

**Barrier Option Pricing under SABR Model Using Monte Carlo
Methods**

by

Junling Hu

A Project Report

Submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Master of Science

in

Financial Mathematics

May 2013

APPROVED:

Professor Stephan Sturm

Professor Bogdan Vernescu, Head of Department

ACKNOWLEDGMENTS

This project could not have been possible without the support of my professors and my friends. In particular, I would like to express my gratitude towards Prof. Stephan Sturm for his guidance throughout the study.

Abstract

The project investigates the prices of barrier options from the constant underlying volatility in the Black-Scholes model to stochastic volatility model in SABR framework. The constant volatility assumption in derivative pricing is not able to capture the dynamics of volatility. In order to resolve the shortcomings of the Black-Scholes model, it becomes necessary to find a model that reproduces the smile effect of the volatility. To model the volatility more accurately, we look into the recently developed SABR model which is widely used by practitioners in the financial industry.

Pricing a barrier option whose payoff to be path dependent intrigued us to find a proper numerical method to approximate its price. We discuss the basic sampling methods of Monte Carlo and several popular variance reduction techniques. Then, we apply Monte Carlo methods to simulate the price of the down-and-out put barrier options under the Black-Scholes model and the SABR model as well as compare the features of these two models.

Contents

1	Introduction of Options	1
1.1	Definition of Options	1
1.2	The Development of Option Trading and Option Pricing	2
1.3	Summary of Vanilla Option Pricing under the Black Scholes	3
2	Barrier Option Pricing under the Black Scholes	7
2.1	Valuation Methods	8
2.2	Analytical Solution under the Black-Scholes Model	8
2.2.1	A Continuity Correction to Discrete Barrier Option	14
2.3	Monte Carlo Methods	15
2.3.1	The Main Framework of Monte Carlo	16
2.3.2	Variance Reduction Techniques	18
2.3.3	Path Generation of Barrier Options	29
2.3.4	Application of Variance Reduction to Barrier Options	31
3	Barrier Option Pricing under SABR Model	41
3.1	The Need for a Stochastic Volatility Model	41
3.2	Building SABR Model Step by Step	43
3.2.1	Black Model and Implied Volatility	43
3.2.2	Local Volatility Models	46
3.2.3	The SABR Model	49
3.3	Estimating the SABR Parameters	51
3.3.1	Choice of β	53
3.3.2	Estimate α , ρ and ν directly	54
3.3.3	α from σ_{ATM}	55
3.3.4	Algorithm for Calibration	56
3.4	Dynamics of Parameters	60
3.4.1	α	60
3.4.2	β	60
3.4.3	ρ	61
3.4.4	ν	62
3.4.5	f	62
3.5	Monte Carlo Methods under SABR	64
3.5.1	Path Generation of Barrier Options	65

3.5.2	Result Discussion	66
Appendices		69
A	Mathematical Appendix	70
A.1	Probability Space	70
A.2	Conditional Expectation	72
A.3	Stochastic Processes	72
A.3.1	Brownian Motions	73
A.3.2	Martingales	73
A.4	Convergence and Central Limit Theorem	74
A.4.1	Convergence	74
A.4.2	The Law of Large Numbers	75
A.4.3	Central Limit Theorem	75
A.5	Change of Measure	75
A.5.1	Radon-Nikodym Derivative	76
A.5.2	Girsanov's Theorem	76
A.5.3	Equivalent Martingale Measure	77

List of Figures

2.1	Antithetic Path Example	19
3.1	$\beta = 0.399$ $\alpha = 1.1649$ $\rho = 0.1659$ $\nu = 1.2543$	56
3.2	α from σ_{ATM} approach	57
3.3	α dynamics	61
3.4	β dynamics	61
3.5	ρ dynamics	62
3.6	ν dynamics	62
3.7	f dynamics	63
3.8	One Simulation Path Example	66
3.9	50 Simulation Paths Example	67

Chapter 1

Introduction of Options

1.1 Definition of Options

Options are traded both on exchanges and in the over-the counter market. There are two types of options. A call option gives the holder the right to buy the underlying asset by a certain date for a certain price. A put option gives the holder the right to sell the underlying asset by a certain date for certain price. The price in the contract is known as the exercise price or strike price; the date in the contract is known as the expiration date or maturity. [15]

Options can be either American or European, a distinction that has nothing to do with geographical location. American options can be exercised at any time up to the expiry date, whereas European options can be exercised only on the expiration date itself.

Options has been considered to be the most dynamic segment of the derivative markets since the inception of the Chicago Board Options Exchange (CBOE) in April 1973, with more than 1 million contracts per day, CBOE is the largest option exchange in the world. After that, several other option exchanges such as London

International Financial Futures and Options Exchange had been set up.

1.2 The Development of Option Trading and Option Pricing

The history of stock options trading began with the 1973 establishment of the CBOE and the development of the Black-Scholes option pricing model. Over the last a few decades due to the famous work of Black and Scholes, the option valuation problem has gained a lot of attention. In Black and Scholes 1973 seminal paper [4], the assumption of log-normality on stock price was applied. Moreover, in the same year, Robert Merton [19] extended the Black-Scholes model in several important ways. The application of the Black-Scholes-Merton option pricing model for valuing various range of financial instruments and derivatives is considered essential.

Options form the foundation of innovative financial instruments, which are extremely versatile securities that can be used in many different ways. Over the past decade, option has been utilized to speculation – leverage of pay-off on the underlying asset and hedging – reducing risk or providing portfolio insurance for financial institutions.

Following Black-Scholes, a number of other popular approaches was developed to pricing options with payoff depending on the value of the underlying asset at a single time horizon, including solving PDEs, binomial tree model, the numerical methods such as Monte-Carlo Simulation, Finite Differences Method and Replication. These work pushed option pricing into a very fascinating position in

quantitative finance.

The Black-Scholes formula is still around, even though it depends on several unrealistic assumptions. In special cases, though, we can improve the formula by making more realistic assumptions. However, we haven't produced a formula that works better across a wide range of circumstances. One theory of the 1987 crash relies on incorrect beliefs, held before the crash, about the extent to which investors were using portfolio insurance, and about how changes in stock prices cause changes in expected returns.

The complexity of innovative derivatives came up with the complexity of option pricing analytical formulas. At the same time, the demand of speed in derivatives trading require fast ways to process these calculations. As a result, the development of computational methods for option pricing models can be a better solution.

1.3 Summary of Vanilla Option Pricing under the Black Scholes

To price the option, we denote the value of the option C , on an underlying asset S_t which pays a function $f(S_T)$ at maturity time T . The interest rate, which is constant, to be r .

The payoff at maturity of an European call option with strike price K is defined by

$$f(S_T) = \max(S_T - K, 0) \tag{1.1}$$

The payoff at maturity of an European put option with strike price K is defined by

$$f(S_T) = \max(K - S_T, 0) \quad (1.2)$$

The Black-Scholes Assumption

In order to incorporate above factors as variables into one consistent model. Black and Scholes made some explicit assumptions on the market for a particular stock:

- There is no arbitrage opportunity (no way to make riskless profits).
- It is possible to borrow and lend cash at a known constant risk-free interest rate r .
- The stock price follows a geometric Brownian motion with constant drift μ and volatility σ .
- It is possible to buy and sell any amount, even fractional, of stock (this includes short selling).
- The above transactions do not incur any fees or costs (i.e., frictionless market).
- The underlying security does not pay a dividend.

Now we can define the process of the stock price to be a stochastic differential equations below:

$$dS_t = uS_t dt + \sigma S_t dW_t \quad (1.3)$$

Apply Itô's lemma to function $\ln S_t$ where S_t is driven by the diffusion process above. Then $\ln S_t$ follows the SDE:

$$d \ln S_t = (u - \frac{1}{2}\sigma^2)dt + \sigma dW_t \quad (1.4)$$

Integrating from t to T , we have:

$$S_T = S_t e^{(u - \frac{1}{2}\sigma^2)(T-t) + \sigma(W_T - W_t)} \quad (1.5)$$

\mathbb{Q} denotes a new probability measure where there is a \mathbb{Q} Brownian motion that

$$W_t^Q = W_t + \frac{u-r}{\sigma}t$$

Its dynamics under the \mathbb{Q} measure is:

$$dS_t = rS_t dt + \sigma S_t dW_t^Q \quad (1.6)$$

Apply Itô's lemma to \tilde{S}_t , the discounted stock price $\tilde{S}_t = \frac{S_t}{e^{rt}}$ is driven by a simple SDE:

$$d\tilde{S}_t = \sigma \tilde{S}_t dW_t^Q \quad (1.7)$$

$$\tilde{S}_t = \tilde{S}_0 e^{-\frac{1}{2}\sigma^2 t + \sigma W_t^Q} \quad (1.8)$$

Therefore, from the risk neutral valuation argument, the arbitrage-free price under risk neutral measure \mathbb{Q} :

$$C_{BS}(S_t, t) = e^{-r(T-t)} \mathbb{E}^Q[\max(S_T - K, 0) | F_t] \quad (1.9)$$

$$C_{BS}(S_t, t) = e^{-r(T-t)} \mathbb{E}^Q\{[S_t e^{(r - \frac{1}{2}\sigma^2)(T-t) + \sigma(W_T^Q - W_t^Q)} - K]^+\} \quad (1.10)$$

$$C_{BS}(S_t, t) = \Phi(d_1)S_t - \Phi(d_2)Ke^{-r(T-t)} \quad (1.11)$$

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

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

$$(1.12)$$

$\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

Similarly, the value of the put option is given by

$$P_{BS} = \Phi(-d_2)Ke^{-r(T-t)} - \Phi(-d_1)S_t \quad (1.13)$$

The Black-Scholes formula gives the price of European put and call options. This price is consistent with the Black-Scholes PDE: this follows since the formula can be obtained by solving the PDE for the corresponding terminal and boundary conditions.

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \quad (1.14)$$

Chapter 2

Barrier Option Pricing under the Black Scholes

A barrier option is a type of exotic option, in which the payoff demands reaching or crossing of a barrier (predetermined price) by the underlying product. They include call options and put options, and are similar to common options in many aspects. Barrier options become active/inactive when the underlying product crosses the barrier.

Barrier options can be grouped into two as knock-in options and knock-out options. Knock-in barrier options are inactive options at the beginning, but become active on reaching the barrier. On the other hand knock-out options starts as active but become inactive on reaching the barrier. Many barrier options carry rebates, which are paid off to the holder on reaching the barrier.

Barrier options are available in both European and American forms. In barrier options trading, premiums are paid in advance. Barrier options come in 4 types like up & out, up & in, down & out, and down & in. Of these four types 4 we can

take either call or put- giving us total 8 single barrier types. Barriers also come in various other forms including double barriers, parisians, and partial time barriers. Here, our discussion only target on single barrier types.

2.1 Valuation Methods

The valuation of barrier options can be tricky, because unlike other simpler options they are path-dependent — that is, the value of the option at any time depends not just on the underlying at that point, but also on the path taken by the underlying since if the barrier is crossed the option is initiated or exterminated. Although the classical Black–Scholes approach does not directly apply, several more complex methods can be used.

An approach is to study the law of the maximum (or minimum) of the underlying. This approach gives explicit prices to barrier options under the Black-Scholes framework. Also when an exact formula is difficult to obtain, barrier options can be priced with the Monte Carlo path simulation. However, computing the Greeks (sensitivities) using this approach is numerically unstable.

2.2 Analytical Solution under the Black-Scholes Model

Using the methods of equivalent martingale pricing (risk neutral valuation principle), the price of a down-and-out put barrier option at time zero with respect to maturity T , strike K , barrier S_b , initial stock price S is given by:

$$P_{dop}(S_T, K, S_b) = e^{-rT} E^Q[(K - S_T) \mathbb{1}_{\{K \geq S_T, m_0^T \geq S_b\}}] \quad (2.1)$$

Note:

1. The realized minimum value of underlying stock from time zero to time t is defined as $m_0^t = \min_{0 < u < t} S_u$.

2. The indicator function $\mathbb{1}_A(X)$ where:

$$\mathbb{1}_A(X) = \begin{cases} 1 & \text{if } X \in A \\ 0 & \text{otherwise} \end{cases}$$

is a useful notational device, in probability theory: if \mathbf{X} is a probability space with probability measure \mathbb{P} and A is a measurable set, then $\mathbb{1}_A$ becomes a random variable whose expected value is equal to the probability of A :

$$E(\mathbb{1}_A) = \int_X \mathbb{1}_A(x) d\mathbb{P} = \int_A d\mathbb{P} = P(A) \quad (2.2)$$

In view of linearity of the expectation, the above equation of down-and-out put option price $P_{dop}(S_T, K, S_b)$ can break into four terms:

$$\begin{aligned} & P_{dop}(S_T, K, S_b) \\ &= e^{-rT} (E^Q[(K - S_T) \mathbb{1}_{\{S_T < K\}}] - E[(K - S_T) \mathbb{1}_{\{S_T < S_b\}}] \\ & \quad - E[(K - S_T) \mathbb{1}_{\{S_T > S_b; m_0^T < S_b\}}] + E[(K - S_T) \mathbb{1}_{\{S_T > K; m_0^T < S_b\}}]) \end{aligned} \quad (2.3)$$

In the above equation, the first term is the price of a vanilla European put option under the Black-Scholes model.

The second term is similar to the Black Scholes put price.

$$e^{-rT} E^Q[(K - S_T) \mathbb{1}_{\{S_T < S_b\}}] = \Phi(-d_2(S_b)) K e^{-rT} - \Phi(-d_1(S_b)) S_0 \quad (2.4)$$

$$d_{1,2}(S_b) = \frac{\ln \frac{S_t}{S_b} + (r \pm \frac{\sigma^2}{2})(T - t)}{\sigma \sqrt{T - t}}$$

The value of the last two terms in price function $P_{dop}(S_T, K, S_b)$ requires the determination of the joint distribution function of S_T and m_0^T . Therefore, we apply the following theorems to solve the problem.

- **Theorem 1 (Reflection Principle)**

Let W_t^0 denote the Brownian motion that starts at zero, with constant volatility σ and zero drift rate. $\xi = \inf\{t : W_t^0 = h\}$.

$$\hat{W}_t^0 = \begin{cases} W_t^0, & t < \xi \\ 2h - W_t^0, & \xi < t < T \end{cases}$$

Then \hat{W}_t^0 is also a standard Brownian motion.

- **Theorem 2 (Girsanov's Theorem)**

Girsanov's Theorem begin with a Brownian motion under measure \mathbb{P} and then construct a new measure \mathbb{Q} under which a “translated” process is a Brownian motion.

So assume that $\{W_t^P\}_{t \geq 0}$ defined on (Ω, F, \mathbb{P}) is a Brownian motion starting from zero. Now define for an process $\{\gamma_s\}_{s \geq 0}$ adapted to the filtration $\{F_t\}_{t \geq 0}$

$$Z_t = e^{\int_0^t \gamma_s dW_s^P - \frac{1}{2} \int_0^t \gamma_s^2 ds} \quad (2.5)$$

$E[Z_t] = 1$ and for each $T > 0$ define a new probability measure \mathbb{Q} on $\{F_t\}_{t \geq 0}$ by

$$\mathbb{Q} = E(\mathbb{1}_A Z_T) \quad \text{for } A \in F_T \quad (2.6)$$

Define a process $W_t^Q = W_t^P - \int_0^t \gamma_s ds$. Then for each fixed $T < \infty$ the process W_t^Q is a Brownian motion on $(\Omega, F_T, \mathbb{Q})$

We illustrate how the reflection principle is applied to derive the joint law of the minimum value over $[0, T]$ and terminal value of a Brownian motion. Regarding to pricing a barrier option, we would like to find $P(W_t^u > k, m_0^t < h)$. To figure out the last two terms, we construct three transformation equations to apply the reflection principle and the Girsanov's Theorem.

$$g = \frac{r - \frac{1}{2}\sigma^2}{\sigma} \quad (2.7)$$

$$h = \frac{1}{\sigma} \ln\left(\frac{S_b}{S_0}\right) \quad (2.8)$$

$$k = \frac{1}{\sigma} \ln\left(\frac{S_b}{S_0}\right) \quad (2.9)$$

To calculate the last term of (2.4):

$$\begin{aligned}
& e^{-rT} E^Q[(K - S_T) \mathbb{1}_{\{S_T > K; m_0^T < S_b\}}] \\
&= e^{-rT} (K E^Q[e^{gW_T^Q - \frac{1}{2}g^2T} \mathbb{1}_{\{W_T^Q > k; \inf_{0 \leq t \leq T} W_t^Q < h\}}] \\
&\quad - E^Q[e^{gW_T^Q - \frac{1}{2}g^2T} S_0 e^{\sigma W_T^Q} \mathbb{1}_{\{W_T^Q > k; \inf_{0 \leq t \leq T} W_t^Q < h\}}]) \\
&= e^{-rT} (K e^{-\frac{1}{2}g^2T} E^Q[e^{\{g(2h - W_T^Q)\}} \mathbb{1}_{\{2h - W_T^Q > k; \inf_{0 \leq t \leq T} W_t^Q < h\}}] \\
&\quad - e^{-\frac{1}{2}g^2T} E^Q[(e^{\{(g+\sigma)(2h - W_T^Q)\}} \mathbb{1}_{\{2h - W_T^Q > k; \inf_{0 \leq t \leq T} W_t^Q < h\}})]) \\
&= e^{-rT} (K e^{-\frac{1}{2}g^2T + 2hg} E^Q[e^{\{-gW_T^Q\}} \mathbb{1}_{\{2h - W_T^Q > k\}}] \\
&\quad - e^{-\frac{1}{2}g^2T + 2h(g+\sigma)} E^Q[(e^{\{(g+\sigma)(-W_T^Q)\}} \mathbb{1}_{\{2h - W_T^Q > k\}})]) \\
&= e^{-rT} [K e^{2hg} \Phi(\frac{(2h - k) + (g + \sigma)T}{\sigma\sqrt{T}})] \\
&\quad - S_0 e^{2h(g+\sigma) + g\sigma T + \frac{1}{2}\sigma^2T} \Phi(\frac{(2h - k) + (g + \sigma)T}{\sigma\sqrt{T}})
\end{aligned}$$

Now we can re-plug in g, h, k , thus get the expression of the last term of (2.4)

Then we put all derivations into a formula that represents P_{dop} to be the value of this down-and-out put barrier option:

$$\begin{aligned}
p_{dop}(S_0, K, S_b) &= K e^{-rT} \Phi(-d_2) - S_0 \Phi(-d_1) + S_0 \Phi(-x_1) \\
&\quad - K e^{-rT} \Phi(-x_1 + \sigma\sqrt{T}) - S_0 \left(\frac{S_b}{S_0}\right)^{2\eta} [\Phi(y) - \Phi(y_1)] \\
&\quad + K e^{-rt} \left(\frac{S_b}{S_0}\right)^{2\eta-2} [\Phi(y - \sigma\sqrt{T}) - \Phi(y_1 - \sigma\sqrt{T})]
\end{aligned}$$

with

$$\begin{aligned}
\eta &= \frac{r + \frac{1}{2}\sigma^2}{\sigma^2}, \quad y = \frac{\ln(\frac{S_b^2}{S_0 K})}{\sigma\sqrt{T}} + \eta\sigma\sqrt{T}, \\
x_1 &= \frac{\ln(\frac{S_0}{S_b})}{\sigma\sqrt{T}} + \eta\sigma\sqrt{T}, \quad y_1 = \frac{\ln(\frac{S_b}{S_0})}{\sigma\sqrt{T}} + \eta\sigma\sqrt{T}
\end{aligned}$$

(2.10)

This analytical solution for continuously monitored down-and-out put barrier option is also in accordance with the formula given in [13]:

$$\begin{aligned}
P_{dop}(S_o, K, S_b) = & -S_0 e^{-rT} (\Phi(d_3) - \Phi(d_1) - b[\Phi(d_8) - \Phi(d_6)]) \\
& + K e^{-rT} (\Phi(d_4) - N(d_2) - a[\Phi(d_7) - \Phi(d_5)])
\end{aligned} \tag{2.11}$$

with

$$\begin{aligned}
a &= \left(\frac{S_b}{S_0}\right)^{-1+\frac{2r}{\sigma^2}} \\
b &= \left(\frac{S_b}{S_0}\right)^{1+\frac{2r}{\sigma^2}} \\
d_1 &= \frac{\ln \frac{S_0}{K} + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_2 &= \frac{\ln \frac{S_0}{K} + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_3 &= \frac{\ln \frac{S_0}{S_b} + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_4 &= \frac{\ln \frac{S_0}{S_b} + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_5 &= \frac{\ln \frac{S_0}{S_b} - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_6 &= \frac{\ln \frac{S_0}{S_b} - (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_7 &= \frac{\ln \frac{S_0 K}{S_b^2} - (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \\
d_8 &= \frac{\ln \frac{S_0 K}{S_b^2} - (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}
\end{aligned}$$

2.2.1 A Continuity Correction to Discrete Barrier Option

The barrier option pricing formulas presented so far assume continuous monitoring of the barrier. In practice, the barrier is normally monitored only at discrete points in time. An exception is the currency options market, where the barrier is frequently monitored almost continuously. For equity option in our case, the barrier is typically monitored against an periodical time point price. Discrete monitoring will naturally lower the probability of barrier hits compared with continuous barrier monitoring. Broadie, Glasserman and Kou [7] showed theoretically and through examples that discrete barrier options can be priced with remarkable accuracy using the following simple correction.

S_b is the continuous barrier options formulas with a discrete barrier level S_d equal to

$$S_d = S_b e^{\beta \sigma \sqrt{\Delta t}}$$

if the barrier is above the underlying security, and to

$$S_d = S_b e^{-\beta \sigma \sqrt{\Delta t}}$$

if the barrier is below the underlying security. Δt is the time between monitoring evens, and $\beta \approx 0.5826$.

2.3 Monte Carlo Methods

Monte Carlo methods are based on the analogy between probability and volume. The mathematics of measure formalizes the intuitive notion of probability, associating an event with a set of outcomes and defining the probability of the event to be its volume or measure relative to that of a universe of possible outcomes. Monte Carlo uses this identity in reverse, calculating the volume of a set by interpreting the volume as a probability. In the simplest case, this means sampling randomly from a universe of possible outcomes and taking the fraction of random draws that fall in a given set as an estimate of the set's volume. The law of large numbers ensures that this estimate converges to the correct value as the number of draws increases. The central limit theorem provides information about the likely magnitude of the error in the estimate after a finite number of draws.

In mathematical finance, a Monte Carlo option model uses Monte Carlo methods to calculate the value of an option with multiple sources of uncertainty or with complicated features.

The term 'Monte Carlo method' was coined by Stanislaw Ulam in the 1940s. The first application to option pricing was by Phelim Boyle in 1977 (for European options). In 1996, M. Broadie and P. Glasserman showed how to price Asian options by Monte Carlo. In 2001 F. A. Longstaff and E. S. Schwartz developed a practical Monte Carlo method for pricing American-style options.

We look at a simple example of European call option. Consider the option granting the holder the right to buy the stock at a fixed price K at a fixed time T in

the future; the current time is $t = 0$. The payoff to option holder is

$$(S_T - K)^+ = \max[0, S_T - K] \quad (2.12)$$

To get the present value of this payoff we multiply by a discount factor e^{-rT} . We denote the expected present value by $E[e^{-rT}(S_T - K)^+]$

As we know we apply the risk neutral measure into the calculation of the above expectation. The solution to the Black-Scholes SDE $dS_t/S_t = rd_t + \sigma dW_t^Q$ under risk neutral measure is

$$S_T = S_0 e^{[(r - \frac{1}{2}\sigma^2)T + \sigma W_T]} \quad (2.13)$$

The random variable W_T is normally distributed with mean 0 and variance T . We may therefore represent the terminal stock price as

$$S_T = S_0 e^{[(r - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}Z]} \quad (2.14)$$

where Z is a standard normal variable.

Then we can simulate the path of stock price and apply the Monte Carlo sampling methods to estimate price of the option by the expectation form $E[e^{-rT}(S_T - K)^+]$.

2.3.1 The Main Framework of Monte Carlo

Suppose we want to estimate some quantity $\theta = \mathbb{E}[h(\mathbf{X})]$, where $\mathbf{X} = \{X_1, X_2, \dots, X_n\}$ is a random vector in \mathbf{R}^n . $h(\cdot)$ is a function from \mathbf{R}^n to \mathbf{R} , and $\mathbb{E}[|h(\mathbf{X})|] < \infty$. Note that \mathbf{X} could represent the values of a stochastic process at different points in time. For example, X_i might be the price of a particular stock at time i and

$h(\cdot)$ might be given by:

$$h(\mathbf{X}) = \frac{X_1 + X_2 + \dots + X_n}{n} \quad (2.15)$$

so then θ is the expected average value of the stock price. To estimate θ we use the following algorithm:

Monte Carlo Algorithm

```

for i = 1 to n
  generate  $\mathbf{X}_i$ 
  set  $h_i = h(\mathbf{X}_i)$ 
set  $\hat{\theta}_n = \frac{h_1 + h_2 + \dots + h_n}{n}$ 

```

There are two reasons which show why $\hat{\theta}$ to be a good estimator:

1. $\hat{\theta}_n$ is an unbiased estimator.

$$\mathbb{E}[\hat{\theta}_n] = \frac{\sum_{i=1}^n h_i}{n} = \frac{\sum_{i=1}^n h(\mathbf{X}_i)}{n} = \frac{n\theta}{n} = \theta \quad (2.16)$$

2. $\hat{\theta}_n$ is consistent. That is:

$$\hat{\theta}_n \rightarrow \theta \quad \text{with probability 1 as } n \rightarrow \infty \quad (2.17)$$

This is followed by Strong Law of Large Numbers.

Note that we can also estimate probabilities this way by representing them as expectations. In particular, if $\theta = \mathbb{P}(X \in A)$, then $\theta = \mathbb{E}[\mathbb{1}_A(X)]$.

We consider simple forms of barrier options. A down-and-out European put op-

tion.

Its payoff at the expiry is given by the product:

$$\mathbb{1}_{\{K \geq S_T, m_0^T \geq H\}} \cdot (K - S_T)^+ \quad (2.18)$$

The risk-neutral price of the down-and-out European call option at time 0 is the expected discounted payoff

$$P_{dop} = \mathbb{E}[\mathbb{1}_{\{K \geq S_T, m_0^T \geq H\}}(K - S_T)^+] \quad (2.19)$$

2.3.2 Variance Reduction Techniques

We have seen in last section that one way to improve the accuracy of an estimate is to increase the number of replications n , since $\text{Var}(\bar{X}(n)) = \text{Var}(X_i)/n$. However, this brute-force approach may require an excessive computational effort. An alternative is to work on the numerator of this fraction and to reduce the variance of the samples X_i directly. This may be accomplished in different ways, more or less complicated, and more or less rewarding as well. [14]

Antithetic sampling

Suppose as usual that we would like to estimate $\theta = \mathbb{E}[h(\mathbf{X})] = \mathbb{E}[Y]$, and that we have generated two samples, Y_1 and Y_2 . Then an unbiased estimate of θ is given by

$$\hat{\theta} = \frac{Y_1 + Y_2}{2} \quad (2.20)$$

and

$$\text{Var}(\hat{\theta}) = \frac{\text{Var}(Y_1) + \text{Var}Y_2 + 2\text{Cov}(Y_1, Y_2)}{4} \quad (2.21)$$

In the antithetic path plot example: we construct the estimator by using two

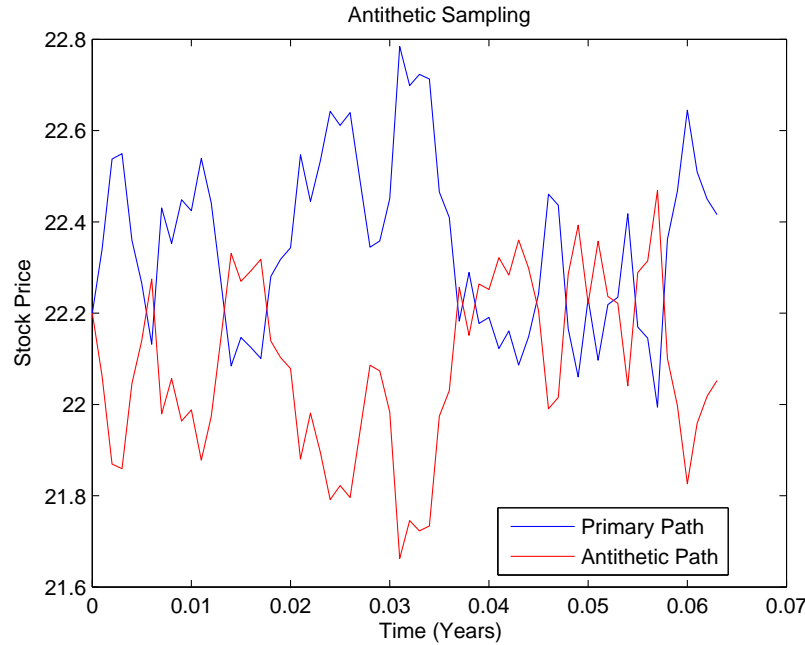


Figure 2.1: Antithetic Path Example

Brownian motion trajectories that are mirror images of each other. This causes cancellation of dispersion.

If Y_1 and Y_2 are IID, then $\text{Var}(\hat{\theta}) = \frac{\text{Var}(Y)}{2}$. However, we could reduce $\text{Var}(\hat{\theta})$ if we could arrange it so that $\text{Cov}(Y_1, Y_2) < 0$. We now describe the method of antithetic variates for doing this. We will begin with the case where Y is a function of IID $U(0, 1)$ random variables so that

$$\theta = E[h(\mathbf{U})]$$

where $\mathbf{U} = (U_1, \dots, U_m)$ and the U_i 's are IID $\sim U(0, 1)$. The crude Monte Carlo algorithm, assuming we use $2n$ samples, is as follows: In the above algorithm,

Crude Simulation Algorithm for Estimating θ

```

for  $i = 1$  to  $2n$ 
  generate  $\mathbf{U}_i$ 
  set  $Y_i = h(\mathbf{U}_i)$ 
end for
set  $\hat{\theta}_{2n} = \bar{Y}_{2n} = \sum_{i=1}^{2n} Y_i / 2n$ 
set  $\hat{\sigma}_{2n}^2 = \sum_{i=1}^{2n} (Y_i - \bar{Y}_{2n})^2 / 2n - 1$ 
set approx.  $100(1 - \alpha)\%$  CI =  $\hat{\theta}_{2n} \pm z_{1-\alpha/2} \frac{\hat{\sigma}_{2n}}{\sqrt{2n}}$ 

```

however, we could also used the $1 - U_i$'s to generate sample Y values, since if $U_i \sim U(0, 1)$, then so too is $1 - U_i$. We can use this fact to construct another estimator θ as follows. As before, we set $Y_i = h(\mathbf{U}_i)$, where $\mathbf{U}_i = (U_1^{(i)}, \dots, U_m^{(i)})$. We now also set $\tilde{Y}_i = h(1 - \mathbf{U}_i)$, where we use $1 - \mathbf{U}_i$ to denote $(1 - U_1^{(i)}, \dots, 1 - U_m^{(i)})$. Note that $\mathbb{E}[Y_i] = \mathbb{E}[\tilde{Y}_i]$ so that in particular, if

$$Z_i := \frac{Y_i + \tilde{Y}_i}{2},$$

then $\mathbb{E}[Z_i] = \theta$. This means that Z_t is an also unbiased estimator of θ . If the U_i 's are IID, then so too are the Z_i 's and we can use them as usual to compute approximate confidence interval for θ . We say that U_i and $1 - U_i$ are *antithetic variates* and we then have the following *antithetic variate* simulation algorithm.

As usual, $\hat{\theta}_{a,n}$ is an unbiased estimator of θ and the Strong Law of Large Numbers implies that $\hat{\theta}_{a,n} \rightarrow \theta$ almost surely as $n \rightarrow \infty$. Each of the two algorithms we have described above uses $2n$ samples so the question naturally arises as to which algorithm is better. Note that both algorithms require approximately the same amount of effort so that comparing the two algorithms amounts to comput-

Antithetic Variate Simulation Algorithm for Estimating θ

```

for  $i = 1$  to  $n$ 
  generate  $\mathbf{U}_i$ 
  set  $Y_i = h(\mathbf{U}_i)$  and  $\tilde{Y}_i = h(1 - \mathbf{U}_i)$ 
  set  $Z_i = (Y_i + \tilde{Y}_i)/2$ 
end for
set  $\hat{\theta}_{n,a} = \bar{Z}_n = \sum_{i=1}^n Z_i/n$ 
set  $\hat{\sigma}_{n,a}^2 = \sum_{i=1}^n (Z_i - \bar{Z}_n)^2/n - 1$ 
set approx.  $100(1 - \alpha)\%$  CI =  $\hat{\theta}_{a,n} \pm z_{1-\alpha/2} \frac{\hat{\sigma}_{n,a}}{\sqrt{n}}$ 

```

ing which estimator has a smaller variance.

It is easy to see that

$$\text{Var}(\hat{\theta}_{2n}) = \text{Var}\left(\frac{\sum_{i=1}^{2n} Y_i}{2n}\right) = \frac{\text{Var}(Y)}{2n}$$

and

$$\begin{aligned} \text{Var}(\hat{\theta}_{n,a}) &= \text{Var}\left(\frac{\sum_{i=1}^n Z_i}{n}\right) = \frac{\text{Var}(Z)}{n} \\ &= \frac{\text{Var}(Y + \hat{Y})}{4n} = \frac{\text{Var}(Y)}{2n} + \frac{\text{Cov}(Y, \hat{Y})}{2n} \\ &= \text{Var}(\hat{\theta}_{2n}) + \frac{\text{Cov}(Y, \hat{Y})}{2n} \end{aligned}$$

Therefore, $\text{Var}(\hat{\theta}_{n,a}) < \text{Var}(\hat{\theta}_{2n})$ if and only if $\text{Cov}(Y, \hat{Y}) < 0$

Control Variate

Suppose again that we wish to estimate $\theta := \mathbb{E}[Y]$ where $Y = h(\mathbf{X})$ is the output of a simulation experiment. Suppose that Z is also an output of the simulation or

that we can easily output it if we wish. Finally, we assume that we know $E[Z]$.

Then we can construct many unbiased estimation of θ :

1. $\hat{\theta} = Y$, our estimator
2. $\hat{\theta}_c = Y + c(Z - \mathbb{E}[Z])$

where c is some real number. It is clear that $\mathbb{E}[\hat{\theta}_c] = \theta$. The question is whether or not $\hat{\theta}_c$ has a lower variance than $\hat{\theta}$. To answer this question, we compute $\text{Var}(\hat{\theta}_c)$ and obtain

$$\text{Var}(\hat{\theta}_c) = \text{Var}(Y) + c^2 \text{Var}(Z) + 2c \text{Cov}(Y, Z) \quad (2.22)$$

Since we are free to choose c , we should choose it to minimize $\text{Var}(\hat{\theta}_c)$. Simple calculus then implies that the optimal value of c is given by

$$c^* = -\frac{\text{Cov}(Y, Z)}{\text{Var}(Z)} \quad (2.23)$$

Substituting for c^* in () we see that

$$\begin{aligned} \text{Var}(\hat{\theta}_c) &= \text{Var}(Y) - \frac{\text{Cov}(Y, Z)^2}{\text{Var}(Z)} \\ &= \text{Var}(\hat{\theta}) - \frac{\text{Cov}(Y, Z)^2}{\text{Var}(Z)} \end{aligned}$$

So we see that in order to achieve a variance reduction it is only necessary that $\text{Cov}(Y, Z) \neq 0$. In this case, Z is called a *control variate* for Y . To use the control variate Z in our simulation, we would like to modify our algorithm so that after generating n samples of Y and Z we would simply set

$$\hat{\theta}_{c^*} = \frac{\sum_{i=1}^n (Y_i + c^*(Z_i - \mathbb{E}[Z]))}{n} \quad (2.24)$$

Control Variate Simulation Algorithm for Estimating θ

```

/*Do pilot simulation first*/
for  $i = 1$  to  $p$ 
    generate  $(Y_i, Z_i)$ 
end for
compute  $\hat{c}^*$ 
/*Now do main simulation*/
for  $i = 1$  to  $n$ 
    generate  $(Y_i, Z_i)$ 
    set  $V_i = Y_i + \hat{c}^*(Z_i - \mathbb{E}[Z])$ 
end for
set  $\hat{\theta}_{\hat{c}^*} = \bar{V}_n = \sum_{i=1}^n V_i / n$ 
set  $\hat{\sigma}_{n,v}^2 = \sum_{i=1}^n (V_i - \hat{\theta}_{\hat{c}^*})^2 / (n - 1)$ 
set approx.  $100(1 - \alpha)\%CI = \hat{\theta}_{\hat{c}^*} \pm z_{1-\alpha/2} \frac{\hat{\sigma}_{n,v}}{\sqrt{n}}$ 

```

There is a problem with this, however, as we usually do not know $\text{Cov}(Y, Z)$. We overcome this problem by doing p pilot simulation and setting

$$\widehat{\text{Cov}}(Y, Z) = \frac{\sum_{j=1}^p (Y_j - \bar{Y}_p)(Z_j - \mathbb{E}[Z])}{p - 1}$$

If it is also the case that $\text{Var}(Z)$ is unknown, then we also estimate it with

$$\widehat{\text{Var}}(Z) = \frac{\sum_{j=1}^p (Z_j - \mathbb{E}[Z])^2}{p - 1}$$

and finally set

$$\hat{c}^* = \frac{\widehat{\text{Cov}}(Y, Z)}{\widehat{\text{Var}}(Z)}$$

Assuming we can find a control variate, our control variate simulation algorithm is as follows. Note that the V_i 's are IID, so we can compute approximate confidence interval as before.

Conditional Monte Carlo

We now consider the conditional Monte Carlo variance reduction technique. The idea here is very simple. As was the case with control variates, we use our knowledge about the system being studied to reduce the variance of our estimator. As usual, suppose we wish to estimate $\theta = \mathbb{E}[h(\mathbf{X})]$ where $\mathbf{X} = (X_1, \dots, X_m)$. If we could compute θ analytically, then this would be equivalent to solving an m -dimensional integral. However, maybe it is possible to evaluate part of the integral analytically. If so, then we might be able to use simulation to estimate the other part and thereby obtain a variance reduction.

Let \mathbf{X} and \mathbf{Z} be random vectors, and let $Y = h(\mathbf{X})$ be a random variable. Suppose we set

$$V = \mathbb{E}[Y|\mathbf{Z}] \quad (2.25)$$

Then V is itself a random variable that depends on \mathbf{Z} , so that we may write $V = g(\mathbf{Z})$ for some function, $g(\cdot)$. We also know that

$$\mathbb{E}[V] = \mathbb{E}[\mathbb{E}[Y|\mathbf{Z}]] = \mathbb{E}[Y] \quad (2.26)$$

so that if we are trying to estimate $\theta = \mathbb{E}[Y]$, one possibility would be to simulate V 's instead of simulating Y 's. In order to decide which would be a better estimator of θ , it is necessary to compare the variances of Y and $\mathbb{E}[Y|\mathbf{Z}]$. To do this, using the conditional variance formula:

$$\text{Var}(Y) = \mathbb{E}[\text{Var}(Y|\mathbf{Z})] + \text{Var}(\mathbb{E}[Y|\mathbf{Z}])$$

Now $\text{Var}(Y|\mathbf{Z})$ is also a random variable that depends on \mathbf{Z} , and since a variance is always non-negative, it must follow that $\mathbb{E}[\text{Var}(Y|\mathbf{Z})] \geq 0$. But then implies

$$\text{Var}(Y) \geq \text{Var}(\mathbb{E}[Y|\mathbf{Z}])$$

so we can conclude that V is a better estimator of θ than Y .

To see this from another perspective, suppose that to estimate θ we first have to simulate \mathbf{Z} and then simulate Y given \mathbf{Z} . If we can compute $\mathbb{E}[Y|\mathbf{Z}]$ exactly, then we can eliminate the additional noise that comes from simulating Y given \mathbf{Z} , thereby obtaining a variance reduction. Finally, we point out that in order for the conditional expectation method to be worthwhile, it must be the case that Y and \mathbf{Z} are dependent.

We want to estimate $\theta := \mathbb{E}[h(\mathbf{X})] = \mathbb{E}[Y]$ using conditional Monte Carlo. To do so, we must have another variable or vector, \mathbf{Z} , that satisfies the following requirements:

- 1. \mathbf{Z} can be easily simulated
- 2. $V := g(\mathbf{Z}) := \mathbb{E}[Y|\mathbf{Z}]$ can be computed exactly.

This means that we can simulate a value of V by first simulating a value of \mathbf{Z} and then setting $V := g(\mathbf{Z}) := \mathbb{E}[Y|\mathbf{Z}]$. We then have the following algorithm for estimating θ .

Conditional Monte Carlo Algorithm for Estimating θ

```

for  $i = 1$  to  $n$ 
  generate  $\mathbf{Z}_i$ 
  compute  $g(\mathbf{Z}_i) = \mathbb{E}[Y|\mathbf{Z}_i]$ 
  set  $V_i = g(\mathbf{Z}_i)$ 
end for
set  $\hat{\theta}_{n,cm} = \bar{V}_n = \sum_{i=1}^n V_i / n$ 
set  $\hat{\sigma}_{n,cm}^2 = \sum_{i=1}^n (V_i - \bar{V}_n)^2 / (n - 1)$ 
set approx. 100(1 -  $\alpha$ )% CI =  $\hat{\theta}_{n,cm} \pm z_{1-\alpha/2} \frac{\hat{\sigma}_{n,cm}}{\sqrt{n}}$ 

```

It is also possible that other variance reduction methods could be used in conjunction with conditioning.

Importance Sampling

Suppose we wish to estimate $\theta = \mathbb{E}_f[h(X)]$ where X has PDF f . Let g be another PDF with the property that $g(x) \neq 0$ whenever $f(x) \neq 0$. That is, g has the same support as f . Then

$$\theta = \mathbb{E}_f[h(X)] = \int h(x)f(x)dx = \int h(x)\frac{f(x)}{g(x)}g(x)dx = \mathbb{E}_g\left[\frac{h(X)f(X)}{g(X)}\right]$$

where $\mathbb{E}_g[\cdot]$ denotes an expectation with respect to the density g . This has very important implications for estimating θ . The original simulation method is to generate n samples of X from the density, f , and set $\hat{\theta}_n = \sum h(x_j)/n$. An alternative method, however, is to generate n values of X from the density, g , and set

$$\hat{\theta}_{n,is} = \sum_{j=1}^n \frac{h(X_j)f(X_j)}{g(X_j)}$$

$\hat{\theta}_{n,is}$ is then an importance sampling estimator of θ . We often define

$$h^*(X) := \frac{h(X)f(X)}{g(X)}$$

so that $\theta = \mathbb{E}_g[h^*(X)]$. We refer to f and g as the original and importance sampling densities, respectively. We also refer to f/g as the likelihood ratio. The general formation

Let $\mathbf{X} = (X_1, \dots, X_n)$ be a random vector with joint PDF $f(x_1, \dots, x_n)$ and suppose we wish to estimate $\theta = \mathbb{E}_f[h(\mathbf{X})]$. Let $g(x_1, \dots, x_n)$ be another PDF such that

$g(x) \neq 0$ whenever $f(x) \neq 0$. Then

$$\begin{aligned}\theta = \mathbb{E}_f[h(\mathbf{X})] &= \int_{x_1} \dots \int_{x_n} h(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= \int_{x_1} \dots \int_{x_n} h(x_1, \dots, x_n) \frac{h(x_1, \dots, x_n)}{f(x_1, \dots, x_n)} g(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= \mathbb{E}_g[h^*(\mathbf{X})]\end{aligned}$$

where $h^*(\mathbf{X}) := h(\mathbf{X})f(\mathbf{X}) = g(\mathbf{X})$. Again have two methods for estimating θ : the original method where we simulate with respect to the density function, f , and the importance sampling method where we simulate with respect to the density, g

Then we know

$$\theta = \mathbb{E}_f[h(\mathbf{X})] = \mathbb{E}_g[h^*(\mathbf{X})]$$

and this give rise to two estimators

1. $h(\mathbf{X})$ where $\mathbf{X} \sim f$ and
2. $h^*(\mathbf{X})$ where $\mathbf{X} \sim g$

The variance of the importance sampling estimator is given by

$$\begin{aligned}\text{Var}_g(h^*(\mathbf{X})) &= \int h^*(x)^2 g(x) dx - \theta^2 \\ &= \int \frac{h(x)^2 f(x)}{g(x)} f(x) dx - \theta^2\end{aligned}$$

while the variance of the original estimator is given by $\text{Var}_f(h(\mathbf{X})) = \int h(x)^2 f(x) dx - \theta^2$. So the reduction in variance is then given by

$$\text{Var}_f(h(\mathbf{X})) - \text{Var}_g(h^*(\mathbf{X})) = \int h^*(x)^2 \left(1 - \frac{f(x)}{g(x)}\right) f(x) dx \quad (2.27)$$

In order to receive a variance reduction, the integral in (1) should be positive. For this to happen, we would like

- 1 $f(x)/g(x) > 1$ when $h(x)f(x)$ is small and
- 2 $f(x)/g(x) < 1$ when $h(x)f(x)$ is large.

Now the important part of the density, f , could plausibly be defined to be that region, A say, in the support of f where $h(x)f(x)$ is large. But, by the above observation, we would like to choose g so that $f(x)/g(x)$ is small whenever x is in A . That is, we would like a density, g , that puts more weight on A : hence the name importance sampling. Note that when h involves a rare event so that $h(x) = 0$ over “most” of the state space, it can then be particularly valuable to choose g so that we sample often from that part of the state space where $h(x) \neq 0$. This is why importance sampling is most useful for simulating rare events. Further guidance on how to choose g is obtained from the following observation.

As we are free to choose g , let's suppose we choose $g(x) = h(x)f(x) = \theta$. Then it is easy to see that

$$\text{Var}_g(h^*(X)) = \theta^2 - \theta^2 = 0$$

so that we have a zero variance estimator!

This means that if we sample with respect to this particular choice of g , then we would only need one sample and this sample would equal θ with probability one. Of course this is not feasible in practice. After all, since it is θ that we are trying to estimate, it does not seem likely that we could simulate a random variable whose density is given by $g(x) = h(x)f(x)/\theta$.

However, all is not lost and this observation can often guide us towards excellent choices of g that lead to extremely large variance reductions.

2.3.3 Path Generation of Barrier Options

The starting point for the application of Monte Carlo methods to option pricing is the generation of sample paths of the underlying stocks. Because the payoff of barrier options depend explicitly on the values of the underlying assets, we would better know the whole paths, or at least, a sequence of values at given instants. we may not know how to sample transitions of the underlying assets exactly and thus need to divide a time interval $[0, T]$ into smaller subintervals δ_t to obtain a more accurate approximation to sampling from the distribution at time T . Under the assumption that the stock price follows a geometric Brownian motion, we are facing a very special case:

$$dS_t = rS_t dt + \sigma S_t dW_t^Q \quad (2.28)$$

with W_t^Q denotes a standard Brownian motion under the risk neutral measure \mathbb{Q} . The Euler scheme [11] yields

$$S_{t+\delta_t} = (1 + r\delta_t)S_t + \sigma S_t \sqrt{\delta_t} Z \quad (2.29)$$

By applying Ito's lemma, we transform equation(12) into:

$$d\ln S_t = (r - \frac{1}{2}\sigma^2)dt + \sigma dW_t^Q \quad (2.30)$$

Integrating the equation above, we obtain:

$$S_t = S_0 e^{[(r - \frac{1}{2}\sigma^2)t + \sigma W_t^Q]} \quad (2.31)$$

Using lognormal distribution of S_t and letting $\nu = r - \frac{1}{2}\sigma^2$, we can obtain: By the properties of standard Wiener process, we get:

$$S_{t+\delta t} = S_t e^{\nu\delta t + \sigma\sqrt{\delta t}Z} \quad (2.32)$$

where $Z \sim N(0, 1)$

Based on the equation above, we can generate the sample paths for the asset price.

The value of an option is the present value of the expected payoff under a risk-neutral random walk.

- 1. Simulate the risk-neutral random walk starting at S_0 over the required time horizon. This gives one realization of the underlying price path.
- 2. For this realization calculate the option payoff.
- 3. Perform many more such realizations over the time horizon.
- 4. Calculate the average payoff over all realizations.
- 5. Take the present value of this average, this is the value of the option.

Essentially, the payoff is calculated for each generated path. An accurate estimation of the down-and-out put barrier option is obtained by discounting the average of all payoffs calculated.

2.3.4 Application of Variance Reduction to Barrier Options

Crude Monte Carlo

The result of Crude Monte Carlo is listed here and the input parameters are $S_0 = 22.2$, $r = 0.04$, $\sigma = 0.18$, $n = \text{Number of Simulations} = 1000, 10000$, S_b is the barrier level

In those tables, SE stands for standard error, corrected stands for the price after discrete monitoring correction.

The computation results on these tables are obtain from MATLAB.

These results mainly showed that the accuracy of estimated value is improved when the number of simulation is increased and the variance reduction technique is applied. The value from Monte Carlo approaches and converges to the analytical solution under the Black Scholes framework.

Table 2.1 shows the result of analytical solution in continuous setting and we made the a continuity corrected solution to discreetly monitored barrier option with respect to the every pair of barrier for 1000 or 10000 trials in MATLAB. We compare price and the standard error for different K and S_b (Barrier), the above results meet our expectation the payoff for a down-and-out put barrier options, the more the strike away from the barrier, the more it is valued which converges to a vanilla put option. By repeating more and more simulations, the standard error is substantially reduced as a result from the larger sample size. The Monte-Carlo prices are fairly close to the analytical solution, which demonstrate the credibility of Monte Carlo methods as well.

Table 2.1: **Crude Monte Carlo**

K	Barrier	Ncrossed/Npath	Analytical	Corrected	MC	Std. Error-10 ³	Std. Error-10 ⁴
22.5	22	830/1000		0.0043	0.0043	0.0011	
	21.5	452/1000		0.0637	0.0779	0.0062	
	21	228/1000		0.2117	0.202	0.0113	
25	22	801/1000		0.3131	0.32	0.0231	
	21.5	505/1000		1.0593	0.9781	0.0364	
	21	234/1000		1.7829	1.8071	0.04	
20	19.5	9/1000		0.0019	0.0031	0.0010	
	19	0/1000		0.0047	0.0043	0.0012	
	18.5	0/1000		0.0059	0.0073	0.0019	
22.5	22	8030/10000	0.0023		0.0047		0.0003697
	21.5	4725/10000	0.0509		0.0641		0.0018
	21	2388/10000	0.1904		0.2076		0.0036
25	22	8028/10000	0.229		0.3126		0.0071
	21.5	4699/10000	0.961		1.0703		0.0116
	21	2322/10000	1.7026		1.8014		0.0126
20	19.5	73/10000	0.0015		0.0022		0.0002508
	19	13/10000	0.0044		0.0047		0.0004754
	18.5	1/10000	0.0058		0.0068		0.0006449

Antithetic Variate

To apply the antithetic variate technique, we generate standard normal random variables \mathbf{Z}_j $j=1,2,...,n$ and define two set of paths in a paired style:

$$S_{t+\delta_t} = S_t e^{\nu\delta_t + \sigma\sqrt{\delta_t}Z}$$

$$S_{t+\delta_t} = S_t e^{\nu\delta_t + \sigma\sqrt{\delta_t}(-Z)}$$

Similarly, we define two set of payoffs as V_j^+ and V_j^- .

Now we construct our down-and-out put barrier option price estimator by averaging the above two set of sampling payoffs.

$$V = \frac{1}{2} \frac{1}{n} \sum_{j=1}^n (V_j^+ + V_j^-)$$

Table 2.2 shows the result adding Antithetic Variate to simulate the payoffs. We have seen the standard error apparently lowers than the crude Monte Carlo error without losing the accuracy of the price simulated.

Table 2.2: **Antithetic Variate**

K	Barrier	Analytical	Corrected	MC	AV SE-3	AV SE-4	Crude SE-3	Crude SE-4
22.5	22		0.0043	0.0037	0.00072186		0.0011	
	21.5		0.0637	0.0640	0.004		0.0062	
	21		0.2117	0.1904	0.008		0.0113	
25	22		0.3131	0.3361	0.0161		0.0231	
	21.5		1.0593	1.0312	0.0257		0.0364	
	21		1.7829	1.7967	0.028		0.04	
20	19.5		0.0019	0.0027	0.00059972		0.0010	
	19		0.0047	0.0056	0.0011		0.0012	
	18.5		0.0059	0.0052	0.0012		0.0019	
22.5	22	0.0023		0.0043		0.00024456		0.00036973
	21.5	0.0509		0.0640		0.0012		0.0018
	21	0.1904		0.2145		0.0025		0.0036
25	22	0.229		0.3048		0.005		0.0071
	21.5	0.961		1.0475		0.0081		0.0116
	21	1.7026		1.7743		0.0088		0.0126
20	19.5	0.0015		0.0018		0.0001523		0.00025079
	19	0.0044		0.0050		0.00034028		0.00047543
	18.5	0.0058		0.0053		0.00039607		0.00064491

Control Variate

In financial applications of the control variate method, the emphasis has traditionally been on using a closely related financial instrument whose value is known analytically to compute the value of another instrument by simulation. The combination of the two instruments allows us to construct an estimator with much less variance if both instruments are closely related. [21]

As used in other derivatives pricing, control variates may be used into barrier path simulation. The estimator includes a problem highly correlated with the one we want to solve. We must know the expectation of the correlated problem either analytically or numerically. The combined problem has less variance. Hence, we

must know the expectation of the control variate very well, because any uncertainty in the control variate will contaminate our desired result.

A reasonable candidate could be the control variate is the price of a vanilla put, which may be computed by the Black-Scholes formula.

Table 2.3 shows the result after using Control Variate technique. We observe that with vanilla put as control variate yield a good variance reduction. The standard error is significantly reduced and the method has done a variance reduction nearly as good as for the antithetic variate.

Table 2.3: **Control Variate**

K	Barrier	Analytical	Corrected	MC	CV SE-3	CV SE-4	Crude SE-3	Crude SE-4
22.5	22		0.0043	0.0046	0.0010		0.0011	
	21.5		0.0637	0.0556	0.005		0.0062	
	21		0.2117	0.1813	0.0107		0.0113	
25	22		0.3131	0.3359	0.0229		0.0231	
	21.5		1.0593	1.0389	0.0348		0.0364	
	21		1.7829	1.796	0.04		0.04	
20	19.5		0.0019	0.0015	0.0004828		0.0010	
	19		0.0047	0.0054	0.0007203		0.0012	
	18.5		0.0059	0.0046	0.0014		0.0019	
22.5	22	0.0023		0.0043		0.0003492		0.0003697
	21.5	0.0509		0.0645		0.0018		0.0018
	21	0.1904		0.2086		0.0035		0.0036
25	22	0.229		0.3108		0.0068		0.0071
	21.5	0.961		1.0533		0.0109		0.0116
	21	1.7026		1.7781		0.0125		0.0126
20	19.5	0.0015		0.0014		0.000174		0.0002508
	19	0.0044		0.0051		0.0002989		0.0004754
	18.5	0.0058		0.006		0.0001619		0.0006449

Conditional Monte Carlo

We may know that antithetic sampling may be not superb effective, as the payoff is non-monotonic with respect to the asset price at expiration. We try a different approach: variance reduction by conditioning.

We see that it is convenient to consider the price P_{di} of the down-and-in put. Pricing this knock-in option is equivalent to pricing the corresponding knock-out option, since we know that

$$P_{dop} = P_{bsp} - P_{dip}$$

Here P_{bsp} represents the price of a European put option under the Black-Scholes model.

Assume that we discretize the option life in time intervals of width δt , so that $T = M\delta t$, and consider the asset price path for $i, i = 1, \dots, M$:

$$\mathbf{S} = \{S_1, S_2, \dots, S_M\}.$$

Based on this path, we estimate the option price as

$$P_{dip} = e^{-rT} \mathbb{E}[\mathbb{1}_S (K - S_M)^+],$$

Here we define the indicator function $\mathbb{1}_S$ for the following discussion

$$\mathbb{1}_S = \begin{cases} 1 & S_j < S_b \text{ for some } j \\ 0 & \text{otherwise} \end{cases}$$

Now let j^* be the index of the time instant at which the barrier is first crossed; by convention, let $j^* = M + 1$ if the barrier is not crossed during the option life. At time $j^*\delta t$ the option is activated, and from now on it behaves just like a vanilla put. So, conditional on the crossing time $t^* = j^*\delta t$ and the price S_{j^*} at which we detect barrier crossing, we may use the Black-Scholes formula to estimate the expected value of the payoff. Hence, if the barrier is crossed before maturity, we

have

$$\mathbb{E}[\mathbb{1}_S(K - S_M)^+ | j^*, S_{j^*}] = e^{r(T-t^*)} d_{bsp}(S_{j^*}, K, T - t^*) \quad (2.33)$$

where $d_{bsp}(S_{j^*}, K, T - t^*)$ is the Black-Scholes price for a vanilla European put option with strike price K , initial underlying price S_{j^*} , and time to maturity $T - t^*$; the exponential term takes discounting into account, from maturity back to crossing time. Given a simulated path \mathbf{S} , this suggests using the following estimator of the price of down and in put option:

$$\mathbb{1}_S e^{-rt^*} d_{bsp}(S_{j^*}, K, T - t^*).$$

Table 2.4 shows the simulation results of the Conditional Monte Carlo compared to crude Monte Carlo method. The variance reduction did not work well as the standard errors we got from conditional Monte Carlo are all larger than the crude simulation.

Table 2.4: **Conditional Monte Carlo**

K	Barrier	Ncrossed	Analytical	Corrected	MC	CE SE-3	CE SE-4	Crude SE-3	Crude SE-4
22.5	22	804		0.0043	0.0031	0.009		0.0011	
	21.5	485		0.0637	0.0539	0.0169		0.0062	
	21	254		0.2117	0.1817	0.0212		0.0113	
25	22	807		0.3131	0.3068	0.0375		0.0231	
	21.5	487		1.0593	1.0158	0.0556		0.0364	
	21	201		1.7829	1.919	0.0509		0.04	
20	19.5	7		0.0019	0.0021	0.0015		0.0010	
	19	1		0.0047	0.0052	0.001		0.0012	
	18.5	1		0.0059	0.0046	0.0016		0.0019	
22.5	22	8044	0.0023		0.0044		0.0028		0.00036973
	21.5	4852	0.0509		0.0521		0.0054		0.0018
	21	2388	0.1904		0.2051		0.0066		0.0036
25	22	8095	0.229		0.2987		0.0118		0.0071
	21.5	4719	0.961		1.0686		0.0175		0.0116
	21	2407	1.7026		1.7587		0.0172		0.0126
20	19.5	78	0.0015		0.0018		0.000497		0.00025079
	19	12	0.0044		0.0049		0.000369		0.00047543
	18.5	3	0.0058		0.0057		0.000268		0.00064491

Importance Sampling

The last run in Monte Carlo simulation shows that variance reduction by conditioning may not be helpful. Even worse, we have run 10000 replications, but the barrier has been crossed only in 78, 12, 3 replications. This means that most of the replications are a wasted effort. In other words, with the data for this option, crossing the barrier is a rare event. This is a typical case in which importance sampling may help.

One possible idea is changing the drift of the asset price in such a way that crossing the barrier is more likely. We should go a step back and consider what we do in path generation. For each time step, we generate a normal variate Z_j with expected value $\nu = (r - \frac{\sigma^2}{2})\delta_t$ and variance $\sigma^2\delta_t$. All these variables are mutually independent, and the asset price is generated by setting

$$\log S_j - \log S_{j-1} = Z_j$$

Let \mathbf{Z} be the vector of the normal variables, and let $f(\mathbf{Z})$ be its joint density.

If we use the modified expected value $\nu - b$. We expect that the barrier will be crossed more often. Let $g(\mathbf{Z})$ be the joint density for the normal variables generated with this modified expected value. Combining importance sampling with the conditional expectation we have just described, as the case in conditional

Monte Carlo if the barrier is crossed before maturity we have:

$$\begin{aligned}
& E_g\left[\frac{f(\mathbf{Z})\mathbb{1}_S(K - S_M)^+}{g(\mathbf{Z})} | j^*, S_{j^*}\right] \\
&= \frac{f(z_1, z_2, \dots, z_j^*)}{g(z_1, z_2, \dots, z_j^*)} E_g\left[\frac{f(Z_{j^*+1}, \dots, Z_M)}{g(Z_{j^*+1}, \dots, Z_M)} \mathbb{1}_S(K - S_M)^+ | j^*, S_{j^*}\right] \\
&= \frac{f(z_1, z_2, \dots, z_j^*)}{g(z_1, z_2, \dots, z_j^*)} E_f[\mathbb{1}_S(K - S_M)^+ | j^*, S_{j^*}] \\
&= \frac{f(z_1, z_2, \dots, z_j^*)}{g(z_1, z_2, \dots, z_j^*)} e^{-r(T-t^*)} d_{bsp}(S_j^*, K, T - t^*)
\end{aligned}$$

We should generate the normal variables with expected value $(\nu - b)$, and multiply the conditional estimator by the likelihood ratio. The only open problem is how to compute the likelihood ratio as the drift in change of measure.

We consider the joint distribution of a multivariate normal with expected value u and covariance matrix Σ

$$f(z) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{\frac{1}{2}(\mathbf{Z}-\mathbf{u})^\top \Sigma^{-1}(\mathbf{Z}-\mathbf{u})}$$

In our case, due to the mutual independence of the random variables Z_j , the covariance matrix is a diagonal matrix with elements $\sigma^2 \delta_t$, and the vector of the expected values has components $u = (r - \frac{\sigma^2}{2})\delta_t$ for the density f and $u - b$ for the density g .

Therefore,

$$\begin{aligned}
& \frac{f(z_1, z_2, \dots, z_{j^*})}{g(z_1, z_2, \dots, z_{j^*})} \\
&= \exp \left(\frac{1}{2} \sum_{k=1}^{j^*} \left(\frac{z_k - u}{\delta_t} \right)^2 \right) \exp \left(\frac{1}{2} \sum_{k=1}^{j^*} \left(\frac{z_k - u + b}{\delta_t} \right)^2 \right) \\
&= \exp \left(-\frac{1}{2\sigma^2\delta_t} \sum_{k=1}^{j^*} [(z_k - u)^2 - (z_k - u + b)^2] \right) \\
&= \exp \left(-\frac{1}{2\sigma^2\delta_t} \sum_{k=1}^{j^*} [-2(z_k - u)b - b^2] \right) \\
&= \exp \left(-\frac{1}{2\sigma^2\delta_t} [-2b \sum_{k=1}^{j^*} z_k + 2j^*u_b - j^*b^2] \right)
\end{aligned}$$

We assume that one can provide a percentage b_p , here

$$b_p = \frac{b}{r - \frac{1}{2}\sigma^2}$$

Table 2.5 gives the result of using importance sampling in calculating the corrected payoff combined with a likelihood ratio. The variance is significantly reduced. By increasing b_p , we see the barrier is crossed many more times with in same numbers of replications compare to other methods, thus the quality of the estimate is improved. This does not necessarily imply that the larger b_p , the better, find b_p is a matter of trial and error which could apply a similar logic in bisection method; suggestions for setting this parameter are given in [6].

Table 2.5: **Importance Sampling**

K	Barrier	Ncrossed	Analytical	Corrected	MC	IS SE-3	IS SE-4	Crude SE-3	Crude SE-4	Bp
25	21	353	1.7026	1.7829	1.7732	0.0408		0.04		10
	21	632			1.7578	0.0241				30
	21	840			1.7724	0.0187				50
	21	909			1.76	0.0217				60
	21	890			1.7855	0.0188				55
	21	3519			1.7764		0.0129		0.0126	10
25	21	6127	1.7026	1.7829	1.7931		0.0076			30
	21	8427			1.7887		0.0058			50
	21	9140			1.7803		0.0066			60
	21	8798			1.7922		0.006			55
	21									

Chapter 3

Barrier Option Pricing under SABR Model

3.1 The Need for a Stochastic Volatility Model

The Black-Scholes formula [4] led to a boom in options trading and legitimised scientifically the activities of the Chicago Board Options Exchange and other options markets around the world [18]. It has been widely used by option market participants to estimate the value of options [5]. In this model, there is a one-to-one relation between the price of the option and the volatility parameter σ . And under the assumptions of the Black-Scholes-Merton world, volatility is constant. But in reality, options with different strikes require different volatilities to match their market prices. The plot of the implied volatility against the strike price is often referred to as market smiles and skews.

Handling these market skews and smiles correctly is critical for hedging. One would like to have a coherent estimate of volatility risk, across all the different strikes and maturities of the positions in the book.

To resolve this problem, in [12], the SABR model is derived. The model allows the market price and the market risks to be obtained immediately from Black's formula. It also provides good, and sometimes spectacular fits to the implied volatility curves observed in the marketplace. More importantly, the SABR model captures the correct dynamics of the smile, and thus yields stable hedges.

3.2 Building SABR Model Step by Step

3.2.1 Black Model and Implied Volatility

The Black formula [3] is similar to the Black–Scholes formula for valuing equity options except that the spot price of the underlying stock is replaced by a discounted futures price F_t . As the model supposes the underlying asset price follows a log-normal distribution and assumes the volatility σ is constant:

$$dF_t = \sigma F_t dW, \quad F(0) = f \quad (3.1)$$

Then we have a similar PDE:

$$\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 F^2 \frac{\partial^2 V}{\partial S^2} - rV = 0 \quad (3.2)$$

Note that the term $rS \frac{\partial V}{\partial F}$ which is present in the Black-Scholes PDE has been dropped here [16]. In Black’s world, the term is missing because the value of a futures contract is zero, while the value of the asset is positive. Futures contracts do not cost anything to enter into, hence the term does not appear in the derivation, unlike a share under Black-Scholes which does cost something to buy. Futures are not worthless though - once a futures contract is made, the value of that contract will change as the market moves.

Solving the equation above, gives a formula similar to the Black-Scholes one:

$$C(F, K, \sigma, T) = \Phi(d_1)f - \Phi(d_2)Ke^{-rT} \quad (3.3)$$

$$P(F, K, \sigma, T) = \Phi(-d_2)Ke^{-rT} - \Phi(-d_1)f \quad (3.4)$$

$$d_1 = \frac{\ln \frac{f}{K} + \frac{\sigma^2}{2}T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

All parameters above in the formula are observable except the volatility σ . However, in general, the value of an option depends on an estimate of the future realized price of the underlying. Or, mathematically:

$$C = f(\sigma, \cdot)$$

where C is the theoretical value of a option, and f is a pricing model like here the Black model that depends on σ along with other parameters.

The function f is monotonically increasing in σ , which means that a higher value of the volatility results in a higher theoretical value of the option. Conversely, by the inverse function theorem, there can be at most one value for σ that, when applied as an input to $f(\sigma, \cdot)$, will result in a particular value for C .

Putting in other terms, assume that there is some inverse function $g = f^{-1}$, such that:

$$\sigma_{\overline{C}} = g(\overline{C}, \cdot)$$

where \overline{C} is the market price for an option. The value $\sigma_{\overline{C}}$ is the volatility **implied** by the market price \overline{C} , or the **implied volatility**.

In general, it is not possible to give a closed form formula for the implied volatility in terms of call price. However, in some cases it is possible to give an asymptotic expansion of implied volatility in terms of call price.

Since $\frac{\partial C(F)}{\partial \sigma} = f\Phi(d_1)\sqrt{T} > 0$ in which $\Phi(d_1)$ is the density function of normal distribution. Then we know that call option values in the Black model are increasing function of the volatility σ and this volatility implied by the market price is unique for given parameters. But for different strike prices K , we will have different σ , which leads to the following three kinds of problems:

- 1 The first problem happens in pricing exotic options. For instance, one needs to price a down-and-out put option with strike price K_1 and barrier $S_b < K_1$. Should we use the implied volatility $\sigma(S_b)$ or $\sigma(K_2)$ or some combinations between them? Clearly, under volatility framework above, we cannot price this option for all possible strike prices without adjustments.
- 2 The second problem is hedging. Since implied volatility varies with different strike prices, it is not clear that delta and vega calculated at one strike price is consistent with results respect to other strike prices.
- 3 The third problem concerns the dynamics of the implied volatility curve with respect to the movement of strike prices. Since the implied volatility depends on K and f as $\sigma_B = \sigma_B(f, K)$. Some of the vega risks of Black's model would actually be due to changes in the price of underlying asset, and should be hedged more properly as delta risks.

3.2.2 Local Volatility Models

An apparent solution to the above problems is provided by the local volatility model. The concept of a local volatility was developed by [9] and [8] noted that there is a unique diffusion process consistent with the risk neutral densities derived from the market prices of European options of Dupire. Dupire assumed that forward price the underlying is given as:

$$dF_t = \sigma_{loc}(t, F_t)F_t dW, \quad F(0) = f \quad (3.5)$$

Dupire argued that instead of theorizing about the unknown local volatility function $\sigma_{loc}(t, F_t)$, one should obtain the function directly from the marketplace by "calibrating" the local volatility model to market prices of liquid European options. The method of the calibration is provided through [12].

Once the $\sigma_{loc}(t, F_t)$ has been obtained by calibration, the first problem of pricing exotic options is solved without ambiguity because local volatility model correctly reproduce the market prices for all strike prices and maturities. The second problem is solved because the model gives out consistent delta and vega risks for all options. Unfortunately, the third problem remains as the local volatility model predicts the wrong dynamics of implied volatility curve $\sigma_B(K, F)$, which causes inaccurate and unstable hedges.

Consider a special case that local volatility function is given as $\sigma_{loc}(F)$ to be a function of F only. It was found through singular perturbation methods in [6] and [7] that the implied volatility under Black's model is :

$$\sigma_B(K, f) = \sigma_{loc}\left(\frac{1}{2}[f + K]\right) \left(1 + \frac{1}{24} \frac{\sigma_{loc}''\left(\frac{1}{2}[f + K]\right)}{\sigma_{loc}\left(\frac{1}{2}[f + K]\right)} (f - K)^2 + \dots\right) \quad (3.6)$$

The last term in the big bracket is usually omitted because it is so small that less than 1% of the first term.

Suppose forward price today is f_0 and the implied volatility curve seen in the marketplace is $\sigma_B^0(K)$. As the local volatility is a function of F . In the equation above, $F = \frac{1}{2}[f_0 + K]$, hence:

$$\sigma_{loc}(F) = \sigma_B^0(2F - f_0)\{1 + \dots\} \quad (3.7)$$

If forward value changes from f_0 to f , given strike price K . $\sigma_B(K) = \sigma_B^0(2F - f_0)$, where $K = 2F - f$. We see that the new model predicts the volatility curve $\sigma_B(K, f)$ to be:

$$\sigma_B(K, f) = \sigma_B^0(K + f - f_0)\{1 + \dots\} \quad (3.8)$$

From the equation above, we know that in particular, if the forward price f_0 increases to f , the curve of $\sigma_B(K)$ moves to the left which indicates the underlying shifts to lower prices; if the forward price f_0 decreases to f , the curve of $\sigma_B(K)$ moves to the right which indicates the underlying shifts to higher prices. Local volatility models predict that the market smile or skew moves in the opposite direction as the price of the underlying asset. This is opposite to typical market behavior, in which smiles and skews move in the same direction as the underlying.

Moreover, hedges calculated from the local volatilities models are wrong. To see this, we assume the value of a call option is given by the Black's formula:

$$C = BS(f, K, \sigma_B(K, f), T)$$

with the volatility $\sigma_B(K, f)$ given by (6). Differentiating with respect to f yields the Δ risk:

$$\Delta = \frac{\partial C}{\partial f} = \frac{\partial BS}{\partial f} + \frac{\partial BS}{\partial \sigma_B} \frac{\sigma_B(K, f)}{\partial f} \quad (3.9)$$

The first term is clearly the Δ risk calculated from Black's model using the implied volatility from the market. The second term is the local volatility model's correction to the Δ risk, which consists of the Black Vega times the predicted change in σ_B due to changes in the underlying forward price f which is opposite to real dynamics of the market. Hence, the local volatility model causes unstable hedges.

3.2.3 The SABR Model

The failure of the local volatility model means that we cannot use a Markovian model based on a single Brownian motion to capture the correct dynamics of the implied volatility curve, hence SABR model differs from the Black-Scholes model and the local volatility model in terms of the dynamics of the underlying asset. SABR presumes that the volatility evolves with the time t , strike price K and the current forward price f_t whereas in the precedent $c(\cdot)$ function is set to be constant in Black-Scholes and $c(t, St)$ respectively in the local volatility model. Therefore, we see that Hagan et al added a randomness from volatility itself by inserting a specific stochastic process α_t in the dynamics of the underlying thereby a new parameter ν incorporated in the dynamics of α_t . Then, the two Brownian motions that described the dynamics of f_t and α_t are connected by a correlation parameter ρ . Accordingly, the SABR model is structured with four parameters α , β , ρ , ν . Therefore, its named as **S**tochastic **A**lpha **B**eta **R**ho model. It is formulated as:

$$dF_t = \alpha_t F_t^\beta dW_1, \quad F(0) = f \quad (3.10)$$

$$d\alpha_t = \nu \alpha_t dW_2, \quad \alpha(0) = \alpha \quad (3.11)$$

under the forward measure, where the two processes are correlated by:

$$dW_1 dW_2 = \rho dt \quad (3.12)$$

Followed by the single perturbation methods applied in [12], the price of an European option is obtained in closed form (up to the accuracy of a series of expansion). Whereafter the option's implied volatility $\sigma_B(f, K)$ is calculated under

Black's model framework:

$$\sigma_B(f, K) = \frac{\alpha}{(fK)^{\frac{1-\beta}{2}} \{1 + \frac{(1-\beta)^2}{24} \log^2 \frac{f}{K} + \frac{(1-\beta)^4}{1920} \log^4 \frac{f}{K} + \dots\}} \cdot \left(\frac{z}{x(z)} \right).$$

$$\{1 + [\frac{(1-\beta)^4}{24} \frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4} \frac{\rho\beta\nu\alpha}{(fK)^{\frac{1-\beta}{2}}} + \frac{2-3\rho^2}{24} \nu^2]T + \dots\} \quad (3.13)$$

Here

$$z = \frac{\nu}{\alpha} (fK)^{\frac{1-\beta}{2}} \log \frac{f}{K}, \quad (3.14)$$

and $x(z)$ is defined by:

$$x(z) = \log \left\{ \frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right\} \quad (3.15)$$

Two special cases are marked out as $\beta = 1$ representing a stochastic log normal model and $\beta = 0$ representing a stochastic normal model.

Also, to at-the-money options that $K = f$, the formula simplifies to:

$$\sigma_{ATM} = \sigma_B(f, f) = \frac{\alpha}{f(1-\beta)} \{1 + [\frac{(1-\beta)^4}{24} \frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4} \frac{\rho\beta\nu\alpha}{(fK)^{\frac{1-\beta}{2}}} + \frac{2-3\rho^2}{24} \nu^2]T + \dots\} \quad (3.16)$$

Note that in these formulas the omitted terms ”+...” seem to be small, it could be large enough to be required for accurate option price. However, in subsequent analysis, I choose not to implement with these terms due to the complexity concerns.

3.3 Estimating the SABR Parameters

We are going to fitting the SABR parameters to the recorded market data in [2] including maturities, strikes and implied volatilities under the Black-Scholes Model.

Table 3.1: Data

Maturities	Strikes	Volatilities
0.078159208	17.5	0.231583868
0.078159208	20	0.194447616
0.078159208	22.5	0.180948553
0.078159208	25	0.20549825
0.078159208	27.5	0.2286896
0.078159208	30	0.254493656
0.078159208	32.5	0.283817284
0.078159208	35	0.303435636
0.58371	12	0.35215
0.58371	15	0.26899
0.58371	17	0.24382
0.58371	19.5	0.20943
0.58371	20	0.20439
0.58371	22	0.19634
0.58371	22.5	0.19318
0.58371	24.5	0.20162
0.58371	25	0.20537
0.58371	27	0.22337
0.58371	27.5	0.22825
0.58371	29.5	0.23799
0.58371	30	0.25288
0.58371	32	0.27194
0.58371	32.5	0.27655
0.58371	34.5	0.29436
0.58371	37	0.31516
1.59483	15	0.31353
1.59483	17.5	0.26695
1.59483	20	0.23042
1.59483	22.5	0.22082
1.59483	25	0.21535
1.59483	27.5	0.23587
1.59483	30	0.2509
1.59483	35	0.30947
1.59483	40	0.35199

In general, the principle is minimizing the errors between SABR implied volatilities $\sigma_B(\alpha, \beta, \rho, \nu)$ and market observed volatilities σ_{MKT} for 3 different maturities (implied volatility surface). Obviously, the calibration for a fix maturity got substantially less error than for all maturities. If we interpolate with time's evolution into this setting, the result of fitting could be better. However, fitting under each single expiration can bring a efficiently accurate parameter scheme for simulation done by Monte Carlo methods. In order to price the barrier option of short expiration on the OTC market, the SABR model here is calibrated to a set of option prices (volatilities) for a single given expiration.

In general, the calibration can be accomplished by minimizing the square errors between SABR implied volatilities $\sigma_B(\alpha, \beta, \rho, \nu)$ and market observed volatilities σ_{MKT} in identical maturity $T=0.078159208$ years. Thus, this produces an optimization problem:

$$(\hat{\alpha}, \hat{\beta}, \hat{\rho}, \hat{\nu}) = \arg \min_{\alpha, \beta, \rho, \nu} \sum_i (\sigma_i^{MKT} - \sigma_B(f_i, K_i; \alpha, \beta, \rho, \nu))^2 \quad (3.17)$$

Market smiles can be fitted equally well with any specific value of β . In particular, β cannot be determined by fitting a market smile since this would clearly amount to “fitting the noise” because the exponent β and ρ affect the volatility smile in similar ways — they both cause a downward sloping skew in $\sigma_B(f, K)$ as the strike K varies. This is demonstrated by figure 3. There we fit the SABR parameters α, ρ, ν with $\beta = 0$ and then re-fit the parameters α, ρ, ν with $\beta = 1$.

The calibration begins with choosing β either by empirical analysis of the asset price and the σ_{ATM} or by setting $\beta = 0$ for a normal process or $\beta = 1$ for a lognormal process. Next, with β chosen, you have two approaches in calibrating

the other 3 parameters.

3.3.1 Choice of β

In the equation of $\sigma_B(f, K)$. when setting $f = K$ the σ_{ATM} equals:

$$\sigma_{ATM} = \sigma_B(f, f) = \frac{\alpha \{1 + [\frac{(1-\beta)^2}{24} \frac{\alpha^2}{f^{2-2\beta}} + \frac{1}{4} \frac{\rho\beta\nu\alpha}{f^{1-\beta}} + \frac{2-3\rho^2}{24} \nu^2] T\}}{f^{1-\beta}}$$

Taking log produces:

$$\ln \sigma_{ATM} \approx \ln \alpha - (1 - \beta) \ln f \quad (3.18)$$

We usually get σ_{ATM} from the market, however, we don't get this volatility from given fictional market data.

In order to solve a reasonable σ_{ATM} , we incorporate **Cubic** and **Spline** model [1] here. This model treats the implied volatility as a cubic function of moneyness X and a quadratic function of the time to expiration τ . The Cubic model is described by the following equation:

$$\sigma_{impl} = a_0 + a_1X + a_2X^2 + a_3X^3 + a_4\tau + a_5\tau^2 \quad (3.19)$$

The a_i parameters are fitted The parameters using the linear least squares regression analysis. Adding a cubic spline and linear interpolation functions to Excel, we can solve the $\sigma_{ATM} = 0.180$ from the implied volatilities and strikes. The strike $K=22.5$ very near to $F=22.27$ has the implied volatility=0.1809, the

data interpolation works as expected.

Hence, β can be estimated by a linear regression on a time series of logs of ATM volatilities and logs of forward rates.

Therefore, in our computation for $T=0.0781$, $\beta=0.399$.

In practice, the β has little effect on the resulting shape volatility curve established by SABR parameters. Alternatively, β can be chosen from priori beliefs for appropriate models [20].

- $\beta = 1$: stochastic log-normal, for FX option markets
- $\beta = 0$: stochastic normal, for markets with zero or negative f
- $\beta = \frac{1}{2}$: CIR model, for interest rate markets

However, the choice of β can affect Greeks, we usually want to find a β stable in a specific market. With β estimated, there are two possible choices to continue calibration:

- Estimate α , ρ and ν directly, or
- Estimate ρ and ν directly, and infer α from ρ , ν from σ_{ATM} .

3.3.2 Estimate α , ρ and ν directly

In this way, the calibration reduce to estimation of all three parameters α , ρ and ν left by minimization of the sum of squared errors(SSE) between the model and the market volatilities:

$$(\hat{\alpha}, \hat{\rho}, \hat{\nu}) = \arg \min_{\alpha, \rho, \nu} \sum_i (\sigma_i^{MKT} - \sigma_B(f_i, K_i; \alpha, \rho, \nu))^2 \quad (3.20)$$

The calibration result of direct α at maturity $T=0.078159208$ is:

Table 3.2: Parameters

Direct α Approach				
alpha	beta	rho	nu	SSE
3.9984	0	0.2420	1.3062	5.00E-05
1.8458	0.25	0.1948	1.2729	3.48E-05
1.1649	0.399	0.1659	1.2543	3.01E-05
0.8523	0.5	0.1455	1.2430	2.87E-05
0.3936	0.75	0.0944	1.2164	3.06E-05
0.1818	1	0.0417	1.1933	3.97E-05

Applying the above values of parameters for $\beta = 0.399$, we can plot the volatility dynamics to three different maturities in Figure 1.

3.3.3 α from σ_{ATM}

Rearrange the terms of σ_{ATM} , we find that α is a root of the cubic:

$$\sigma_{ATM} f^{1-\beta} = \left[\frac{(1-\beta)^2 T}{24 f^{2-2\beta}} \right] \alpha^3 + \left[\frac{\rho \beta \nu T}{4 f^{1-\beta}} \right] \alpha^2 + \left[1 + \frac{2-3\rho^2}{24} \nu^2 T \right] \alpha$$

West [22] noted that it is possible to receive more than a single real root, and suggested to select the smallest positive root in this case.

Hence, the problem transforms into:

$$(\hat{\alpha}, \hat{\rho}, \hat{\nu}) = \arg \min_{\alpha, \rho, \nu} \sum_i (\sigma_i^{MKT} - \sigma_B(f_i, K_i; \alpha(\rho, \nu), \rho, \nu))^2 \quad (3.21)$$

The calibration result for this approach at maturity $T=0.078159208$ is in table:

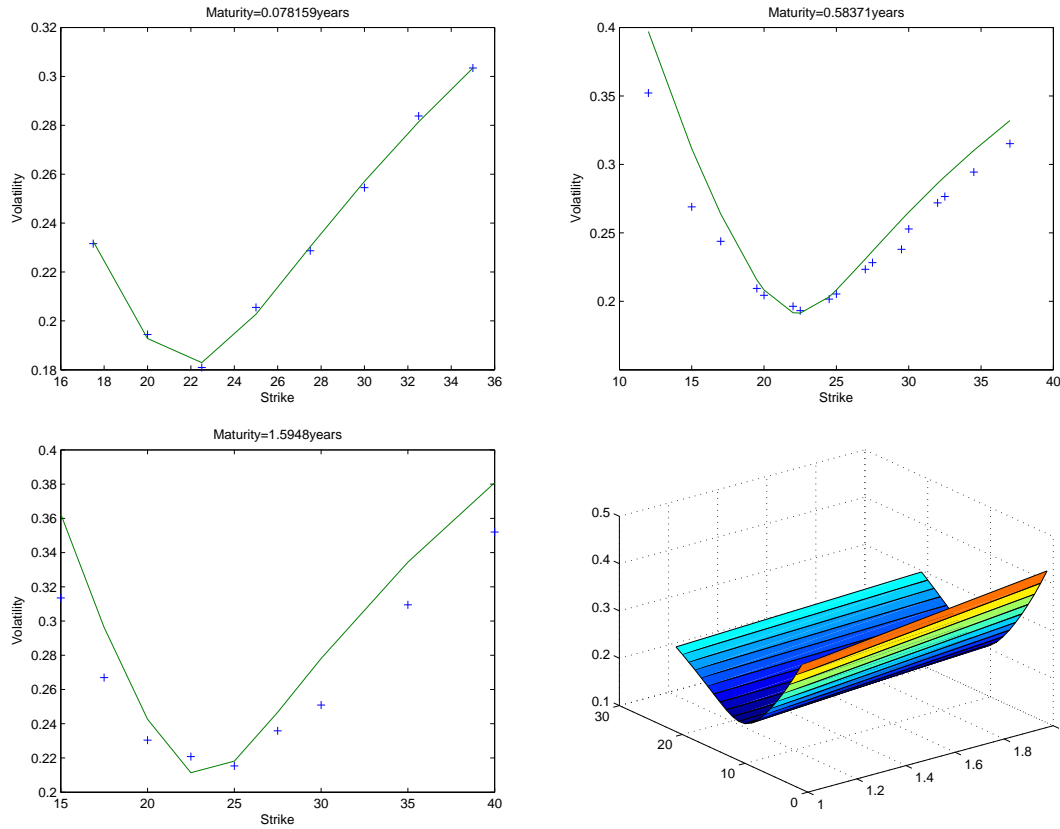


Figure 3.1: $\beta = 0.399$ $\alpha = 1.1649$ $\rho = 0.1659$ $\nu = 1.2543$

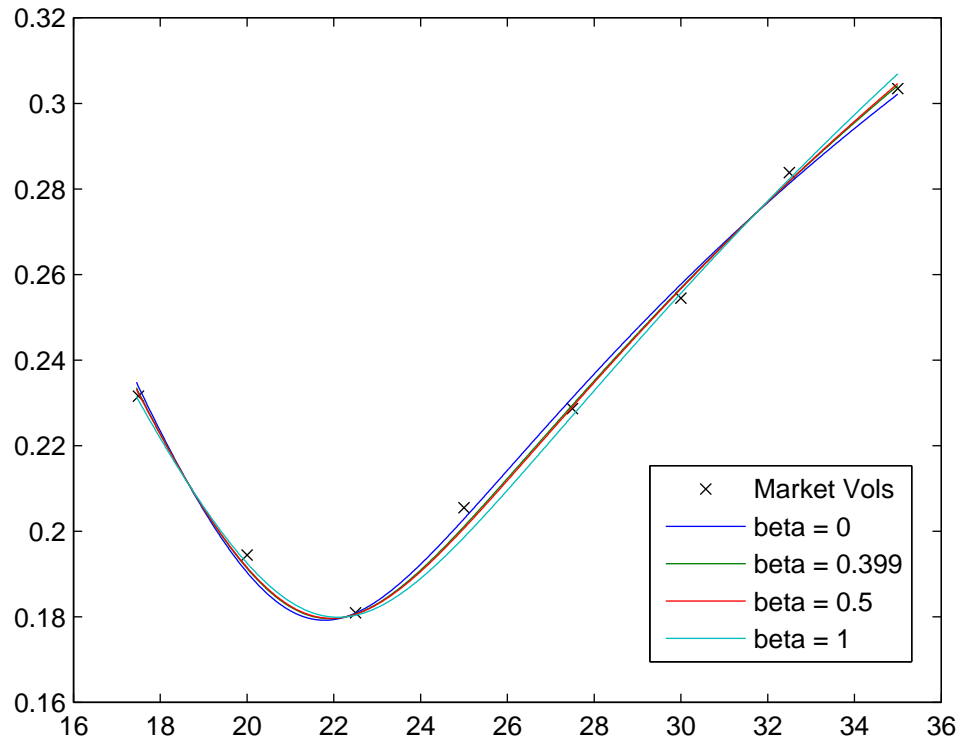
3.3.4 Algorithm for Calibration

Many nonlinear optimization method can be used to carry out the calibration, in our case, we used Levenberg-Marquardt Algorithm [17], [10]. And we will show the SABR model is calibrated to fit a implied volatility smile observed in the marketplace for an option with a fixed maturity time frame, hence the dependence of σ_B on t is not reflected in our analysis.

First, we will discuss how we estimate the parameters. Further, we will investigate how the parameters affect the shape of the curve of implied volatility. Finally, we will do some fix to our methods.

Table 3.3: Parameters

α from σ_{ATM}				
alpha	beta	rho	nu	SSE
3.9668	0	0.2391	1.3227	5.27E-05
1.8259	0.25	0.1902	1.2948	4.20E-05
1.1499	0.399	0.1603	1.2799	4.08E-05
0.8406	0.5	0.1397	1.2705	4.20E-05
0.3870	0.75	0.0879	1.2498	5.20E-05
0.1782	1	0.0350	1.2327	7.11E-05

Figure 3.2: α from σ_{ATM} approach

The primary application of the Levenberg–Marquardt algorithm is in the least squares curve fitting problem: given a set of m empirical datum pairs of independent and dependent variables, (x_i, y_i) , optimize the parameters λ of the model curve $f(x, \lambda)$ so that the sum of the squares of the deviations.

$$S(\boldsymbol{\lambda}) = \sum_{i=1}^m [y_i - f(x_i, \boldsymbol{\lambda})]^2 \quad (3.22)$$

In each iteration step, the parameter vector, λ , is replaced by a new estimate, $\lambda + \delta$. To determine δ , the functions $f(x_i, \boldsymbol{\lambda} + \boldsymbol{\delta})$ are approximated by their linearizations

$$f(x_i, \boldsymbol{\lambda} + \boldsymbol{\delta}) \approx f(x_i, \boldsymbol{\lambda}) + J_i \boldsymbol{\delta} \quad (3.23)$$

where

$$J_i = \frac{\partial f(x_i, \boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \quad (3.24)$$

is the gradient (row-vector in this case) of f with respect to λ .

At the minimum of the sum of squares, $S(\lambda)$, the gradient of S with respect to δ will be zero. The above first-order approximation of $f(x_i, \boldsymbol{\lambda} + \boldsymbol{\delta})$ gives

$$S(\boldsymbol{\lambda} + \boldsymbol{\delta}) \approx \sum_{i=1}^m (y_i - f(x_i, \boldsymbol{\lambda}) - J_i \boldsymbol{\delta})^2. \quad (3.25)$$

Or in vector notation,

$$S(\boldsymbol{\lambda} + \boldsymbol{\delta}) \approx \|\mathbf{y} - \mathbf{f}(\boldsymbol{\lambda}) - \mathbf{J}\boldsymbol{\delta}\|^2. \quad (3.26)$$

Taking the derivative with respect to δ and setting the result to zero gives:

$$(\mathbf{J}^T \mathbf{J}) \boldsymbol{\delta} = \mathbf{J}^T [\mathbf{y} - \mathbf{f}(\boldsymbol{\lambda})] \quad (3.27)$$

where \mathbf{J} is the Jacobian matrix whose i th row equals J_i , and where \mathbf{f} and \mathbf{y} are vectors with i th component $f(x_i, \boldsymbol{\lambda})$ and y_i , respectively. This is a set of linear equations which can be solved for δ .

The upshot here is that in cases with multiple minima that the algorithm converges only if the initial guess is already somewhat close to the final solution. As we want to do a least square regression involving all 4 parameters, its convenient to set a good initial guess such that the times of iteration will be ideally few.

3.4 Dynamics of Parameters

In this subsection, we will examine the effects of different parameters on the shape of implied smile in SABR framework. The calibration of the data lends us a based plot of for researching the evolution of smiles with respect to the dynamics of the related parameters.

$$\alpha = 1.1499, \quad \beta = 0.399, \quad \rho = 0.1603, \quad \nu = 1.2799$$

3.4.1 α

α can be treated as a ‘volatility’ in the process of forward price evolution. We start from the initial ‘volatility’ $\alpha=1.1499$.

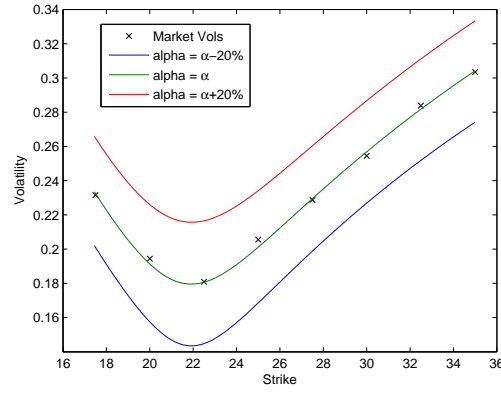
$$\sigma_{ATM} = \frac{\alpha}{f^{1-\beta}}(1 + \dots) \approx \frac{\alpha}{f^{1-\beta}} \quad (3.28)$$

$$Backbone : \frac{\alpha}{f^{1-\beta}} \quad (3.29)$$

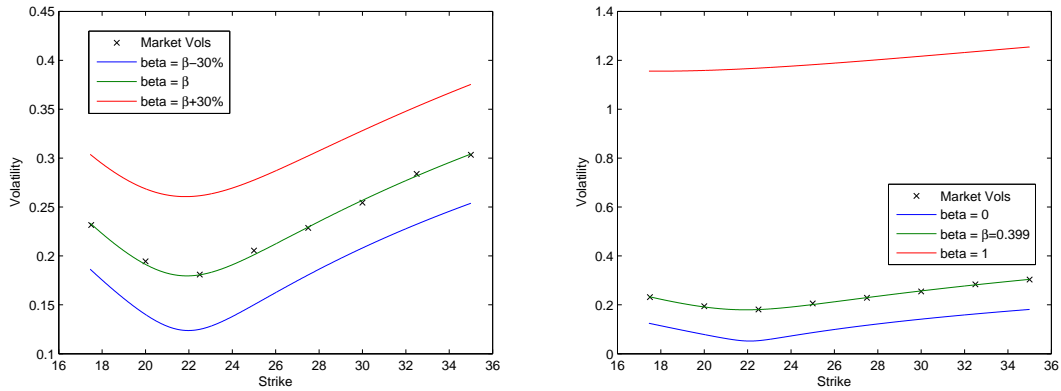
While we shifting the value of α upwards or downwards by 20%, the shape of the smile doesn’t incur any notably change. However, there is a obvious difference between each smile of the vertical location. The change of σ_{ATM} reflects the vertical change as well, as the σ_{ATM} is approximately a multiple of α . Furthermore, the backbone of the smiles is formulated as $\frac{\alpha}{f^{1-\beta}}$ quantify the move of smiles with respect to parameter of α and β .

3.4.2 β

Shifting β by 30%, a significant change emerged in the vertical difference, which is in accordance with its contribution to the backbone with α . β also has little

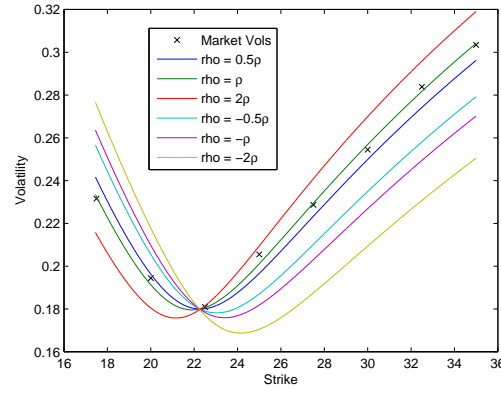
Figure 3.3: α dynamics

effect on the slope of the smile: comparing $\beta = 0, \beta = 0.399, \beta = 1$, we can see the curve steepens when β decreases to 0.

Figure 3.4: β dynamics

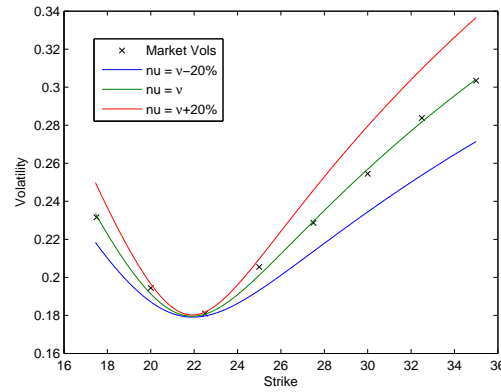
3.4.3 ρ

Changing ρ pivots the curve on the point $K = F$ and $\sigma = \sigma_{ATM}$. In our case, the increase of ρ makes the right part of smile steeper and the left part of smile flatter.

Figure 3.5: ρ dynamics

3.4.4 ν

Figure reveals a simple mechanism that a larger parameter ν will raise the implied volatility and or a smaller one will lower the curve except the at-the-money point. Specially, the further-away points from the at-time-money strike demonstrate a much more relationship with ν than the points around $K=F$.

Figure 3.6: ν dynamics

3.4.5 f

We can see from figure that the change of forward price would not affect the shape of the curve. The smile shifts downwards as the forward price increases. Because of

this downwards backbone, the f dynamics here also proves that market smile/skew in the SABR model moves in the same direction with the forward price.

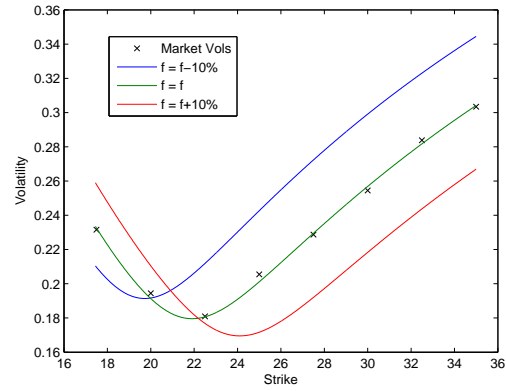


Figure 3.7: f dynamics

3.5 Monte Carlo Methods under SABR

When pricing derivatives, particularly exotic derivatives, analytic solutions are often difficult to obtain and numerical methods are normally employed. Monte Carlo, tree, finite difference and finite element methods are standard numerical methods used for pricing derivatives.

Monte Carlo simulation is an important tool in computational finance: It may be used to price options, to estimate value at risk, to simulate hedging strategies. Its main advantages are generality, relative ease of use and flexibility. As I incorporated SABR stochastic volatility framework, with the features of barrier options into account, and lends itself to treating other problems where the lattice or PDE methods may not be applied. The potential disadvantage of Monte Carlo is its computational burden. An increasing number of simulations is needed to refined the result that we are interested in. The problem may be partially solved by several variance reduction techniques and importance sampling. We will give examples of down out put barrier options and compare it with the previous Monte Carlo pricing of them derived under the Black-Scholes framework.

We will apply Monte Carlo method to pricing of continuously monitor barrier options under the SABR model. There is some difference between path generation method from the Black-Scholes model.

First, the volatility now is not set constant in the path generation as we suppose a stochastic volatility incurred by SABR model. We have done calibrations to SABR model parameters α , β , ν and ρ . We can derive the evolution of underlying stock by inserting these parameters into the path we want.

Second, compared to the Black-Scholes underlying process, we have one more two Brownian motions included in the processes of F_t and α_t is correlated with ρ . In order to solve this problem, we pick two independent samples from standard normal distribution and form a new process with correlation ρ to the another.

3.5.1 Path Generation of Barrier Options

To simulate the path of underlying equity stock S , the fixed time to maturity can be divided into n short intervals Δt and the r is constant, the evolution of S can be approximated by:

$$S_{k+1} = S_k + rS_k\Delta t + \alpha_k e^{r(\beta-1)T} S_k^\beta \sqrt{\Delta t} \epsilon_1(k\Delta t), S_0 = S_0$$

$$\alpha_{k+1} = \alpha_k + \nu \alpha_k (\rho \epsilon_1(k\Delta t) + \sqrt{1-\rho^2} \epsilon_2(k\Delta t)) \sqrt{\Delta t}, \alpha_0 = \alpha$$

where S_k and α_k are the values of S and α at time $k\Delta t$, the total number of time steps is n . $\epsilon_1(\Delta t)$ and $\epsilon_2(\Delta t)$ are two samples from a standard normal distribution, both independent of each other and independent with respect to values of k . The new process $\rho \epsilon_1(k\Delta t) + \sqrt{1-\rho^2} \epsilon_2(k\Delta t)$ has correlation ρ with $\epsilon_1(k\Delta t)$ and its also a sample of standard normal distribution.

To get an approximation of the price of a European style option under the SABR model with the Monte Carlo method, we should follow the same steps as that of the Black-Scholes model. The only difference is that there should be sample $2n$ instead of n values of a standard normal distribution to simulate the payoff.

3.5.2 Result Discussion

We have been using the parameter value from in SABR calibration section. The number of simulation is set to be 1000 and 10000. The payoff is set to 0 whenever the barrier is crossed. Here note that we simulate the complete path even if the barrier is crossed during the life of option.

The ‘SABR’ prices are lower than continuous corrected(for discrete monitoring) Black-Scholes Analytical solution. Because the volatilities we incorporated for simulation at this time is larger than $\sigma_{BS} = 0.18$ for simulation under the Black-Scholes Model. Considering a down-and-out put option, the probability to cross the barrier raises as a result of the added stochastic volatility framework in the implied smile.

As usual, we simulation times increasing, the standard error is substantially

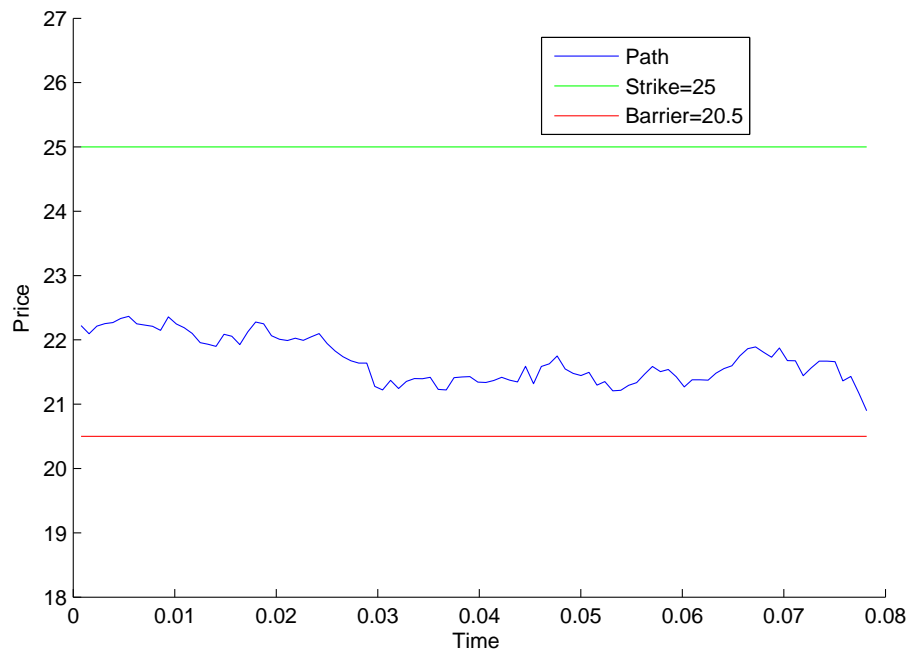


Figure 3.8: One Simulation Path Example

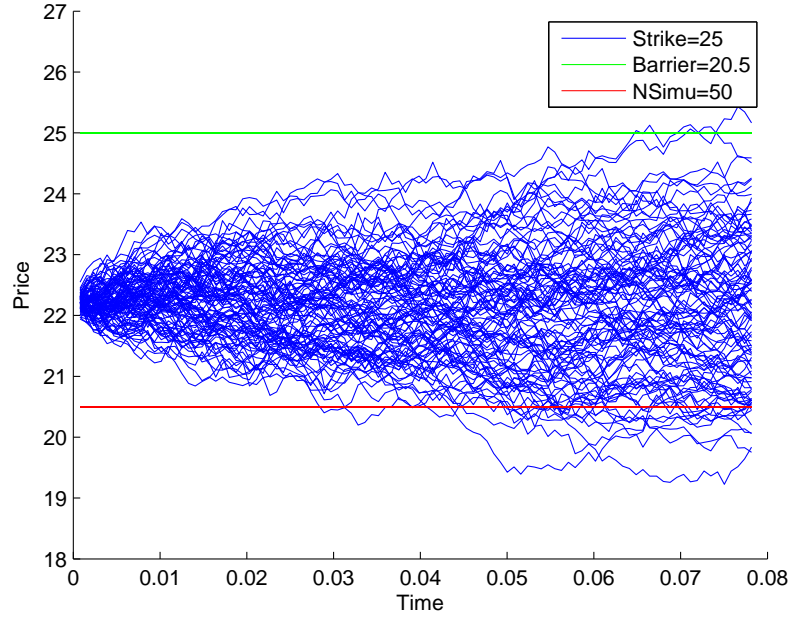


Figure 3.9: 50 Simulation Paths Example

reduced. When $K = 25$ and $S_b = 20.5$, with $S_0 = 22.2$, there are appropriate number of paths crossed the barrier, the variance seems to be the optimal lowest in the calculation of this down-and-out barrier price. We can expect a large enough probability to construct the indicator function part in our product of indicator function with final profit. In such situations, the barrier option gain some reasonable value to trade on the market and provide its functions as a special type of path-dependent exotic options.

For $K = 20$ and $S_b = 19$ or 19.5 , with $S_0 = 22.2$, the value of such down-and-out options become nearly 0 because of the very low probability to cross the barrier. Moreover, these options is similar to vanilla put options, which have a price converges to 0 when the barrier is enough far away from the spot price as well.

Furthermore, we may research the use of stratified sampling and importance sam-

pling into the variance reduction in barrier option pricing case by case.

Table 3.4: **SABR Crude MC**

K	Barrier	Ncrossed	BS	Corrected	SABR MC	Crude SE-3	Crude SE-4
22.5	22	1000		0.0043	na	na	
	21.5	904		0.0637	0.026	0.0036	
	21	466		0.2117	0.1739	0.0096	
	20.5	206		0.3794	0.3445	0.014	
	20	89		0.4949	0.3847	0.0152	
25	22	1000		0.3131	na	na	
	21.5	877		1.0593	0.3414	0.0291	
	21	497		1.7829	1.3794	0.0449	
	20.5	213		2.2975	2.1605	0.0401	
	20	70		2.5730	2.5426	0.0316	
20	19.5	22		0.0019	0.000051657	0.000040152	
	19	9		0.0047	0.00054057	0.00033036	
K	Barrier	Ncrossed	BS	Corrected	SABR MC	Std. Error-3	Std. Error-4
22.5	22	10000	0.0023		na		na
	21.5	8758	0.0509		0.0371		0.0013
	21	4809	0.1904		0.193		0.0032
	20.5	2060	0.3612		0.3352		0.0043
	20	793	0.4848		0.4099		0.0049
25	22	10000	0.2290		na		na
	21.5	8779	0.9610		0.335		0.0091
	21	4665	1.7026		1.4624		0.0142
	20.5	2110	2.2490		2.1632		0.0126
	20	776	2.5511		2.5387		0.0102
20	19.5	300	0.0015		0.00011346		0.000050439
	19	88	0.0044		0.00039651		0.00012048

Appendices

Appendix A

Mathematical Appendix

In this section, a brief summary of several concepts and theorems is given. An understanding of these will provide a foundation to construct the financial models employed in this report. Certain additional conditions applied for the completeness of theorems, such as existence and uniqueness, will be taken as understood without proof. The first two sections A.1 - A.2 give very basic definitions in classic probability theory. Section A.3 introduces several stochastic processes which are of great importance in the field of financial mathematics. Section A.4 presents the fundamental convergence theory and section A.5 gives a more specific introduction to stochastic calculus embedded in option pricing theory. A more advanced introduction to this area may be found in any textbook related.

A.1 Probability Space

Suppose that Ω is a set. Then a collection of subsets of Ω , \mathcal{F} , is called a σ - algebra (σ -field) if:

1. $\emptyset \in \mathcal{F}$;
2. if $A \in \mathcal{F}$, then so is the complement of A (i.e. $A^c \in \mathcal{F}$)

3. if A_i for $i=1,2,\dots$ is a family of subsets such that $A_i \in \mathcal{F}$, then

$$A = \bigcup_{i=1}^n A_i \in \mathcal{F} \quad (\text{A.1})$$

A probability measure \mathbb{P} is a real-valued function defined as:

1. $0 \leq \mathbb{P}(A) \leq 1, \forall A \in \mathcal{F}$;
2. $\mathbb{P}(\Omega)=1$, where Ω is a sample space;
3. if A_i for $i=1,2,\dots$ is a family of subsets such that $A_i \in \mathcal{F}$, and $A_i \cap A_j = \emptyset$ for any $i \neq j$, then

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i). \quad (\text{A.2})$$

Then a probability space is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ such that

1. Ω is non-empty set (called sample space);
2. \mathcal{F} is a family of subsets of Ω with the property of a σ -algebra (a set of “events”);
3. \mathbb{P} is a probability measure such that $\mathbb{P} : \mathcal{F} \rightarrow \mathbb{R}$

Measurability

A random variable X is a \mathcal{F} -measurable if the value of X is completely determined by the information in \mathcal{F} . Formally speaking:

A random variable $X : \Omega \rightarrow \mathbb{R}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called \mathcal{F} -measurable if

$$X^{-1}(U) = \{\omega \in \Omega : X(\omega) \in U\} \in \mathcal{F},$$

for all open sets $U \in \mathbb{R}$

A.2 Conditional Expectation

The conditional expectation of X given σ -algebra $\mathcal{G} \subset \mathcal{F}$ is a random variable $\mathbb{E}[X \mid \mathcal{G}] : \Omega \rightarrow \mathbb{R}$ satisfying:

- i:) $\mathbb{E}[X \mid \mathcal{G}]$ is \mathcal{G} -measurable;
- ii:) $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mathbb{1}_G] = \mathbb{E}[X \mathbb{1}_G]$, where $\mathbb{1}_{(\cdot)}$ is an indicator function.

Conditional expectation is the essence of option price modeling, especially for the options with early exercise features, Option prices are expectations conditioned on the information given at the present time.

A.3 Stochastic Processes

A stochastic process is a family of random variables $X_t(\omega), t \in T$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with a set T which is called the index set of the process. Given any $t \in T$ fixed, the possible values of X_t are called the states of the process at t . Whereas, given $\omega \in \mathcal{F}$ fixed, $X(\omega)$ is called its sample path of the stochastic process, and the family of all sample paths is a path space. This path space is the probability space.

If T is discrete, then the stochastic process is referred to as a discrete-time process, and it is sometimes called a “sequence”. If T is an interval of \mathbb{R} , then the stochastic process is a continuous-time process. Note that continuous-time stochastic processes are more general than discrete-time stochastic processes. Therefore, in theoretical finance, continuous-time stochastic processes are widely-used, and these processes are of practical importance. For instance, PDEs or SDEs may be built up on a continuous-time platform. The most well-known numerical method approach which is applied on a discrete-time platform is binomial tree method. Note that properties presented in this chapter are for continuous-time processes by default,

which are then applicable to discrete-time cases, in the limit of small time steps. The stochastic processes are basic building blocks for financial models. Below, two fundamental processes are considered: Brownian Motions and martingales.

A.3.1 Brownian Motions

A stochastic process $(W_t)_{t \geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a Brownian motion (i.e. a Wiener process) if:

- i:) the random variables $(W_{t_i} - W_{t_{i-1}}), i = 1, 2, \dots, n$ are independent for any given $0 \leq t_0 < t_1 < \dots < t_n$ (independent increment);
- ii:) $W_t - W_s \sim W_{t-s}$ for any $0 \leq s \leq t$ (stationary increment);
- iii:) W_t is continuous in t with \mathbb{P} -a.s.;
- iv:) $W_0=0$, with \mathbb{P} -a.s.

With the above four conditions satisfied, a useful result can be obtained:

$$W_t \sim N(\mu t, \sigma^2 t), \forall t > 0 \quad (\text{A.3})$$

In a discrete-time platform, a Brownian motion is also known as a simple or scaled symmetric random walk, which is used in binomial tree models etc.

A.3.2 Martingales

Martingales are very important and useful in the study of stochastic processes.

A martingale is a stochastic process whose future movements are always unpredictable –it is a model of a fair game. A formal definition is given below.

A process $(M_t)_{t \geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a martingale with respect to a filtration $\{\mathcal{F}_t\}_{t \geq 0}$, with $\mathcal{F}_t \subset \mathcal{F}$ if the following conditions are satisfied:

- i:) M_t is adapted to $\{\mathcal{F}_t\}_{t \geq 0}$, (i.e. M_t is \mathcal{F}_t -measurable for all t);
- ii:) $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$
- iii:) $\mathbb{E}[M_t | \mathcal{F}_s] = M_s$ for all $0 \leq s \leq t$, with \mathbb{P} -a.s.

Brownian motion is the most trivial example of both a martingale and a Markov process, and is one of the reasons that Brownian motion performs a key role in mathematical finance.

A.4 Convergence and Central Limit Theorem

A.4.1 Convergence

- **Convergence almost surely \mathbb{P} -a.s**

Suppose that X and $\{X_n, n = 1, 2, \dots\}$ are real-valued random variables.

Then X_n converges to X \mathbb{P} -a.s. if

$$\mathbb{P}\left(\lim_{n \rightarrow +\infty} X_n = X\right) = 1.$$

- **Convergence in probability \mathbb{P}**

X_n converges to X in probability for every $\epsilon > 0$ if

$$\mathbb{P}(|X_n - X| > \epsilon) \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (\text{A.4})$$

- **Convergence in distribution**

Suppose X and $\{X_n, n = 1, 2, \dots\}$ are real-valued random variables with distribution F and $F_n, n=1,2,\dots$ respectively. Then X_n converges to X in distribution if

$$F_n(x) \rightarrow F(x) \text{ as } n \rightarrow \infty \quad (\text{A.5})$$

for all $x \in \mathbb{R}$ at which F is continuous.

A.4.2 The Law of Large Numbers

Suppose X_n , $n=1,2,\dots$ is a sequence of independent and identically distributed random variables with mean μ . Then

$$\frac{1}{n} \left(\sum_{i=1}^n X_i \right) \rightarrow \mu \text{ as } n \rightarrow \infty, \quad \mathbb{P} - a.s. \quad (\text{A.6})$$

This is called strong law of large numbers. It provides the theoretical basis for stochastic simulations, such as the Monte Carlo method.

A.4.3 Central Limit Theorem

Suppose X_n , $n=1,2,\dots$ is a sequence of independent and identically distributed random variables with mean μ and variance σ^2 . Then

$$\frac{\sum_{i=1}^n (X_i - \mu)}{\sigma\sqrt{n}} \rightarrow N(0,1) \text{ as } n \rightarrow \infty, \quad \mathbb{P} - a.s. \quad (\text{A.7})$$

Central limit theorem implies that no matter what distribution X_i has, the sum of X_i (properly normalised) has a normal distribution when n is large enough.

A.5 Change of Measure

The idea of changing probability measure is of central importance in derivative pricing theory. As mentioned before, a derivative is contingent on one or several underlying assets whose uncertainties do not affect the price of the derivative. Therefore, changing the probability measure have many other applications, for instance, “importance sampling” in the Monte Carlo method.

A.5.1 Radon-Nikodym Derivative

Probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ are called equivalent if

$$\mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0, \forall A \in \mathcal{F}. \quad (\text{A.8})$$

This is often written as $\mathbb{P} \sim \mathbb{Q}$.

Suppose $\mathbb{P} \sim \mathbb{Q}$ on space (Ω, \mathcal{F}) . The random variable R defined on (Ω, \mathcal{F}) is called the Radon-Nikodym derivative of \mathbb{P} with respect to \mathbb{Q} if

i:) R is strictly positive;

ii:) R is unique with \mathbb{P} -a.s.

iii:) $\mathbb{Q}(A) = \mathbb{E}_{\mathbb{P}}[R\mathbb{1}_A], \forall A \in \mathcal{F}$

It is customary to write $R = \frac{d\mathbb{P}}{d\mathbb{Q}}|_{\mathcal{F}}$, which is defined as the Radon-Nikodym derivative of \mathbb{P} with respect to \mathbb{Q} ; it is also called a numeraire in the financial world.

This generalizes the concept of numeraire, as the numeraires are normally risk-free assets, however, when the model setup becomes more sophisticated, the numeraire can also be a stochastic process, which is sometimes called the stochastic discount factor.

A.5.2 Girsanov's Theorem

Girsanov's theorem establishes a link between two probability measures. Assume $(W_t)_{t \geq 0}$ is a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to filtration $\mathcal{F}_t, \mathcal{F}_t \subset \mathcal{F}$. Then define

$$d\widetilde{W} = \theta dt + dW_t \quad (\text{A.9})$$

and

$$R_t = \exp\left(-\frac{1}{2} \int_0^t \theta_u^2 du - \int_0^t \theta_u dW_u\right), \quad (\text{A.10})$$

where $E[R_t] = 1$ and $\theta = \theta(t)$ is adapted to filtration $\{\mathcal{F}_t\}$. Assume that a new probability measure is defined by $\mathbb{Q}(F) = \int_A R_t d\mathbb{P}$ for all $A \subset \mathcal{F}$, then under \mathbb{Q} , the process $(\widetilde{W}_t)_{t \geq 0}$ is a Brownian motion. Basically, this theorem implies that a Brownian motion process with any drift can be converted to another Brownian motion process with the same variance but with different drift.

A.5.3 Equivalent Martingale Measure

Assume M is a continuous martingale defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to filtration $\{\mathcal{F}_t\}, \mathcal{F}_t \subset \mathcal{F}$. The set of equivalent martingale measures for M is the set of probability measure \mathbb{Q} satisfying:

- i:) $\mathbb{P} \sim \mathbb{Q}$ with respect to \mathcal{F} ;
- ii:) \mathbb{P} and \mathbb{Q} agree on \mathcal{F}_0 ;
- iii:) M is a \mathcal{F}_t -measurable martingale on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

One of the most important concepts in finance is risk-neutral measure (martingale measure), which is any probability measure, equivalent to the market measure (i.e. the real world measure), which makes all discounted asset prices martingales. This property of the risk-neutral measure makes it more desirable in option pricing, as the risk-neutral does not require investors' preference towards risk which is very difficult to quantify and it is the essence of arbitrage pricing theory which guarantees arbitrage-free pricing.

Bibliography

- [1] Carol Alexander. *Market Risk Analysis, Quantitative Methods in Finance (The Wiley Finance Series)*, volume 1. May 2008.
- [2] Leif B.G. Andersen and Vladimir V. Piterbarg. *Interest Rate Modeling. Volume 2: Term Structure Models*. 2010.
- [3] Fischer Black. *The pricing of commodity contracts*, volume 3. January-March 1976.
- [4] Fischer Black and Myron Scholes. *The pricing of option and corporate liabilities*, volume 81. May-June 1973.
- [5] Zvi Bodie, Alex Kane, and Alan J. Marcus. *Investments*. McGraw-Hill/Irwin, 2008.
- [6] Paolo Brandimarte. *Numerical Methods in Finance and Economics: A MATLAB-Based Introduction (Statistics in Practice)*. Wiley-Interscience, 2006.
- [7] Mark Brodie, Paul Glasserman, and Steven Kou. *A Continuity Correction For Discrete Barrier Options*, volume 7. October 1997.
- [8] Emanuel Derman and Iraj Kani. *Riding on a Smile*, volume 2. February 1994.
- [9] Bruno Dupire. *Pricing with a Smile*. July 1994.

- [10] Henri Gavin. The levenberg-marquardt method for nonlinear least squares curve-fitting problems. 2011.
- [11] Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003 edition edition, 2003.
- [12] Patrick S. Hagan, Deep Kumar, Andrew S. Lesniewski, and Diana E. Woodward. Managing smile risk, 2002.
- [13] Espen Gaarder Haug. *The Complete Guide to Option Pricing Formulas*. McGraw-Hill, 2nd edition, 2006.
- [14] Martin Haugh. *IEOR E4703 Lecture Notes: Variance Reduction Techniques I*. Columbia University, Fall 2004.
- [15] John C. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, 8th edition, 2011.
- [16] Robert V. Kohn and Steve Allen. *Lecture Notes: Derivative Securities*. Courant Institute of Mathematical Sciences, NYU, Fall 2012.
- [17] Kenneth Levenberg and Donald Marquardt. *A Method for the Solution of Certain Non Linear Problems in Least Squares*, volume 2. 1944.
- [18] Donald Angus MacKenzie. *An Engine, Not a Camera: How Financial Models Shape Markets*. MIT Press, 2006.
- [19] Robert C. Merton. *Theory of Rational Option Pricing*. 1973.
- [20] Peter Nowak and Patrick Sibetz. *Volatility Smile(Heston, SABR)*. 2012.
- [21] Domingo Tavella. *Quantitative Methods in Derivatives Pricing: An Introduction to Computational Finance*. John Wiley & Sons, 2002.
- [22] Graeme West. Calibration of the sabr model in illiquid markets, 2005.