X_t, t)dt + b(X_t, t)dW_t, \quad (2.8)$$

and $M(t)$ is an exponential martingale under a measure P and defined by

$$M(t) = \exp\left(-\frac{1}{2} \int_0^t \gamma^2(s)ds + \int_0^t \gamma(s)dW_t^M(s)\right) \quad (2.9)$$

with $E^P[M(t)] = 1$ and $\gamma(t)$ is a function of t .

W_t and W_t^M are Brownian motions on the same probability space (Ω, \mathcal{F}, P) and correlated with $dW_t dW_t^M = \rho dt$.

Then we can have the following results,

- $M(t)$ defines the a Radon Nikodym derivative $\frac{dP^*}{dP}$ and
- the new Brownian motion W^* under the probability space $(\Omega, \mathcal{F}, P^*)$ can be

written as

$$dW_t^* = dW_t - \gamma \sigma dt. \quad (2.10)$$

2.4 Jump Process

2.4.1 Poisson and Compound Poisson Process

A Poisson process is one of the simplest discontinuous processes and will be the main building box to create other jump processes. The standard Poisson process is a stochastic process $(N_t)_{t \geq 0}$ with jump size of one unit and is constant between two jumps.

Definition 2.10. (Poisson Process). A stochastic process $N = (N_t)_{t \geq 0}$ is said to be a *Poisson process* with parameter λ if it satisfies the following conditions;

- $N_0 = 0$,
- For every $s \leq t \leq u$, $N_u - N_t$ is independent with $N_t - N_s$,
- The path N_t is an increasing function with jump size of one unit.

- For any real number $t \geq 0$ and $h \geq 0$, the process $N_{t+h} - N_t$ has a stationary Poisson distribution with a parameter λh . That is,

$$P((N_{t+h} - N_t) = n) = e^{-\lambda h} \frac{(\lambda h)^n}{n!}, \text{ for } n = 0, 1, 2, \dots \quad (2.11)$$

The process N is usually referred as a Poisson process with intensity λ which is the expectation of the number of jumps per unit time. Since the jump size of a Poisson process is fixed at 1 unit, the application in financial modeling is rather limited. The compound Poisson process is introduced to give more flexibility on the jump size specification.

Definition 2.11. (Compound Poisson Process). Let process N_t be a Poisson process with intensity λ and Y_n a sequence of independent and identically distributed (i.i.d.) random variables with distribution η and independent of N_t . The *compound Poisson* X_t is defined as

$$X_t = \sum_{n=1}^{N_t} Y_n, \text{ with } t \geq 0. \quad (2.12)$$

The jumps of the compound Poisson process X_t in (2.12) occur at the same times as the jumps on N_t . While the jumps of N_t have one unit in size, the jumps of the compound Poisson process have the jump of Y_n which have the distribution η .

The Poisson and compound Poisson processes are examples in the class of jump processes. In financial application, these processes are used as the components in the model to represent a dynamic of asset price. By the concept of no-arbitrage pricing (Chapter III), the cadlag version (right continuous with left limit) of these jump processes is considered.

Definition 2.12. (Cadlag). A real valued non-anticipating stochastic process $(X_t)_{t \geq 0}$ on a filtered probability space is said to be a *cadlag version* stochastic process if for $\forall t \in [0, T]$:

- *Right continuous:* $X_{t+} = X_t$,
- *Left limit:* left limit of the process exists, i.e. $\lim_{s \uparrow t} X_s = X_{t-}$.

An asset price, in financial modeling, is normally assumed to be a cadlag process. Because at time $t -$, we know the value of the process before time t but do not know the value at time t . But at time $t +$, we know the value of time t which is the past information. Since in the no-arbitrage market, the investors are not able to see the future of the price movement, therefore, the definition of a cadlag process is equivalent to the definition of a *non-anticipating* process which is defined to be a stochastic process $(X_t)_{t \geq 0}$ with respect to the filtration $\{F_t\}_{t \geq 0}$ where the value of X_t is revealed at time t for each $t \in [0, T]$.

With this rationale of the cadlag process, a jump of a process X_t at time $t \geq 0$ is defined to be

$$\Delta X_t = X_t - X_{t-},$$

$$\text{where } X_{t-} = \lim_{s \uparrow t} X_s.$$

Lemma 2.13. (Characteristic Function of a Compound Poisson Process). Let X_t be a compound Poisson process with intensity λ and the distribution of jumps η . Then the *characteristic function of the process* X_t , denotes by $\phi_{X_t}(u)$, is given by

$$\phi_{X_t}(u) \equiv E[e^{iuX_t}] = \exp(t\lambda \int_{-\infty}^{\infty} (\exp(iux) - 1)d\eta(x)). \quad (2.13)$$

Proof:

$$\begin{aligned}
 E[e^{iuX_t}] &= E[\exp(iu \sum_{n=1}^{N_t} Y_n)] = E[\prod_{n=1}^{N_t} \exp(iuY_n)] \\
 &= E[E[\prod_{n=1}^{N_t} \exp(iuY_n) / N_t]] = E[E[\exp(iuY_1)]^n] \\
 &= \sum_{n=1}^{\infty} E[\exp(iuY_1)]^n P(N_t = n)
 \end{aligned}$$

$$\text{from } P(N_t = n) = \exp(-\lambda t) \frac{(\lambda t)^n}{n!}$$

$$E[e^{iuX_t}] = \sum_{n=0}^{\infty} \exp(-\lambda t) \frac{(\lambda t)^n E[(\exp(iuY_1)]^n)}{n!}.$$

$$\begin{aligned}
 \text{From } \exp(x) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\
 &= \exp(\lambda t (E[\exp(iuY_1)] - 1)) \\
 &= \exp(\lambda t \int_{-\infty}^{\infty} (\exp(iux) - 1) d\eta(x)).
 \end{aligned}$$

Defining a new measure $\nu(x) = \lambda \eta(x)$, the above equation can be written as

$$E[e^{iuX_t}] = \exp(t \int_{-\infty}^{\infty} (\exp(iux) - 1) d\nu(x).$$

The measure ν is called the *Levy measure* of the process X_t . (2.21)

In financial modeling, the asset price under a risk neutral measure (Chapter 3) is a martingale but both Poisson and compound Poisson processes are not martingales i.e. the mean of the process condition with the information up to time t is not equal to zero. Then a mean correction is needed to make them martingales.

Definition 2.14. (Compensated Compound Poisson Process). A *compensated compound Poisson process* \tilde{X} is defined to be a compound Poisson process from which the mean has been subtracted. That is

$$\tilde{X}_t = \sum_{n=1}^{N_t} Y_t - \lambda k_y t \text{ where } k_y = \int_{-\infty}^{\infty} y \eta(dy). \quad (2.14)$$

Lemma 2.15. (Characteristic Function of a Compensated Compound Poisson Process). For the compensated compound Poisson process $\tilde{X}_t = \sum_{n=1}^{N_t} Y_t - \lambda k_y t$, the characteristic function of \tilde{X}_t is given by

$$\begin{aligned} \phi_{\tilde{X}_t}(u) &= E[\exp(iu(X_t - \lambda k_y t))], \\ &= \exp(-iu \lambda k_y t) \exp(\lambda t \int_{-\infty}^{\infty} (\exp(iux) - 1) d\eta(x), \\ &= \exp(\lambda t \int_{-\infty}^{\infty} (\exp(iux) - 1 - iux) d\eta(x)). \end{aligned} \quad (2.15)$$

2.4.2 Stochastic Integral for a Jump Process and its Integral

Definition 2.16. (Stochastic Integral for a Jump Process). Given a stochastic process $f(t, X_t)$ that is adapted and left continuous, the *stochastic integral with respect to the jump process* N_t is defined by

$$\int_0^T f(t, X_t) dN_t = \sum_{k=1}^{N_T} f(T_k, X_{T_k}) \Delta N_k.$$

In this definition, the jump process is restricted to the jump process that has a finite number of jumps in the interval $[0, T]$ with $N_0 = 0$. The *jump time* T_k is defined at the time when the jumps occur and $\Delta N_k = N_k - N_{k-1}$.

Theorem 2.17. (Ito Formula for a Jump Process).

Let $dX_t = a(t, X_t)dt + b(t, X_t)dW_t + J(t, X_t)dN_t$ be an Ito process with jumps where $a(t, X_t)$, $b(t, X_t)$ are adapted processes and $J(t, X_t)$ is a left continuous adapted process. Define the process Y by $Y_t = f(t, X_t)$, $C^{1,2}$ function. Then

$$dY_t = \frac{\partial f(t-, X_{t-})}{\partial t} dt + \frac{\partial f(t-, X_{t-})}{\partial x} dX_t^c + \frac{1}{2} \frac{\partial^2 f(t-, X_{t-})}{\partial x^2} dX_t^c dX_t^c + [f(t, X_t) - f(t-, X_{t-})] dN_t,$$

where $f(t-, X_{t-})$ represents the value of the function just before the jump occurs and X_t^c is the continuous part of X_t , i.e. X_t^c is obtained by removing the jump parts from X_t .

2.5 Levy Process

Both Brownian motions and jump processes share the same properties that the increments of their paths are stationary and independent from the past. The class of Levy process includes both BM and jump process and extends to all the processes that have both stationary and independent increment properties and includes all the independent and linear combination of them.

2.5.1 Levy Process and its Properties

Definition 2.18. (Levy process). A right continuous with left limits (cadlag) stochastic process X on a space (Ω, \mathcal{F}, P) is said to be a *Levy process* if it satisfies the

following conditions:

- $X_0 = 0$,
- Its increments are independent,
- Its increments are stationary,
- The process is stochastically continuous; that is

$$\forall \varepsilon > 0, \lim_{h \rightarrow 0} P(|X_{t+h} - X_t| \geq \varepsilon) = 0.$$

The last condition implies that the probability of the jump at any time t is zero. That is jumps occur at random times i.e. one cannot predict when jumps will occur.

By having a Levy process to be a cadlag process, then the number of jumps in a compact interval $[0,T]$ will be countable by the following lemma.

Lemma 2.19. (Lemma 13.12 Pascucci (2011)). Let X be a cadlag process on a compact interval $[0,T]$. Then for any $n \in \mathbb{N}$, the number of jumps of X of size greater than $\frac{1}{n}$ is finite:

$$\#\left\{t \in (0,T], |\Delta X_t| \geq \frac{1}{n}\right\} < \infty.$$

In particular, X has at most a countable number of jumps. By this lemma, we can say that a Levy process can have only finite number of large jumps (This is due to the last condition of definition 2.18). On the other hand, a Levy process can have a countably infinite number of small jumps.

2.5.2 Levy Ito Decomposition and Levy Kitchen Theorem

Once a model gets more complicated, the probability density function of the log-return of the asset price may not have an analytical form or is difficult to derive. However there is a one to one relationship of probability density function and characteristic function to uniquely determine the probability law of that variable.

Definition 2.20. (Fourier Transform and Inversion). Let $f(x)$ be a piecewise continuous function in the *absolute integrable space* L^1 over \mathbb{R} , that is

$$\int_{-\infty}^{\infty} |f(x)| dx < \infty.$$

The *Fourier transform* of $f(x)$ is defined by

$$Ff(u) \equiv \int_{-\infty}^{\infty} e^{iux} f(x) dx,$$

where u is the *transform variable* defined in the real domain and $i = \sqrt{-1}$ is the imaginary unit. The *inverse Fourier transform* is defined as;

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} Ff(u) du.$$

In some cases, we will allow u to be a complex number and call the inverse transform the *Generalized Fourier transform*:

$$f(x) = \frac{1}{2\pi} \int_{z_i - \infty}^{z_i + \infty} e^{-izx} Ff(z) dz \quad \text{where } z \text{ is a complex number.} \quad (2.16)$$

Definition 2.21. (Characteristic Function). The *characteristic function* for a random variable X_i with probability density function $f_X(x)$, denoted by $\phi_X(u)$, is defined as the Fourier transform of the probability density function of the random variable, so we can write the characteristic function as

$$\phi_X(u) = E[e^{iuX}] = \int_{-\infty}^{\infty} e^{iuX} f_X(x) dx = F[f_X(x)], \quad (2.17)$$

where $F[f_X(x)]$ denotes the Fourier transform of the probability density function.

Using the Euler formula, the characteristic function can also be expressed as

$$\phi_X(u) = E[e^{iuX}] = E[\cos(uX)] + iE[\sin(uX)]. \quad (2.18)$$

By the inverse formula of Fourier's transform, the density function is derived by

$$f_X(x) = F^{-1}[\phi_X(u)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \phi_X(u) du.$$

Below are some of the important properties of characteristic functions (Schmelzle 2010).

- $\phi_X(u)$ always exists since $|e^{iuX}|$ is a continuous and bounded function for all finite real u and X .
- $\phi_X(0) = 1$ for any continuous distribution,
- $\phi_X(u)$ is a continuous function of u ,
- $|\phi_X(u)| \leq 1$,
- If X_1 and X_2 are independent, then the characteristic function $\phi_Y(u)$ of the new random variable $Y = X_1 + X_2$ is the product of the characteristic functions of each random variable, $\phi_{X_1}(u)\phi_{X_2}(u)$.

One of the most important characterizations of a Levy process is its infinitely divisible distribution property. This property depicts the fine structure of a Levy process when we want to study the minute detail of the distribution of an asset.

Definition 2.22. (Infinitely Divisible). A random variable X is said to be an infinitely divisible if for $n \geq 2$, there exist independent identically distributed variables X_1, \dots, X_n such that

$$X \stackrel{d}{=} X_1^n + \dots + X_n^n. \quad (2.19)$$

This means the process X can be decomposed into a sum of an infinite number of i.i.d. random variables. From this it follows that the characteristic function of X_t has a Levy-Khintchine representation.

Theorem 2.23. (Levy Khintchine Theorem). If X is a Levy process, there exists a unique function $\psi(u): \mathbb{R} \rightarrow \mathbb{C}$ such that $\psi(0) = 0$ and

$$E[\exp(iuX)] = \exp(-t\psi(u)).$$

The function $\psi(u)$ is called the *characteristic exponent* of the process X . And $\psi(u)$ has the following *Levy-Khinchine representation*,

$$\psi(u) = iau + \frac{1}{2}\sigma u^2 - \int_{-\infty}^{\infty} (\exp(iux) - 1 - iux1_{|x|\leq 1})v(dx), \quad (2.20)$$

where a is the drift coefficient, σ is the diffusion coefficient and ν is the Levy measure. The *Levy measure* is defined by

$$\nu(A) = E \left[\sum_{s \leq 1} 1_A(\Delta X_s(\omega)) \right], \quad (2.21)$$

where 1_A is the indicator function and A is an arbitrary interval on $R \setminus \{0\}$ such that

$\int_R (|x|^2 \wedge 1) \nu(dx) < \infty$ i.e. the Levy measure is the expected number of jumps before

time $t = 1$ with a size which belong to A .

As a consequence of this theorem, the law of the Levy process X is uniquely determined by the continuous function $\psi(u)$ which is generated by the triplet (a, σ, ν) .

This triplet is called the *Levy triplet* of the Levy process X .

The Levy-Ito decomposition indicates that, given a Levy triplet (a, σ, ν) , one can find the unique corresponding Levy process which has four decomposed independent Levy processes.

Theorem 2.24. (Levy Ito Decomposition Theorem). Consider the triplet (a, σ, ν) where $a \in R$, $\sigma \in R^+ \cup \{0\}$ and ν is the Levy measure satisfying $\nu(\{0\}) = 0$ and

$\int_R (|x|^2 \wedge 1) \nu(dx) < \infty$. There exists a unique Levy process X which can be decomposed

into four independent processes:

$$X = X_1 + X_2 + X_3 + X_4,$$

where the characteristic exponent of each process can be expressed by,

$$\psi_{X_1}(u) = iau,$$

$$\psi_{X_2}(u) = \frac{\sigma u^2}{2},$$

$$\psi_{X_3}(u) = \int_{|x| \geq 1} (e^{iux} - 1)v(dx),$$

$$\psi_{X_4}(u) = \int_{|x| < 1} (e^{iux} - 1 - iux)v(dx).$$

The process X_1 represents the drift process with parameter a and the process X_2 represents the diffusion process with parameter σ . The process X_3 represents the compound Poisson process with intensity $\lambda = v[|x| \geq 1]$ and jump distribution given by $u_j[A] = v[A \cap \{|x| \geq 1\}] / \lambda$. By Lemma 2.19, the number of jumps of this component is finite. The last process X_4 is the limit of compensated compound Poisson process: $X_4 = \lim_{\varepsilon \rightarrow 0} (\sum_{s \leq t} 1_{\varepsilon \leq |\Delta X_s| < 1} \Delta X_s - v[[\varepsilon, 1]]t)$. The number of jumps in this component can be infinite but is compensated by the term $v[[\varepsilon, 1]]t$, the Levy measure of the jump sizes between a small number $\varepsilon > 0$ and 1, and this makes the term X_4 finite. Then we can write this Levy process as

$$X_t = at + \sigma W_t + \sum_{s < t} 1_{\{|\Delta X_s| \geq 1\}} \Delta X_s + \lim_{\varepsilon \rightarrow 0} (\sum_{s \leq t} 1_{\varepsilon \leq |\Delta X_s| < 1} \Delta X_s) - v[[\varepsilon, 1]]t. \quad (2.22)$$

Definition 2.25. (Finite and Infinite Activity). Let X be a Levy process with Levy triplet (a, σ, ν) then

- if $\nu(\mathbb{R}) < \infty$, the number of jumps of all paths of the Levy process is almost surely finite in any fixed interval and we say that the Levy process has *finite activity*.

- If $\nu(\mathbf{R}) = \infty$, the number of jumps of all paths of the Levy process is almost surely infinite on any fixed interval and we say that the Levy process has *infinite activity*.

2.5.3 Subordination

By the fact that the volatility of an asset price process exhibits the correlation between time intervals (Volatility Clustering (Section 3.5)), it cannot be modeled by a pure Levy process. A popular solution to this problem in financial modeling is by subordinating a pure process by a suitable subordinator. A *subordinator* is a Levy process that has almost surely non-decreasing paths. And by *subordination*, we mean to change a pure Levy process $(Y_t)_{t \geq 0}$ to a new process $Z_t = Y_{X(t)}$ where $X(t)$ is an independent subordinator. The new process Z is called a *subordinated Levy process*.

Instead of characterizing this process by its characteristic exponent, it is much easier to characterize this process by its Laplace exponent.

Definition 2.26. (Laplace Transform). Let X_t be a subordinator and for each $t \geq 0$ let $L_x(\lambda)$ denote the Laplace transform of X_t , i.e. $L_x(\lambda) = E[e^{-\lambda X_t}]$. Then there is a continuous and non-decreasing function $\ell(\lambda)$ such that for all $t \geq 0$ and all $\lambda \geq 0$,

$$L_x(\lambda) = \exp[-t\ell(\lambda)],$$

where $\ell(\lambda)$ is the *Laplace exponent* of X_t . Similar to the characteristic function, the Laplace exponent holds a unique relationship with the Levy process X . The Levy Khinchine representation also holds for the Laplace exponent under the form

$$\ell(\lambda) = a\lambda + \int (1 - e^{-\lambda x})\nu(dx),$$

where a again denotes the drift coefficient and ν is the Levy measure.

Theorem 2.27. (Bochner). Winkel (n.d.). Let $(Y_t)_{t \geq 0}$ be a Levy process and $(X_t)_{t \geq 0}$ be an independent subordinator. Then the process Z_t defined as $Z_t = Y_{X_t}$ is a Levy process and we have

$$E(e^{iuZ_t}) = e^{-t\ell_X(\psi_Y(u))},$$

$$\text{where } E(e^{iuY_t}) = e^{t\psi_Y(u)} \text{ and } E(e^{\lambda X_t}) = e^{-t\ell_X(\lambda)}, \quad (2.23)$$

where $\psi_Y(u)$ denotes the characteristic exponent of the Levy process Y and $\ell_X(\lambda)$ denotes the Laplace exponent of the subordinator X .



CHAPTER III

DERIVATIVE PRICING

3.1 No-Arbitrage Principle

The main question here is how to value an option. The very first idea is to set up the portfolio to replicate the payoff of that option. As long as the value of the portfolio equals the payoff of the option, the cost of the portfolio should be the price of that option. This is called “*The Law of One Price*”. In the real world we may not be able to replicate some options due to an insufficient number of traded underlying assets, or from market imperfections. This first and second section of this chapter will describe a way to handle this situation.

Consider a market model based on $d+1$ assets denoted by $S = (S^0, S^1, \dots, S^d)$, with $S^i \in \mathbb{R}^+ \cup \{0\}$, under a given probability space (Ω, \mathcal{F}, P) endowed with a filtration F . Assume our assets can be modeled as regular diffusion processes. The asset S^0 is normally taken to be a numeraire asset which is often but not limited to the money market asset B_t , defined by $B_t = \exp\left(\int_0^t r_u du\right)$ with r_u the instantaneous risk free interest rate.

We will also define a discounted process $D_t = 1/B_t$. The asset S^1 will be the asset whose dynamic we want to design and we call it here *the underlying asset*. The rest of S will be option price processes that we try to relate to S^0 and S^1 .

A trading strategy is a stochastic process $(h_t)_{t \geq 0} = (h_t^0, h_t^1, \dots, h_t^d)_{t \geq 0}$ with $h_t^i \in \mathbb{R}$ which represents the number of units of each asset in the portfolio. So that the value of a portfolio at time t denoted by V_t , is given by

$$V_t = \sum_m h_t^m S_t^m. \quad (3.1)$$

A portfolio with a strategy $(h_t)_{t \geq 0} = (h_t^0, h_t^1, \dots, h_t^d)_{t \geq 0}$ is called a *self-financing strategy* if

$$V_t = V_0 + \sum_m \int_0^t h_s^m ds_s^m, \quad (3.2)$$

where V_0 denotes the amount invested at time $t = 0$. The term $\sum_m \int_0^t h_s^m ds_s^m$ represents the profit or loss up from time 0 to time t . The element of the process h_t can be a positive or negative number with a positive number representing the number of that asset in the portfolio and a negative number representing the number of the asset being short sell. The above equation indicates that the change of value of a self-financing portfolio comes only from the changes of the prices of all securities in the portfolio without any injection or withdrawal of funds.

A general assumption for any model is that the market has no arbitrage opportunity. In the real market, there might be some arbitrage opportunities but by the mechanism of the efficient market, these opportunities will quickly disappear. The *arbitrage opportunity* is defined by,

- The initial value of the portfolio is zero and at anytime $t > 0$,
- The value of the portfolio has no negative value with probability one

- The expected value of the portfolio is positive.

Simply speaking, the arbitrage opportunity is an event when a zero cost portfolio can generate some positive expected return.

A model is called *complete* if every asset in the market can be replicated by a self-financing portfolio. That is, with the strategy as in (3.2), one can design a portfolio whose value equals to any asset in the market.

We will define an *Equivalent Martingale Measure (EMM)* by: a probability measure Q is an *equivalent martingale measure* with respect to P (typically, we take P as the historical measure that we obtain from empirical studies) if

- Both P and Q have the same null sets.
- The discounted price process $D_t S_t$ is a martingale under Q . That is

$$D_s S_s^m = E^Q[D_t S_t^m | F_s] \quad 0 \leq s \leq t. \quad (3.3)$$

The notation E^Q denotes the expectation under the equivalence martingale measure Q . And F_s is the information from time 0 to time s . When the discounted process is defined to be $D_t = 1/B_t$, we call this measure as “*the risk neutral measure*” as all assets have the same return (risk free return). That is, under the equivalence martingale measure, investors are indifferent in any asset because it has the same return.

The connection between no arbitrage strategy and existence of equivalent martingale measure was presented in Harrison and Pliska (1981), summarized in the following theorem:

First Fundamental Theorem of Asset Pricing (FTAP): A model admits no arbitrage if and only if there exists an equivalent martingale measure (EMM).

A consequence of this theorem is that the price of any option can be represented by

$$D_s C_s^m = E^Q[D_t C_t^m | F_s] \quad 0 \leq s \leq t. \quad (3.4)$$

When the market is complete and admits no arbitrage, the price of an option can be calculated from the above formula. We call this method to derive the price of an option “*the risk neutral expectation pricing method*” which we will use extensively in later chapters.

However the theorem does not guarantee the uniqueness of this equivalent martingale measure if the model is not complete.

3.2 Incomplete Market

When the market is not complete, not every option can be replicated by a self-financing portfolio. By the assumption of no-arbitrage, the market will admit non-uniqueness of an equivalent martingale measure. The prices of an option at time s , which has payoff at time t denoted by S_t , are bounded by

$$V_s = (\inf_{Q \in M} E^Q[D_t S_t / F_s], \sup_{Q \in M} E^Q[D_t S_t / F_s]),$$

where M is the set of equivalent martingale measures. This is often the case, when the model has more than one source of risk and we have only the underlying asset and money market asset in the self-financing portfolio. Most of the models used in the industry contain more than one source of risk to better explain the dynamic of the asset. So these models are generally incomplete and need to be assumed how investors value the asset prices for additional risks. The more complicated models serve both financial institutions and their clients. Whereas the financial institutions need a better model to describe the dynamic of an underlying asset, the clients need

some fancier structures to hedge their increasing risks. The requirements of the models are mainly for;

- Valuation: The model should produce the price of an option that is consistent with the market price and correctly produce the prices of illiquid options in terms of the payoff and maturities.
- Hedging: The model should better guide the seller of an option to mitigate the risks (the process to mitigate the risk is referred to as the *hedging process*) arrived from the uncertainties due to changes of underlying asset prices and other state variables.

For a complete model, the valuation problem is the same as the hedging problem. That is, the value of the option will match with the designed riskless self-financing portfolio. But in an incomplete market, this replication is not possible. The determination of price somehow depends on the risk preference, endowment, and views of the investors which can be incorporated into a *utility function*. The model, that includes the utility function, is referred as the *equilibrium model*. Even though this model is well accepted in the field of econometrics as it can fit the values of options to the past historical prices, in financial modeling, it is difficult to pin down the exact specification of this utility function. In the practitioner community, traders prefer to have a complete model whose dynamic of the underlying asset will be represented by the parametric model that implicitly includes the risk premium. We show in section 3.6 and 3.7 how this risk premium is included in the most common models.

3.3 Derivatives and Their Uses

The liberalization of financial markets all over the world has eased the barriers for movements of fund flows to seek for better returns in different markets and different geographical locations. The free market system has made markets more efficient and, at the same time, causes more volatility in the financial markets. The market participants are exposed more to the financial risks while they can enjoy more rewards for their investments. The derivatives market grows hand in hand with the increase of the financial risks. The use of derivatives has increased to provide the market participants with a way to manage their exposures.

The simplest derivative instrument is a *forward* which is an over-the-counter contract to exchange for an underlying asset at fixed price at a future date between two parties. The fair price of the forward for the underlying asset that is not perishable and requires small cost of carrying, is derived by the no-arbitrage principle which can be formulated as

$$F_t = S_0 e^{rt},$$

where F_t denotes the fair price of the underlying asset to be delivered at time t , S_0 denotes the price of the underlying asset at time 0 and r_t is the risk free rate. An almost similar instrument that is traded in the established exchanges is called a *future*. The difference between these two instruments is that a future is traded in a more orderly fashion i.e. fixed amount, fixed delivery date, while a forward is traded between two parties where normally one party is a financial institution, and the specification of the contract can be negotiated.

A little more advanced instrument is an option. An option gives a right to the option buyer, but not an obligation, to buy or sell the underlying asset at a specific price and time from the option seller (or the writer). The option buyer compensates this right by paying the seller the *premium* which is the price of the option. A call option gives the right to buy an asset while a put option gives the right to sell. So the profit to the option buyer at maturity of the option is $\max(S_T - K, 0)$ for the call option and $\max(K - S_T, 0)$ for the put option where S_T denotes the price of the asset at the maturity and K is the value of the strike. A *European option* allows the buyer to exercise the contract at only the maturity date while an *American option* allows the buyer to exercise any time from start to the maturity date. These types of options are also called *plain vanilla options*. There are more different payoff structures and different features of options which are called *exotic options*. The determination of a fair value of an option is more complicated than a forward and requires an assumption on the dynamic of an asset price which is called the *pricing model*. The very first and successful model was developed by Black and Scholes (1973) which is referred as the BS model here. The dynamic of an underlying asset in the BS model is a geometric Brownian motion with a drift,

$$dS_t = \mu_t S_t dt + \sigma S_t dW_t,$$

where S_t denotes the price of an underlying asset, μ_t is the expected return of that asset, σ is referred to as the volatility of the underlying asset and W_t denotes a standard Brownian motion. By constructing the riskless portfolio Π_t which consists of Δ_t units of the underlying asset and a short selling one unit of option whose price is denoted by $V(t, S_t)$, the value of the portfolio can be written as

$$\Pi_t = \Delta_t S_t - V(t, S_t) \quad (3.5)$$

By applying the Ito formula, setting $\Delta_t = \frac{\partial V(t, S_t)}{\partial S_t}$ and equating the return of this riskless portfolio to the risk free rate, the price of an option V_t is the solution of the following partial differential equation,

$$\frac{\partial V_t}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V_t}{\partial S_t^2} + r_t S_t \frac{\partial V_t}{\partial S_t} - r_t V_t = 0.$$

For a European call option with the payoff function at the maturity given by $\max(S_T - K, 0)$, where S_T is the price of the underlying asset at time T and K is the strike price, the value of the call option at time $t = 0$, V_0 , can be solved by the above partial differential equation with the boundary condition $V_T = \max(S_T - K, 0)$. We write here C_0 , instead of V_0 in the following formula to signify that this is the value of a call option,

$$C_0(S_0, T) = S_0 N(d_1) - Ke^{-rt} N(d_2),$$

$$\text{with } d_1 = \frac{\ln \frac{S_0}{K} + (r_t + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}, \quad d_2 = d_1 - \sigma\sqrt{T}, \quad (3.6)$$

where $C_0(S_0, T)$ is the value of a European call option at time $t = 0$ for the option maturing at time T . Here $N(d)$ is the standard normal cumulative density function.

The BS model is complete because we can replicate the value of an option with the underlying asset and the money market account by always keeping the number of units of the underlying asset, $\Delta_t = \frac{\partial V_t}{\partial S_t}$ and the rest of the portfolio in the money market account. Observe that the pricing formula of a European call option in the BS model does not depend on the value of the expected return of the underlying asset, μ_t ,

which is subject to a risk preference of each investor where a risk averse investor requires higher μ_t to compensate for risk assuming while a risk seeker requires smaller μ_t . This observation asserts, what we have discussed, that a complete model has a unique equivalent martingale measure and this measure transforms expected physical returns of all assets to a risk free return, r_t . This is why we call this measure the *risk-neutral measure* because, under this measure, all the assets have the same expected return that make all investors to be neutral to a holding of any asset as these assets generate the same return.

3.4 Dynamic Hedging and Option Trading

Based on (3.5), the change of a riskless portfolio due to the change of underlying asset price is expressed as

$$d\Pi_t = \Delta_t dS_t - \frac{\partial V(t, S_t)}{\partial S} dS_t.$$

By setting $\Delta_t = \frac{\partial V(t, S_t)}{\partial S_t}$, the riskless portfolio will be immune to the change of the underlying asset price if we assume that the other variables that determine the option price are fixed. We call this process the *delta hedging*. As in (3.6), the value of a European call option, according to the BS model, depends on 5 BS variables, S_t , K , r , t and σ , so the delta hedging keeps the value of the option unchanged due to the change of the underlying asset price. There are some other changes of the option value respect to other variables and are call sensitivities of the option value. Here are some important sensitivities on the BS model;

$$\begin{aligned} \text{Delta} = \Delta &= \frac{\partial C}{\partial S} & \text{Gamma} = \Gamma &= \frac{\partial^2 C}{\partial S^2} \\ \text{Vega} = \Lambda &= \frac{\partial C}{\partial \sigma} & \text{Theta} = \Theta &= \frac{\partial C}{\partial t} \end{aligned}$$

$$\text{Rho} = \rho = \frac{\partial C}{\partial r}$$

$$\text{Vanna} = \frac{\partial \Lambda}{\partial S} \quad \text{Volga} = \frac{\partial \Lambda}{\partial \sigma}$$

Note: The first five sensitivities are the primary sensitivities that are derived by the first derivative of the option value with respect to the BS variables. The last two sensitivities are the second derivative of the option value with respect to the BS variables. Most of plain vanilla options tend to have very small second derivative as a first derivative is smooth and continuous. But for a complicated structure, these two sensitivities are not small and very crucial in determining the hedging process.

The job of the options trader is to keep some sensitivities due to BS variables fixed or “*neutral*” in the trader’s jargon, and trade on his views on some specific sensitivities. All the sensitivities are local variables which means they vary to the large changes and even some sensitivities are not smooth depending on the payoff structures of the option. Also some of these variables are correlated to the other variables which requires the trading to be adjusted dynamically. The process of the adjustment is called the *dynamic hedging*. In practice, a trader will not manage options one by one but rather a portfolio of options with different types, different time to maturities and many payoff structures which is called the *option book*.

Even though the BS model rests on unrealistic assumptions, the model has a simple closed form solution which explains the changes of an option price by its small number of variables and allows for the simple mechanism for the dynamic hedging

for small changes in BS variables. It is a necessary tool for the option trading in practice. However the model is not perfect and faces a lot of criticisms which we will explore in the next sections.

3.5 Empirical Surveys

As discussed in the previous section on the features of the BS model, and by the fact that it is simple and has nice features: closed form formula, allowing for dynamic hedging, it has become popular in the industry at least for a way of quoting the option prices. By these advantages, the market normally quotes the prices by the implied volatility of the BS model. In the BS formula for a plain vanilla European call option, the formula is given by

$$C(S_t, T) = S_t N(d_1) - Ke^{-rt} N(d_2),$$

$$\text{with } d_1 = \frac{\ln \frac{S_t}{K} + (r_t + \frac{\sigma^2}{2})T}{\sigma \sqrt{T}}, \quad d_2 = d_1 - \sigma \sqrt{T}.$$

There are five parameters determining the options prices, namely: the price of the underlying asset S_t , the risk free rate r_t , the strike price K , the time to maturity T and the asset volatility σ . All the parameters, except the volatility, are directly observable in the market. If the traded option prices are available, the volatility can be uniquely recovered by finding the root of the BS formula since the formula is an increasing function with the volatility. The volatility obtained by this way is called the *implied volatility*. In any specific time to maturity, there are usually a limited number of standard strikes quoted while the unquoted strikes can be interpolated from the available implied volatilities.

Since the job of the option trader, unlike the underlying asset traders, is to trade on the level of volatilities, it is more convenient to look at the term structure of the implied volatility which is the three dimensional graph of volatilities level against the *time to maturity* or call it “*term*” in Figure 3.1 and the strike. Figure 3.1 is the example of the volatility term structure of S&P index options on September 27, 1995.

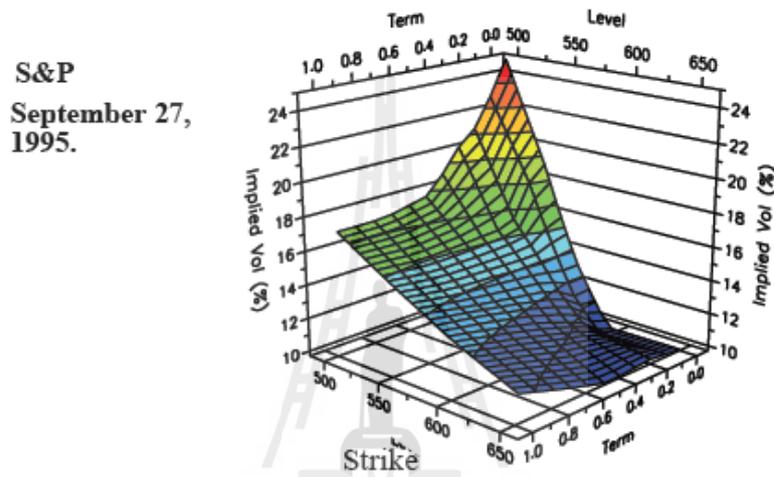


Figure 3.1 The S&P index option implied volatility term structure.

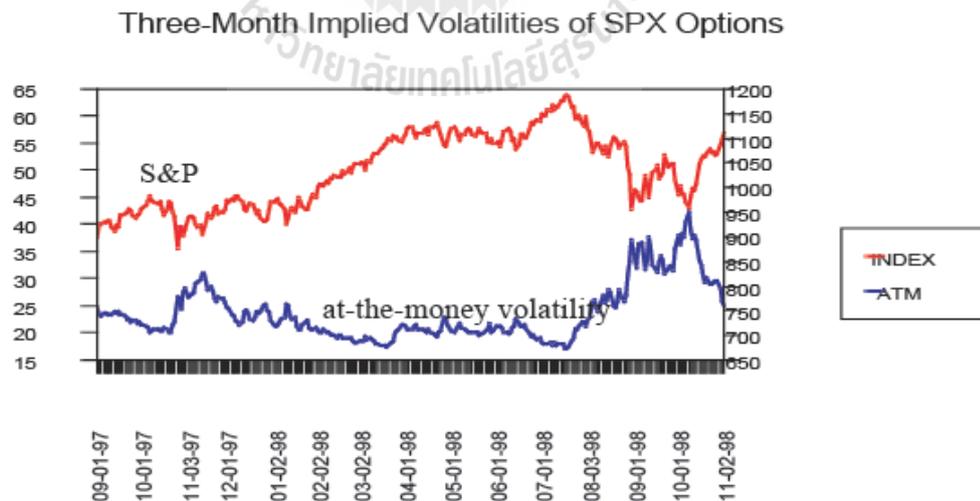


Figure 3.2 Three months implied volatilities of SPX Options vs. the S&P index price.

If we observe the charts in Figure 3.1 and Figure 3.2 (taken from Derman (2007)), we can see the volatility is not constant as assumed by the BS model. Here we give the common characteristics of the equity index option implied volatility. (All the characteristics except the last one are also observable in the Figure 3.1 and Figure 3.2)

- Volatilities are steepest for the short date options and shallower for the longer date options.
- The minimum of the volatility for each tenor occurs around the strike that corresponds to the at the money forward level (the forward price of the underlying asset that corresponds to the time to maturity).
- Lower strike options are priced at higher implied volatilities than the higher strikes at the same time to maturity.
- There is a negative correlation between the movement of the underlying asset and the change in implied volatilities i.e. the volatility tends to go up when the price of index moves lower and vice versa.
- The implied volatility tends to rise faster and decline slowly.

For other asset classes such as single stocks, currencies or commodities, the implied volatilities share the common characteristics with the implied volatilities of the equity index. Such common characteristics of the time series analysis across the asset classes is called the *stylized fact*.

Schoutens (2003) has also pointed out 2 main characteristics of the dynamics of the asset prices which are

- The log returns of the asset price do not behave like a normal distribution.

- The volatility of an asset price changes stochastically over time and is clustered.

Here we summarize the measures to the behaviors of the dynamic of asset price and its volatility.

Skewness: Skewness measures the degree of asymmetry in the distribution. This measure is defined by the third moment about the mean divided by the third power of the standard deviation:

$$\frac{E[(X - \mu_X)^3]}{\text{var}[X]^{3/2}} \quad (3.7)$$

For a symmetric distribution, the skewness is zero. We say that a distribution has a negative skewness if the distribution has longer tail to the left than to the right and vice versa.

Fat Tails and Excess Kurtosis: Fat Tails and Excess Kurtosis measure the shape of the distribution compared to the normal distribution. The Kurtosis is defined by

$$\frac{E[(X - \mu_X)^4]}{\text{var}[X]^2} \quad (3.8)$$

For the normal distribution, the Kurtosis is at 3. For a distribution that has a higher peak and more fat tails than the normal distribution, the Kurtosis is bigger than 3.

Volatility clustering: Volatility clustering refers to the positive autocorrelation of asset returns, when higher returns tend to follow higher returns, and so are the lower returns. A common indicator to this property is the autocorrelation function of the squared returns

$$\text{corr}(|r(t + \tau, \Delta t)|^2, |r(t, \Delta t)|^2) \quad (3.9)$$

where $r(t, \Delta t) = X(t + \Delta t) - X(t)$.

Here $X(t)$ denotes the logarithm of the asset price at time t , i.e. $X(t) = \log S(t)$ so that $r(t, \Delta t)$ represents the logarithm return of the asset price for the time lapsing Δt and τ represent the time lag, so that $\text{corr}(|r(t+\tau, \Delta t)|^2, |r(t, \Delta t)|^2)$ denotes the correlation of return at time t and at time $t + \tau$.

3.6 Jump-Diffusion Model

The need to have a jump component is apparent if we look into high frequency data of asset price paths that show that the asset prices can jump in small and large scales. Also as indicated by Birge and Linetsky (2007) that a diffusion model, including a stochastic volatility model, tends to lose the kurtosis property in a high sampling data, e.g. a time series of daily data has less kurtosis than a time series of weekly or monthly data. But this is not true for a jump-diffusion process as a jump is dependent on other stochastic factors. Related to this observation, several pieces of evidence have indicated that a jump-diffusion process is needed to explain the steepness of short term implied volatilities. Branger (2004) shows with the analytical models to confirm that a jump-diffusion model tends to reveal the characteristic of short date options smile (the skewness of the distribution of asset price returns) better than a stochastic volatility model.

The typical jump-diffusion model consists of two components, a continuous component as in the BS model and the jump component. The jump component is assumed to be distributed independently and identically. The continuous component will account for the normal movement of the underlying asset while the jump component will account for rare events such as a jump on the new information or a

crash of the underlying asset. Thus one often sees a jump process or a like process included in every modern model of the asset price dynamic.

The dynamic of an asset in the jump-diffusion model under risk neutral measure is represented by

$$\frac{dS_t}{S_t} = (r_t - \lambda k_Y)dt + \sigma dW + (Y_t - 1)dN_t \quad (3.10)$$

and $\ln Y_t$ distributed as i.i.d. $N(u_J, v_J)$,

where r_t is the risk free rate, σ is the volatility of the underlying asset. The process Y_t denotes the jump ratio upon the arrival of a jump event. That is S_t jumps to $Y_t S_t$ when a jump occurs. In this case, the jump ratio will have the log-normal distribution with mean u_J and variance v_J . The parameter λ denotes the jump intensity of the Poisson process N_t and k_Y is defined to be the expectation of the term $(Y_t - 1)$. The term λk_Y is added in the asset dynamic equation to compensate for the drift in the jump term to make this asset price dynamic to be a martingale. The jump-diffusion model will have 3 more parameters over the BS model which are λ , u_J and v_J . For example when the estimation of λ , u_J and v_J reads 0.2, -0.3 and 0.12 respectively, this means the rare jump will happens 0.2 time per year on average with the average of jump ratio of -0.3 and the variance of the jump ratio is 0.12 percent.

As the term $\ln Y_t$ has the normal probability density function with mean u_J and variance v_J , then the probability density function of $\ln Y_t$ is given by

$$f(\ln Y_t; u_J, v_J) = \frac{1}{\sqrt{2\pi v_J}} \exp\left(-\frac{(\ln Y_t - u_J)^2}{2v_J}\right).$$

We can calculate $k_Y \triangleq E[Y_t - 1]$ from the moment generating function $m(t)$ for random variable $Z = \ln Y_t$ which is given by

$$\begin{aligned} m(t) &= E[\exp(tZ)] \\ &= \int_{-\infty}^{\infty} \exp(tz) f(Z; u_J, v_J) dz \\ &= \exp(u_J t + \frac{1}{2} v_J t^2). \end{aligned} \quad (3.11)$$

By setting $t = 1$, we have

$$k_Y = E(Y_t - 1) = \exp(u_J + \frac{1}{2} v_J) - 1.$$

The price of an option under the jump-diffusion process can be derived by setting up the riskless portfolio which consists of a holding of Δ units of the underlying asset and a short selling in a unit of option $V(S_t, t)$. The portfolio value Π_t is then given by

$$\Pi_t = \Delta S_t - V(S_t, t). \quad (3.12)$$

By the Ito formula for a jump process, the differential of the price of the option can be expressed as

$$\begin{aligned} dV(S_t, t) &= \frac{\partial V(S_{t-}, t)}{\partial t} dt + \frac{\partial V(S_{t-}, t)}{\partial S_t} dS_{t-} + \frac{1}{2} \frac{\partial^2 V(S_{t-}, t)}{\partial S_t^2} dS_{t-}^2 \\ &\quad + [V(S_t, t) - V(S_{t-}, t)] dN_t, \end{aligned} \quad (3.13)$$

with $dS_{t-} = (r - k_Y \lambda) S_t dt + \sigma S_t dW_t$, $dS_t = (r - k_Y \lambda) S_t dt + \sigma S_t dW_t + (Y_t - 1) dN_t$,

$dS_{t-}^2 = \sigma^2 S_t^2 dt$ and $V(S_t, t) - V(S_{t-}, t) = V(Y_t S_{t-}, t) - V(S_{t-}, t)$.

Then the differential of the portfolio is given by

$$\begin{aligned} d\Pi_t &= -\left(\frac{\partial V(S_t, t)}{\partial t} + \frac{1}{2} \sigma^2 S_t^2 \frac{\partial^2 V(S_t, t)}{\partial S_t^2}\right) dt + \left(\Delta - \frac{\partial V(S_t, t)}{\partial S}\right) ((r - k_Y \lambda) S_t dt + \sigma S_t dW_t) \\ &\quad + \Delta (Y_t - 1) S_t - [V(Y_t S_t, t) - V(Y_t S_t, t)] dN_t. \end{aligned} \quad (3.14)$$

To hedge the diffusion risk, we set $\Delta = \frac{\partial V(S_t, t)}{\partial S}$. For the jump risk, Merton (1976) argued that the jump risk is a non-systematic risk and that it is diversifiable. By this argument, we can assume this portfolio is riskless and should have the same expected return as a riskless return. That is

$$E[d\Pi] = r\Pi dt. \quad (3.14)$$

From **Error! Reference source not found.** and (3.14), we will obtain the stochastic differential equation governing $V(S_t, t)$ under the jump-diffusion,

$$\begin{aligned} \frac{\partial V(S_t, t)}{\partial t} + \frac{\sigma^2}{2} S_t^2 \frac{\partial^2 V(S_t, t)}{\partial S_t^2} + (r - k_Y \lambda) S_t \frac{\partial V(S_t, t)}{\partial S_t} - rV(S_t, t) + \\ \lambda E[V(Y_t S_t, t) - V(S_t, t)] = 0. \end{aligned} \quad (3.15)$$

A more realistic assumption is that the jump risk is systematic and not diversifiable which makes the model incomplete. That is, the model needs to involve the investor risk preference. Some empirical evidences in the investor risk preference are studied in Pan (2002), Santa-Clara and Yan (2004) and many authors. The most common method to relate the investor risk preference is called the *equilibrium model method* which assumes that investors are rational and risk averse. In words, the model assumes that the investors try to maximize their utility function and require a risk premium to take more risk. The simplest version to include the risk premium into the jump-diffusion model is shown in Bates (1988) where the equation (3.15) is modified to be

$$\frac{\partial V(S_t, t)}{\partial t} + \frac{\sigma^2}{2} S_t^2 \frac{\partial^2 V(S_t, t)}{\partial S_t^2} + (r - k_y^\varrho \lambda^\varrho) S_t \frac{\partial V(S_t, t)}{\partial S_t} - rV(S_t, t) + \lambda^\varrho E^\varrho[V(Y_t S_t, t) - V(S_t, t)] = 0,$$

with

$$\lambda^\varrho = \lambda E[Y_t^{(\gamma-1)}], \quad k^\varrho = E^\varrho[Y_t - 1] \quad \text{where} \\ E^\varrho[X] = E[Y_t^{\gamma-1} X] / E[Y_t^{\gamma-1}].$$

(3.16)

The parameter γ is a *constant risk-aversion parameter* which is less than 1 for the risk averse case. The adjustment by the equilibrium method reveals what happens in the real market where investors price the option more expensive than the estimation from the historical price data which can be explained below.

Here we will derive u_j^ϱ and λ^ϱ which are the mean of jump ratio and frequency of the equilibrium model.

From $E^\varrho[X] = E[Y_t^{\gamma-1} X] / E[Y_t^{\gamma-1}]$ in (3.16), we then have

$$f^\varrho(y) \propto Y_t^{\gamma-1} f(y) \\ \propto Y_t^{\gamma-1} \exp\left[-\frac{(x - u_j)^2}{2v_j}\right] \\ \propto \exp\left[-\frac{(x - u_j^\varrho)^2}{2v_j}\right],$$

where $u_j^\varrho = u - (1 - \gamma)v_j$.

From (3.16) and by (3.11) with $t = \gamma - 1$

$$\lambda^\varrho = \lambda E[Y_t^{(\gamma-1)}] = \lambda \exp\left[(\gamma - 1)u_j + \frac{1}{2}(\gamma - 1)^2 v_j\right].$$

Suppose the estimation from the historical price of the underlying asset gives $\lambda = 0.10$, $u_j = -0.25$ and $v_j = 0.05$. With $\gamma = -1.0$, the mean of jump ratio and the frequency of jumps can be computed from the two above equations as -0.35 and 0.18

respectively. The lower number of γ will result in the higher mean of jump ratio and frequency of jumps. This means, the higher the risk aversion, the higher the risk premium for option prices.

3.7 Stochastic Volatility Model

While the jump-diffusion model can correct the behavior of short term skew feature of option implied volatility, the skew of volatility tends to flatten out as time goes by. The stochastic volatility model tends to do better to explain the skew from the medium term according to Branger (2004). In addition, the stochastic volatility model can also explain the volatility clustering and leverage effect properties. One of the most popular stochastic volatility models is the Heston model in Heston (1993). This model can be expressed in terms of the historical measure as

$$\begin{aligned} dS_t &= \mu_t S_t dt + \sqrt{v_t} S_t dW_t, \\ dv_t &= k(\theta - v_t)dt + \beta \sqrt{v_t} dW_t^v, \\ dW_t dW_t^v &= \rho dt, \end{aligned} \quad (3.17)$$

where S_t denotes the underlying asset price at time t , μ_t denotes the expected asset return rate, $\sqrt{v_t}$ denotes the volatility of the asset price and W_t is the Brownian motion associated with the underlying asset price process. v_t is the variance process, k is the speed of mean reversion, θ is the mean of long term variance, β is the volatility of variance process and W_t^v is the Brownian motion associated with the variance process. From empirical studies, there exists some correlation between these two processes and here we write this relation in the third equation.

As there are two stochastic factors that are W_t and W_t^v , the Heston model is not complete. The completion of the model is done by adding another security that has the W_t^v component. This is normally done by adding another option, that has longer maturity than the target option which we want to price, in the riskless portfolio.

So here we have 3 assets that are the underlying asset price S_t , the money market asset B_t and the option C_t that has time to maturity longer than the target option C_t^* whose value we want to find. We will set up the riskless portfolio Π_t , to replicate the value of this option C_t^* . Here we set up Π_t by

$$\Pi_t = C_t^*(S_t, v_t, t) - \Delta S - \Delta_1 C_t(S_t, v_t, t) - \Delta_2 B_t. \quad (3.18)$$

That is, the riskless portfolio consists of 1 unit of option C_t^* , short of Δ units of underlying asset, short of Δ_1 units of option C_t and short of Δ_2 units of money market asset. The change of this portfolio over t to $t+dt$ is given by (By the multidimensional Ito formula)

$$\begin{aligned} d\Pi_t = & \left[\frac{\partial C_t^*}{\partial t} + \frac{1}{2} v_t S_t^2 \frac{\partial^2 C_t^*}{\partial S_t^2} + \rho \beta v_t S_t \frac{\partial^2 C_t^*}{\partial S_t \partial v_t} + \frac{1}{2} \beta^2 v_t \frac{\partial C_t^*}{\partial v_t^2} \right] dt - \Delta_2 r B dt \\ & - \left[\frac{\partial C_t}{\partial t} + \frac{1}{2} v_t S_t^2 \frac{\partial^2 C_t}{\partial S_t^2} + \rho \beta v_t S_t \frac{\partial^2 C_t}{\partial S_t \partial v_t} + \frac{1}{2} \beta^2 v_t \frac{\partial C_t}{\partial v_t^2} \right] dt \\ & + \left[\frac{\partial C_t^*}{\partial S_t} - \Delta_1 \frac{\partial C_t}{\partial S_t} - \Delta \right] dS_t - \left[\frac{\partial C_t^*}{\partial v_t} - \Delta_1 \frac{\partial C_t}{\partial v_t} \right] dv_t. \end{aligned}$$

To make this portfolio riskless, we set $\frac{\partial C_t^*}{\partial S_t} - \Delta_1 \frac{\partial C_t}{\partial S_t} - \Delta$ and $\frac{\partial C_t^*}{\partial v_t} - \Delta_1 \frac{\partial C_t}{\partial v_t}$ to 0 so

that there is no dependence on these stochastic variables.

Then the portfolio will be riskless and leave us with only deterministic part

$$\begin{aligned}
d\Pi_t = & \left[\frac{\partial C_t^*}{\partial t} + \frac{1}{2} v_t S_t^2 \frac{\partial^2 C_t^*}{\partial S_t^2} + \rho \beta v_t S_t \frac{\partial^2 C_t^*}{\partial S_t \partial v_t} + \frac{1}{2} \beta^2 v_t \frac{\partial^2 C_t^*}{\partial v_t^2} \right] dt \\
& - \Delta_2 r B_t dt - \left[\frac{\partial C_t}{\partial t} + \frac{1}{2} v_t S_t^2 \frac{\partial^2 C_t}{\partial S_t^2} + \rho \beta v_t S_t \frac{\partial^2 C_t}{\partial S_t \partial v_t} + \frac{1}{2} \beta^2 v_t \frac{\partial^2 C_t}{\partial v_t^2} \right] dt
\end{aligned}
\tag{3.19}$$

As the portfolio is riskless, the return should be the riskless interest rate. That is

$$d\Pi_t = r\Pi_t dt = r[C_t^* - \Delta S - \Delta_1 C_t - \Delta_2 B_t] dt .$$

Substituting $d\Pi_t$ from (3.19) and after some rearranging, we obtain

$$\begin{aligned}
\left(\frac{\partial C_t^*}{\partial t} + AC_t^* + rS \frac{\partial C_t^*}{\partial S} - rC_t^* \right) / \frac{\partial C_t^*}{\partial v_t} = & \\
\left(\frac{\partial C_t}{\partial t} + AC_t + rS \frac{\partial C_t}{\partial S} - rC_t \right) / \frac{\partial C_t}{\partial v_t} & \tag{3.20}
\end{aligned}$$

with
$$AC_t^* = \frac{1}{2} v_t S_t^2 \frac{\partial^2 C_t^*}{\partial S_t^2} + \rho \beta v_t S_t \frac{\partial^2 C_t^*}{\partial S_t \partial v_t} + \frac{1}{2} \beta^2 v_t \frac{\partial^2 C_t^*}{\partial v_t^2} .$$

Since the left hand side is a function of C_t^* and the right hand side is a function of C_t , the possible way is each side has to be a function of the independent variables S_t, C_t and t . So we write this as

$$\frac{\partial C_t}{\partial t} + AC_t + rS \frac{\partial C_t}{\partial S} - rC_t = -f(S_t, v_t, t) \frac{\partial C_t}{\partial v_t} .$$

Here we set it as the negative of function f . Heston (1993) shows that the function $f(S_t, v_t, t)$ is $(\alpha - \lambda_v \beta \sqrt{v_t})$ where $\alpha = k(\theta - v_t)$ and λ_v is called the *market price of volatility risk*. This is the partial differential equation that governs the price of an option under the Heston stochastic volatility model. Together with the payoff function, we can find an option price formula for that payoff.

CHAPTER IV

MODEL SPECIFICATIONS

4.1 Introduction

In this chapter we present a pricing derivation method for a jump-diffusion model with stochastic volatility and stochastic interest rate (JDSVSI), where the main component is the jump-diffusion process with Brownian motion part which is subordinated with a time integral of a CIR process (Cox, Ingersoll Jr. and Ross, 1985) and the interest rate in the model is stochastic. We mainly derive the pricing formula of a European call option by the combination of the Lewis Fourier transform method (Lewis, 2001) and modular Pricing method (Zhu, 2010). Our stochastic volatility is generated by the mentioned subordinated Brownian motion which is a special case of a time changed Levy process. We construct this time changed Brownian motion following the work of Carr and Wu (2004) which allows the leverage effect to be incorporated in the model. By having this leverage feature, Carr and Wu (2004) have handled the impact of the correlation of the asset price process and variance process by employing the leverage-neutral measure in the complex domain.

4.2 The Typical Model

The typical risk neutral model for jump-diffusion with stochastic volatility and stochastic interest rate can be described by

$$\begin{aligned}
\frac{dS_t}{S_t} &= (r_t - \lambda k_Y)dt + \sqrt{v_t}dW_t + (Y_t - 1)dN_t, \\
&\text{with } \ln Y_t \text{ distributed as } Normal(\mu_J, v_J), \\
E[Y_t - 1] &= \exp(\mu_J + \frac{1}{2}v_J) - 1 \equiv k_Y, \\
&\text{and } dv_t = k(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^v, \\
&\text{with } dW_t dW_t^v = \rho dt, \\
dr_t &= \alpha(\omega - r_t)dt + \beta\sqrt{r_t}dW_t^r.
\end{aligned} \tag{4.1}$$

Here W_t , W_t^v and W_t^r are the Brownian motions associated to the underlying asset process, the variance process and the interest rate process, respectively. The process S_t is the underlying asset price process and r_t is the instantaneous risk free rate process. The process N_t is a Poisson process with jump frequency λ and independent from the other processes. The jump size is $Y_t - 1$ and Y_t is log-normally distributed with mean μ_J and variance v_J . The jump component is included in the model to make the short term implied volatility curve steep as indicated by empirical studies. The process v_t is the variance process, k is the speed of mean reversion, θ is the mean of long term variance and σ is the volatility of the variance process. The variance process is the mean reverting square root process known as the CIR process. To explain the leverage effect, a negative correlation is usually introduced between the underlying asset and the variance as shown above. The interest rate process is also the CIR process but with different parameters and independent from the other processes.

4.3 The Lewis Fourier Transform Method

We derive the pricing formula of this model using the Lewis Fourier transform method. There is a variety of Fourier Transform Pricing methods but we choose this method as its integration domain is on the complex plane. This complex domain will

correspond to the domain for the time changed Levy process. Another nice feature of this method is that it produces a formula in a single integration form compared with the typical approaches which produce two integrations. This single integral reduces computation time of option prices in the calibration process. During our calculation, we also apply the Modular Pricing method, introduced in Zhu (2010), which employs the rule of independence of characteristic functions to write the characteristic function as product of each characteristic function of an independent stochastic factor. This approach will help us to handle each stochastic factor independently which results in the reduction in the dimensions of problem.

Our dynamic of asset price will be an exponential Levy process which is driven as

$$\begin{aligned} S_t &= S_0 \exp\left(\int_0^t r_s ds + L_t\right) \\ &= \exp\left(X_0 + \int_0^t r_s ds + L_t\right). \end{aligned} \tag{4.2}$$

Here S_t is the stock price at time t , $S_0 = \exp X_0$ is the price of stock at time $t=0$, r_t is the risk free rate process and L_t is a Levy process with e^{L_t} being a martingale. Let us assume that the characteristic function of the process L_t , $\phi_{L_t}(z) = E^Q[\exp(izL_t)]$ is well defined for $\alpha < \text{Im}(z) < \beta$ where α and β are real numbers and z is a complex number. For a European call option strike at K with a payoff function at the maturity, $\max(S_T - K, 0)$ or denoted by $(S_T - K)^+$, the Fourier transform of the payoff with transform $\hat{H}(z)$ can be computed as

$$\begin{aligned}
\hat{H}(z) &= \int_{-\infty}^{\infty} \exp(izx)(\exp(x) - K)^+ dx \\
&= \int_{\ln K}^{\infty} \exp(izx)(\exp(x) - K) dx \\
&= \left(\frac{\exp(iz+1)x}{iz+1} - K \frac{\exp(izx)}{iz} \right) \Big|_{\ln K}^{\infty} \\
&= \left[\lim_{x \rightarrow \infty} \left(\frac{\exp(iz+1)x}{iz+1} - K \frac{\exp(izx)}{iz} \right) \right] - \\
&\quad \left[\frac{\exp(iz+1) \ln K}{iz+1} - K \frac{\exp(izx) \ln K}{iz} \right] \\
&= -\frac{K^{iz+1}}{z^2 - iz}, \quad \text{given } \text{Im}(z) > 1.
\end{aligned} \tag{4.3}$$

The first term on the fourth line above is zero by substituting $z = z_r + iz_i$ and employing the Euler identity as follows,

$$\begin{aligned}
&\left[\lim_{x \rightarrow \infty} \left(\frac{\exp(iz+1)x}{iz+1} - K \frac{\exp(izx)}{iz} \right) \right] \\
&= \lim_{x \rightarrow \infty} \left(\frac{\exp(iz_r x - (z_i - 1)x)}{iz+1} - K \frac{\exp(iz_r x - xz_i)}{iz} \right) \\
&= \lim_{x \rightarrow \infty} (\cos z_r x - \cos z_r x) \exp((1 - z_i)x) - K \lim_{x \rightarrow \infty} (\cos z_r x - \cos z_r x) \exp(-z_i x) \\
&= 0 \text{ if } z_i > 1 \text{ and not defined if } z_i \leq 1.
\end{aligned}$$

In (4.3), $\hat{H}(z)$ is defined in the region where the imaginary part of Fourier transform variable z , is greater than 1. The corresponding *generalized inverse Fourier transform* $H(x)$ for the payoff function is defined below,

$$H(x) = \frac{1}{2\pi} \int_{iz_i - \infty}^{iz_i + \infty} e^{-izx} \hat{H}(z) dz. \tag{4.4}$$

In (4.3) and (4.4), we extend the transform variable z to take a value in the complex domain that is defined in the generalized Fourier transform sense. Given, $\hat{H}(z)$ is well defined on the plane where the imaginary part of z is greater than 1, the integration in (4.4) is just integration in the complex plane on a line paralleled to the real axis with

any $z_i > 1$. From the Fundamental Theorem of Asset Pricing, the no arbitrage condition is equivalent to the existence of a risk neutral measure where a discount asset price is a martingale. Based on this Fundamental Theorem, we can write the value of a European call option at time $t = 0$ as the risk neutral expectation of the discount payoff,

$$\begin{aligned}
 V_0 &= E^Q[\exp(\int_0^T -r_t dt)(S_T - K)^+] \\
 &= E^Q[\exp(\int_0^T -r_t dt)(e^x - K)^+] \\
 &= \frac{1}{2\pi} E^Q[\int_{iz_i - \infty}^{iz_i + \infty} \exp(\int_0^T -r_t dt) \exp(-iz(x_0 + \int_0^T r_t dt + L_T)) \hat{H}(z) dz] \\
 &= \frac{1}{2\pi} \int_{iz_i - \infty}^{iz_i + \infty} E^Q[\exp(-iz(x_0 + \int_0^T r_t dt))] \exp(-izx_0) \phi_{L_T}(-z) \hat{H}(z) dz.
 \end{aligned} \tag{4.5}$$

Here $\exp(X_T) = S_T$ or $X_T = X_0 + \int_0^T r_t dt + L_T$. The expectation E^Q is the expectation under a risk neutral measure. The third line is derived from the second line by replacing the payoff function with the corresponding generalized Fourier transform in (4.4). In (4.5), here we suppose the interest rate process is independent from the other processes, therefore we can write the expectation out from the other terms.

4.4 Derivation of the Pricing Formula

The time changed Levy process was introduced in financial modeling by Clark (1973) where he used a subordinated Brownian motion to interpret the relationship between the return of an asset price and the volume traded in the market instead of the usual time variable. Monroe (1978) has later proved that any semi-martingale (a drift process plus a martingale process) can be written as a time changed Brownian motion.

This proof has validated the uses of a time changed Brownian motion as an asset price dynamic. Since then the use of a time changed Levy process as an asset price dynamic has grown substantially.

The idea to create a stochastic volatility by the *Time Changed Levy method* is related to a subordinated Levy process (cf. Chapter2) where we change the *calendar time* of a pure Levy process to the *operational time* and expressed as a process T_t , a function of calendar time t , which is a subordinator process. So the stochastic volatility of the model is generated by the speed of the change of this process. The most common of a subordinator process is the integral of the variance process i.e.

$T_t = \int_0^t v_s ds$ where T_t is the subordinator to represent the random nature of operational time and v_s is the variance process as in (4.1). The process v_s is called the *instantaneous activity rate* in Carr and Wu (2004).

Here we will apply the time changed Levy method to generate the stochastic volatility in the model. That is our model is driven by the time changed Levy process and the compensated jump-diffusion process with the stochastic interest rate. Therefore the asset price dynamic can be written as

$$S_t = S_0 \exp\left(\int_0^t r_s ds + X_{T_t} + \sum_{k=1}^{N_t} \ln Y_k - \lambda k_y t\right),$$

where

$$T_t = \int_0^t v_s ds \text{ with} \tag{4.6}$$

$$dv_t = k(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^v, \text{ and}$$

$$dr_t = \alpha(\omega - r_t)dt + \beta\sqrt{r_t}dW_t^r.$$

The variables $r_t, Y_k, \lambda, k_y, v_t, k, \theta, \sigma, \alpha, \omega$ and β are defined as in (4.1) and X_{T_t} is a subordinated Brownian process which is defined later. Compared with (4.2), the Levy part in (4.6) is

$$\begin{aligned} L_t &= X_{T_t} + \sum_{k=1}^{N_t} \ln Y_k - \lambda k_y t, \\ &= X_{T_t} + J_t. \end{aligned} \quad (4.7)$$

The term $\sum_{k=1}^{N_t} \ln Y_k - \lambda k_y t := J_t$ is a compensated compound Poisson process. Our time changed Levy process X_{T_t} is constructed by two stochastic processes; a subordinator and an underlying Levy process, The subordinator T_t is defined to be a process

$$T_t = \int_0^t v_s ds, \quad (4.8)$$

where v_s is the variance at time t defined as in (4.1). The underlying Levy process is the risk neutral Brownian motion with drift rate equaled to a risk free rate r_t ,

$$dS_t = S_t(r_t dt + dW_t), \quad (4.9)$$

whose log return can be written by (by applying the Ito formula to the function $f(S_t) = \ln S_t$)

$$\ln S_t / S_0 = rt + W_t - \frac{1}{2}t = rt + X_t. \quad (4.10)$$

As mentioned previously, the time changed Levy process X_{T_t} can be generated by substituting T_t for t in (4.10). And X_{T_t} has the form

$$X_{T_t} = W_{T_t} - \frac{1}{2}T_t. \quad (4.11)$$

Remark: Even though our original Levy part has two components; the Brownian motion part and the Jump process part, we reserve to subordinate only the Brownian part by the following rationale. The jump process part will only account for the rare events and is independent from the other process. Then if we subordinate the jump part, we need to address the problem of correlation between a jump process and the variance process. This will break down the assumption of independent rare events.

We can now start to derive the pricing formula of this model by employing the formula (4.5)

$$V_0 = \frac{1}{2\pi} \int_{iz_i - \infty}^{iz_i + \infty} E^Q[\exp(-(iz + 1) \int_0^T r_t dt)] \exp(-izx_0) \phi_{L_T}(-z) \hat{H}(z) dz. \quad (4.12)$$

By assumption of the independence between the time changed Levy process and the jump process, we may write the characteristic function of the Levy process $\phi_{L_T}(-z)$ from (4.12) as the product of the characteristic function of the time changed Levy process and the characteristic function of the compensated compound jump process,

$$\phi_{L_T}(-z) = \phi_{X_{T_i}}(-z) \phi_{J_T}(-z). \quad (4.13)$$

We denote $\phi_{X_{T_i}}(u)$ and $\phi_{J_T}(u)$ as the characteristic functions of the time changed Levy process and of the compensated Poisson process respectively. So we now need to calculate each component of (4.5) that are $E^Q[\exp(-(iz + 1) \int_0^T r_t dt)]$, $\phi_{X_{T_i}}(u)$ and $\phi_{J_T}(u)$.

The characteristic function of the time changed Levy process is derived by

$$\begin{aligned}
\phi_{X_{T_t}}(z) &= E^Q[e^{iu(W_{T_t} - \frac{1}{2}T_t)}] \\
&= E^Q[e^{iu(W_{T_t} - \frac{1}{2}T_t) + \psi_X(z)T_t - \psi_X(z)T_t}] \\
&= E^M[\exp(-\psi_X(z)T_t)] \\
&= E^M[\exp(\int_0^T -\psi_X(z)v_s ds)].
\end{aligned} \tag{4.13}$$

From the second line to the third line above, we apply the measure change defined as the complex valued Radon-Nikodym derivative and this measure is called the *leverage neutral measure* M (the detail of this measure can be found in Carr and Wu (2004)) and has the form,

$$\frac{dM}{dQ} \Big|_t = \exp(izX_{T_t} + T_t\psi_X(z)), \tag{4.14}$$

where $\psi_X(z)$ denotes the characteristic exponent of the underlying process X_{T_t} . This measure allows us to write the characteristic function of the correlated processes (in our case, the underlying process is designed to correlate with the variance process) as the Laplace transform under a new measure. That is

$$E^Q[e^{izX_{T_t}}] = E^M[e^{-\psi_X(z)T_t}]. \tag{4.15}$$

Here the characteristic exponent of the Levy process $\psi_X(z)$ of the process $W_t - \frac{1}{2}t$ is

$\frac{1}{2}(iz + z^2)$. This result is close to Theorem 2.27 (Bochner) but the Bochner law

applies to the case where the subordinator is not correlated to the base process. To

calculate $T_t = \int_0^t v_s ds$ in (4.15), we need to find the dynamic of the variance process

under this new measure M . Based on the Girsanov theorem (cf. Theorem 2.9), given

a measurable space (Ω, \mathcal{F}, P) , the Ito process for the dynamic of v_t is

$$dv_t = k(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^v. \quad (4.16)$$

Denote M_t as an exponential martingale under measure Q defined by

$$M_t = \frac{dM}{dQ} \Big|_t = \exp(izX_{T_t} + \int_0^t v_s \psi_X(z) ds),$$

with $E^Q[M_t] = 1$.

Substitute $X_{T_t} = W_{T_t} - \frac{1}{2}T_t = \int_0^T \sqrt{v_s} dW_s - \int_0^T v_s ds$ and $\psi_X(z) = \frac{1}{2}(iz + z^2)$. Then M_T can

be expressed as,

$$\begin{aligned} \frac{dM}{dQ} \Big|_T &= \exp\left(iz\left(\int_0^T \sqrt{v_s} dW_s - \frac{1}{2}\int_0^T v_s ds\right) + \int_0^T v_s \left(\frac{1}{2}(iz + z^2) ds\right)\right) \\ &= \exp\left(\int_0^T iz\sqrt{v_s} dW_s - \frac{1}{2}\int_0^T v_s z^2 ds\right) \\ &= \exp\left(\int_0^T \gamma dW_s + \frac{1}{2}\int_0^T \gamma^2 ds\right) \text{ with } \gamma = iz\sqrt{v_t}. \end{aligned}$$

By the assumption that the underlying asset is correlated with the variance process or $dW_t dW_t^v = \rho dt$, then we have the following results,

- M_t defines the Radon Nikodym derivative. That is $M_t = \frac{dM}{dQ}$.
- The new Brownian motion W_t^{vM} under the measure M is defined by

$$dW_t^{vM} = dW_t^v - \gamma dW_t dW_t^v = dW_t^v - iz\sqrt{v_t} \rho dt.$$

Substituting dW_t^v in the third equation of (4.6), we have

$$\begin{aligned} dv_t &= k(\theta - v_t)dt + iz\rho\sigma v_t dt + \sigma\sqrt{v_t}dW_t^{vM}, \\ &= (k\theta - k^M v_t)dt + \sigma\sqrt{v_t}dW_t^{vM}, \text{ with} \\ k^M &= k - iz\rho\sigma. \end{aligned} \quad (4.17)$$

The variable $W_t^{v^M}$ is the Brownian motion associated to the variance process under this new measure. Then we can solve for the characteristic function of the time changed Levy process as

$$\begin{aligned}\phi_{X_{T_t}} &= \exp(-C(t) - D(t)v_0), \\ \text{with } C(t) &= \frac{k\theta}{-\sigma^2} \left[(k^M + d)t - 2 \ln \left[\frac{1 - ge^{dt}}{1 - g} \right] \right], \\ D(t) &= \frac{k^M + d}{-\sigma^2} \left[\frac{1 - e^{dt}}{1 - ge^{dt}} \right], \\ d &= \sqrt{(k^M)^2 + 2\psi_X \sigma^2}, \\ g &= \frac{(k^M + d)}{(k^M - d)}.\end{aligned}\tag{4.18}$$

Proof:

From (4.16) and according to the Feynman Kac theorem (cf. Theorem 2.8), the characteristic function of X_{T_t}

$$\phi_{X_{T_t}}(z) = E^M \left[\exp \left(\int_0^t -\psi_X(z) v_s ds \right) \right]$$

will satisfy the following partial differential equation (PDE),

$$\frac{\partial \phi(v_t, t)}{\partial t} = (k\theta - k^M v_t) \frac{\partial \phi(v_t, t)}{\partial v_t} + \frac{1}{2} \sigma \sqrt{v_t} \frac{\partial^2 \phi(v_t, t)}{\partial v_t^2} - \psi(v_t) \phi(x, t),$$

$$\text{with } \phi(v_0, t = 0) = 1 \tag{4.19}$$

$$\text{given that } dv_t = (k\theta - k^M v_t) dt + \sigma \sqrt{v_t} dW_t^v.$$

With the assumption that the characteristic function of the time changed Levy process has the form, $\phi_{X_{T_t}}(z) = \exp(-C(t) - D(t)v_t)$, we can substitute $\phi_{X_{T_t}}(z)$ in the above PDE where (Here we write $\phi(v_t, t)$ for $\phi_{X_{T_t}}(z)$ to signify the dependence of the characteristic function on the variables v_t and t .)

$$\begin{aligned}\frac{\partial \phi(v_t, t)}{\partial t} &= \phi(v_t, t) \left(-\frac{\partial C(t)}{\partial t} - \frac{\partial D(t)}{\partial t} v_t \right), \\ \frac{\partial \phi(v_t, t)}{\partial v_t} &= \phi(v_t, t) (-D(t)), \\ \frac{\partial \phi^2(v_t, t)}{\partial v_t^2} &= \phi(v_t, t) (D^2(t)),\end{aligned}$$

which produces

$$\begin{aligned}\left(-\frac{\partial C(t)}{\partial t} - \frac{\partial D(t)}{\partial t} v_t \right) \phi(v_t, t) &= k^M (\theta - v_t) (-D(t)) \phi(v_t, t) + \frac{1}{2} \sigma \sqrt{v_t} (D^2(t)) \phi(v_t, t) \\ &\quad - \psi_x(z) \phi(v_t, t).\end{aligned}$$

The above PDE can be reduced to two differential equations

$$\begin{aligned}\frac{dC(t)}{dt} &= D(t) k^M \theta, \\ \frac{dD(t)}{dt} &= \psi(z) - k^M D(t) - \frac{1}{2} \sigma^2 D^2(t),\end{aligned}\tag{4.20}$$

and by the initial condition $\phi(v_0, T=0) = 1$ then $C(0) = 0$ and $D(0) = 0$.

The differential equation for $D(t)$ as in (4.20) is called the *Riccati equation* which is a nonlinear ordinary differential equation. To simplify it, we will write the equation for $D(t)$ as

$$\begin{aligned}\frac{\partial D(t)}{\partial t} &= P - QD(t) + RD^2(t), \\ \text{where } P &= \psi_x, \quad Q = k^M \text{ and } R = -\frac{1}{2} \sigma^2.\end{aligned}$$

The solution for $D(t)$ will be $D(t) = -\frac{1}{R} \frac{u'}{u}$ where u satisfies the following auxiliary

differential equation

$$u'' + \left[\frac{P'}{P} + Q \right] u' + PR = 0.$$

The general solution for the above equation is

$$u(t) = Ae^{\alpha t} + Be^{\beta t},$$

where A and B are constant,

$$\text{with } \alpha = \frac{-[\frac{P'}{P} + Q] + \sqrt{[\frac{P'}{P} + Q]^2 - 4PR}}{2},$$

$$\text{and } \beta = \frac{-[\frac{P'}{P} + Q] - \sqrt{[\frac{P'}{P} + Q]^2 - 4PR}}{2}.$$

Therefore,

$$D(t) = -\frac{A\alpha e^{\alpha t} + B\beta e^{\beta t}}{Ae^{\alpha t} + Be^{\beta t}} \cdot \frac{1}{R(t)}.$$

Substitute the values of P,Q,R and we have

$$\alpha = \frac{-k^M + d}{2} \quad \text{and} \quad \beta = \frac{-k^M - d}{2}$$

$$\text{where } d = \sqrt{(k^M)^2 + 2\psi_x \sigma^2}.$$

With $D(0) = 0$, we can solve for

$$D(t) = \frac{k^M + d}{-\sigma^2} \left[\frac{1 - e^{dt}}{1 - ge^{dt}} \right],$$

$$\text{with } g = \frac{(k^M + d)}{(k^M - d)}.$$

The solution for $C(t)$ can be solved by integrating (4.20). Therefore,

$$\begin{aligned} C(T) &= \int_0^T k\theta \frac{k^M + d}{-\sigma^2} \left[\frac{1 - e^{dt}}{1 - ge^{dt}} \right] dt \\ &= k\theta \frac{k^M + d}{-\sigma^2} \int_0^T \frac{1 - e^{dt}}{1 - ge^{dt}} dt. \end{aligned}$$

By having $x = \exp(dt)$ which makes $dx = xddt$ and $dt = \frac{1}{xd} dx$, substitute this dt in

the above equation, thus

$$\begin{aligned}
C(T) &= \frac{k\theta}{d} \frac{k^M + d}{-\sigma^2} \int_1^{\exp(d.T)} \frac{1-x}{1-gx} \frac{1}{x} dx \\
&= \frac{k\theta}{d} \frac{k^M + d}{-\sigma^2} \int_1^{\exp(d.T)} \left[\frac{1}{x} - \frac{1-g}{1-gx} \right] dx \\
&= \frac{k\theta}{d} \frac{k^M + d}{-\sigma^2} \left[\ln x + \frac{1-g}{g} \ln(1-gx) \right]_1^{\exp(d.T)} \\
&= \frac{k\theta}{d} \frac{k^M + d}{-\sigma^2} \left[d.T + \frac{1-g}{g} \ln\left(\frac{1-g \exp(d.T)}{1-g}\right) \right].
\end{aligned}$$

By rearranging the above equation, we have

$$C(T) = -\frac{k\theta}{\sigma^2} \left[(k^M + d)T - 2 \ln\left[\frac{1-g \exp(d.T)}{1-g}\right] \right]. \quad (4.21)$$

The computation of the stochastic interest rate part is similar to the calculation of the characteristic function of the time changed Levy part. We can write

$$\begin{aligned}
E^Q[\exp(-(iz+1) \int_0^T r_t dt)] &= \exp(G(t)r_0 + H(t)), \\
\text{with } G(t) &= \frac{\alpha + d}{\beta^2} \left(\frac{1 - e^{-dt}}{1 - ge^{-dt}} \right), \\
H(t) &= \frac{\alpha\omega}{\beta^2} \left[(\alpha + d)t - 2 \ln\left(\frac{1 - ge^{-dt}}{1 - g}\right) \right], \quad (4.22)
\end{aligned}$$

where

$$d = \sqrt{\alpha^2 + 2(iz+1)\beta^2} \quad \text{and} \quad g = \frac{\alpha + d}{\alpha - d}.$$

The last part is the characteristic of the compensated compound Poisson process which is computed below,

$$\begin{aligned}
\phi_{J_r}(z) &= E^Q[\exp(iz(\sum_{k=1}^{N_r} \ln Y_k - \lambda k_Y t))] \\
&= E^Q[\exp(iz \sum_{k=1}^{N_r} \ln Y_k)] \exp(-iz \lambda k_Y t).
\end{aligned}$$

By (2.13)

$$\phi_{J_T}(z) = \exp(-iz\lambda k_Y t) \exp\left(\lambda t \int_{-\infty}^{+\infty} (\exp(izx) - 1) \eta(dx)\right),$$

with $k_Y = \exp(u_J + \frac{1}{2}v_J) - 1$ and in (3.11) for $m(t)$ with $t = iz$,

$$\begin{aligned} \phi_{J_T}(z) &= \exp(-iz\lambda t (\exp(u_J + \frac{1}{2}v_J) - 1)) \exp(\lambda t (\exp(izu_J - \frac{1}{2}z^2v_J) - 1)) \\ &= \exp(-iz\lambda t (\exp(u_J + \frac{1}{2}v_J) - 1) + (\lambda t (\exp(izu_J - \frac{1}{2}z^2v_J) - 1))). \end{aligned} \quad (4.23)$$

4.5 Numerical Integration

4.5.1 The Trapezoidal Rule

As in (4.5), the value of a European call option, which is in the semi-closed form formula, requires an integration over the complex-valued Fourier transform variable. We use a numerical integration method to find the value of an option as the closed-form formula is hard to derive. Since this semi-closed form formula is in a single Fourier integral form, the computation for the price of the option is more efficient than the typical formula which has two integrals. The only difficulty is that this integral is on the complex plane, but it does not require special treatment as this integral is only a contour integration on the complex plane. We present here the algorithm for the numerical integration.

The normal way of a numerical integration is to represent an integral as the rule of approximated calculation of an integral denoted by $I[f]$,

$$\int_a^b f(x) dx \approx I[f], \quad (4.24)$$

where $f(x)$ is the integrand and $[a,b]$ is finite integral interval. Here we will employ the Trapezoidal rule for the numerical integration. The trapezoidal rule divides the

integral interval in to small subintervals along the x-axis. In each subinterval, the integrand $f(x)$ is approximated by a straight line joining the coordinate $(x_i, f(x_i))$ and $(x_{i+1}, f(x_{i+1}))$ as shown in the Figure 4.1, the area in the small interval is computed as

$$\frac{1}{2}(x_i - x_{i-1})[f(x_i) + f(x_{i+1})] = \frac{\Delta x}{2}[f(x_i) + f(x_{i+1})]$$

Therefore the area under the graph from a to b is calculated as the sum of the small trapezoids from a to b as

$$I[f] = \frac{1}{2} \sum_{i=0}^{n-1} (x_{i+1} - x_i)[f(x_{i+1}) + f(x_i)] \quad (4.25)$$

with $a = x_0 < x_1 < \dots < x_n = b$

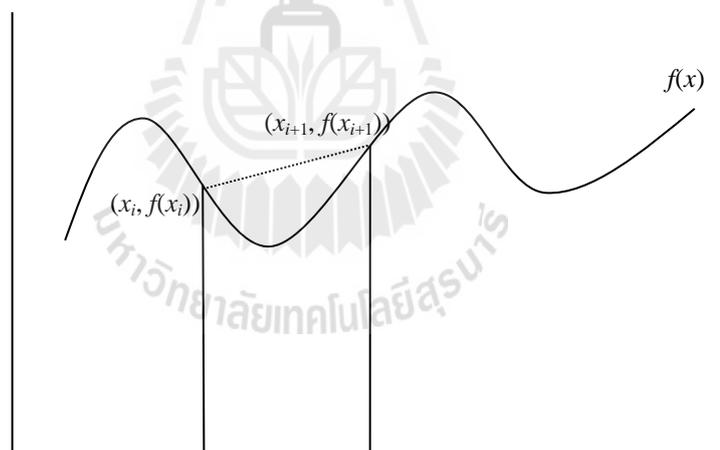


Figure 4.1 Illustration of Trapezoid rule.

4.5.2 Numerical Calculation of the Pricing Formula

The integration, in the formula (4.5), is an indefinite contour integral in the complex plane where the transform variable z is well defined in the area where the imaginary part of z is greater than 1 as described in (4.3). In actual calculation, the integration interval of the infinite integral is truncated to an appropriate finite interval

lying between the upper and lower bounds of integration interval. There are 2 problems to be addressed for the numerical integration of the inverse Fourier transform,

- The set up of upper and lower bounds of truncated integration interval and
- The stability of the integrand.

For the first problem, generally these upper and lower bounds of integration are set to be high enough for that the integral part beyond this limit is below a small

tolerance number ε i.e. $\int_{iz+b_{\max}}^{iz+\infty} |I(z)| dz < \varepsilon$ and $\int_{iz-\infty}^{iz-b_{\max}} |I(z)| dz < \varepsilon$. The considerations for

setting up the interval for the Fourier integral are mentioned in Schmelzle (2010) and Chourdakis (2008) whose main concept is related to the fat tail of the distribution of the asset price dynamic e.g. for a BS model whose fat tail distribution is low, the interval can be narrower than for a model that has higher fat tail distribution.

For a reason of convenience and by observing the charts of the integrands of the pricing formula when $T = 0.5$ and 1.0 which are shown in Figure 4.2 and 4.3 respectively, the integrands of the pricing formula reach a level of close to zero rapidly and have quite symmetric distributions. Therefore, we set the same for lower and upper bounds as $-b_{\max}$ and b_{\max} . Then we set this b_{\max} at 500 and this is validated by Table 4.1 which shows the option prices calculation when we set b_{\max} at 200, 500 and 700. The calculation shows that there is no difference in the prices when b_{\max} is greater than 500.

To handle the second problem, we need to observe the characteristic of the integrand. The integrand of the pricing formula is the product of the characteristic

function $\phi_L(-z)$ and the term $\hat{H}(z)$ as in (4.5). The characteristic function of the random variable is always continuous (Chapter II) and the term $\hat{H}(z) = \frac{-K^{iz+1}}{z^2 - iz}$ is a smooth function as long as $\text{Im}(z) > 1$. By this fact, we can apply the Trapezoid rule to the semi closed form formula (4.5) without any special treatment, and the result is shown in Table 4.1. The MATLAB code for this pricing formula is listed in the Appendix.

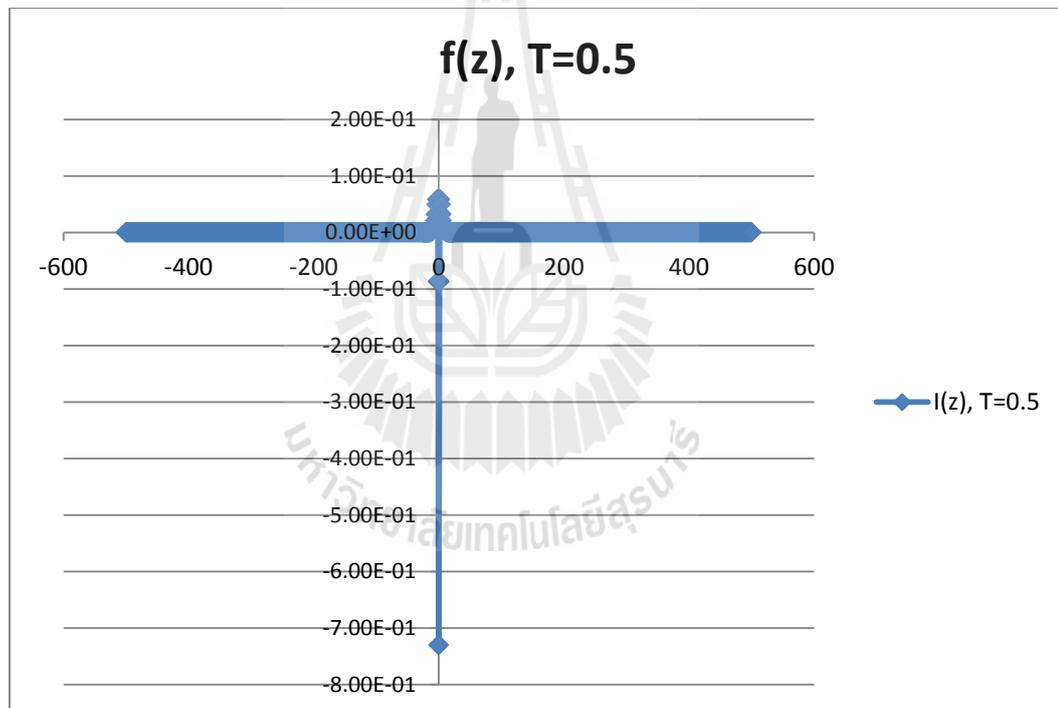


Figure 4.2 The integrand function $f(z)$ with $T=0.5$.

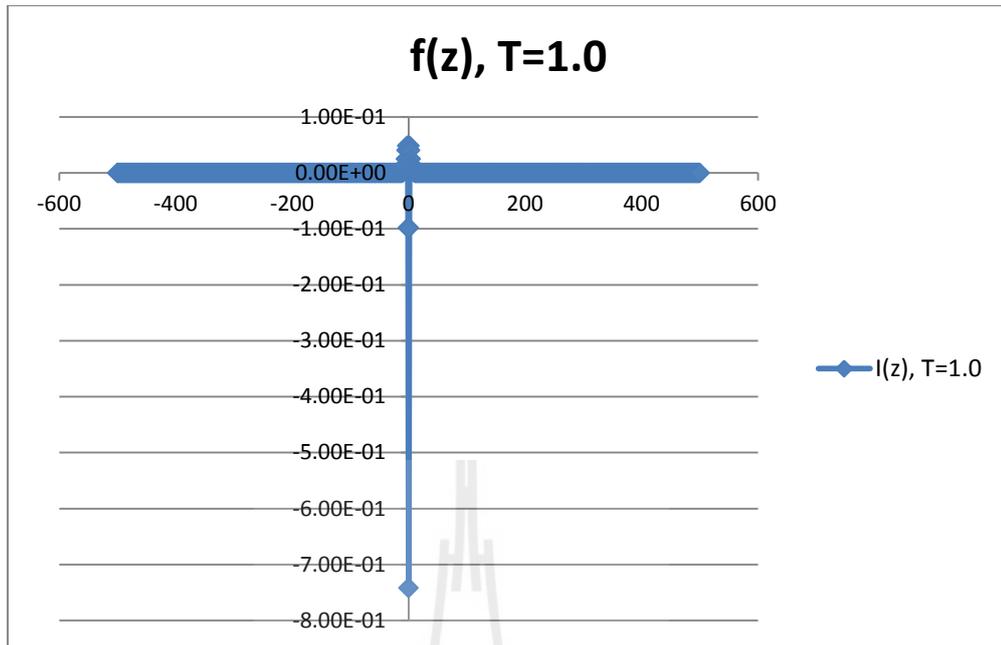


Figure 4.3 The integrand function $f(z)$ with $T=1.0$.

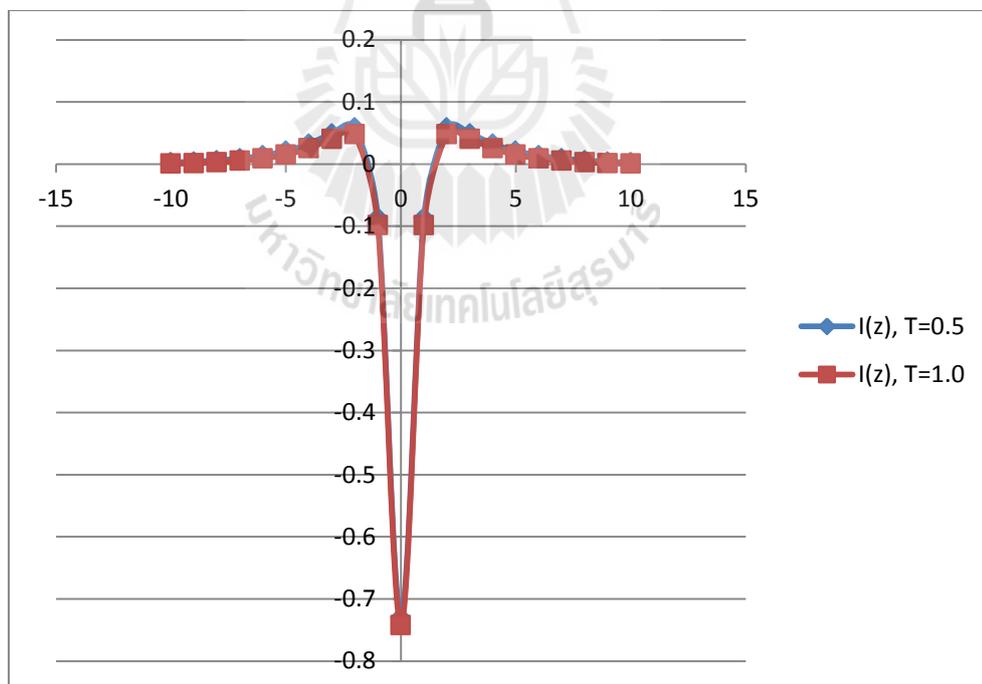


Figure 4.4 The integrand function $f(z)$ near the origin.

Table 4.1 Option prices and execution times.

Maturity	Option Price $b_{\max} = 200$	Execution Time (sec.)	Option Price $b_{\max} = 500$	Execution Time (sec.)	Option Price $b_{\max} = 700$	Execution Time (sec.)
T=0.5	426.0795	0.0313	426.1857	0.0625	426.1857	0.1563
T=1.0	596.0210	0.0313	596.0221	0.0625	596.0221	0.1563
T=2.0	857.8093	0.0313	857.8099	0.0625	857.8099	0.1563

CHAPTER V

CALIBRATION

5.1 Calibration Problem

Calibration is the process to obtain a model's parameters that match to the present market prices of options. These model's parameters generally will be different from parameters estimated by the statistical methods which usually derive parameters by estimating from statistical characteristics of time series data of underlying assets prices. The parameters extracted from statistical methods more or less reflect the past characteristics of the dynamic of the underlying asset, as discussed in the earlier chapter. These parameters do not reflect the present risk premium implied in the option prices. Thus the statistical estimation does not guarantee that the models built out of these parameters are arbitrage free. Contrary to the statistical method, the parameters from the calibration are arrived at with the principles of no arbitrage, and the model's parameters will match at least to the prices of traded options that are included in the calibration process. The differences of these two methods are not only the inclusion of investors risk preferences, but also hedging costs and views of the participants in the market which cannot be captured by a statistical method.

As we can observe the prices of the options from the market, the equation for the value of the option based on the risk-neutral valuation is

$$V_0(\theta_p, T, K) = E^Q[\exp(\int_0^T -r_t dt)(S_T - K)^+]. \quad (5.1)$$

Here θ_p is the set of the model parameters. The constants T and K are the time to maturity and strike of the observed option respectively. The dynamic of S_T is described as a parametric model under a risk neutral measure. If we can obtain the prices of options at any time T for all strikes K , we can uniquely determine the parameters of the dynamic of S_T by solving the above equations that correspond to the available option prices. But this is impossible in the real market where we have limited prices of option in any single maturity. One possible way to derive the model's parameter is to minimize the discrepancies between the available market prices and model prices generated from a parametric model. Therefore, in this case, the calibration problem has been transformed into an optimization problem for the least square of the discrepancies.

The scenario is that the market prices consist of the prices of European call options spanning a set of expiration dates T_1, \dots, T_N and for each T_i , the market quotes for strikes K_{i1}, \dots, K_{iM} . The least squares method is to find the minimum of the difference of the market prices and the model prices and can be described by

$$\theta_p^* = \arg \min_{\theta} \sum_{i=1}^N \sum_{j=1}^M w_{ij} [C^M(\theta_p, T_i, K_{ij}) - C(T_i, K_{ij})]^2. \quad (5.2)$$

The value θ_p^* is the set of calibrated parameters. The function $F(\theta_p) = \sum_{i=1}^N \sum_{j=1}^M w_{ij} [C^M(\theta_p, T_i, K_{ij}) - C(T_i, K_{ij})]^2$ represents the objective function with the set of parameters θ_p . The functions $C^M(\theta_p, T_i, K_{ij})$ and $C(T_i, K_{ij})$ are the value of the call option generated by parameters θ_p and the observed price at maturity T_i and strike K_{ij} respectively. As the function $C^M(\theta_p, T_i, K_{ij})$ is normally a nonlinear function, the

problem to find the calibrated parameter θ_p^* is therefore a nonlinear least square problem. The variable w_{ij} is the weight associated to the confidence of the observed price which varies with the value of option vega (The rationale of this weight is described in the following).

The rationale to set the value of w_{ij} to reflect the confidence in the individual data point was proposed by Cont and Tankov (2004) and Cont and Tankov (2002) where w_{ij} is defined as

$$w_{ij} = \frac{1}{|C_{ij}^{bid} - C_{ij}^{ask}|^2},$$

where C_{ij}^{bid} and C_{ij}^{ask} represent the bid and offer prices of option at maturity T_i and strike K_{ij} . This weighing scheme will outweigh the options that are liquid over the illiquid ones as the illiquid ones have a wider bid offer spread and it is hard to locate the real price compared to the liquid options that have a narrow bid offer spread. However the collection of the bid ask prices is sometimes not available practically in the high frequency data. Furthermore the options with the strike that are not too far from the money have the bid-ask spread close to one percent difference of option implied volatility (i.e. the spread of implied volatility are more systematic than the premium price). By this reason, instead of minimizing the function of option prices as in (5.2), it is more reasonable to minimize the implied volatilities of the option prices instead. Then the alternative problem can be expressed as

$$\begin{aligned}
\sum_i^n \sum_j^m (I(C^\theta(T_i, K_{ij})) - I_{ij})^2 &\approx \sum_i^n \sum_j^m \left(\frac{\partial I(I_{ij})}{\partial C} (C^\theta(T_i, K_{ij}) - C_{ij}) \right)^2 \\
&= \frac{\sum_i^n \sum_j^m (C^\theta(T_i, K_{ij}) - C_{ij})^2}{Vega^2(I_{ij})},
\end{aligned} \tag{5.3}$$

where $I(I_{ij})$ denotes the BS implied volatility corresponding with maturity T_i and strike K_{ij} and $Vega(I_{ij})$ is the Vega value of the option with maturity T_i and strike K_{ij} . The Vega of BS option can be computed by this formula

$$\begin{aligned}
Vega(I_{ij}) &= K_{ij} \exp(-rT_i) \phi(d_2) \sqrt{T_i} \\
\text{with } \phi(x) &= \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}.
\end{aligned} \tag{5.4}$$

5.2 Calibration as an Inverse Problem

The difficulties to find the parameters to minimize the objective function in (5.2) or (5.3) can be described by

- Prices in the financial markets are quoted as bid/offer prices and sometimes, due to the liquidity of the market, the spreads are relatively wide. So we cannot determine the exact price but only the range of exact price. Also actual prices may implicitly include hedging costs, tax and other costs.
- From the number of observed prices, there might be many sets of parameters that produce the local minimum objective functions that depend on the algorithm to search for the minimum point. Figure 5.1 shows the graph of the objective function with our calibrated parameters with varying kappa k and sigma σ (kappa and sigma are the JDSVSI model parameters as in

Chapter 4). From this figure, we see a few of regions that produce local minima of objective functions. In a real calibration problem which has a higher numbers of dimensions it can be even more difficult to locate the global minimum.

- The landscape of the domain may be flat and non-convex where the flat domain causes a stability problem for the solution and the non convexity will give the difficulty to locate the global minimum of the objective function. Figure 5.2 shows the landscape of the flat landscape near the minimum of the objective function when only kappa κ and sigma σ change values.
- There are some other factors and statistical properties that may influence the market prices of options but are not included in the models. These factors and properties make the model less than perfect in describing the true dynamic of the option prices.

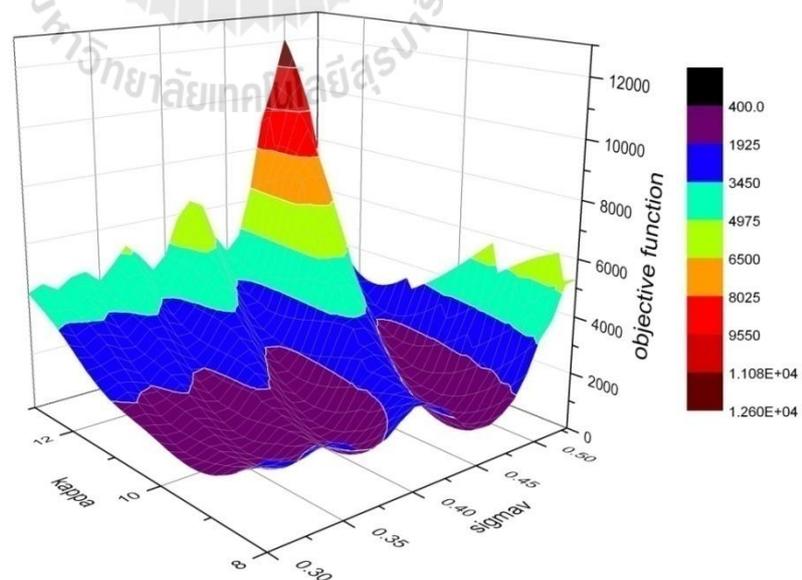


Figure 5.1 The objective function graph with varied kappa and sigma.

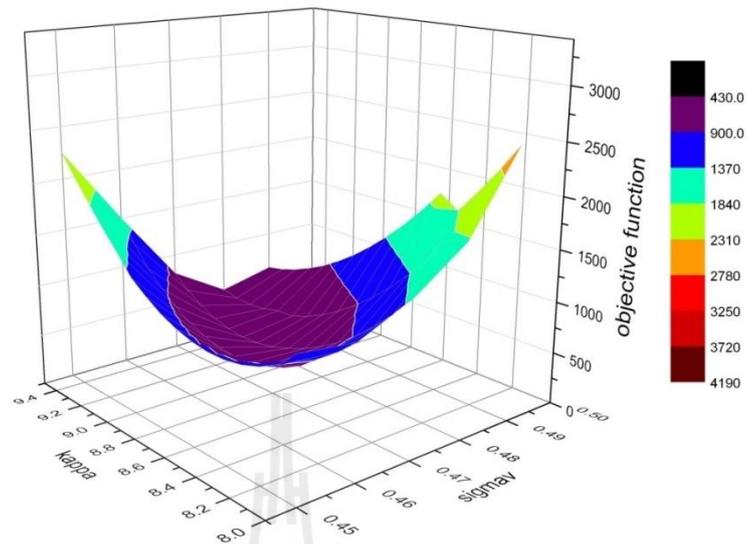


Figure 5.2 The objective function graph near the minimum point.

The field of Mathematics to solve this kind of problem is called the *Inverse Problem and Regularization methods*. We will briefly touch this subject on the area we need to solve our problem.

Definition 5.1. (Direct and Inverse Problem). Let X and Y be separable Hilbert spaces of finite or infinite dimensions and $T : X \rightarrow Y$ a compact operator (T is said to be compact if for any bounded set $B \subset X$ the image $T(B) \subset Y$ is pre-compact). Consider the problem of finding a solution $x \in X$ given $y \in Y$ satisfying the equation

$$T(x) = y. \quad (5.5)$$

The problem of finding y given x in $y = T(x)$ is called the *direct problem*. The *inverse problem* is to find a collection of solutions, x , given we have a collection of data, y . In many applications, y represents measured data which contains some

noises. Then instead of solving for x given y , our problem becomes to solve for an approximate solution x^δ given the perturbed data y^δ or

$$\begin{aligned} T(x^\delta) &= y^\delta \\ \text{given } |y - y^\delta| &\leq \delta \end{aligned} \quad (5.6)$$

where δ is called the *noise level*. Mathematically, if the operator T is not well-behaved there is no guarantee that x^δ is close to x . Here we like to point out the characteristics of an inverse problem.

Definition 5.2. (Well-Posed Problem). Let X and Y be normed spaces, $T : X \rightarrow Y$ a compact operator. The problem of finding a solution $x \in X$ given $y \in Y$ in the equation $T(x) = y$ is called *well-posed* if

1. *Existence*: For every $y \in Y$, there is at least one $x \in X$ such that $T(x) = y$.
2. *Uniqueness*: For every $y \in Y$, there is at most one $x \in X$ with $T(x) = y$.
3. *Stability*: The solution x depends continuously on $y \in Y$ i.e. if $T(x_n) \rightarrow T(x)$ then $x_n \rightarrow x$.

If one of these properties is violated, the problem is called an *ill-posed problem*.

For our least square calibration problem as in (5.2), a solution always exists, as the least square problem always has at least one solution in a compact domain. But considering data, observed option prices that are discrete and overdetermined i.e. there are a few numbers of option prices for each maturity but the numbers of option prices are not significantly greater than the number of parameters of the models. This problem may be solved for different sets of parameters that fit for a given data set. Therefore, this problem is an ill-posed problem. In addition to the mentioned characteristics of an ill-posed problem, if an objective function is not a convex

function, a solution to the inverse problem may not be the global solution. A method to handle a nonconvex function by a global optimization will be treated in the next section.

In order to overcome the problems of uniqueness and stability, a regularization technique is introduced here. The most popular regularization method is the Tikhonov regularization which replaces the original inverse problem by a family of neighborhood well-posed problem. Though we discuss only the principle of linear regularization, the generalization to the nonlinear case carries the same idea.

Definition 5.3. (Regularization Strategy) Kirsch (2011) : A regularization for the equation $Tx = y$ as in (5.6) is a family of linear and bounded operators $R_\alpha : Y \rightarrow X$, with $\alpha > 0$ as *the regularization parameter*, such that

$$\lim_{\alpha \rightarrow 0} R_\alpha T x = x, \text{ for all } x \in X. \quad (5.7)$$

That is, the operators $R_\alpha T$ converge pointwise to the identity. By this definition, R_α performs as a sequence of operators that approximates the inverse operator T^{-1} . When the operator T is compact and the dimension of X is infinite, we have the following theorem.

Theorem 5.4. Kirsch (2011) Let R_α be a regularization function for a compact operator $T : X \rightarrow Y$ with $\dim X = \infty$. Then we have

- The operators R_α are not uniformly bounded; that is, there exist a sequence (α_j) with $\alpha_j \rightarrow 0$ and $\|R_{\alpha_j}\| \rightarrow \infty$ for $j \rightarrow \infty$.
- The sequence $(R_\alpha T x)$ does not converge uniformly on bounded subsets of X ; that is, there is no convergence of $R_\alpha T$ to the identity I in the operator norm.

However the equation (5.7) is based on the assumption that we know the exact data y , that is $R_\alpha y$ converges exactly to x . In practice, we know y within the precision of $\delta \geq \|y - y^\delta\|$. With the regularization, our approximate solution is $x^{\alpha, \delta} = R_\alpha y^\delta$.

Then we can approximate the error by

$$\begin{aligned} \|x^{\alpha, \delta} - x\| &\leq \|R_\alpha y^\delta - R_\alpha y\| + \|R_\alpha y - x\| \\ &\leq \|R_\alpha\| \|y^\delta - y\| + \|R_\alpha T x - x\| \\ &\leq \delta \|R_\alpha\| + \|R_\alpha T x - x\|. \end{aligned} \quad (5.8)$$

The first term is the error due to the noise level δ multiplied by the norm of the operator $\|R_\alpha\|$ which is often referred to as “the data noise error” which goes to infinity as $\alpha \rightarrow 0$. The second term is the error due to the regularization which goes to zero as $\alpha \rightarrow 0$. The characteristics of these 2 errors are illustrated in Figure 5.3. In practice, we need to select the proper value of α to keep the balance between these two errors.

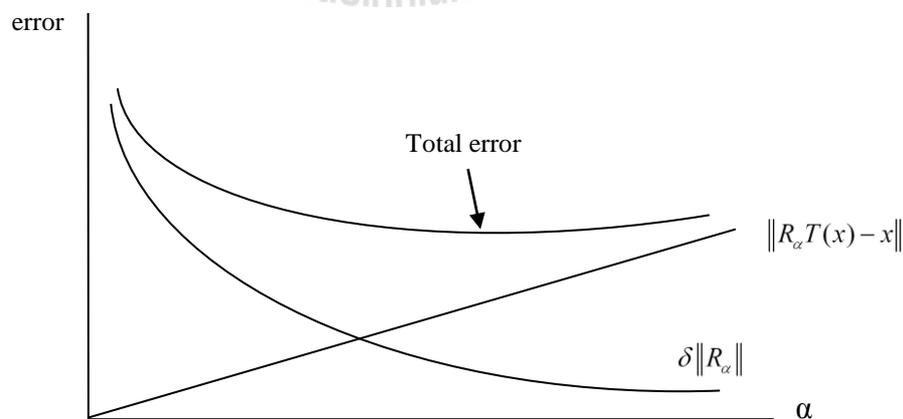


Figure 5.3 Characteristic of the error function.

Theorem 5.5. (Tikhonov Regularization). Let a constant $\alpha > 0$ be given. The Tikhonov solution x_α is the minimizer of the *Tikhonov functional* $F_\alpha(x)$, given by

$$F_\alpha(x) = \|T(x) - y^\delta\|^2 + \alpha \|G(x)\|^2, \quad (5.9)$$

provided that a minimizer x_α exists. The term $\|T(x) - y^\delta\|^2$ determines the accuracy of the solution. And the term $\alpha \|G(x)\|^2$ is known as the *penalty term* defined by $G(x) = x - x_0$, where x_0 is the initial guess of the solution. The penalty term is generally convex. The convexity of this term makes the Tikhonov functional more convex, thus enhancing the uniqueness and stability of the solution. That is, the bigger α makes the solution more unique and stabilized while the smaller α makes the solution more precise. The choice of α plays a crucial role in determining the success of the regularization. One of the most well-known method of determining the value of the regularization parameter α is known as Morozov's discrepancy principle which suggests the largest regularization parameter $\alpha = \alpha(\delta, y^\delta)$ such that the residual $\|T(x^{\alpha, \delta}) - y^\delta\|$ is lower or equal $c\delta$ for a fixed parameter $c \geq 1$ i.e.

$$\alpha(\delta, y^\delta) = \sup\{\alpha > 0 : \|T(x^{\alpha, \delta}) - y^\delta\| \leq c\delta\}. \quad (5.10)$$

This means that the choice of α should not try to solve for the accuracy of the solution than up to the noise level. In practice, the value of c is chosen as a number just slightly above 1 to prevent the optimization of the Thikonov functional to be more accurate than the noise level. The value of $\alpha(\delta, y^\delta)$ can be solved numerically by a simple gradient based algorithm as most of the functional $\alpha \mapsto \|T(x^{\alpha, \delta}) - y^\delta\|$ is monotonely increasing. The proof is due to (Engl, Hanke and Neubauer 1996).

5.3 Optimization for Continuous Variables

For a nonlinear least square problem in (5.2), there exist two major optimization methods that can be applied to this problem. They are local search methods and global search methods.

Local search methods typically locate the optimization point by an iteration scheme. In each iteration, a new point is derived from the old point to make a new point get closer to the optimal point. Though these methods are efficient, they depend on the starting point when the objective function is not convex. As illustrated in Figure 5.4, where the objective function is not convex, if the starting points are in the neighborhood of point A or point B, the local search method will locate the optimal points at point A and point B respectively.

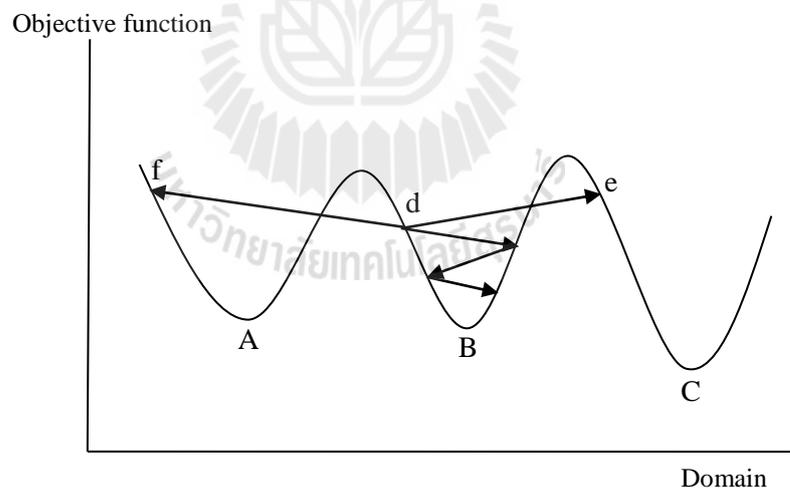


Figure 5.4 Local and global minimums of the objective function.

On the other hand, a global search method is normally less efficient and one needs to carefully choose the control parameters that suit the nature of the problem. Contrary to a local search method, a global search method sometimes allows, with

some probability, a new candidate point to have a worse objective function compared with an existing solution in order to explore other prospective areas. By this idea, the method can avoid trapping to local optimum points even when a starting point is located in the neighborhood of a local optimal point.

For our calibration problem, we apply the simulated annealing algorithm which allows accepting a new candidate point even when the new candidate point has a worse objective function. In Figure 5.4, suppose that the starting point to find the minimum solution is at point d. By local search methods, the algorithm will generate a new solution given the point has a better objective function and will finally reach point B as shown by the dark arrow line. By the simulated annealing method, the algorithm, at a high temperature (The temperature is a control parameter for this method), allows to accept a worse new candidate point, as in Figure 5.4, so a new point can be point e or point f or any other point, depending on the random search generating function, in the neighborhoods of a different minimum point (*Basin of Attractions*) as shown by the arrow lines. This mechanism allows a search to explore in other areas of the domain. By slowly lowering the temperature, the probability to accept the new worse point is reduced so that the candidate point does not wander over all the places, and starts to converge to the global optimum. However, by pursuing the random search method and permitting a worse objective function, a simulated annealing method will take a lot longer time to reach the global optimum point.

5.4 Simulated Annealing (SA)

5.4.1 General Simulated Annealing (GSA)

The simulated annealing method is one of the most popular methods to find the global optimum. The method was first implemented by Kirkpatrick, Jr. and Vecchi (1983) for a discrete optimization problem applied to a salesman travelling problem. Since then, this method has gained some popularity and there are a number of modifications to the original algorithm to make the algorithm more efficient and be able to solve a variety of problems. The attractiveness of this method over other methods is that, in order to implement an SA algorithm, it is not necessary to obtain the derivatives of the objective function. And it also does not require a lot of knowledge on the objective function. Although, with the increase in its popularity, the method is often claimed for the difficulty in getting a robust solution, especially when the problem is a high dimensional problem or has a very flat domain landscape, and this method requires unreasonable time for an optimum to be reached.

The simulated annealing method is the global optimization method that replicates the way metal is heated to a suitable temperature and cooled down slowly to get the optimum structure. The suitable temperature is called the *initial temperature* T_0 and the way the temperature is reduced is called the *annealing schedule*. The steps of the generic simulated annealing algorithm (GSA) are summarized below,

Simulated Annealing Algorithm

To minimize $f(x)$ subject to $x_i \in [A_i, B_i]$

where $f(x)$ denotes the *objective function*, sometimes called the *cost function*. The solution x is a vector of D-dimensional continuous variables in space R^D .

Step 1: initialization

- Set the initial solution x_0 .
- Set the initial temperature T_0 that is high enough so that the optimal solution can be reached.
- Set the round $k=0$ and k will count the number of changes in temperature,

Step 2: Perturbation

- Generate the new candidate point x_i according to the *generating function* denoted by $g(\Delta x)$. In GSA the generating function is the Gaussian probability distribution which is defined as

$$g(\Delta x) = \frac{1}{(2\pi T_k)^{D/2}} \exp(-(\Delta x)^2 / 2T_k),$$

where $\Delta x = x_{i+1} - x_i$ denotes the deviation of a new candidate point from an existing solution and T_k denotes the temperature at round k .

- Determine the difference of objective function $\Delta f = f(x_{i+1}) - f(x_i)$.

Step 3: Acceptance Determination

- The new candidate point is accepted as a new solution if the *acceptance probability (accepting function)* denoted by P is greater than the random number denoted by R which is distributed uniformly between 0 and 1.

$$P = \min\{1, \exp(-\Delta f / T_k)\}.$$

Replace x_i with x_{i+1} if $P > R$.

- Repeat step 2 and 3 until the equilibrium is reached (where there is not much improvement in the objective function at this temperature). In practice, for each level of temperature, we may fix to run step 2 and step 3 for a fixed number of times and denote this number as the number of rounds at temperature T_k by N_{T_k} .

Step 4: Annealing

- Set $k = k + 1$.
- The temperature is reduced by the following function,

$$T_k = T_0 (\ln k_0 / \ln k).$$

- Repeat step 2 to step 4 until the objective function is lower than a specified goal or the change in objective function is less than a specified small number.

The simulated annealing method replaces a deterministic acceptance rule in local search based methods by a probability acceptance rule to allow a new candidate point to be accepted even when the objective function of the new candidate point is worse than the existing objective function as shown in step 3. This accepting probability depends on both the temperature and the improvement of the objective function according the description in step 3. For a high temperature or a large improvement in the objective function, the acceptance probability is high. The acceptance probability reduces when the temperature goes down or the improvement in the objective function is low. Theoretically, it can be shown that, under some carefully chosen control parameters, this algorithm converges to the global optimum point.

While the convergence is guaranteed, it is not much of practical value because it may take unreasonable time for a large dimensional problem. Some variances of the method have been studied in order to reduce the convergent time. Two of the most popular variations are the *Fast Simulated Annealing (FSA)* and the *Adaptive Simulated Annealing (ASA)* invented by Ingber (1996).

We like to address the Fast Simulated Annealing algorithm first. As the generic simulated annealing algorithm is claimed to be a notoriously slow algorithm, a lot of modifications have been explored. One of the simple algorithms is the Fast Simulated Annealing. The generating function of FSA is the Cauchy function which is described by

$$g(\Delta x) = \frac{T_k}{(T_k^2 - |\Delta x|^2)^{(D+1)/2}}.$$

We provide here the heuristic proofs for the convergence time for the GSA and the FSA algorithm.

Proof:

The proof (taking the guidance from Szu and Hartley (1987)) is based on the fact that all the points in the D-dimensions can be equally sampled as the sample time goes to infinity. By letting $g(x)$ to be a probability density function to generate a candidate point, we can show that the probability of not visiting any point on the domain equals zero as the sampling times goes to infinity i.e.

$$\prod_{n=0}^{\infty} (1 - g_n(x)) = 0,$$

or in the other words

$$\sum_{n=0}^{\infty} g_n(x) = \infty.$$

In case of GSA, the annealing function is defined by $T_k = T_0 / \ln(k)$ and the probability

density function is defined by $g(\Delta x) = \frac{1}{(2\pi T_k)^{D/2}} \exp(-(\Delta x)^2 / 2T_k)$. We have that

$$\begin{aligned} g_k(\Delta x) &= \frac{1}{(2\pi T_k)^{D/2}} \exp(-(\Delta x)^2 / 2T_k) \\ &= \frac{1}{(2\pi T_k)^{-D/2}} \exp\left(\frac{1}{2T_0} (\Delta x)^2 (-\ln k)\right), \end{aligned}$$

$$\text{as } T_k \rightarrow 0 \text{ and } \frac{1}{2T_0} (\Delta x)^2 \leq 1,$$

$$\begin{aligned} g_k(\Delta x) &\geq \exp(\ln(-k)), \\ &\geq \frac{1}{k}. \end{aligned}$$

So that

$$\sum_{k=0}^{\infty} g_k(x) \geq \sum_{k=0}^{\infty} \frac{1}{k} = \infty.$$

For the case of FSA, the annealing function is defined by $T = T_0 / k$. And the probability density function of the generating function is defined by

$$g(\Delta x) = \frac{T_k}{(T_k^2 - |\Delta x|^2)^{(D+1)/2}}. \text{ We have that}$$

$$\begin{aligned} g_k(\Delta x) &= \frac{T_k}{(T_k^2 + |\Delta x|^2)^{(D+1)/2}} \\ &\approx \frac{T_k}{|\Delta x|^{(D+1)}}. \end{aligned}$$

So that

$$\sum_{k=0}^{\infty} g_k(x) \approx \frac{T_0}{\Delta x^{D+1}} \sum_{k=0}^{\infty} \frac{1}{k} = \infty.$$

The proofs imply that by having $T_k = T_0 / \ln(k)$ for the case of GSA and having $T_k = T_0 / k$ for the case of FSA, the algorithm will converge to the optimum. And it shows that FSA will converge faster than GSA by $k / \ln k$ times.

Even though the algorithm can converge to the optimum point theoretically, it requires too much time to reach the optimum point. There are some ongoing studies of modification of the GSA algorithm. One of the most efficient modified algorithms is the Adaptive Simulated Annealing (ASA) method. The ASA has two distinct features; First, it allows to have a specific annealing schedule depending on the sensitiveness of each variable. Second, after some number of acceptance events, there is a rescaling of control temperature for each variable to help adjust the generating function of each variable.

5.4.2 Adaptive Simulated Annealing

In the Adaptive Simulated Annealing algorithm, there are two controlled temperatures which are the parameters temperature T_i associated with the i^{th} parameter and the cost parameter T_c . The element x^i , the i^{th} of the new candidate D-dimensions point $x^T = (x^1, \dots, x^D)$ is generated by

$$x_{k+1}^i = x_k^i + \lambda^i (B_i - A_i),$$

where B_i and A_i are the upper bound and the lower bound of the x^i variable. The variable λ^i is determined by

$$\lambda^i = \text{sgn}(u^i - 0.5) T_i [(1 + 1/T_i)^{(2^{u^i-1})} - 1],$$

where u^i is a uniformly distributed random variable in $[0,1]$. The function $\text{sgn}(\)$ represents the function with three values defined by

$$\text{sgn}(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$$

In case the new x_{k+1}^i falls outside the range $[A_i, B_i]$, the new point is regenerated until it falls within the range.

The cooling schedule of T_i is defined by

$$T_i(k) = T_{0i} \exp(-c_i k_i^{1/D})$$

where T_{0i} denotes the initial temperature of the i^{th} parameter, k_i is the number of rounds of the i^{th} parameter, and c_i denotes the cooling scaling factor for T_i .

Ingber (1996) suggests to choose c_i such that

$$\begin{aligned} T_{fi} &= T_{0i} \exp(-m_i) \text{ when } k_f = \exp n_i, \\ c_i &= m_i \exp(-n_i / D), \end{aligned}$$

where T_{fi} is the expected final temperature and m_i and n_i are the free parameters to help tune ASA for specific problems

For the cooling schedule of T_{cost} , the schedule function is defined below

$$T_c(k_c) = T_{0,c} \exp(-c_c k_c^{1/D})$$

where $T_{0,c}$ denotes the initial temperature of the acceptance function, k_c is the number acceptance, and c_c denotes the cooling scaling factor for T_c .

5.4.3 Reannealing

For a multi-dimensional optimization problem, an objective function may not depend on each parameter at the same proportion. It is more efficient to rescale the temperature of each variable which is called a *reannealing*. By this temperature

rescaling, a simulated annealing algorithm allows to have larger λ_i for an insensitive parameter and smaller λ_i for a sensitive parameter. By changing a suitable λ_i , the algorithm can speed up the search process. The reannealing may be added into the normal simulated annealing after each number of acceptance-events, N_A . That is for every cycle of N_A , the algorithm evaluates the sensitiveness of each variable given by

$$s_i = \left| \frac{f(t^* + \delta t^* \cdot e) - f(t^*)}{\delta t^*} \right|$$

where $f(t)$ denotes the objective function, t^* is the vector of the best point recorded, δ is the small real number and e is the D-dimensional unit vector.

We define $s_{\max} = \max_{1 \leq i \leq D} \{s_i\}$.

The new value of k and T for each parameter can be calculated by the formula below,

$$\begin{aligned} T_{ik'} &= T_{ik} (s_{\max} / s_i), \\ k_i' &= (\ln(T_{i0} / T_{ik'}) / c_i)^D. \end{aligned} \quad (5.11)$$

And T_{i0} is reset to 1 to begin the search after a reannealing.

5.4.4 Matlab Optimization Toolbox

To code a proper and efficient program to implement a simulated annealing algorithm requires a tremendous work in coding and testing. A poor algorithm will result in not converging to the global optimum or taking an unreasonable time to reach the optimal point. Fortunately, the Matlab Optimization Toolbox provides all the proper coding and varieties of options to be tailored to one's specific problem. We describe briefly in this chapter, the algorithms and options that we apply to our

calibration problem. We refer to the *Global Optimization Toolbox version 2012* in this thesis. The SA algorithm in the Toolbox consist of 5 components,

1. Random Search Generation

This is similar to step 2 in the generic algorithm. The Toolbox provides two distributions in this step

@annealingfast; The distribution is similar to the FSA algorithm

@annealingboltz; The distribution is similar to the GSA algorithm

2. Acceptance determination

This is similar to step 3 in the generic algorithm.

3. Annealing function

There are 3 annealing functions provided in the Toolbox,

@temperatureexp (default); $T = T_0 * 0.95^k$

@temperaturefast; $T = T_0 / k$

@temperatureboltz; $T = T_0 / \log k$

4. Reannealing

The algorithm of this component is similar to the reannealing that we have discussed except that it does now allow to change the control parameter c_i .

5. Ending criteria

Provide the criteria to end the algorithm when the average change in the objective function is relative to stopping criteria. The details are provided as in section 6-14 in the *Global Optimization Toolbox User's Guide*.

5.5 Calibrating the Model

5.5.1 Implementation

In this part, we will calibrate the model to the DAX index option prices on July 5, 2002 which we take from Sepp (2003) as shown in Table 1. Based on (4.18), (4.22) and (4.23), our model has 12 parameters namely v_0 , k , θ , σ , ρ , λ , u_J , v_J , r_0 , α , ω and β .

We run the calibration algorithm in the MATLAB optimization toolbox which provides both simulated annealing algorithm and gradient based optimization algorithms. Then we compare our model with the jump-diffusion model with stochastic volatility model (JDSV).

The simulated annealing tool, in the MATLAB's optimization toolbox, provides a lot of options that are catered to many natures of problems, including the generating of distributions, annealing schedules and stopping criterions. The MATLAB code for our simulated annealing algorithm is listed in the appendix. We have run the optimization a number of times and finally set the initial temperature at 400. Busetti (2003) and Ledesma, Avina and Sanchez (2008) provide a general concept to set these control parameters, including the idea on the setup of initial temperature and annealing schedule for a simulated annealing algorithm. For the generating of distributions option, the Boltzmann generating function seems to be the best choice to obtain a stable minimum point to our problem. We compare our model (JDSVSI) with the jump-diffusion model with stochastic volatility (JDSV) which has 8 parameters namely v_0 , k , θ , σ , ρ , λ , u_J and v_J . The data points that are the in-sample on this calibration are all the data in Table 1 except for the columns of 1 month and 18 months that we use as the out-of-sample points.

Due to the nature of the simulated annealing method, the search for the candidate points being random can locate only the neighborhood of the minimum point (*Basin of Attraction*). We need a *gradient based method*, here called the “*local search*” algorithm, to find the precise minimum point. With both algorithms, the parameters of the precise JDSV and JDSVSI can be obtained.

5.5.2 Calibration Results

After running the Matlab Optimization Toolbox, we will have the result of optimization which produces the calibrated parameters and the value of the minimum of objective function. The examples of the results of both simulated annealing and local search algorithms are shown in the appendix. Table 5.2 presents the sum of square price differences of the model option price and the BS option price. The benchmark for the comparison is the sum of discrepancies of option values due to 1% change in the BS implied volatility generated from Table 5.1. This 1% is the normal spread between bid and offer in the usual two ways market. The result shows the sum of square price differences for the JDSVSI model is lower than the sum of square price differences of the JDSV model. We believe that the reason is that the JDSVSI model has more parameters than the JDSV model so it can fit better. Both models have minimum objective functions slightly higher than the benchmark due to the noise factors from the observed data and some discrepancies of the models. Considering only the noise factor included in the prices collected, these minimum of the sum of price differences are quite in line with our expectation.

The parameters of the calibration of each model are shown in Table 5.3 and Table 5.4 which is the source to produce the implied volatility structure of each model

in Table 5.5 and Table 5.6. In Figure 5.5, we show the data point error of the implied volatility between the BS model and the JDSV model and between the BS model the JDSVSI model. Then we calculate the square error in implied volatility of each model compared to the market implied volatility in Table 5.7 and Table 5.8.

The first row of Table 5.3 and Table 5.4 present the parameters generated from the simulated annealing algorithm and the parameters in the second row are generated from the local search algorithm with the initial point from the first row. For Table 5.7 and Table 5.8, the second last line presents the sum of the square error of each column and the last line is the error of each column adjusted by the sum of square error for all columns. As expected, the errors of the JDSV model are smaller for short maturities, but the errors of JDSVSI model are smaller in the long maturities in accordance with the finding from Bakshi (1997). This is because the effect of the interest rate part cannot take much effect in the short tenors especially for the 2 weeks and 1 month data, where the adjusted error of the JDSVSI model read at 0.5858 and 0.1384 compared to those of JDSV model which read at 0.5499 and 0.1109. In terms of the fitting accuracy, the total errors of the out-of-sample data points shown on the last line of column of 1 month and 18 months of both tables are within the average except for the error for 1 month of the JDSVSI model.

Table 5.2 Implied volatility surface of DAX index options on Jul. 5, 2002.

Spot at 4468.17

Expiry	Jul. 02	Aug. 02	Sep. 02	Dec. 02	Mar. 03	Jun. 03	Dec. 03	Jun. 04
Time	2 weeks	1 month	3 months	6 months	9 months	12 months	18 months	24 months
Date	19/7/2002	16/8/2002	20/9/2002	20/12/2002	21/3/2003	20/6/2003	19/12/2003	18/6/2004
Maturity (T)	0.0389	0.1139	0.2083	0.4583	0.7111	0.9583	1.4556	1.9528
Interest rate (r)	0.0357	0.0349	0.0341	0.0355	0.0359	0.0368	0.0386	0.0401
Strike								
3600	0.6007	0.4543	0.3967	0.3511	0.3279	0.3154	0.2984	0.2921
4000	0.4541	0.3869	0.3492	0.3149	0.2963	0.2926	0.2819	0.2800
4400	0.3726	0.3396	0.3108	0.2871	0.2788	0.2722	0.2661	0.2686
4800	0.3302	0.3062	0.2799	0.2631	0.2573	0.2533	0.2504	0.2544
5200	0.3460	0.2845	0.2624	0.2463	0.2425	0.2385	0.2373	0.2422
5600	0.3976	0.2860	0.2607	0.2356	0.2297	0.2268	0.2241	0.2320

Table 5.2 Comparison of the sum of square price differences in each model.

	Sum of Square Price Differences	Percent over BS
1% Difference in BS Model	419.3167	100.00%
JDSV	503.6367	120.11%
JDSVSI	460.3675	109.79%

As we have done so far, the calibration method is still not regularized. We see from Table 5.3 and Table 5.4 that some parameters have changed in large values from the simulated annealing algorithm to the local search algorithm. This indicates that the domain function is rather flat (This is also confirmed by Figure 5.2) and may cause the stability problem of the solution. Even though the domain of the objective function is well behaved, it is hard to be certain if the data of option prices does not contain some noises. These problems can be solved by the regularization method. Due to lack of time series data of option prices which is necessary for a statistical prior estimate, we will not perform the regularized calibration in this thesis. But we summarize the regularized calibration procedure in the following,

Regularized Calibration Procedure

- The average of time series of calibrated parameters is set as statistical prior θ_{p0} .
- The unregularized parameters are obtained by the method implemented in the beginning of this section. The minimum of the calculated objective function is denoted by $\delta(\alpha = 0)$ which is the unregularized intrinsic error of the model. This error is due to the collection of data by averaging the bid and offer prices

or the bias in prices in the quotations. By this way, the model error may be estimated to be

$$\begin{aligned} \delta(\alpha = 0) &= \|C^{bid} - C^{offer}\| \\ &\equiv \sqrt{\sum_{i=1}^N \sum_{j=1}^M w_{ij} [C^{bid}(T_i, K_{ij}) - C^{offer}(T_i, K_{ij})]^2}. \end{aligned} \quad (5.12)$$

Therefore it is useless to calibrate to get the exact parameters of the model as our data is not the true value but lie in the range of bid offer spread.

- Find the suitable regularization parameter α according to the Morozov's discrepancy principle. This is to solve the root of the following equation,

$$c\delta = \min \sum_{i=1}^N \sum_{j=1}^M w_{ij} [C^M(\theta_p, T_i, K_{ij}) - C(T_i, K_{ij})]^2 + \alpha |\theta_p - \theta_{p0}|^2. \quad (5.13)$$

This equation is the translation of equation (5.10) and in practice c is in the range of around 1.1 to 1.5 according to Cont and Tankov (2004).

- Run the optimization for the regularized equation again with this regularization parameter α derived from (5.13),

$$\theta_p^* = \arg \min_{\theta} \sum_{i=1}^N \sum_{j=1}^M w_{ij} [C^M(\theta_p, T_i, K_{ij}) - C(T_i, K_{ij})]^2 + \alpha |\theta_p - \theta_{p0}|^2. \quad (5.14)$$

The set of θ_p^* is the set of the calibrated parameters from the regularized optimization. According to (5.14), we add the penalty term $\alpha |\theta_p - \theta_{p0}|^2$ to help stabilize the solution as this term is convex. And by adding this term, it will penalize a solution that is far from the statistical prior.

Either unregularized or regularized calibrations are lengthy processes, especially by the global optimization algorithms. A practitioner may not do these

calibrations as often by considering whether the changes in the market variables need a recalibration. Many papers suggest a local search algorithm instead when the market does not move much or the stability of the model variables allows.



Table 5.3 JDSV parameters from calibration.

	v_0	k	θ	σ	ρ	λ	u_j	v_j
Simulated Annealing	0.1205	7.0371	0.0197	0.3937	-0.5852	0.3686	-0.2675	0.0793
Local search	0.1294	8.4472	0.0242	0.4717	-0.6318	0.3387	-0.2832	0.0744

Table 5.4 JDSVSI parameters from calibration.

	v_0	k	θ	σ	ρ	λ	u_j	v_j
Annealing	0.1538	11.3981	0.0188	0.4810	-0.6681	0.3677	-0.5257	0.1793
Local search	0.1369	9.1408	0.0228	0.4788	-0.5882	0.2968	-0.5388	0.1640
	r_0	α	ω	β				
Annealing	0.0330	2.9368	-0.0003	0.7644				
Local search	0.0100	2.9800	-0.0088	0.8600				

Table 5.5 Implied volatility term structure generated from JDSV model.

Strike	2 weeks	1 month	3 months	6 months	9 months	12 months	18 months	24 months
3600	0.5973	0.4574	0.4065	0.3506	0.3259	0.3121	0.3016	0.2975
4000	0.4323	0.3835	0.3540	0.3154	0.2990	0.2912	0.2847	0.2798
4400	0.3623	0.3381	0.3153	0.2856	0.2751	0.2713	0.2678	0.2644
4800	0.3350	0.3097	0.2876	0.2618	0.2550	0.2536	0.2533	0.2521
5200	0.3354	0.2927	0.2688	0.2438	0.2388	0.2391	0.2417	0.2427
5600	0.4087	0.2943	0.2607	0.2319	0.2267	0.2279	0.2329	0.2358

Table 5.6 Implied volatility term structure generated from JDSVSI model.

Strike	2 weeks	1 month	3 months	6 months	9 months	12 months	18 months	24 months
3600	0.5848	0.4514	0.4041	0.3503	0.3286	0.3149	0.3001	0.2922
4000	0.4287	0.3807	0.3517	0.3145	0.3014	0.2947	0.2874	0.2793
4400	0.3662	0.3379	0.3135	0.2837	0.2759	0.2737	0.2708	0.2647
4800	0.3417	0.3117	0.2871	0.2597	0.2546	0.2546	0.2552	0.2523
5200	0.3416	0.2968	0.2700	0.2425	0.2381	0.2392	0.2427	0.2428
5600	0.4158	0.3001	0.2639	0.2318	0.2262	0.2276	0.2334	0.2360

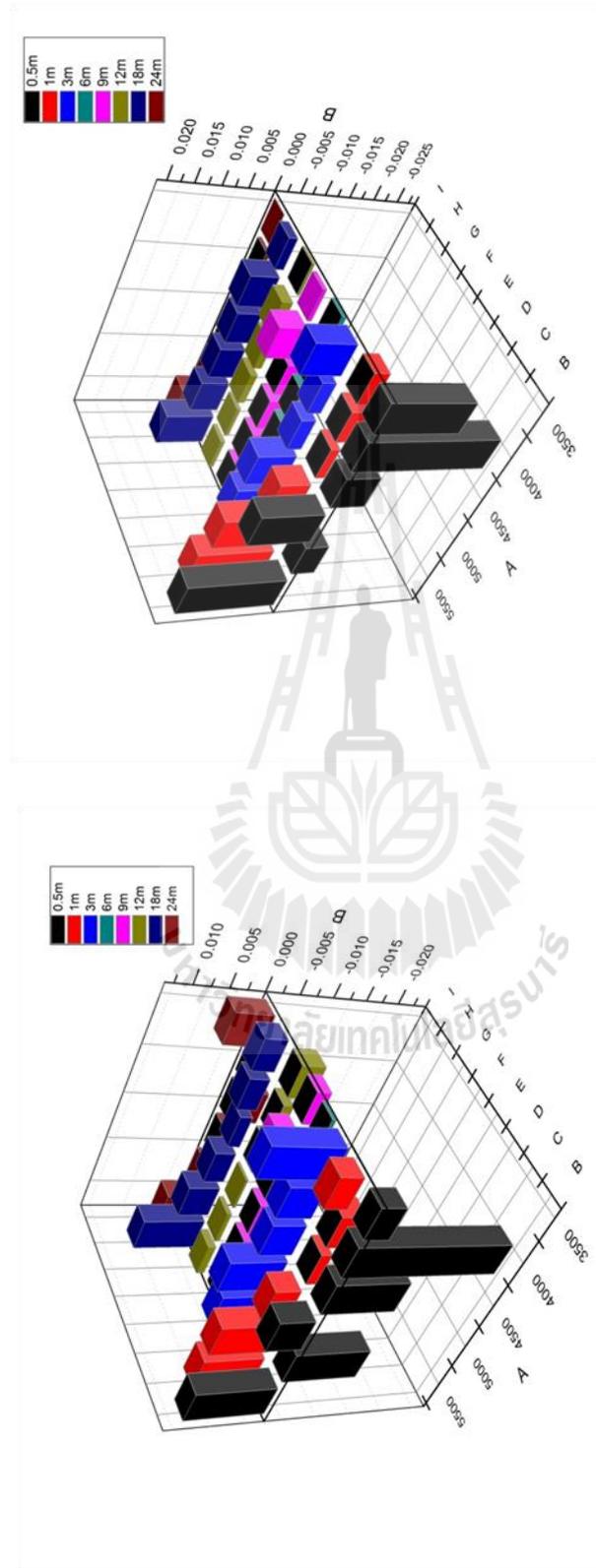


Figure 5.1 Error of each data point for JDSV (left) and JDSVSI (right).

Table 5.7 Square of errors of implied volatilities of JDSV compared to market prices.

Strike	2 weeks	1 month	3 months	6 months	9 months	12 months	18 months	24 months	$\Sigma(\text{error})^2$	$\% \Sigma(\text{error})^2$
3600	1.16E-05	9.61E-06	9.60E-05	2.50E-07	4.00E-06	1.09E-05	1.02E-05	2.92E-05	1.72E-04	0.1109
4000	4.75E-04	1.16E-05	2.30E-05	2.50E-07	7.29E-06	1.96E-06	7.84E-06	4.00E-08	5.27E-04	0.3405
4400	1.06E-04	2.25E-06	2.03E-05	2.25E-06	1.37E-05	8.10E-07	2.89E-06	1.76E-05	1.66E-04	0.1071
4800	2.30E-05	1.22E-05	5.93E-05	1.69E-06	5.29E-06	9.00E-06	8.41E-06	5.29E-06	1.15E-04	0.0745
5200	1.12E-04	6.72E-05	4.10E-05	6.25E-06	1.37E-05	3.60E-07	1.94E-05	2.50E-07	2.60E-04	0.1682
5600	1.23E-04	6.89E-05	0.00E+00	1.37E-05	9.00E-06	1.21E-06	7.74E-05	1.44E-05	3.08E-04	0.1988
$\Sigma(\text{error})^2$	0.00085	0.00017	0.00024	0.00002	0.00005	0.000015	0.000126	0.000067	1.55E-03	
$\% \Sigma(\text{error})^2$	0.5499	0.1109	0.1547	0.0157	0.0342	0.0099	0.0815	0.0432		

Table 5.8 Square of errors of implied volatilities of JDSVSI compared to market prices.

Strike	2 weeks	1 month	3 months	6 months	9 months	12 months	18 months	24 months	$\Sigma(\text{error})^2$	$\% \Sigma(\text{error})^2$
3600	2.53E-04	8.41E-06	5.48E-05	6.40E-07	4.90E-07	2.50E-07	2.89E-06	1.00E-08	3.20E-04	0.1331
4000	6.45E-04	3.84E-05	6.25E-06	1.60E-07	2.60E-05	4.41E-06	3.03E-05	4.90E-07	7.51E-04	0.3121
4400	4.10E-05	2.89E-06	7.29E-06	1.16E-05	8.41E-06	2.25E-06	2.21E-05	1.52E-05	1.11E-04	0.0460
4800	1.32E-04	3.02E-05	5.18E-05	1.16E-05	7.29E-06	1.69E-06	2.30E-05	4.41E-06	2.62E-04	0.1090
5200	1.94E-05	1.51E-04	5.78E-05	1.44E-05	1.94E-05	4.90E-07	2.92E-05	3.60E-07	2.92E-04	0.1214
5600	3.31E-04	1.99E-04	1.02E-05	1.44E-05	1.22E-05	6.40E-07	8.65E-05	1.60E-05	6.70E-04	0.2784
$\Sigma(\text{error})^2$	0.00142	0.00043	0.00019	0.00005	0.00007	0.00001	0.00019	0.00004	2.41E-03	
$\% \Sigma(\text{error})^2$	0.5907	0.1787	0.0782	0.0219	0.0307	0.0040	0.0806	0.0152		

CHAPTER VI

CONCLUSION AND RESEARCH POSSIBILITIES

6.1 Conclusion

This thesis has proposed the asset price dynamic by the jump-diffusion process with stochastic volatility and stochastic interest rate, where the stochastic volatility is generated from the time changed Levy process. This stochastic volatility has similar characteristics as the CIR process which exhibits both volatility clustering and leverage properties as indicated by several empirical studies. The derivation of a price of a European call option follows the combination of Lewis Fourier transform method and modular pricing method which are both based on the risk neutral expectation pricing method. This combination produces the formula as a single generalized Fourier transform integral that results in computation efficiency.

By the Modular Pricing method, our derivation has reduced the four dimensional stochastic factors problem into each one or two dimensional problem. And by the Lewis Fourier Transform method, we have a pricing formula as a single integral on the complex domain that corresponds to the domain of the Carr and Wu leverage neutral measure. The numerical computation for the option pricing is handled with the direct integration method, which although not a sophisticated method, produces a stable and accurate price.

The most difficult part of this thesis is the calibration method, as the concept of global optimization is still very recent and most algorithms are difficult to

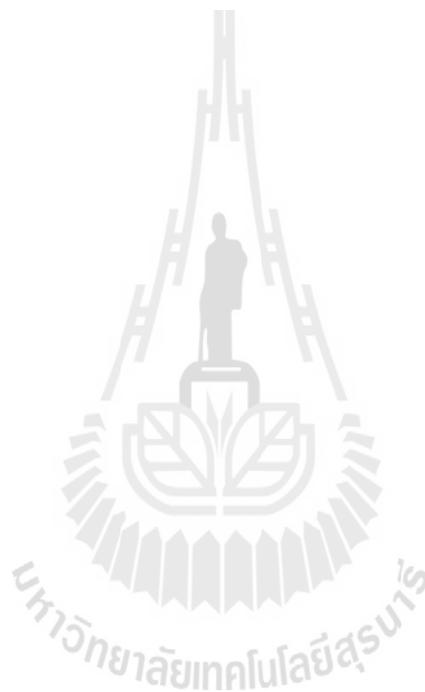
guarantee the convergence in practice. Also most algorithms are not efficient and take a very long time to produce a reasonable and satisfactory result. In this thesis, we apply both simulated annealing algorithm to locate the neighborhood of the optimum point and local search algorithm to search for the precise optimum point. The fitness to the one single day market prices is quite satisfactory but we do not test with a time series of the market data.

6.2 Research Possibilities

Although the model provides a very good fit for market price data, this is done for one particular day and only for European option prices. We believe this model and calibration is sufficient as a plain vanilla option pricer. However, as discussed in chapter 2 and chapter 3, a good model should provide a correct hedging scheme and can give a good fit to an option price with complicated structure. So we give some clues here for a further research.

1. The good fit of the model parameters through time series data of market prices is needed in order to test if the model will predict a set of correct hedging parameters. Even though we believe our model is capable, this test should give some further knowledge of the characteristics of this model. And if the model can fit to a time series of market prices without unstable calibrated parameters, then we can be sure that the model can describe the dynamic of the asset price well. Then it is not necessary to waste a lot of time to calibrate the model as often as it should be.
2. With the development to provide more complicated structures with different pay off structures or some other features to better hedge the clients' exposure, new

structures are constantly launched. The dynamics of the asset price may suit some particular structures or some asset classes but provide poor description to others. A study how this model can give a right price and right hedging parameters to some complicated structures is worth as clue to validate or improve the model.





REFERENCES

REFERENCES

- Bakshi G., Cao C. and Chen. Z. (1997). Empirical performance of alternative option pricing models. **Journal of Finance**. 52: 2003-2049.
- Bates, D. S. (1988). **Pricing options under jump-diffusion processes**. Working Paper 37-38. University of Pennsylvania, Rodney L. White Center.
- Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. **Review of Financial Studies**. 9: 69-107.
- Black F. and Scholes M. (1973). The pricing of options and corporate liabilities. **Journal of Political Economy**. 81: 637-654.
- Branger N. (2004). **An Anatomy of Option Pricing Models**. Arbeitspapier, Dezember. Goethe University, Frankfurt, Germany.
- Busetti, F. (2003). **Simulated annealing overview** [On-Line]. Available: <http://163.18.62.64/wisdom/Simulated%20annealing%20overview.pdf>.
- Carr P. and Wu L. (2004). Time-changed Lévy processes and option pricing. **Journal of Financial Economics**. 71: 113-141.
- Carr P., Geman H., Madan D. B. and Yor M. (2003). Stochastic volatility for Lévy processes. **Mathematical Finance**. 13: 345-382.
- Clark P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. **Econometrica: Journal of the Econometric Society**. 41: 135-155

- Cont R. and Tankov P. (2004). **Financial modelling with jump processes**. CRC Press.
- Cox J. C., Ingersoll Jr J. E. and Ross S. A. (1985). A theory of the term structure of interest rates. **Econometrica: Journal of the Econometric Society**. 385-407.
- Derman E. (2007). **Volatility smile modeling** [On-Line].
Available: <http://www.ederman.com/new/docs/laughter.html>.
- Engl H. W., Hanke M. and Neubauer A. (1996). **Regularization of inverse problems**. Kluwer Academic Pub.
- Harrison J. M. and Pliska S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading. **Stochastic Processes and their Applications**. 11: 215-260.
- Heston S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. **Review of Financial Studies**. 6: 327-343.
- Kirkpatrick S., Jr D. G. and Vecchi M. P. (1983). Optimization by simulated annealing. **Science**. 220: 671-680.
- Kirsch A. (2011). **An introduction to the mathematical theory of inverse problems**. Springer Science+ Business Media.
- Ledesma S., Aviña G. and Sanchez R. (2008). **Practical considerations for simulated annealing implementation** [On-Line].
Available: http://cdn.intechopen.com/pdfs/4631/InTech-Practical_considerations_for_simulated_annealing_implementation.pdf.
- Lewis A. (2001). **A simple option formula for general jump-diffusion and other exponential Lévy processes** [On-Line].

Available: <http://www.optioncity.net/pubs/ExpLevy.pdf>.

Merton R. C. (1976). Option pricing when underlying stock returns are discontinuous.

Journal of Financial Economics. 3: 125-144.

Mikhailov S. and Nogel U. (2006). Heston's Stochastic Volatility Model:

Implementation, Calibration and Some Extensions. **The Best of Wilmott 2.**

401.

Monroe I. (1978). Processes that can be embedded in Brownian motion. **Annals of**

Probability. 42-56.

Moodley N. (2005). **The Heston model: A practical approach with Matlab code**

[On-Line]. Available: <http://math.nyu.edu/~atm262/fall06/compmethods/a1/nimalinmoodley.pdf>.

Øksendal B. (1998). **Stochastic differential equations.** Springer.

Pan J. (2002). The jump-risk premia implicit in options: Evidence from an integrated

time-series study. **Journal of Financial Economics.** 63: 3-50.

Pascucci A. (2011). **PDE and martingale methods in option pricing.** Springer.

Pinkham S. and Sattayatham P. (2011). European Option Pricing for a Stochastic

Volatility Lévy Model with Stochastic Interest Rates. **Journal of**

Mathematical Finance. 1: 98-108.

Schmelzle M. (2010). **Option pricing formulae using Fourier transform: Theory**

and application [On-Line]. Available: <http://pfadintegral.com/articles>.

Schoutens W. (2003). Lévy process in finance. **Pricing financial derivatives,** John

Wiley & Sons Ltd.

Sepp A. (2003). **Fourier transform for option pricing under affine jump-**

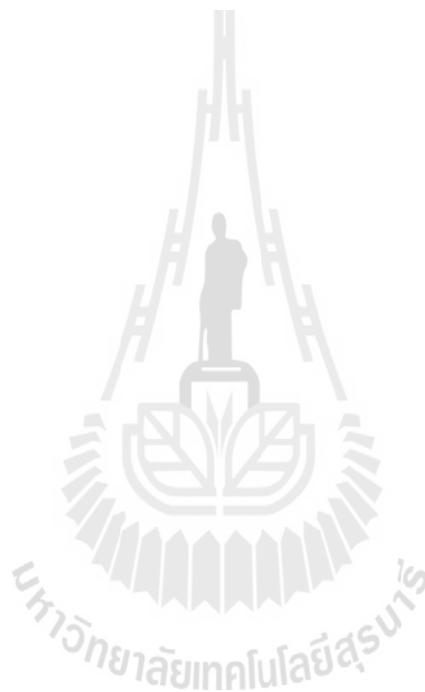
diffusions: An overview [On-Line]. Available: <http://math.ut.ee/~spartak/>.

Szu H. H. and Hartley R. L. (1987). Nonconvex optimization by fast simulated annealing. **Proceedings of the IEEE**. 75: 1538-1540.

Winkel M. (n.d.). **MS3b/MScMCF Lévy processes and finance** [On-Line].

Available: <http://www.stats.ox.ac.uk/~winkel/ms3b10.pdf>.

Zhu J. (2010). **Applications of fourier transform to smile modeling**. Springer.





APPENDIX

APPENDIX

MATLAB CODES

MATLAB codes for a European call option pricing - JDSVSI model

```
%Numerical Integration for JDSVSI
%hctrans2 is the function of an option value.

function y =
hctrans2(u,s,k,t,v0,kappa,theta,sigmav,rho,lamda,uj,vj,r0
,alpha,omega,beta)
kmax=round(max(500,10/(v0*t)^0.5));
cnt = 0;
for phi = -(kmax / 10):0.2: kmax
cnt = cnt + 1;
int_x(cnt,1) = phi;
pass_phi = phi+u*i;
int_y(cnt,1) = inth2(pass_phi, s, k, t,v0, kappa, theta, sigmav, rho,
lamda, uj, vj,r0,alpha,omega,beta);
end
y=-(1 / (2 * pi)) * k* trapnumint(int_x, int_y)
end

%The Integrand of JDSVSI
function y =
inth2(u,s,k,t,v0,kappa,theta,sigmav,rho,lamda,uj,vj,r0,alpha,omega,be
ta)
exy=exy2(u,s,k);
char1=(si(u,t,r0,alpha,omega,beta)*phihestsi(-
u,s,k,t,v0,kappa,theta,sigmav,rho))/exp(t*purejump(-u,lamda,uj,vj));
y=real(exy*char1);
end

%Calculation of  $\exp(-izx_0)H(z)$  as in Eq. 4.5
function y = exy2(u,s,k)
k1=log(s/k);
y=exp(-i*u*k1)/(u^2-i*u);
end

%Calculation of Characteristic Function Stochastic Interest Part
%as in Eq. 4.21
function y = si(u,t, r0,alpha,omega,beta )
p=-((u*i)+1);
q=alpha;
r=0.5*(beta^2);
```

```

d=(q^2+(4*p*r))^0.5;
g=(q+d)/(q-d);
gg1=(q+d)/(2*r);
gg2=(1-exp(d*t))/(1-(g*exp(d*t)));
gg=gg1*gg2;
hh1=alpha*omega/(beta)^2;
hh2=(q+d)*t;
hh3=2*log((1-(g*exp(d*t)))/(1-g));
hh=hh1*(hh2-hh3);
y=exp((gg*r0)+hh);
end

%Calculation of Characteristic Function Time Change Levy Part
%as in Eq. 4.17
function y= phihestsi(u,s,k,t,v0,kappa,theta,sigmav,rho )
km=kappa-(u*sigmav*rho*i);
alp=0.5*(i*u+u^2);
d=(km^2+(2*alp*sigmav^2))^0.5;
g=(km+d)/(km-d);
cc1=-kappa*theta/(sigmav)^2;
cc2=(km+d)*t;
cc3=2*log((1-(g*exp(d*t)))/(1-g));
c=cc1*(cc2-cc3);
dd1=-(km+d)/sigmav^2;
dd2=(1-exp(d*t))/(1-(g*exp(d*t)));
dd=dd1*dd2;
y=exp(-c-(dd*v0));
end

% Calculation of Characteristic Function of Compensated Compound
Poisson
% as in Eq. 4.22
function y = purejump(u,lamda,uj,vj)
y1=lamda*(1-exp((i*u*uj)-(0.5*u^2*vj)));
y2=i*u*lamda*(exp(uj+0.5*vj)-1);
y=y1+y2;
end

% A Trapezoidal Procedure
function z= trapnumint(x,y )
[n,m]=size(x);
a=0;
for t=2:1:n
    a=a+(0.5*(x(t)-x(t-1))*(y(t-1) + y(t)));
end
z=a;
end

```

MATLAB code for a Simulated Annealing and Local Search

```

% Simulated Annealing Procedure for JDSVSI
v0=0.120;
kappa=7.0385;
theta= 0.0190;
sigmav=0.3954;
rho=-0.7505;
lamda=0.3616;
uj=-0.2786;
vj=0.2600
r0=0.0313
alpha=1.8785
omega=0.005
beta=0.2989
options =
saoptimset('PlotFcns',{@saplotbestf,@saplottemperature,@saplotf,@saplotstopping},'TemperatureFcn',
@temperatureboltz,'annealingFcn',@annealingboltz,'InitialTemperature',400);
ObjectiveFunction = @(x)
evaluatevegal(x,v0,kappa,theta,sigmav,rho,lamda,uj,vj,r0,alpha,omega,beta);
x0 = [0.12 4.2 0.02 0.22 -0.5 0.4 -0.4 0.14 0.02 1.7 0.01 0.3];
lb=[0.001 0.1 0.001 0.01 -0.95 0.01 -0.6 0.01 0.01 0.1 -0.05 0.001];
ub=[0.3 12.0 0.3 0.57 0.1 0.8 0.1 0.4 0.20 4.0 0.12 0.9];

[x,fval,exitFlag,output] =
simulannealbnd(ObjectiveFunction,x0,lb,ub,options)

% A Local Search for JDSVSI
v0=0.07;
kappa=2.2;
theta= 0.078;
sigmav=0.324;
rho=-0.72;
lamda=0.017;
uj=-0.155
vj=0.151
r0=0.05
alpha=0.2
omega=0.05
beta=0.20
options = optimset('Display','iter','Algorithm','active-set');
ObjectiveFunction = @(x)
evaluatevegaalpha1(x,v0,kappa,theta,sigmav,rho,lamda,uj,vj,r0,alpha,omega,beta);
x0 = [0.1293 8.1756 0.0154 0.3055 -0.7472 0.3810 -0.5676 0.1907
0.0351 1.9501 0.0101 0.1388 0.00]; % Starting point
lb=[0.001 0.1 0.001 0.01 -0.95 0.01 -0.6 0.01 0.01 0.1 -0.05 0.01
0.000]
ub=[0.6 12.0 0.6 0.55 0.1 0.8 0.1 0.4 0.20 4.0 0.7 0.86 0]
[x,fval] = fmincon(ObjectiveFunction,x0,[],[],[],[],lb,ub,[],options)

```

Sample of Optimization Result from MATLAB – Simulated Annealing

Optimization terminated: change in best function value less than
options.TolFun.

x =

Columns 1 through 6

0.1583 11.3981 0.0188 0.4810 -0.6681 0.3677

Columns 7 through 12

-0.5257 0.1793 0.0330 2.9368 -0.0003 0.7644

fval = 0.8796

exitFlag = 1

output =

iterations: 12345

funccount: 12610

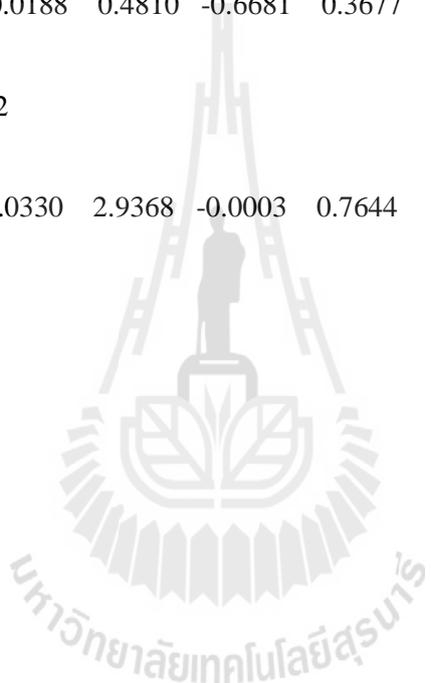
message: [1x80 char]

rngstate: [1x1 struct]

problemtype: 'boundconstraints'

temperature: [12x1 double]

totaltime: 3.0140e+04



Sample of Optimization Result from MATLAB – Local Search

Local minimum possible. Constraints satisfied.

fmincon stopped because the predicted change in the objective function is less than the default value of the function tolerance and constraints are satisfied to within the default value of the constraint tolerance.

<stopping criteria details>

Active inequalities (to within options.TolCon = 1e-06):

lower	upper	ineqlin	ineqnonlin
9	12		
13	13		

x =

Columns 1 through 7

0.1369	9.1408	0.0228	0.4788	-0.5882	0.2968	-0.5388
--------	--------	--------	--------	---------	--------	---------

Columns 8 through 13

0.1640	0.0100	2.9800	-0.0088	0.8600	0
--------	--------	--------	---------	--------	---

fval = 0.4559

CURRICULUM VITAE

Name : Paiboon Peeraparp

Education :

International University of Japan, Niigata, Japan 1991-1993

MBA. Finance

Ramkhamhaeng University, Bangkok, Thailand 1987-1990

B.Econ. International Economics

Chulalongkorn University, Bangkok, Thailand 1986-1989

B.Eng. Electrical Engineering

Work experiences :

Citibank, Bangkok, Thailand 2001-2009

Vice President

Dai-Ichi Kangyo Bank, Bangkok, Thailand 1998-2000

Senior Assistant Manager

HSBC, Bangkok, Thailand 1994-1997

Interbank Dealer

Publications :

P. Peeraparp and P. Sattayatham (2013). A jump-diffusion with stochastic volatility and interest rate. **Journal of Mathematics and Statistics**. 9(1): 43-50.