

Robust Option Pricing - An ϵ -arbitrage Approach

by

Si Chen

B.Eng. in Industrial and Systems Engineering (2008)
National University of Singapore

Submitted to the School of Engineering
in partial fulfillment of the requirements for the degree of
Master of Science in Computation for Design and Optimization
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

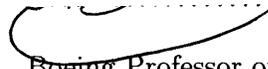
September 2009

© Massachusetts Institute of Technology 2009. All rights reserved.

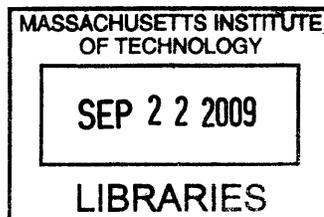
Author
School of Engineering
August 6, 2009



Certified by
Dimitris Bertsimas
Boeing Professor of Operations Research
Thesis Supervisor



Accepted by
Jaime Peraire
Professor of Aeronautics and Astronautics
Director, Computation for Design and Optimization Program



ARCHIVES

Robust Option Pricing - An ϵ -arbitrage Approach

by

Si Chen

Submitted to the CDO Program
on August, 2009, in partial fulfillment of the
requirements for the degree of
Master of Science in Computation for Design and Optimization

Abstract

This research aims to provide tractable approaches to price options using robust optimization. The pricing problem is reduced to a problem of identifying the replicating portfolio which minimizes the worst case arbitrage possible for a given uncertainty set on underlying asset returns. We construct corresponding uncertainty sets based on different levels of risk aversion of investors and make no assumption on specific probabilistic distributions of asset returns.

The most significant benefits of our approach are (a) computational tractability illustrated by our ability to price multi-dimensional options and (b) modeling flexibility illustrated by our ability to model the “volatility smile”. Specifically, we report extensive computational results that provide empirical evidence that the “implied volatility smile” that is observed in practice arises from different levels of risk aversion for different strikes. We are able to capture the phenomenon by appropriately finding the right risk-aversion as a function of the strike price. Besides European style options which have fixed exercising date, our method can also be adopted to price American style options which we can exercise early. We also show the applicability of this pricing method in the case of exotic and multi-dimensional options, in particular, we provide formulations to price Asian options, Lookback options and also Index options. These prices are compared with market prices, and we observe close matches when we use our formulations with appropriate uncertainty sets constructed based on market-implied risk aversion.

Thesis Supervisor : Dimitris Bertsimas
Title : Boeing Professor of Operations Research

Acknowledgment

I would like to express my deepest gratitude to my supervisor, Professor Dimitris Bertsimas, who has carefully guided me and inspired me through my research. I treasure this opportunity of studying an interesting topic and bringing something new to the field. Moreover, his passion, persistence and commitment have influenced me and taught me many great lessons of life.

I also want to thank Mr Chaitanya Bandi, who closely collaborated with me on this topic. I am always amazed by his creative ideas and his quick and clear mind. Thanks for sharing with me the good and bad times of this research experience.

I would like to thank my parents for their continuous support. And my last thanks goes to Chris, who has always been there for me when needed and who makes everyday precious and meaningful.

Contents

1	Introduction	7
1.1	Motivation	8
1.2	Thesis Overview	10
2	General Pricing Model	11
2.1	The Underlying Primitive for Price Dynamics	11
2.2	The ϵ – arbitrage Robust Optimization Model	12
3	Pricing European Options	14
3.1	A Nominal Formulation	14
3.2	A Linear Optimization Re-formulation	16
3.3	Modeling the Implied Volatility Smile for European Options	18
4	Pricing European Style Exotic Options	19
4.1	Asian Options	19
4.2	Lookback Options	22
5	Pricing American Options	24
5.1	Pricing with General Exercising Policy	24
6	Options on multiple assets	29
6.1	Pricing Dynamics for Multiple Assets	31
6.2	Payoff Function	33
6.3	A Linear Optimization Re-formulation	34
7	Computational Results	37
7.1	Experiment 1 : Comparison with market prices for an European call option	40

7.2	Experiment 2 : Comparison with prices of Asian call options obtained from Monte Carlo simulation	43
7.3	Experiment 3 : Comparison with prices of Lookback call options obtained from Monte Carlo simulation	47
7.4	Experiment 4 : Comparison with market prices for American put option	50
7.5	Experiment 5 : Comparison with market prices for European Style Index option	53
7.6	Remarks	56
8	Conclusion	58

List of Figures

7.1	$\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for a European Call option.	40
7.2	Comparison of Model Price and Market Price for a European Call option.	41
7.3	$\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for an Asian Call option.	44
7.4	Comparison of Model Price and Monte Carlo Price for an Asian Call option.	45
7.5	$\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for a Lookback Call option.	48
7.6	Comparison of Model Price and Monte Carlo Price for a Lookback Call option.	49
7.7	$\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for an American Put option.	51
7.8	Comparison of Model Price and Market Price for an American Put option.	52
7.9	$\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for an Index Call option.	54
7.10	Comparison of Model Price and Market Price for an Index Call option. .	55

List of Tables

1.1	Computational complexity of our model for different types of options. . .	9
7.1	Finding the Implied Gamma, the Quadratic Relationship, and Error for a European Call option.	42
7.2	Finding the Implied Gamma, the Quadratic Relationship, and Error for an Asian Call option.	46
7.3	Finding the Implied Gamma, the Quadratic Relationship, and Error for a Lookback Call option.	50
7.4	Finding the Implied Gamma, the Quadratic Relationship, and Error for an American Put option.	53
7.5	Finding the Implied Gamma, the Quadratic Relationship, and Error for an Index Call option.	56
7.6	Analysis on errors of our model prices compared with market prices (or simulation prices)	57

Chapter 1

Introduction

The problem of pricing and hedging derivative securities has been one of the most well studied problems in Financial Economics. The most important breakthrough, to this date, has been the celebrated Black and Scholes [1] and Merton [11] option-pricing formula, which is based on the principle of dynamic replication that allows one to look for a portfolio of simpler securities which is self-financing, i.e., no cash inflows or outflows except at the start and the end, and whose value at the end of the time horizon matches the payoff of the option. Such a portfolio of simpler securities is known as a replicating portfolio, and the value of this portfolio at the beginning is the no-arbitrage price of the financial instrument. For example, under geometric Brownian motion for price dynamics they show that a European call-option on a stock can be replicated exactly by a dynamic-hedging strategy involving only stocks and riskless borrowing and lending. This replication allowed them to give the first closed-form expression for the price of a European Option.

However, the assumptions required for perfect replication are often not observed in the market and in the price dynamics of the stocks. For example, the assumption of geometric Brownian motion is not supported by empirical data. Rubinstein [14] among others, documents evidence that implied volatilities tend to rise for options that are deeply in- or out-of-the-money. It has been suggested that this discrepancy between theory and market observables arises due to the assumption of deterministic or constant volatilities in Black-Scholes formula, and attempts have been made to model the volatility of the underlying asset as a stochastic quantity. The notable models include the Merton's Mixed Jump Diffusion model [12], Cox and Ross' Constant elasticity of variance [7], Hull and White's model [9], Madan, Carr and Chang's Variance Gamma model [10]. As noted by Taylor [16], there is no economic intuition behind this generalization

of volatility being stochastic. Also, as a result of increase in complexity, estimation of unobserved volatility process from market data is very difficult which makes it very hard to use it in practice.

Apart from the problems with the price-dynamics, there are other factors that arise mostly due to institutional rigidities, transaction costs, market closures etc. that makes it impossible for one to look for an exact replicating portfolio. This suggests that we need to consider the natural trade-off between exactness and tractability and to look for nearly exact replications which can be found in a tractable manner.

1.1 Motivation

Motivated by the inability to exactly replicate different types of options, Bertsimas et. al. [3] have proposed the idea of ϵ -arbitrage using stochastic dynamic programming (SDP) under which we seek a self-financing dynamic portfolio strategy that most closely approximates the payoff of the option. In all prior work in asset pricing the key primitive is the underlying stochastic process for the price dynamics. There are difficulties with assuming a specific stochastic process as the price dynamics: (1) The only available information is really return data. Fitting specific stochastic process to the data is, in our view, an approximation of reality, not reality itself; (2) Even if the stochastic processes of the the underlying price dynamics are known, they may lead to very high dimensional SDPs that are not really solvable even under the ϵ -arbitrage approach.

The second motivation for this work is the success robust optimization has enjoyed in solving high dimensional optimization problems under uncertainty. The first step in robust optimization was taken by Soyster [15]. Ben-Tal and Nemirovski [2], El-Ghaoui and Lebret [8] have moved a step forward to introduce less conservative models with ellipsoidal uncertainties. Bertsimas and Sim [4] propose ways to construct linear robust optimization models with flexible level of conservativeness and later extend the results to uncertainty in general norms [6]. The key philosophical reason for success of robust optimization is to replace probability distributions as the underlying model of randomness with uncertainty sets. The resulting robust optimization problems becomes mathematical programming models that scale with dimension much better (typically polynomial time solvable) compared to SDPs (typically exponential time solvable).

In this thesis, we propose to model the underlying price dynamics with uncertainty sets, and then apply robust optimization as opposed to SDP to solve the ϵ -arbitrage problem. We use the L_1 - norm to measure the error in replication which when combined with polyhedral uncertainty sets results in linear programming problems. This

approach also allows us to easily model transaction costs, very general pricing dynamics, accommodate high dimensional problems that today can only be handled by time consuming simulation methods.

In addition, we adopt our approach to model the “implied volatility smile” that characterizes the classical Black-Scholes model. While the original Black-Scholes model assumes volatility that is independent of the stock price, if one uses the Black-Scholes formula to price an option, one finds that the volatility needed to match prices from empirical data needs to depend on the strike price which is implausible. Our approach to explain this phenomenon is to assume that the uncertainty set that models the underlying price dynamics depends on the risk aversion of the modeler. By using this extra degree of freedom by varying a single parameter, we find a very close match between empirically observed prices of options and prices produced by our approach.

The key features of our approach are:

1. **Computational Tractability** : We combine ϵ -arbitrage and robust optimization to solve, via linear optimization methods, option pricing models that can model transaction costs, high-dimensional options, high-dimensional price dynamics, etc. The key here is that unlike current SDP and simulation methods, our approach scales in a polynomial way (as opposed to exponential way) with the dimension of the original pricing problem. Table 1.1 below summarizes the computational complexity of our model for a variety of options. T is the number of discretization steps while M is the number of underlying assets. As evidence of computational tractability and accuracy of the method, we report results for a variety of options (European, Asian, Lookback, American, Index) using empirical data, which show that our approach produces prices that are close to those observed in options market.

Options	European	Asian	Lookback	American	Index	American Index
Complexity	$O(T)$	$O(T)$	$O(T^2)$	$O(T^2)$	$O(M \cdot T)$	$O(M \cdot T^2)$

Table 1.1: Computational complexity of our model for different types of options.

2. **Modeling Flexibility** : Our approach allows to model additional information on return dynamics (by adjusting the uncertainty sets used), and risk aversion of the investor (by adjusting the parameter Γ). The flexibility of selecting different

Γ 's for different strike prices allows us to capture the “implied volatility smile” that is observed in the market.

1.2 Thesis Overview

In Chapter 2, we describe our way to model the price dynamics and propose the generic pricing model for options based on a single asset, which can be easily extended to accommodate multiple assets and complex option types in later chapters. In Chapter 3, we introduce the problem of pricing a European Option - the simplest type of option. In Chapter 4, we use our model to price more complex options such as Asian and Lookback Options. In Chapter 5, we apply our method to price American options whose exercising date is flexible. In Chapter 6, we extend our results to price options which are based on multiple assets. In Chapter 7, we conduct various experiments to compare our model to market observables and simulation results. Chapter 8 describes briefly our conclusions.

Chapter 2

General Pricing Model

In this section, we explain our approach for the case of a general derivative instrument with a predetermined payoff function and whose payoff is based on the price of underlying security.

2.1 The Underlying Primitive for Price Dynamics

In this section, we propose a method to describe the price dynamics. A key result from probability theory is the Central Limit Theorem (CLT). If \tilde{r}_t is the single period return of the underlying asset at time t, the CLT would suggest that

$$\frac{\sum_{i=1}^t \log(1 + \tilde{r}_i) - t \cdot \mu_{\log}}{\sigma_{\log} \cdot \sqrt{t}}$$

obey a normal distribution, where μ_{\log} , σ_{\log} are mean and standard deviation of $\log(1 + \tilde{r}_i)$ respectively.

Rather than assuming the primitives of probability theory, namely the Kolmogorov axioms that imply the CLT, we assume as a primitive that a CLT type of behavior happens deterministically. That is:

$$\left| \frac{\log \tilde{R}_t - t\mu_{\log}}{\sigma_{\log} \cdot \sqrt{t}} \right| \leq \Gamma, \quad \forall t \tag{2.1}$$

where $\tilde{R}_t = \prod_{i=1}^t (1 + \tilde{r}_i)$, is the cumulative return at time t. The parameter Γ determines how conservative or risk averse we want the solution to be. Inequality (2.1) is equivalent to

$$e^{t\mu_{log}-\Gamma\sqrt{t}\sigma_{log}} \leq \tilde{R}_t \leq e^{t\mu_{log}+\Gamma\sqrt{t}\sigma_{log}}, \quad \forall t \quad (2.2)$$

In addition, we assume some boundaries on single period return \tilde{r}_t and by definition $\tilde{r}_t = \frac{\tilde{R}_t}{\tilde{R}_{t-1}}$, we have:

$$\mu - \Gamma_t\sigma \leq \frac{\tilde{R}_t}{\tilde{R}_{t-1}} \leq \mu + \Gamma_t\sigma, \quad \forall t \quad (2.3)$$

where μ and σ are the mean and standard deviation of \tilde{r}_t respectively and Γ_t are constants (in our experiments, we let $\Gamma_t = \Gamma$). Notice that the uncertainty sets are defined using cumulative returns \tilde{R}_t instead of single period return \tilde{r}_t . This is because in the later sections where we simplify our model, we transform our model to use cumulative returns only. The values of μ , σ , μ_{log} , σ_{log} can be obtained from empirical data on single periods returns.

2.2 The ϵ – arbitrage Robust Optimization Model

An option has a predetermined payoff function $P(\tilde{\mathbf{S}}, \mathbf{K})$ that is determined by the price of the underlying securities at exercising date and the strike price K . For example, a European Call option's payoff is $\max\{(\tilde{S}_T - K), 0\}$ or for simplicity of notation $(\tilde{S}_T - K)^+$, where \tilde{S}_t denotes the price of the underlying security at time t .

The idea of our approach is to find a portfolio that consists of underlying security S and risk-free asset B so that the wealth W_T of this portfolio at exercising date T should match the payoff of the option within a error of ϵ . We refer to this error as arbitrage which is calculated by:

$$\left| P(\tilde{\mathbf{S}}, \mathbf{K}) - W_T \right|$$

In a robust optimization setting, our goal is to find a portfolio that minimizes the worst case arbitrage (denoted by ϵ) between the portfolio wealth and the option payoff over the uncertainty set of security return defined in Section 2.1. The portfolio found by the robust optimization model will have payoff that is within $\pm\epsilon$ from the actual option payoff for all random security returns under the uncertainty sets (2.2) and (2.3). The current price of the option would thus be the initial value of the best matching portfolio.

We define the following:

Decision variables:

- x_t^S : amount invested in the underlying security during the period $[t, t + 1]$;

- x_t^B : amount invested in the riskless asset during the period $[t, t + 1]$;
- y_t : amount traded from the underlying security to the riskless asset at the starting of the period $[t, t + 1]$. It can be either positive (means transfer money from underlying security to riskless asset) or negative (the other way around);
- p_0 : amount of wealth we start out with. This will be the price of the instrument computed by the method.

Data:

- \tilde{r}_t^S : rate of return from underlying security during the period $[t, t + 1]$. It is subject to uncertainty;
- r_t^B : rate of return from riskless asset during the period $[t, t + 1]$, which assumes to be constant;
- $\mathbf{U} = \left\{ \tilde{R}_t^S \mid e^{t\mu_{log} - \Gamma\sqrt{t}\sigma_{log}} \leq \tilde{R}_t \leq e^{t\mu_{log} + \Gamma\sqrt{t}\sigma_{log}}, \mu - \Gamma_t\sigma \leq \frac{\tilde{R}_t}{\tilde{R}_{t-1}} \leq \mu + \Gamma_t\sigma, \forall t \right\}$.

The objective is to minimize over the portfolio weight the maximum arbitrage over random returns \tilde{r}_t^S :

Generic Option Pricing Formulation (P1)

$$\begin{aligned}
& \min_{\{x_t^S, x_t^B\}} \max_{\{\tilde{r}_t^S\}} \left| P(\tilde{\mathbf{S}}, \mathbf{K}) - W_T \right| \\
& \text{s.t.} \quad W_T = x_T^S + x_T^B \\
& \quad x_t^S = (1 + \tilde{r}_{t-1}^S) (x_{t-1}^S + y_{t-1}), \forall t = 1 \dots T, \\
& \quad x_t^B = (1 + r_{t-1}^B) (x_{t-1}^B - y_{t-1}), \forall t = 1 \dots T,
\end{aligned}$$

Our proposed price is then $p_0 = x_0^S + x_0^B$, where x_0^S, x_0^B are the optimal solutions of the above LP.

In the following chapters, we explain how we apply this model to price different types of options, and how we reformulate this problem to become a linear programming problem.

Chapter 3

Pricing European Options

In this section, we present our approach in the context of a European option which is the foundation of the options universe. This option allows the holder of the option the ability to exercise the option at (and only at) the expiry date T with a given price K . The payoff function for a European Call is given by $(\tilde{S}_T - K)^+$.

3.1 A Nominal Formulation

Using the same set of decision variables and data in General Pricing Problem (P1) and with payoff function $(\tilde{S}_T - K)^+$, we write the following:

European Call Option Pricing Formulation (ECO-1)

$$\begin{aligned} \min_{\{x_t^S, x_t^B\}} \max_{\{\tilde{r}_t^S\}} & \left| (\tilde{S}_T - K)^+ - W_T \right| \\ \text{s.t.} & W_T = x_T^S + x_T^B \\ & x_t^S = (1 + \tilde{r}_{t-1}^S) (x_{t-1}^S + y_{t-1}), \forall t = 1 \dots T, \\ & x_t^B = (1 + r_{t-1}^B) (x_{t-1}^B - y_{t-1}), \forall t = 1 \dots T, \end{aligned}$$

Our proposed price is then $p_0 = x_0^S + x_0^B$. In the above formulation, the uncertainty lies in return values. The following transformation of variables allows us to reduce the number of constraints with uncertain coefficients :

- $\alpha_t^S = \frac{x_t^S}{R_t^S}$
- $\alpha_t^B = \frac{x_t^B}{R_t^B}$

- $\beta_t = \frac{y_t}{R_t^S}$

where cumulative returns are defined as:

- $\tilde{R}_t^S = \prod_{i=0}^{t-1} (1 + \tilde{r}_i^S)$

- $R_t^B = \prod_{i=0}^{t-1} (1 + r_i^B)$

These transformations lead to the following formulation:

European Call Option Pricing Formulation (ECO-2)

$$\begin{aligned} \min_{\{\alpha_t^S, \alpha_t^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} & \left| \left(S_0 \tilde{R}_T^S - K \right)^+ - \left(\tilde{R}_T^S \alpha_T^S + R_T^B \alpha_T^B \right) \right| \\ \text{s.t.} & \alpha_t^S = \alpha_{t-1}^S + \beta_{t-1}, \quad \forall t = 1 \dots T \\ & \alpha_t^B = \alpha_{t-1}^B - \beta_{t-1} \frac{\tilde{R}_t^S}{R_t^B}, \quad \forall t = 1 \dots T \end{aligned}$$

Substituting all intermediate α_t^B , α_t^S , (ECO-2) is equivalent as (ECO-3) shown below:

European Call Option Pricing Formulation (ECO-3)

$$\min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} \left| \left(S_0 \tilde{R}_T^S - K \right)^+ - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_t \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right|$$

Model (ECO-3) is equivalent to the following problem considering the sign of the non-negative payoff function and the modulus:

European Call Option Pricing Formulation (ECO-4)

$$\begin{aligned} \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} & \epsilon \\ \text{s.t.} & \left(S_0 \tilde{R}_T^S - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & - \left(\left(S_0 \tilde{R}_T^S - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \\ & - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & - \left(- \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \end{aligned}$$

3.2 A Linear Optimization Re-formulation

We provide linear optimization formulations that when solved give the optimal solution to the min-max problem discussed above. We consider the uncertainty sets (2.2) - (2.3).

Theorem 3.1 (Pachamanova [13]) If the polyhedral uncertainty set is $P^A = \{vec(\tilde{\mathbf{A}}) | \mathbf{G} \cdot vec(\tilde{\mathbf{A}}) \leq \mathbf{d}\} \neq \emptyset$, then a given $\hat{\mathbf{x}}$ satisfies $(\mathbf{a})'\hat{\mathbf{x}} \leq b$ for all $\tilde{\mathbf{A}} \in P^A$ if and only if there exist a vector $\mathbf{p} \in \mathfrak{R}^{l \times 1}$ such that:

$$\begin{aligned} (\mathbf{p})'\mathbf{d} &\leq b \\ (\mathbf{p})'\mathbf{G} &= (\hat{\mathbf{x}}) \\ \mathbf{p} &\geq 0 \end{aligned}$$

Where $\mathbf{p} \in \mathfrak{R}^{l \times 1}$ is the vector of dual variables associated with the constraint $\mathbf{G} \cdot vec(\mathbf{A}) \leq \mathbf{d}$. The proof of this theorem is in Appendix A.

With regards to the uncertainty set defined by (2.2)-(2.3), we introduce the following four sets of parameters:

1. $\underline{R}_t^S = e^{t\mu_{log} - \Gamma\sqrt{t}\sigma_{log}} \forall t = 1 \dots T$ (the lower bound of cumulative returns)
2. $\overline{R}_t^S = e^{t\mu_{log} + \Gamma\sqrt{t}\sigma_{log}} \forall t = 1 \dots T$ (the upper bound of cumulative returns)
3. $\underline{r}_t^S = \mu_r - \Gamma_t\sigma_r \forall t = 2 \dots T$ (the lower bound of one period returns)
4. $\overline{r}_t^S = \mu_r + \Gamma_t\sigma_r \forall t = 2 \dots T$ (the upper bound of one period returns)

We rewrite uncertainty sets (2.2) - (2.3) into constraints in the form of $\mathbf{G} \cdot vec(\mathbf{A}) \leq \mathbf{d}$:

$$-\tilde{R}_t^S \leq -\underline{R}_t^S \quad \forall t \quad (3.1)$$

$$\tilde{R}_t^S \leq \overline{R}_t^S \quad \forall t \quad (3.2)$$

$$-\tilde{R}_t^S \leq -\underline{r}_t^S R_{t-1}^S \quad \forall t \quad (3.3)$$

$$\tilde{R}_t^S \leq \overline{r}_t^S R_{t-1}^S \quad \forall t \quad (3.4)$$

We introduce \mathbf{p} , \mathbf{q} , \mathbf{m} , \mathbf{n} as vectors of dual variables associated with the four series of constraints specified in (3.1)-(3.4). z is the dual variable with $\tilde{R}_T^S \geq \frac{K}{S_0}$ or $\tilde{R}_T^S \leq \frac{K}{S_0}$ depending on the condition specified in that constraint in (ECO-4). Apply Theorem 3.1 to each constraint in (ECO-4), treating \tilde{R}_t^S as random data A and $\alpha_0^S, \alpha_0^B, \beta_t$ as

decision variables \mathbf{x} which assume certain given value. We yield linear optimization model (ECO-LP) below.

European Call Option Pricing - Linear Formulation (ECO-LP)

min ϵ

s.t.

$$\frac{K}{S_0} z_1 + \sum_{t=1}^T R_t^S p_{t,1} + \sum_{t=1}^T \overline{R}_t^S q_{t,1} - K - \alpha_0^B R_T^B \leq \epsilon$$

$$p_{1,1} + q_{1,1} - \underline{r}_1^S m_{2,1} - \overline{r}_1^S n_{2,1} = \beta_1 \frac{R_T^B}{R_1^B}$$

$$p_{t,1} + q_{t,1} + m_{t,1} + n_{t,1} - \underline{r}_t^S m_{t+1,1} - \overline{r}_t^S n_{t+1,1} = \beta_t \frac{R_T^B}{R_t^B}, \quad \forall t = 2 \dots T$$

$$z_1 + p_{T,1} + q_{T,1} + m_{T,1} + n_{T,1} = S_0 - \alpha_0^S - \sum_{t=1}^T \beta_t + \beta_T$$

$$\frac{K}{S_0} z_2 + \sum_{t=1}^T R_t^S p_{t,2} + \sum_{t=1}^T \overline{R}_t^S q_{t,2} + K + \alpha_0^B R_T^B \leq \epsilon$$

$$p_{1,2} + q_{1,2} - \underline{r}_1^S m_{2,2} - \overline{r}_1^S n_{2,2} = -\beta_1 \frac{R_T^B}{R_1^B}$$

$$p_{t,2} + q_{t,2} + m_{t,2} + n_{t,2} - \underline{r}_t^S m_{t+1,2} - \overline{r}_t^S n_{t+1,2} = -\beta_t \frac{R_T^B}{R_t^B}, \quad \forall t = 2 \dots T$$

$$z_2 + p_{T,2} + q_{T,2} + m_{T,2} + n_{T,2} = S_0 + \alpha_0^S + \sum_{t=1}^T \beta_t - \beta_T$$

$$\frac{K}{S_0} z_3 + \sum_{t=1}^T R_t^S p_{t,3} + \sum_{t=1}^T \overline{R}_t^S q_{t,3} - \alpha_0^B R_T^B \leq \epsilon$$

$$p_{1,3} + q_{1,3} - \underline{r}_1^S m_{2,3} + \overline{r}_1^S n_{2,3} = \beta_1 \frac{R_T^B}{R_1^B}$$

$$p_{t,3} + q_{t,3} + m_{t,3} + n_{t,3} - \underline{r}_t^S m_{t+1,3} + \overline{r}_t^S n_{t+1,3} = \beta_t \frac{R_T^B}{R_t^B}, \quad \forall t = 2 \dots T$$

$$z_3 + p_{T,3} + q_{T,3} + m_{T,3} + n_{T,3} = -\alpha_0^S - \sum_{t=1}^T \beta_t + \beta_T$$

$$\frac{K}{S_0} z_4 + \sum_{t=1}^T R_t^S p_{t,4} + \sum_{t=1}^T \overline{R}_t^S q_{t,4} + \alpha_0^B R_T^B \leq \epsilon$$

$$p_{1,4} + q_{1,4} - \underline{r}_1^S m_{2,4} + \overline{r}_1^S n_{2,4} = -\beta_1 \frac{R_T^B}{R_1^B}$$

$$p_{t,4} + q_{t,4} + m_{t,4} + n_{t,4} - \underline{r}_t^S m_{t+1,4} + \overline{r}_t^S n_{t+1,4} = -\beta_t \frac{R_T^B}{R_t^B}, \quad \forall t = 2 \dots T$$

$$z_4 + p_{T,4} + q_{T,4} + m_{T,4} + n_{T,4} = \alpha_0^S + \sum_{t=1}^T \beta_t - \beta_T$$

$$p_{t,i} \leq 0, q_{t,i} \geq 0, m_{t,i} \leq 0, n_{t,i} \geq 0, \quad \forall t = 1 \dots T, \quad \forall i = 1 \dots 4$$

$$z_1 \leq 0, z_2 \leq 0, z_3 \geq 0, z_4 \geq 0$$

The size of linear optimization problem ECO-LP scales linearly as the number of periods - which is the number of discretization time steps increases. The number of decision variables is $16T + 4$. The number of constraints is $4T + 4$.

One could find the price for European put option using the put-call parity or simply replacing the payoff function with $(K - \tilde{S}_T)^+$ in the above formulation. In the last part of this section, we describe our approach to model the implied volatility smile that is observed in the market.

3.3 Modeling the Implied Volatility Smile for European Options

The volatility smile is a long-observed pattern in which at-the-money options tend to have lower implied volatilities than in- or out-of-the-money options. The traditional way the research and practice community deals with this phenomenon is to assume that the volatility of the underlying security varies with the strike K , which is not a sound concept. Our proposal is to assume that the parameter Γ to define the uncertainty set (2.2)-(2.3) depends on $\frac{K}{S_0}$. The intuition here is that we want to select a larger Γ , that is to make our approach more robust, when $\frac{K}{S_0}$ is away of one (that is in-the-money or out-of-the-money). In other words, the parameter Γ attempts to capture the risk-aversion of the modeler to define the uncertainty set. For more details, see for example the discussion surrounding equation (7.1).

Chapter 4

Pricing European Style Exotic Options

In this section, we extend our proposed methodology and present our solution for Asian and Lookback options.

4.1 Asian Options

An Asian option (also called an average option) is an option whose payoff is linked to the average value of the underlying securities on a specific set of dates during the life of the option. Exact analytic formulas for average rate options do not exist. This is primarily due to the fact that the arithmetic average of a set of log-normal random variables has a distribution that is largely intractable. We present here a pricing mechanism that is computationally tractable and allows one to capture different levels of risk-aversion.

Suppose the Asian option has an expiry date T and strike price K . If we assume discrete time steps, then the payoff of an Asian call option is $(S_{ave} - K)^+$ with $S_{ave} = \frac{\sum_{t=1}^T S_t}{T}$, where S_t is the price of the security observed at these t . With this setup, one can formulate our optimal replication problem using the same decision variables and uncertainty set of underlying security returns in Chapter 3.

Asian Call Option Pricing Formulation (ACO-1)

$$\begin{aligned} & \min_{\{\alpha_t^S, \alpha_t^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} \left| \left(S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} - K \right)^+ - \left(\tilde{R}_T^S \alpha_T^S + R_T^B \alpha_T^B \right) \right| \\ & \text{s.t.} \quad \alpha_t^S = \alpha_{t-1}^S + \beta_{t-1}, \quad \forall t = 1 \dots T \\ & \quad \quad \alpha_t^B = \alpha_{t-1}^B - \beta_{t-1} \frac{\tilde{R}_t^S}{R_t^B}, \quad \forall t = 1 \dots T \end{aligned}$$

Substituting all intermediate α_t^B , α_t^S , (ACO-1) is equivalent as (ACO-2) shown below:

Asian Call Option Pricing Formulation(ACO-2)

$$\min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} \left| \left(S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} - K \right)^+ - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right|$$

(ACO-2) is equivalent to the following problem (ACO-3) by considering case by case the sign of the non-negative payoff function and the modulus:

Asian Call Option Pricing Formulation (ACO-3)

$$\begin{aligned} & \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} \epsilon \\ & \text{s.t.} \quad \left(S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & \quad - \left(\left(S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \\ & \quad \quad - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & \quad \quad - \left(- \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \end{aligned}$$

We proceed as in Section 3.2, and we introduce \mathbf{p} , \mathbf{q} , \mathbf{m} , \mathbf{n} as vectors of dual variables associated with the four series of constraints specified in (3.1)-(3.4). z is the dual variable with associated with the constraints $S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} \geq \frac{K}{S_0}$ or $S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} \leq \frac{K}{S_0}$ depending on the case. Applying Theorem 3.1 to each constraint in (ACO-3), treating \tilde{R}_t^S

as random data A and $\alpha_0^S, \alpha_0^B, \beta_t$ as decision variables, we obtain the linear optimization model (ACO-LP).

Asian Call Option Pricing Linear Formulation(ACO-LP)

min ϵ

s.t.

$$\frac{K}{S_0} z_1 - \sum_{t=1}^T \underline{R}_t^S p_{t,1} + \sum_{t=1}^T \overline{R}_t^S q_{t,1} - K - \alpha_0^B R_T^B \leq \epsilon$$

$$\frac{z_1}{T} - p_{1,1} + q_{1,1} + \underline{r}_1^S m_{2,1} - \overline{r}_1^S n_{2,1} = \beta_0 \frac{R_T^B}{R_1^B} + \frac{S_0}{T}$$

$$\frac{z_1}{T} - p_{t,1} + q_{t,1} - m_{t,1} + n_{t,1} + \underline{r}_t^S m_{t+1,1} - \overline{r}_t^S n_{t+1,1} = \beta_{t-1} \frac{R_T^B}{R_t^B} + \frac{S_0}{T}, \quad \forall t = 2 \dots T$$

$$\frac{z_1}{T} - p_{T,1} + q_{T,1} - m_{T,1} + n_{T,1} = \frac{S_0}{T} - \alpha_0^S - \sum_{t=1}^T \beta_{t-1} + \beta_{T-1}$$

$$\frac{K}{S_0} z_2 - \sum_{t=1}^T \underline{R}_t^S p_{t,2} + \sum_{t=1}^T \overline{R}_t^S q_{t,2} + K + \alpha_0^B R_T^B \leq \epsilon$$

$$\frac{z_2}{T} + p_{1,2} + q_{1,2} + \underline{r}_1^S m_{2,2} - \overline{r}_1^S n_{2,2} = -\beta_0 \frac{R_T^B}{R_1^B} + \frac{S_0}{T}$$

$$\frac{z_2}{T} - p_{t,2} + q_{t,2} - m_{t,2} + n_{t,2} + \underline{r}_t^S m_{t+1,2} - \overline{r}_t^S n_{t+1,2} = -\beta_{t-1} \frac{R_T^B}{R_t^B} + \frac{S_0}{T}, \quad \forall t = 2 \dots T$$

$$\frac{z_2}{T} - p_{T,2} + q_{T,2} - m_{T,2} + n_{T,2} = -\frac{S_0}{T} + \alpha_0^S + \sum_{t=1}^T \beta_{t-1} - \beta_{T-1}$$

$$\frac{K}{S_0} z_3 + \sum_{t=1}^T \underline{R}_t^S p_{t,3} + \sum_{t=1}^T \overline{R}_t^S q_{t,3} - \alpha_0^B R_T^B \leq \epsilon$$

$$\frac{z_3}{T} - p_{1,3} + q_{1,3} - \underline{r}_1^S m_{2,3} + \overline{r}_1^S n_{2,3} = \beta_0 \frac{R_T^B}{R_1^B}$$

$$\frac{z_3}{T} - p_{t,3} + q_{t,3} - m_{t,3} + n_{t,3} + \underline{r}_t^S m_{t+1,3} - \overline{r}_t^S n_{t+1,3} = \beta_{t-1} \frac{R_T^B}{R_t^B}, \quad \forall t = 2 \dots T$$

$$\frac{z_3}{T} - p_{T,3} + q_{T,3} - m_{T,3} + n_{T,3} = -\alpha_0^S - \sum_{t=1}^T \beta_{t-1} + \beta_{T-1}$$

$$\frac{K}{S_0} z_4 + \sum_{t=1}^T \underline{R}_t^S p_{t,4} + \sum_{t=1}^T \overline{R}_t^S q_{t,4} + \alpha_0^B R_T^B \leq \epsilon$$

$$\frac{z_4}{T} - p_{1,4} + q_{1,4} - \underline{r}_1^S m_{2,4} + \overline{r}_1^S n_{2,4} = -\beta_0 \frac{R_T^B}{R_1^B}$$

$$\frac{z_4}{T} - p_{t,4} + q_{t,4} - m_{t,4} + n_{t,4} + \underline{r}_t^S m_{t+1,4} - \overline{r}_t^S n_{t+1,4} = -\beta_{t-1} \frac{R_T^B}{R_t^B}, \quad \forall t = 2 \dots T$$

$$\frac{z_4}{T} - p_{T,4} + q_{T,4} - m_{T,4} + n_{T,4} = \alpha_0^S + \sum_{t=1}^T \beta_{t-1} - \beta_{T-1}$$

$$p_{t,i} \leq 0, q_{t,i} \geq 0, m_{t,i} \leq 0, n_{t,i} \geq 0, \quad \forall t = 1 \dots T, \quad \forall i = 1 \dots 4$$

$$z_1 \leq 0, z_2 \leq 0, z_3 \geq 0, z_4 \geq 0$$

The price of put options can be obtained replacing $(S_{ave} - K)^+$ with $(K - S_{ave})^+$. Note that same as European option, the number of decision variables in (ACO-LP) is $16T + 4$. The number of constraints is $4T + 4$.

4.2 Lookback Options

A Lookback option is a path dependent option settled based upon the maximum or minimum underlying value achieved during the entire life of the option. Essentially, at expiration, the holder can look back over the life of the option and exercise based upon the optimal underlying value achieved during that period.

Here we consider a fixed strike Lookback call option. Such an option is described by strike price K and time of exercise T . The payoff of a call option is given by $(S_{max} - K)^+$ where $S_{max} = \max \{S_t | t \in (0, T)\}$. Similarly, the payoff of a put option is $(K - S_{max})^+$. In order to find S_{max} ($= S_0 R_{max}$), we consider a discretization of the time into T periods and take note of the price in each period. Let \tilde{R}_t^S be the cumulative return at time t . Then one can write the optimization problem as follows:

Lookback Call Option Pricing Formulation (LCO-1)

$$\begin{aligned} \min_{\{\alpha_t^S, \alpha_t^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} & \left| (S_0 R_{max} - K)^+ - \left(\tilde{R}_T^S \alpha_T^S + R_T^B \alpha_T^B \right) \right| \\ \text{s.t.} & \alpha_t^S = \alpha_{t-1}^S + \beta_{t-1}, \quad \forall t = 1 \dots T \\ & \alpha_t^B = \alpha_{t-1}^B - \beta_{t-1} \frac{\tilde{R}_t^S}{R_t^B}, \quad \forall t = 1 \dots T \\ & R_{max} = \max_{t=1 \dots T} \left\{ \tilde{R}_t^S \right\} \end{aligned}$$

Substituting all intermediate α_t^B , α_t^S , (LCO-1) is equivalent to (LCO-2) shown below:

Lookback Call Option Pricing Formulation (LCO-2)

$$\min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} \left| (S_0 R_{max} - K)^+ - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right|$$

where R_{max} is largest of all \tilde{R}_t^S . Since the objective is to minimize the worst case arbitrage, we divide the uncertainty set into T subsets, where $R_{max} = \tilde{R}_t^S$ in the t^{th} subset. The resulting mathematical program is (LCO-3):

Lookback Call Option Pricing Formulation (LCO-3)

$$\begin{aligned}
 & \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} \epsilon \\
 & \text{s.t} \\
 & \text{When } \tilde{R}_1^S \geq \tilde{R}_t^S, \forall t : \\
 & \left(S_0 \tilde{R}_1^S - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\
 & - \left(\left(S_0 \tilde{R}_1^S - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \\
 & - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\
 & - \left(- \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \\
 & \text{When } \tilde{R}_2^S \geq \tilde{R}_t^S, \forall t : \\
 & \vdots \\
 & \vdots \\
 & \text{When } \tilde{R}_T^S \geq \tilde{R}_t^S, \forall t : \\
 & \left(S_0 \tilde{R}_T^S - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\
 & - \left(\left(S_0 \tilde{R}_T^S - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \\
 & - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\
 & - \left(- \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon
 \end{aligned}$$

There are $4 \times T$ constraints in (LCO-3), and for each of the constraint, we can apply Theorem 3.1 to obtain a linear formulation. The number of decision variables in (ACO-LP) is $80T^2 + 16T$. The number of constraints is $4T^2 + 4T$.

Chapter 5

Pricing American Options

American style options are options which can be exercised early at a date before expiry. Such an option is described by strike price K and time of exercise τ , $\tau \in (0, T]$. The payoff of a call option is given by $(S_\tau - K)^+$. Similarly, the payoff of a put option is $(K - S_\tau)^+$. Under the assumption of no dividend, the price of American call option is the same as the equivalent European call option because it is optimal to wait till the expiry date to exercise. However, the option of early exercising does bring value to the put options. Thus, the interest of study in this section is American put options. Consider a discretization of the time into T equal periods and we assume the option can be exercised only at the end of these time periods. All other notations remain the same.

5.1 Pricing with General Exercising Policy

An American option is a contract that allows the holder to exercise before the contract's maturity. Because of the flexibility of choosing the exercise time, the fair value of the contract is calculated as the value of the option in the worst case for the issuer among all feasible exercise strategies that the option holder may choose. This is a safe strategy from the perspective of the option writer as he is delta-hedged even for the optimal option exerciser. The problem, then, is to come up with a dynamic hedging strategy that tries to replicate the payoff at every time step at which there is a possibility for the option exerciser to exercise. If the option writer believes that the option holder will exercise only at certain time steps, he can then choose a strategy that allows him to replicate the payoff at those time steps.

Without loss of generality, we assume that the option holder can exercise at any

of the time steps $\tau = 1 \dots T$. The problem then reduces to determining the dynamic hedging strategy that minimizes the worst case difference between the replicating portfolio and the payoff (arbitrage or ϵ) accounting for the possibility that the exercising time can be any time from 1 to T.

We define (APO-1) below:

American Put Option Pricing Formulation(APO-1)

$$\begin{aligned} \min_{\{\alpha_t^S, \alpha_t^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} \max_{\tau=1 \dots T} & \left| \left(K - S_0 \tilde{R}_\tau \right)^+ - \left(\tilde{R}_\tau^S \alpha_\tau^S + R_\tau^B \alpha_\tau^B \right) \right| \\ \text{s.t.} & \alpha_t^S = \alpha_{t-1}^S + \beta_{t-1}, \quad \forall t = 1 \dots T \\ & \alpha_t^B = \alpha_{t-1}^B - \beta_{t-1} \frac{\tilde{R}_t^S}{R_t^B}, \quad \forall t = 1 \dots T \end{aligned}$$

Substituting all intermediate α_t^B , α_t^S , (APO-1) is equivalent as (APO-2) shown below:

American Put Option Pricing Formulation(APO-2)

$$\min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} \max_{\tau=1 \dots T} \left| \left(K - S_0 \tilde{R}_\tau \right)^+ - \left(\alpha_0^S + \sum_{t=1}^{\tau} \beta_{t-1} \right) \tilde{R}_\tau^S - \alpha_0^B R_\tau^B + \sum_{t=1}^{\tau} \beta_{t-1} \frac{R_\tau^B}{R_t^B} \tilde{R}_t^S \right|,$$

which is equivalent to (APO-3) where we enumerate the possible values of τ and consider case by case the sign of the payoff function and the modulus.

American Put Option Pricing Formulation(APO-3)

$$\min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} \epsilon$$

s.t.

$\tau = 1 :$

$$\begin{aligned} & \left(K - S_0 \tilde{R}_1^S \right) - \alpha_0^S \tilde{R}_1^S - \alpha_0^B R_1^B \leq \epsilon \\ & - \left(\left(K - S_0 \tilde{R}_1^S \right) - \alpha_0^S \tilde{R}_1^S - \alpha_0^B R_1^B \right) \leq \epsilon \\ & - \alpha_0^S \tilde{R}_1^S - \alpha_0^B R_1^B \leq \epsilon \\ & - \left(-\alpha_0^S \tilde{R}_1^S - \alpha_0^B R_1^B \right) \leq \epsilon \end{aligned}$$

\vdots

$\tau = 2 \dots (T-1)$

$$\begin{aligned} & \left(K - S_0 \tilde{R}_\tau^S \right) - \left(\alpha_0^S + \sum_{t=1}^{\tau} \beta_{t-1} \right) \tilde{R}_\tau^S - \alpha_0^B R_\tau^B + \sum_{t=1}^{\tau} \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & - \left(\left(K - S_0 \tilde{R}_\tau^S \right) - \left(\alpha_0^S + \sum_{t=1}^{\tau} \beta_{t-1} \right) \tilde{R}_\tau^S - \alpha_0^B R_\tau^B + \sum_{t=1}^{\tau} \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \\ & - \left(\alpha_0^S + \sum_{t=1}^{\tau} \beta_{t-1} \right) \tilde{R}_\tau^S - \alpha_0^B R_\tau^B + \sum_{t=1}^{\tau} \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & - \left(- \left(\alpha_0^S + \sum_{t=1}^{\tau} \beta_{t-1} \right) \tilde{R}_\tau^S - \alpha_0^B R_\tau^B + \sum_{t=1}^{\tau} \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \end{aligned}$$

\vdots

$\tau = T :$

$$\begin{aligned} & \left(K - S_0 \tilde{R}_T^S \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & - \left(\left(K - S_0 \tilde{R}_T^S \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \\ & - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \leq \epsilon \\ & - \left(- \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \right) \leq \epsilon \end{aligned}$$

There are $4 \times T$ constraints in APO-3. We introduce \mathbf{p} , \mathbf{q} , \mathbf{m} , \mathbf{n} as vectors of dual variables associated with the four sets of constraints specified in (3.1)-(3.4). \mathbf{z} is the dual variable with $\tilde{R}_\tau^S \geq \frac{K}{S_0}$ or $\tilde{R}_\tau^S \leq \frac{K}{S_0}$ depending on the condition specified in that constraint. Applying Theorem 3.1 to each constraint in (APO-3) by treating \tilde{R}_t^S as random data A and $\alpha_0^S, \alpha_0^B, \beta_t$ as decision variables, we obtain (APO-LP).

American Put Option Pricing - Linear Optimization formulation (APO-LP)

min ϵ

s.t.

$\forall \tau = 1 \dots T :$

$$\begin{aligned}
& \frac{K}{S_0} z_{\tau,1} - \sum_{t=1}^T \underline{R}_t^S p_{t,\tau,1} + \sum_{t=1}^T \overline{R}_t^S q_{t,\tau,1} + K - \alpha_0^B R_T^B \leq \epsilon \\
& -p_{1,\tau,1} + q_{1,\tau,1} + \underline{r}_1^S m_{2,\tau,1} - \overline{r}_1^S n_{2,\tau,1} = \beta_0 \frac{R_T^B}{R_1^B} \\
& -p_{t,\tau,1} + q_{t,\tau,1} - m_{t,\tau,1} + n_{t,\tau,1} + \underline{r}_t^S m_{t+1,\tau,1} - \overline{r}_t^S n_{t+1,\tau,1} = \beta_{t-1} \frac{R_T^B}{R_t^B} \quad \forall t = 2 \dots (\tau - 1) \\
& z_{\tau,1} - p_{\tau,\tau,1} + q_{\tau,\tau,1} - m_{\tau,\tau,1} + n_{\tau,\tau,1} + \underline{r}_\tau^S m_{\tau+1,\tau,1} - \overline{r}_\tau^S n_{\tau+1,\tau,1} \\
& \quad = -S_0 - (\alpha_0^S + \sum_{t=1}^T \beta_{t-1}) + \beta_{\tau-1} \\
& -p_{t,\tau,1} + q_{t,\tau,1} - m_{t,\tau,1} + n_{t,\tau,1} + \underline{r}_t^S m_{t+1,\tau,1} - \overline{r}_t^S n_{t+1,\tau,1} = 0 \quad \forall t = (\tau + 1) \dots T - 1 \\
& -p_{T,\tau,1} + q_{T,\tau,1} - m_{T,\tau,1} + n_{T,\tau,1} = 0
\end{aligned}$$

$$\begin{aligned}
& \frac{K}{S_0} z_{\tau,2} - \sum_{t=1}^T \underline{R}_t^S p_{t,\tau,2} + \sum_{t=1}^T \overline{R}_t^S q_{t,\tau,2} - K + \alpha_0^B R_T^B \leq \epsilon \\
& -p_{1,\tau,2} + q_{1,\tau,2} + \underline{r}_1^S m_{2,\tau,2} - \overline{r}_1^S n_{2,\tau,2} = -\beta_0 \frac{R_T^B}{R_1^B} \\
& -p_{t,\tau,2} + q_{t,\tau,2} - m_{t,\tau,2} + n_{t,\tau,2} + \underline{r}_t^S m_{t+1,\tau,2} - \overline{r}_t^S n_{t+1,\tau,2} = -\beta_{t-1} \frac{R_T^B}{R_t^B} \quad t = 2 \dots (\tau - 1) \\
& z_{\tau,2} - p_{\tau,\tau,2} + q_{\tau,\tau,2} - m_{\tau,\tau,2} + n_{\tau,\tau,2} + \underline{r}_\tau^S m_{\tau+1,\tau,2} - \overline{r}_\tau^S n_{\tau+1,\tau,2} \\
& \quad = S_0 + (\alpha_0^S + \sum_{t=1}^T \beta_{t-1}) - \beta_{\tau-1} \\
& -p_{t,\tau,2} + q_{t,\tau,2} - m_{t,\tau,2} + n_{t,\tau,2} + \underline{r}_t^S m_{t+1,\tau,2} - \overline{r}_t^S n_{t+1,\tau,2} = 0 \quad t = (\tau + 1) \dots T - 1 \\
& -y_{T,\tau,2} - p_{T,\tau,2} + q_{T,\tau,2} - m_{T,\tau,2} + n_{T,\tau,2} = 0
\end{aligned}$$

$$\begin{aligned}
& -\frac{K}{S_0} z_{\tau,3} - \sum_{t=1}^T \underline{R}_t^S p_{t,\tau,3} + \sum_{t=1}^T \overline{R}_t^S q_{t,\tau,3} - \alpha_0^B R_T^B \leq \epsilon \\
& -p_{1,\tau,3} + q_{1,\tau,3} + \underline{r}_1^S m_{2,\tau,3} - \overline{r}_1^S n_{2,\tau,3} = \beta_0 \frac{R_T^B}{R_1^B} \\
& -p_{t,\tau,3} + q_{t,\tau,3} - m_{t,\tau,3} + n_{t,\tau,3} + \underline{r}_t^S m_{t+1,\tau,3} - \overline{r}_t^S n_{t+1,\tau,3} = \beta_{t-1} \frac{R_T^B}{R_t^B} \quad t = 2 \dots (\tau - 1) \\
& -z_{\tau,3} - p_{\tau,\tau,3} + q_{\tau,\tau,3} - m_{\tau,\tau,3} + n_{\tau,\tau,3} + \underline{r}_\tau^S m_{\tau+1,\tau,3} - \overline{r}_\tau^S n_{\tau+1,\tau,3} \\
& \quad = -(\alpha_0^S + \sum_{t=1}^T \beta_{t-1}) + \beta_{\tau-1} \\
& -p_{t,\tau,3} + q_{t,\tau,3} - m_{t,\tau,3} + n_{t,\tau,3} + \underline{r}_t^S m_{t+1,\tau,3} - \overline{r}_t^S n_{t+1,\tau,3} = 0 \quad t = (\tau + 1) \dots T - 1 \\
& -y_{T,\tau,3} - p_{T,\tau,3} + q_{T,\tau,3} - m_{T,\tau,3} + n_{T,\tau,3} = 0
\end{aligned}$$

$$\begin{aligned}
& -\frac{K}{S_0} z_{\tau,4} - \sum_{t=1}^T \underline{R}_t^S p_{t,\tau,4} + \sum_{t=1}^T \overline{R}_t^S q_{t,\tau,4} + \alpha_0^B R_T^B \leq \epsilon \\
& -p_{1,\tau,4} + q_{1,\tau,4} + \underline{r}_1^S m_{2,\tau,4} - \overline{r}_1^S n_{2,\tau,4} = -\beta_0 \frac{R_T^B}{R_1^B} \\
& -p_{t,\tau,4} + q_{t,\tau,4} - m_{t,\tau,4} + n_{t,\tau,4} + \underline{r}_t^S m_{t+1,\tau,4} - \overline{r}_t^S n_{t+1,\tau,4} = -\beta_{t-1} \frac{R_T^B}{R_t^B} \quad t = 2 \dots (\tau - 1) \\
& -z_{\tau,4} - p_{\tau,\tau,4} + q_{\tau,\tau,4} - m_{\tau,\tau,4} + n_{\tau,\tau,4} + \underline{r}_\tau^S m_{\tau+1,\tau,4} - \overline{r}_\tau^S n_{\tau+1,\tau,4} = (\alpha_0^S + \sum_{t=1}^T \beta_{t-1}) - \beta_{\tau-1} \\
& -p_{t,\tau,4} + q_{t,\tau,4} - m_{t,\tau,4} + n_{t,\tau,4} + \underline{r}_t^S m_{t+1,\tau,4} - \overline{r}_t^S n_{t+1,\tau,4} = 0 \quad t = (\tau + 1) \dots T - 1 \\
& -p_{T,\tau,4} + q_{T,\tau,4} - m_{T,\tau,4} + n_{T,\tau,4} = 0
\end{aligned}$$

$$\begin{aligned}
& p_{t,i,j} \leq 0, q_{t,i,j} \geq 0, m_{t,i,j} \leq 0, n_{t,i,j} \geq 0, \quad \forall t = 1 \dots T, \quad \forall i = 1 \dots 4, \quad \forall j = 1 \dots 4 \\
& z_1 \leq 0, z_2 \leq 0, z_3 \geq 0, z_4 \geq 0
\end{aligned}$$

Note that when $\tau = 1$ or T the formulation is slightly different. The number of

decision variables in (ACO-LP) is $64T^2 + 16T$. The number of constraints is $4T^2 + 4T$.

Chapter 6

Options on multiple assets

Suppose the option is based on M securities instead of only one security, in order to find the optimal ϵ – *arbitrage* portfolio for this option, we define the following:

Decision variables:

- x_t^m : amount invested in asset m during the period $[t, t + 1]$;
- y_t^m : amount added to asset m from riskless asset 0 at the starting of the period $[t, t + 1]$, can be positive or negative;
- p_0 : amount we start out with and thus the price of the option and it is calculated by $p_0 = \sum_{m=0}^M x_0^m$.

Data:

- \tilde{r}_t^m : rate of return from asset m (excluding the riskless asset) during the period $[t, t + 1]$. It is subject to uncertainty;
- r_t^0 : rate of return from riskless asset during the period $[t, t + 1]$, which assumes to be constant;
- \mathbf{U} : uncertainty set in which the returns lie - will be constructed based on historical data, risk aptitudes of the players etc. It will be explained in Section 6.1;
- $P_f(\mathbf{S}, \mathbf{K})$: payoff function of the option which depends on the prices of the underlying assets and the strike price.

Similarly to single asset case in Chapter 2, we formulate (MAP-1) as follows:

Options on Multiple Assets Generic Formulation (MAP-1)

$$\begin{aligned}
 \min_{\{x_t^m, y_t^m\}} \quad & \max_{\{\tilde{r}_t^m\} \in \mathbf{U}} |P_f(\mathbf{S}, \mathbf{K}) - W_T| \\
 \text{s.t.} \quad & W_T = \sum_{m=0}^M x_T^m \\
 & x_t^m = (1 + \tilde{r}_{t-1}^m) (x_{t-1}^m + y_{t-1}^m) \quad \forall t = 1 \dots T, \forall m = 1 \dots M \\
 & x_t^0 = (1 + \tilde{r}_{t-1}^0) \left\{ x_{t-1}^0 - \sum_{m=1}^M y_{t-1}^m \right\} \quad \forall t = 1 \dots T
 \end{aligned}$$

The price is then $p_0 = \sum_{m=0}^M x_0^m$. We define cumulative returns as in the single asset case. The following transformation of variables allows us to reduce the number of constraints with uncertain coefficients :

- $\alpha_t^m = \frac{x_t^m}{R_t^m}, \forall m = 0 \dots M, t = 0 \dots T$
- $\beta_t^m = \frac{y_t^m}{R_t^m}, \forall m = 1 \dots M, t = 1 \dots T$

where cumulative returns are defined as:

- $\tilde{R}_t^m = \prod_{i=0}^{t-1} (1 + \tilde{r}_i^m), \forall m = 1 \dots M, t = 1 \dots T$
- $R_t^0 = \prod_{i=0}^{t-1} (1 + r_i^0), t = 1 \dots T$
- $R_0^m = 1, \forall m = 0 \dots M$

These transformations lead to (MAP-2):

Options on Multiple Assets Generic Formulation (MAP-2)

$$\begin{aligned}
 \min_{\{\alpha_t^m, \beta_t^m\}} \quad & \max_{\{\tilde{R}_t^m\} \in \mathbf{U}} \left| P_f(\mathbf{S}, \mathbf{K})^+ - \sum_{m=0}^M \alpha_T^m \cdot \tilde{R}_T^m \right| \\
 \text{s.t.} \quad & \alpha_t^m = \alpha_{t-1}^m + \beta_{t-1}^m \quad \forall m = 1 \dots M, t = 1 \dots T \\
 & \alpha_t^0 = \alpha_{t-1}^0 - \sum_{m=1}^M \beta_{t-1}^m \frac{\tilde{R}_t^m}{R_t^0} \quad \forall t = 1 \dots T
 \end{aligned}$$

Substituting all intermediate α_t^m 's, MAP-2 is equivalent as MAP-3 shown below:

Options on Multiple Assets Generic Formulation (MAP-3)

$$\min_{\{\alpha_0^m, \beta_t^m\}} \max_{\{\tilde{R}_t^m\} \in \mathbf{U}} \left| P_f(\mathbf{S}, \mathbf{K})^+ - \sum_{m=1}^M \left(\alpha_0^m + \sum_{t=1}^T \beta_t^m \right) \cdot \tilde{R}_T^m + \left(\alpha_0^0 - \sum_{t=1}^T \sum_{m=1}^M \beta_t^m \frac{\tilde{R}_t^m}{R_t^0} \right) \cdot R_T^0 \right|$$

6.1 Pricing Dynamics for Multiple Assets

For the price dynamics of multiple assets, it is desirable to incorporate information about the variability and the correlations of asset returns. Let \underline{r}_t^m be the lower bound for single period return at time t , \overline{r}_t^m the upper bound, $\underline{R}_t^m, \overline{R}_t^m$ be the lower bound and upper bound for cumulative returns, Σ be the covariance matrix of the single period returns. Given that Σ is symmetric and positive definite, it has a Cholesky decomposition and we can define matrix $\mathbf{C} = (\Sigma)^{-1/2}$. $\tilde{\mathbf{R}}_1$ is the vector of first period's cumulative return (same as single period return). It is a $M \times 1$ vector with \tilde{R}_1^m as its entries. $\check{\mathbf{R}}_1$ is the vectors of means of single period returns. We define the uncertainty sets below (Bertsimas and Pachamanova [5]):

$$\|\mathbf{C}(\tilde{\mathbf{R}}_1 - \check{\mathbf{R}}_1)\| \leq \Gamma \quad (6.1)$$

$$\underline{r}_t^m \tilde{R}_{t-1}^m \leq \tilde{R}_t^m \leq \overline{r}_t^m \tilde{R}_{t-1}^m, \quad \forall m = 1..M, t = 2 \dots T \quad (6.2)$$

$$\underline{R}_t^m \leq \tilde{R}_t^m \leq \overline{R}_t^m, \quad \forall m = 1..M, t = 2 \dots T \quad (6.3)$$

$\|\mathbf{x}\|$ is a general norm of a factor, depending on the modeler's preference, it can be L_1, L_2, L_∞ , or D-norm (Bertsimas et. al. [6]).

L1-norm

If L_1 is used, $\|\mathbf{C}(\tilde{\mathbf{R}}_1 - \check{\mathbf{R}}_1)\|_1 \leq \Gamma$ is equivalent to:

$$\begin{aligned} \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - s^m &\leq \sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\ -\sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - s^m &\leq -\sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\ \sum_{m=1}^M s^m &\leq \Gamma \end{aligned}$$

where C_{ij} corresponds to the $[i][j]^{th}$ entry in the matrix \mathbf{C} .

L2-norm

If L_2 is used, $\|\mathbf{C}(\tilde{\mathbf{R}}_1 - \check{\mathbf{R}}_1)\|_2 \leq \Gamma$ is equivalent to:

$$\begin{aligned} \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - s^m &\leq \sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\ -\sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - s^m &\leq -\sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\ \sum_{m=1}^M (s^m)^2 &\leq (\Gamma)^2 \end{aligned}$$

Infinity-norm

If L_∞ is used, $\|\mathbf{C}(\tilde{\mathbf{R}}_1 - \check{\mathbf{R}}_1)\|_\infty \leq \Gamma$ is equivalent to:

$$\begin{aligned} \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - s^m &\leq \sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\ -\sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - s^m &\leq -\sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\ s^m &\leq \Gamma \quad \forall m = 1 \dots M \end{aligned}$$

D-norm

If D-norm is used (Bertsimas et. al 2004), $\|\mathbf{C}(\tilde{\mathbf{R}}_1 - \check{\mathbf{R}}_1)\|_d \leq \Gamma_t$ with $d \in [1, n]$ and for $\mathbf{y} \in \mathfrak{R}^{n \times 1}$ where

$$\|\mathbf{y}\|_d = \max_{\{S \cup \{t\} | S \subseteq N, |S| \leq \lfloor d \rfloor, t \in N \setminus S\}} \left\{ \sum_{j \in S} |y_j| + (d - \lfloor d \rfloor) |y_t| \right\}$$

The norm $\|\mathbf{y}\|_d$ can be written as:

$$\begin{aligned}
\|y\|_d = & \max \sum_{j=1}^n u_j y_j \\
\text{s.t.} & \sum_{j=1}^n u_j \leq d, \quad 0 \leq u_j \leq 1 \\
& = \min d \cdot r + \sum_{j=1}^n t_j \\
\text{s.t.} & r + t_j \geq |y_j|, \quad t_j \geq 0, \quad j = 1..n, \quad r \geq 0
\end{aligned}$$

The second equality follows by linear programming strong duality. Thus, we can write $\|C(\tilde{\mathbf{R}}_1 - \check{\mathbf{R}}_1)\|_d \leq \Gamma$ as

$$\begin{aligned}
r + t_m & \geq |C(\tilde{\mathbf{R}}_1 - \check{\mathbf{R}}_1)| \\
d \cdot r + \sum_{m=1}^M t_m & \leq \Gamma
\end{aligned}$$

Rearrange the terms and get rid of the vector representation:

$$\begin{aligned}
\sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - r - t_j & \leq \sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\
-\sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - r - t_j & \leq -\sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\
d \cdot r + \sum_{m=1}^M t_m & \leq \Gamma
\end{aligned} \tag{6.4}$$

Notice that \tilde{R}_1^i is the first period's cumulative return (same as single period return) for asset i and it is subject to uncertainty. \check{R}_1^i is the mean of \tilde{R}_1^i .

6.2 Payoff Function

Depending on the nature of the options, the payoff function $P_f(\mathbf{S}, \mathbf{K})^+$ can assume many forms. For example, for European style index call options, the payoff is $(\sum_{m=1}^M w_m \cdot S_0^m \cdot R_T^m - K)^+$ where w_m is the weight (e.g. percentage market volume) of asset m. $S_0^m \cdot R_T^m$ gives the value of asset m at time expiry date T. Also, one could also look at the highest return out of the M assets at the expiry date. In this case, it would give us a payoff function $(\max\{S_0^m \cdot R_T^m\} - K)^+$.

In this section, we will use payoff function of $(\sum_{m=1}^M w_m \cdot S_0^m \cdot R_T^m - K)^+$ to illustrate our model:

European Index Call Options Formulation (EICO-1)

$$\left| \left(\sum_{m=1}^M w_m \cdot S_0^m \cdot R_T^m - K \right)^+ - \sum_{m=1}^M \left(\alpha_0^m + \sum_{t=1}^T \beta_t^m \right) \cdot \tilde{R}_T^m + \left(\alpha_0^0 - \sum_{t=1}^T \sum_{m=1}^M \beta_t^m \frac{\tilde{R}_t^m}{R_t^0} \right) \cdot R_T^0 \right|$$

6.3 A Linear Optimization Re-formulation

In this section, we present the linear re-formulation of (EICO-1) under uncertainty sets (6.1)-(6.3). A linear re-formulation exists for L_1 , L_∞ , and D-norm. We use D-norm in our study because of its proximity to L_2 (Euclidean norm) and its probabilistic guarantee (Bertsimas and Sim [4]). The linear optimization equivalent of other norms can be formulated in similar ways.

(EICO-1) is equivalent to (EICO-2) listed below:

European Index Call Options Formulation (EICO-2)

$$\begin{aligned}
& \min_{\{\alpha_0^m, \beta_t^m\}} \max_{\{\tilde{R}_t^m\} \in \mathbf{U}} \epsilon \\
& \left(\sum_{m=1}^M w_m \cdot S_0^m \cdot R_T^m - K \right) - \left(\sum_{m=1}^M \left(\left(\alpha_0^m + \sum_{t=1}^T \beta_t^m \right) \cdot R_T^m \right) + \left(\alpha_0^0 - \sum_{t=1}^T \sum_{m=1}^M \beta_t^m \frac{R_t^m}{R_t^0} \right) \cdot R_T^0 \right) \leq \epsilon \\
& - \left(\left(\sum_{m=1}^M w_m \cdot S_0^m \cdot R_T^m - K \right) - \left(\sum_{m=1}^M \left(\alpha_0^m + \sum_{t=1}^T \beta_t^m \right) \cdot R_T^m + \left(\alpha_0^0 - \sum_{t=1}^T \sum_{m=1}^M \beta_t^m \frac{R_t^m}{R_t^0} \right) \cdot R_T^0 \right) \right) \leq \epsilon \\
& - \left(\sum_{m=1}^M \left(\alpha_0^m + \sum_{t=1}^T \beta_t^m \right) \cdot R_T^m + \left(\alpha_0^0 - \sum_{t=1}^T \sum_{m=1}^M \beta_t^m \frac{R_t^m}{R_t^0} \right) \cdot R_T^0 \right) \leq \epsilon \\
& - \left(- \left(\sum_{m=1}^M \left(\alpha_0^m + \sum_{t=1}^T \beta_t^m \right) \cdot R_T^m + \left(\alpha_0^0 - \sum_{t=1}^T \sum_{m=1}^M \beta_t^m \frac{R_t^m}{R_t^0} \right) \cdot R_T^0 \right) \right) \leq \epsilon
\end{aligned}$$

Replacing (6.1) with (6.4) and rearranging (6.2)-(6.3), we obtain the following uncertainty sets :

$$\begin{aligned}
& \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - r - t_m \leq \sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\
& - \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i - r - t_m \leq - \sum_{i=1}^M C_{m,i} \cdot \check{R}_1^i \quad \forall m = 1 \dots M \\
& d \cdot r + \sum_{m=1}^M t_m \leq \Gamma \\
& -\tilde{R}_t^m + r_t^S \tilde{R}_{t-1}^m \leq 0 \quad \forall m = 1 \dots M, t = 1 \dots T \\
& \tilde{R}_t^m - r_t^S \tilde{R}_{t-1}^m \leq 0 \quad \forall m = 1 \dots M, t = 1 \dots T \\
& -\tilde{R}_t^m \leq -\check{R}_t^m, \quad \forall m = 1 \dots M, t = 2 \dots T \\
& \tilde{R}_t^m \leq \check{R}_t^m, \quad \forall m = 1 \dots M, t = 2 \dots T
\end{aligned} \tag{6.5}$$

We introduce $\mathbf{p}, \mathbf{q}, \mathbf{u}, \mathbf{m}, \mathbf{n}, \mathbf{a}, \mathbf{b}$ as vectors of dual variables associated with several series of constraints specified in (6.5) respectively. The variable z is the dual variable with $\sum_{m=1}^M w_m \cdot S_0^m \cdot R_T^m \geq K$ or $\sum_{m=1}^M w_m \cdot S_0^m \cdot R_T^m \leq K$ depending on the condition specified in that constraint in (EICO-2). We apply Theorem 3.1 to each constraint in (EICO-2), treating \tilde{R}_t^S as random data A and α_0^m, β_t^m as decision variables \mathbf{x} which assume certain given value. We yield linear optimization model (EICO-LP) below:

European Index Call Options- Linear Formulation (EICO-LP)

$$\begin{aligned}
& \mathbf{max} \ \epsilon \\
& \text{s.t.} \ \epsilon \leq \sum_{m=1}^M \left((p_{m,1} - q_{m,1}) \cdot \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i \right) - \sum_{m=1}^M \sum_{t=1}^T \underline{R}_t^S a_{t,1}^m + \sum_{t=1}^T \overline{R}_t^S b_{t,1}^m \\
& \quad + u_1 \Gamma + z_1 K - R_T^0 \alpha_T^0 - K \\
& -a_{1,1}^i + b_{1,1}^i + \sum_{m=1}^M C_{m,i} \cdot p_{m,1} - \sum_{m=1}^M C_{m,i} \cdot q_{m,1} + r_1^S m_{2,1}^i - \overline{r}_1^S n_{2,1}^i = \beta_t^i \frac{R_T^0}{R_1^0} \quad \forall i = 1 \dots M \\
& -a_{t,1}^i + b_{t,1}^i - m_{t,1}^i + n_{t,1}^i + r_1^S m_{t+1,1}^i - \overline{r}_1^S n_{t+1,1}^i = \beta_t^i \frac{R_T^0}{R_1^0} \quad \forall i = 1 \dots M, t = 2 \dots (T-1) \\
& -a_{T,1}^i + b_{T,1}^i + w_i \cdot S_0^i \cdot z_1 - m_{T,1}^i + n_{T,1}^i = w_i \cdot S_0^i - \left((\alpha_0^i + \sum_{t=1}^T \beta_t^i) - \beta_T^i \right) \quad \forall i = 1 \dots M \\
& -p_{i,1} - q_{i,1} + u_1 \geq 0 \quad i = 1 \dots M \\
& -\sum_{i=1}^M p_{i,1} - \sum_{i=1}^M q_{i,1} + d \cdot u_1 \geq 0 \\
\\
& \epsilon \leq \sum_{m=1}^M \left((p_{m,2} - q_{m,2}) \cdot \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i \right) - \sum_{m=1}^M \sum_{t=1}^T \underline{R}_t^S a_{t,2}^m + \sum_{t=1}^T \overline{R}_t^S b_{t,2}^m \\
& \quad + u_2 \Gamma + z_2 K + R_T^0 \alpha_T^0 + K \\
& -a_{1,2}^i + b_{1,2}^i + \sum_{m=1}^M C_{m,i} \cdot p_{m,2} - \sum_{m=1}^M C_{m,i} \cdot q_{m,2} + r_1^S m_{2,2}^i - \overline{r}_1^S n_{2,2}^i = -\beta_t^i \frac{R_T^0}{R_1^0} \quad \forall i = 1 \dots M \\
& -a_{t,2}^i + b_{t,2}^i - m_{t,2}^i + n_{t,2}^i + r_1^S m_{t+1,2}^i - \overline{r}_1^S n_{t+1,2}^i = -\beta_t^i \frac{R_T^0}{R_1^0} \quad \forall i = 1 \dots M, t = 2 \dots (T-1) \\
& -a_{T,2}^i + b_{T,2}^i + w_i \cdot S_0^i \cdot z_2 - m_{T,2}^i + n_{T,2}^i = -\left(w_i \cdot S_0^i - \left((\alpha_0^i + \sum_{t=1}^T \beta_t^i) - \beta_T^i \right) \right) \quad \forall i = 1 \dots M \\
& -p_{i,2} - q_{i,2} + u_2 \geq 0 \quad i = 1 \dots M \\
& -\sum_{i=1}^M p_{i,2} - \sum_{i=1}^M q_{i,2} + d \cdot u_2 \geq 0 \\
\\
& \omega \leq \sum_{m=1}^M \left((p_{m,3} - q_{m,3}) \cdot \sum_{i=1}^M C_{m,i} \cdot \tilde{R}_1^i \right) - \sum_{m=1}^M \sum_{t=1}^T \underline{R}_t^S a_{t,3}^m + \sum_{t=1}^T \overline{R}_t^S b_{t,3}^m \\
& \quad + u_3 \Gamma + z_3 K - R_T^0 \alpha_T^0 \\
& -a_{1,3}^i + b_{1,3}^i + \sum_{m=1}^M C_{m,i} \cdot p_{m,3} - \sum_{m=1}^M C_{m,i} \cdot q_{m,3} + r_1^S m_{2,3}^i - \overline{r}_1^S n_{2,3}^i = \beta_t^i \frac{R_T^0}{R_1^0} \quad \forall i = 1 \dots M \\
& -a_{t,3}^i + b_{t,3}^i - m_{t,3}^i + n_{t,3}^i + r_1^S m_{t+1,3}^i - \overline{r}_1^S n_{t+1,3}^i = \beta_t^i \frac{R_T^0}{R_1^0} \quad \forall i = 1 \dots M, t = 2 \dots (T-1) \\
& -a_{T,3}^i + b_{T,3}^i + w_i \cdot S_0^i \cdot z_3 - m_{T,3}^i + n_{T,3}^i = -\left((\alpha_0^i + \sum_{t=1}^T \beta_t^i) - \beta_T^i \right) \quad \forall i = 1 \dots M \\
& -p_{i,3} - q_{i,3} + u_3 \geq 0 \quad i = 1 \dots M \\
& -\sum_{i=1}^M p_{i,3} - \sum_{i=1}^M q_{i,3} + d \cdot u_3 \geq 0
\end{aligned}$$

$$\begin{aligned}
\omega &\leq \sum_{m=1}^M \left((p_{m,4} - q_{m,4}) \cdot \sum_{i=1}^M C_{m,i} \cdot \bar{R}_1^i \right) - \sum_{m=1}^M \sum_{t=1}^T \underline{R}_t^S a_{t,4}^m + \sum_{t=1}^T \overline{R}_t^S b_{t,4}^m \\
&\quad + u_4 \Gamma + z_4 K + R_T^0 \alpha_T^0 \\
-a_{1,4}^i + b_{1,4}^i + \sum_{m=1}^M C_{m,i} \cdot p_{m,4} - \sum_{m=1}^M C_{m,i} \cdot q_{m,4} + \underline{r}_1^S m_{2,4}^i - \overline{r}_1^S n_{2,4}^i &= -\beta_1^i \frac{R_T^0}{R_1^0} \quad \forall i = 1 \dots M \\
-a_{t,4}^i + b_{t,4}^i - m_{t,4}^i + n_{t,4}^i + \underline{r}_1^S m_{t+1,4}^i - \overline{r}_t^S n_{t+1,4}^i &= -\beta_t^i \frac{R_T^0}{R_t^0} \quad \forall i = 1..M, t = 2 \dots (T-1) \\
-a_{T,4}^i + b_{T,4}^i + w_i \cdot S_0^i \cdot z_4 - m_{T,4}^i + n_{T,4}^i &= \left(\alpha_0^i + \sum_{t=1}^T \beta_t^i \right) - \beta_T^i \quad \forall i = 1 \dots M \\
-p_{i,4} - q_{i,4} + u_4 &\geq 0 \quad i = 1..M \\
-\sum_{i=1}^M p_{i,4} - \sum_{i=1}^M q_{i,4} + d \cdot u_4 &\geq 0
\end{aligned}$$

$$\mathbf{p}, \mathbf{q}, \mathbf{u}, \mathbf{m}, \mathbf{n}, \mathbf{a}, \mathbf{b} \geq 0, z_j \geq 0, j = 1, 2, z_j \leq 0, j = 3, 4$$

The number of variables in (EICO-LP) is $16MT + 8M + 8$. The number of constraints is $4MT + 4M + 8$.

Chapter 7

Computational Results

In this section, we detail the results obtained in the experiments we performed to compare our approach with previous approaches and with market observables. We perform the following experiments:

- Experiment 1: Comparison with actual market prices for European Call Options.
- Experiment 2: Comparison with prices of Asian Call Options obtained from Monte Carlo Simulation.
- Experiment 3: Comparison with prices of Lookback Call Options obtained from Monte Carlo Simulation.
- Experiment 4: Comparison with actual market prices for American put Options.
- Experiment 5: Comparison with actual market prices for Index call options.

Experiment settings:

For Experiment 1-3:

- The underlying security is MSFT stock.
- The number of periods $T = 18$ weeks.
- The initial price of underlying security $S_0 = 21.4$.
- Strike price of options K : ranges from 2.5 to 30.

For Experiment 4:

- The underlying security is MSFT stock.
- The number of periods $T = 25$ weeks.
- The initial price of underlying security $S_0 = 24.8$.
- Strike price of options K : ranges from 7.5 to 50.

The parameters as inputs to Experiment 1-4 are:

- μ : mean of single period returns of underlying asset.
- σ : standard deviation of single period returns of underlying asset.
- μ_{log} : mean of the logs of single period returns of underlying asset.
- σ_{log} : standard deviation of the logs of single period returns of underlying asset.
- Γ : level of risk aversion (we use $\Gamma_t = \Gamma$ in uncertainty set (2.3)).

With these 5 parameters, we could construct uncertainty sets (2.2) - (2.3).

For Experiment 5:

- The index we used is 1/100 Dow Jones Industrial Average
- Number of periods $T = 8$ weeks.
- Initial quote for Dow-Jones index $S_0 = 90.8$.
- Strike prices of options K : ranges from 74 to 105.

The parameters as inputs to Experiment 5 are:

- μ^i : mean of single period returns of asset i .
- σ^i : standard deviation of single period returns of asset i .
- μ_{log}^i : mean of the logs of single period returns of asset i .
- σ_{log}^i : standard deviation of the logs of single period returns of asset i .
- Σ : the covariance matrix of the single period returns among assets, $C = \Sigma^{-\frac{1}{2}}$.
- Γ : level of risk aversion (we use $\Gamma_t = \Gamma$ in uncertainty set (6.5)).

With these 6 parameters, we could construct uncertainty sets (6.5) for options based on multiple assets.

Implied Gamma vs Implied Volatility

We define Implied Gamma ($\Gamma_{implied}$) as the value of the parameter Γ that is to be used as input to our model so that the price calculated from the model matches the market price, or price from simulation if market price is not obtainable. Implied volatility is defined as the volatility to be substituted into the Black-Scholes formula to get the market price. The observation of implied volatility smile shows that for different strike prices, the volatility to be input to the pricing model is different. However, volatility is a characteristic of the asset returns that should not be affected by the strike price. In our approach, $\Gamma_{implied}$ is a measure of investor's risk appetite. When the strike price is far away from the current stock price, the investor would be more risk averse, thus he wants to use a higher Γ and wants the price of the option to reflect this fact.

Γ vs $\frac{K}{S_0}$ - quadratic dependence

We propose and verify empirically that a quadratic variation of $\Gamma_{implied}$ with $\frac{K}{S_0}$ would be adequate to characterize the risk aversion of investor towards different strike prices. We use the following function to describe the relationship:

$$\Gamma\left(\frac{K}{S_0}\right) = a_0 + a_1\frac{K}{S_0} + a_2\left(\frac{K}{S_0}\right)^2, \quad a_2 \geq 0 \quad (7.1)$$

$\frac{K}{S_0}$ captures the distance between the strike and the spot price. We use a quadratic regression model to compute the coefficients $\{a_0, a_1, a_2\}$ so that the prices given using these Γ 's match the price from the model we want to compare. We can then use this quadratic model $\Gamma\left(\frac{K}{S_0}\right)$ to calculate Γ and input it to yield the price for options with other strike prices.

Sampling

After we obtain the various strike prices K which are observed in the market, we divide them into 2 groups: in-sample and out-of-sample. If it is in-sample, we find $\Gamma_{implied}$ that matches our model price to the market price for this specific K . We then input these K 's and $\Gamma_{implied}$'s to a regression model to find the coefficients of (7.1). After that we use $\Gamma\left(\frac{K}{S_0}\right)$ to find the prices for out-of-sample options.

7.1 Experiment 1 : Comparison with market prices for an European call option

In this experiment, we aim to adjust our “risk aversion” parameter Γ to match the market prices of the MSFT 18 weeks call options. Note that MSFT options are American style, however, since they do not pay dividends, the American call price is equal to its equivalent European call price. We solve (ECO-LP) with different Γ that ranges from 0 to with a step size of 0.01, and find the $\Gamma_{implied}$ so that the model price is the closest to the market prices. We observe from Figure 7.1 that $\Gamma_{implied}$ has a quadratic relationship with the $\frac{K}{S_0}$. Using $\Gamma\left(\frac{K}{S_0}\right)$ to calculate prices for out-of-sample options, we plot the price from our model and the market price- Figure 7.2 and it shows a close matching for both in-sample options and out-of-sample options.

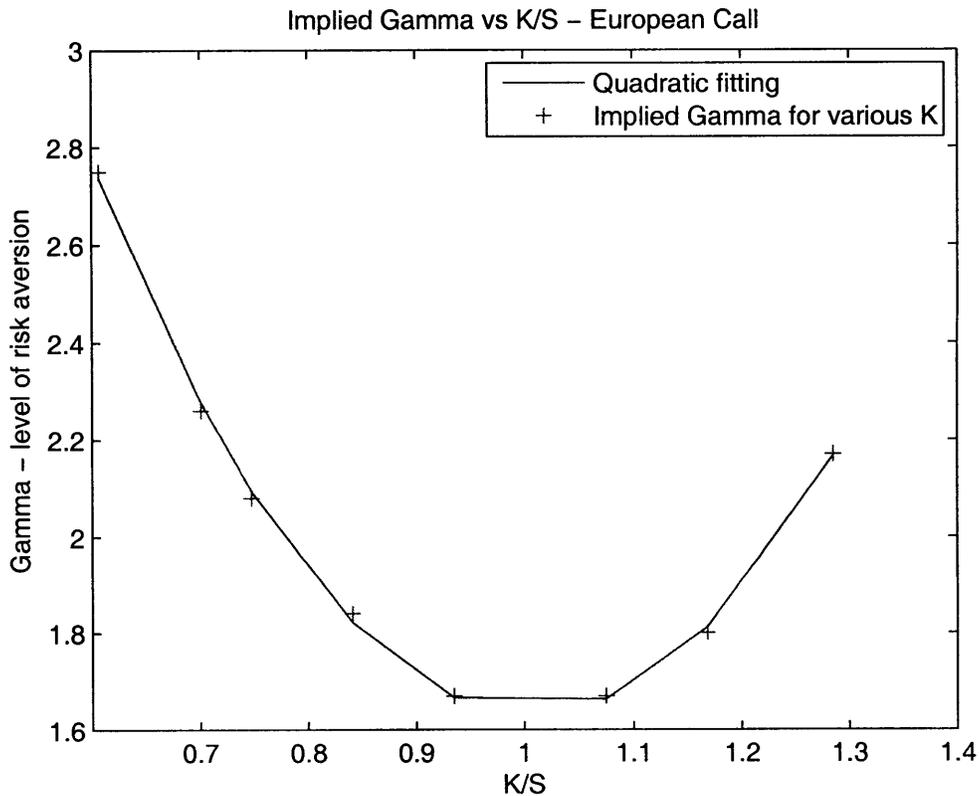


Figure 7.1: $\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for a European Call option.

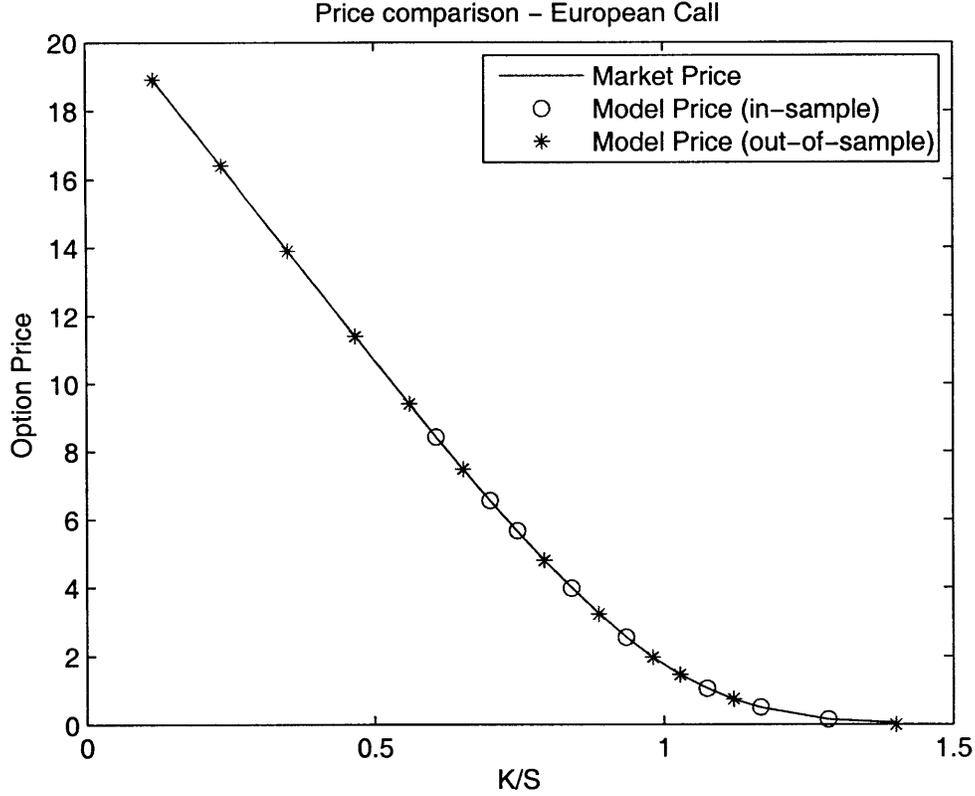


Figure 7.2: Comparison of Model Price and Market Price for a European Call option.

From Table 7.1, it is observed that when $\frac{K}{S_0}$ is very small- that is “deeply in the money”, the Γ that matches our model price to market price is not a unique value but a range. For example, in Table 7.1 row 4, as long as Γ is selected to fall between 0 - 3.92, the price given by the model is always 11.4. And the price matches the one given by Black-Scholes and the Market Price. The following paragraph explains mathematically why $\Gamma_{implied}$ can be an interval instead of a single value.

\tilde{R}_T^S is bounded by $[\underline{R}_T^S, \overline{R}_T^S]$, where $\underline{R}_T^S = e^{T\mu_{log} - \Gamma\sqrt{T}\sigma_{log}}$, $\overline{R}_T^S = e^{T\mu_{log} - \Gamma\sqrt{T}\sigma_{log}}$. Thus as long as $\Gamma \leq \frac{T\mu_{log} - \ln(\frac{K}{S_0})}{\sqrt{T}\sigma_{log}}$, we have $\underline{R}_T^S \geq \frac{K}{S_0}$. This means $\tilde{R}_t^S \geq \frac{K}{S_0}$ always holds and the pricing problem becomes:

$$\epsilon = \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} \left| \left(S_0 \tilde{R}_T^S - K \right) - \left(\alpha_0^S + \sum_{t=1}^T \beta_{t-1} \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_{t-1} \frac{R_t^B}{R_t^B} \tilde{R}_t^S \right|$$

One can make the arbitrage $\epsilon=0$ by the following 2 steps:

1) Let $\beta_t = 0, \forall t$, which eliminate uncertainty for all $t \neq T$, then:

$$\epsilon = \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} \left| S_0 \tilde{R}_T^S - K - \alpha_0^S \tilde{R}_T^S - \alpha_0^B R_T^B \right|$$

2) Make $\alpha_0^S = S_0$ to eliminate uncertainty from $t = T$, then:

$$\epsilon = \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} \left| -K - \alpha_0^B R_T^B \right|$$

In order to make $\epsilon = 0$, we assign $\alpha_0^B = -\frac{K}{R_T^B}$, thus:

$$price = S_0 - \frac{K}{R_T^B}. \text{ And this price is constant when } 0 \leq \Gamma \leq \frac{T\mu_{log} - \ln(\frac{K}{S_0})}{\sqrt{T}\sigma_{log}}.$$

This corresponds to a trading strategy of borrowing $\frac{K}{R_T^B}$ in the riskless asset, buying S_0 in the underlying stock and do nothing for rebalancing. At the end of the day, the investor will get same value from this portfolio and the call option.

No.	T	K/S	$\Gamma_{implied}$	Mkt Price	Model Price	Error	Sampling
1	18	0.117	[0, 10]	18.9	18.9	0	out
2	18	0.234	[0, 7.15]	16.4	16.4	0	out
3	18	0.35	[0, 5.26]	13.9	13.9	0	out
4	18	0.467	[0, 3.92]	11.4	11.4	0	out
5	18	0.561	[0, 3.07]	9.4	9.4	0	out
6	18	0.607	2.75	8.425	8.42	-0.005	in
7	18	0.654	2.48	7.475	7.48	0.005	out
8	18	0.701	2.26	6.55	6.561	0.011	in
9	18	0.748	2.08	5.65	5.667	0.017	in
10	18	0.794	1.94	4.8	4.797	-0.003	out
11	18	0.841	1.84	4	3.984	-0.016	in
12	18	0.888	1.75	3.25	3.232	-0.018	out
13	18	0.935	1.67	2.56	2.556	-0.004	in
14	18	0.981	1.64	1.97	1.968	-0.002	out
15	18	1.028	1.64	1.47	1.462	-0.008	out
16	18	1.075	1.67	1.071	1.062	-0.009	in
17	18	1.121	1.71	0.735	0.749	0.014	out
18	18	1.168	1.8	0.5	0.512	0.012	in
19	18	1.285	2.17	0.155	0.152	-0.003	in
20	18	1.402	2.87	0.055	0	-0.055	out

Table 7.1: Finding the Implied Gamma, the Quadratic Relationship, and Error for a European Call option.

7.2 Experiment 2 : Comparison with prices of Asian call options obtained from Monte Carlo simulation

In this and the next section, we perform experiments to use our model to price exotic options - Asian and Lookback Options. We present below the results obtained when we compared the prices obtained by our model with those obtained by using a Monte Carlo pricer. We note that the Asian Option and Lookback options are a path-dependent options and Monte Carlo is one of the methods used to price them. While Monte Carlo pricing usually leads to accurate pricing, the computational complexity associated makes it a difficult method to use.

Our methodology, as discussed in previous sections, uses a linear optimization approach to price the Asian Option which leads to a highly efficient method. We observe from the experiments that, once we choose the appropriate Γ , our model produces fairly accurate results.

Figure 7.3 presents the $\Gamma_{implied}$ corresponding to an Asian Options whose price is calculated using Monte Carlo Simulation. We could see the $\Gamma_{implied}$ also has a quadratic relationship with the Strike price K . Using Γ that is calculated using the coefficients from the quadratic function (7.1), we plot the price from our model and the price obtained by Monte Carlo simulation- Figure 7.4 and it shows a close matching for “at the money” and “out of the money” options. However, for “in-the-money” options, the price given by our model is smaller than the Monte Carlo simulation.

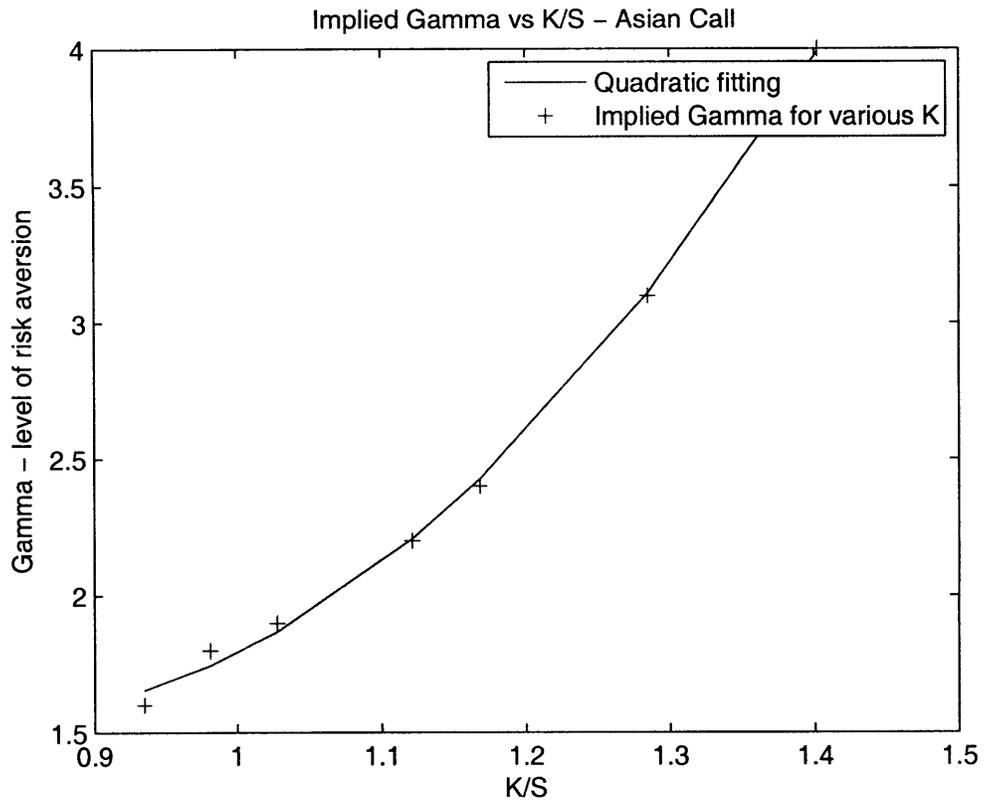


Figure 7.3: $\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for an Asian Call option.

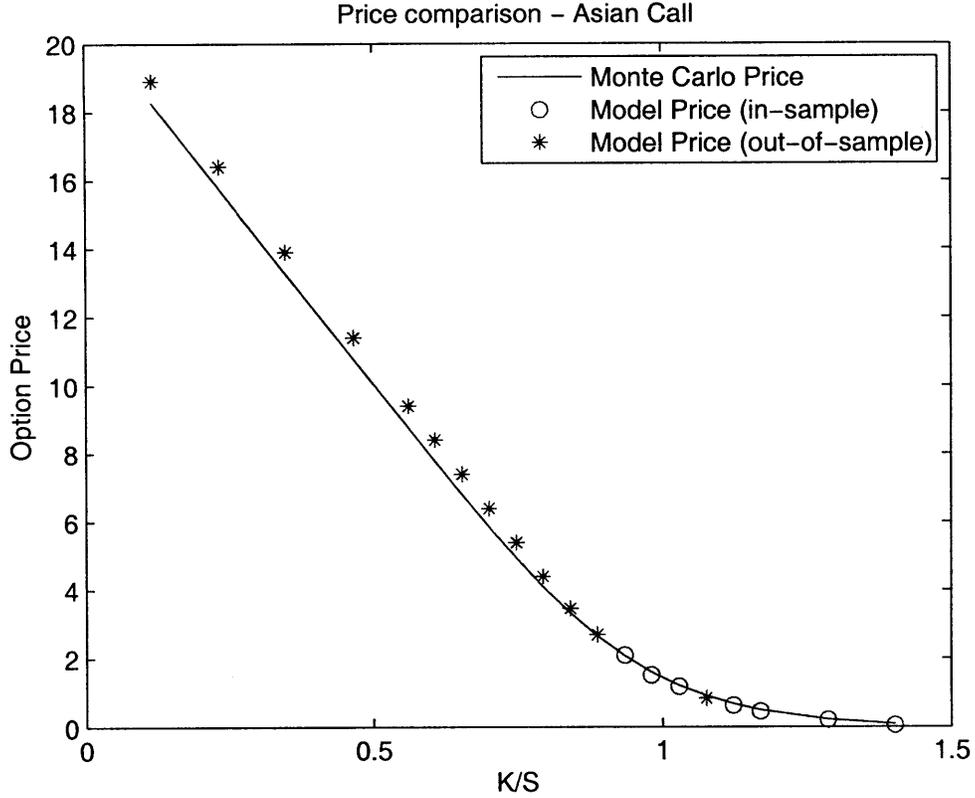


Figure 7.4: Comparison of Model Price and Monte Carlo Price for an Asian Call option.

From Table 7.2, it is observed that when $\frac{K}{S_0}$ is small- that is “in the money”, the Γ that matches our model price to market price is not a unique value but a range. Similarly as European option case in Section 7.1, we will explain mathematically why $\Gamma_{implied}$ can be an interval instead of a single value and why the prices generated by our model make sense.

Using the same argument in section 7.1, it is possible to find a range of Γ denoted by $[\underline{\Gamma}, \bar{\Gamma}]$ so that $S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} \geq \frac{K}{S_0}$ always holds and the pricing problem becomes:

$$\epsilon = \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathcal{U}} \left| S_0 \frac{\sum_{t=1}^T \tilde{R}_t^S}{T} - K - \left(\alpha_0^S + \sum_{t=1}^T \beta_t \right) \tilde{R}_T^S - \alpha_0^B R_T^B + \sum_{t=1}^T \beta_t \frac{R_T^B}{R_t^B} \tilde{R}_t^S \right|$$

One can make the arbitrage $\epsilon = 0$ by the following 2 steps:

- 1) Let $\beta_t = -\frac{S_0}{T}$, $\forall t \neq T$ and $\beta_T = 0$, which gives us:

$$\epsilon = \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} |S_0 R_T^S - K - \alpha_0^S R_T^S - \alpha_0^B R_T^B|$$

2) Make $\alpha_0^S = S_0$ to eliminate uncertainty from $t = T$, then:

$$\epsilon = \min_{\{\alpha_0^S, \alpha_0^B, \beta_t\}} \max_{\{\tilde{R}_t^S\} \in \mathbf{U}} |-K - \alpha_0^B R_T^B|$$

In order to make $\epsilon = 0$, $\alpha_0^B = -\frac{K}{R_T^B}$, thus: $price = S_0 - \frac{K}{R_T^B}$. Notice that the price is the same as the price for European Options but different from the prices of Asian options from Monte Carlo simulation.

This corresponds to a trading strategy of borrowing $\frac{K}{R_T^B}$ in riskless asset, buying S_0 in the underlying stock and selling $\frac{S_0}{T}$ value of stock at each period except for the last.

No.	T	K/S	$\Gamma_{implied}$	Monte Carlo Price	Model Price	Error	Sampling
1	18	0.117	[0, 10]	18.274	18.9	0.626	out
2	18	0.234	[0, 10]	15.774	16.4	0.626	out
3	18	0.35	[0, 7.7]	13.274	13.9	0.626	out
4	18	0.467	[0, 5.6]	10.774	11.4	0.626	out
5	18	0.561	[0, 4.3]	8.776	9.4	0.624	out
6	18	0.607	[0, 3.7]	7.783	8.4	0.617	out
7	18	0.654	[0, 3.2]	6.803	7.4	0.597	out
8	18	0.701	[0, 2.7]	5.85	6.4	0.55	out
9	18	0.748	[0, 2.2]	4.94	5.4	0.46	out
10	18	0.794	[0, 1.8]	4.093	4.4	0.307	out
11	18	0.841	[0, 1.4]	3.326	3.4597	0.133	out
12	18	0.888	1.5	2.653	2.692	0.039	out
13	18	0.935	1.6	2.078	2.0922	0.015	in
14	18	0.981	1.8	1.601	1.5141	-0.087	in
15	18	1.028	1.9	1.22	1.1817	-0.038	in
16	18	1.075	2.1	0.917	0.8414	-0.076	out
17	18	1.121	2.2	0.683	0.6304	-0.053	in
18	18	1.168	2.4	0.504	0.4558	-0.049	in
19	18	1.285	3.1	0.225	0.2103	-0.014	in
20	18	1.402	4	0.092	0.0556	-0.037	in

Table 7.2: Finding the Implied Gamma, the Quadratic Relationship, and Error for an Asian Call option.

7.3 Experiment 3 : Comparison with prices of Lookback call options obtained from Monte Carlo simulation

Similarly as Asian options, we find $\Gamma_{implied}$ that matches the price from our model to Monte Carlo price for Lookback call options (see Figure 7.5). We observe from Figure 7.5 that the $\Gamma_{implied}$ does not have a nice single quadratic relationship with K over the entire span of K. When $\frac{K}{S_0}$ is very small- that is “deeply in the money”, the implied Γ is constant for different K’s ($\frac{K}{S_0}$ between 0 - 0.5). This is because for “deeply in the money” situation, the value of the option is always $S_0 \cdot R_{max} - K$. Monte-Carlo simulation assumes the return distribution is invariant with respect to K. This gives $S_0 \cdot R_{max}$ constant and the price has a linear relationship with K. In our model, R_{max} is a function of Γ , thus if we are finding the best Γ to match the Monte Carlo prices, the Γ will be constant for small K’s (“in the money” situation).

We can also observe that as K becomes larger, the variability factor- that whether the option will assume $S_0 R_{max} - K$ or 0, starts to play a part in the price of the option, and the curve starts to become quadratic.

Thus we can model the relationship of optimal Γ and K as the minimum of constant function and a quadratic function:

$$\Gamma(K) = \min \left\{ \begin{array}{l} \Gamma_{constant} \\ a_0 + a_1 \frac{K}{S_0} + a_2 \left(\frac{K}{S_0} \right)^2 \end{array} \right\}.$$

Figure 7.6 shows the comparison of our model prices and the prices by Monte Carlo simulation, the matching is quite close except for 1 or 2 outliers.

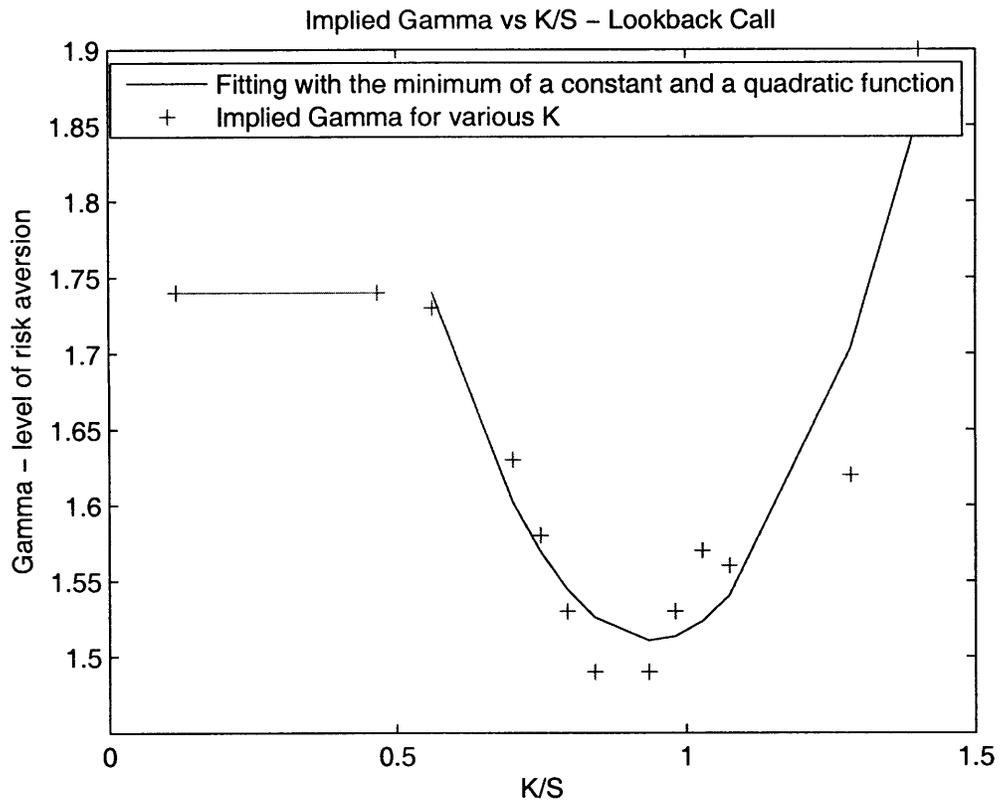


Figure 7.5: $\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for a Lookback Call option.

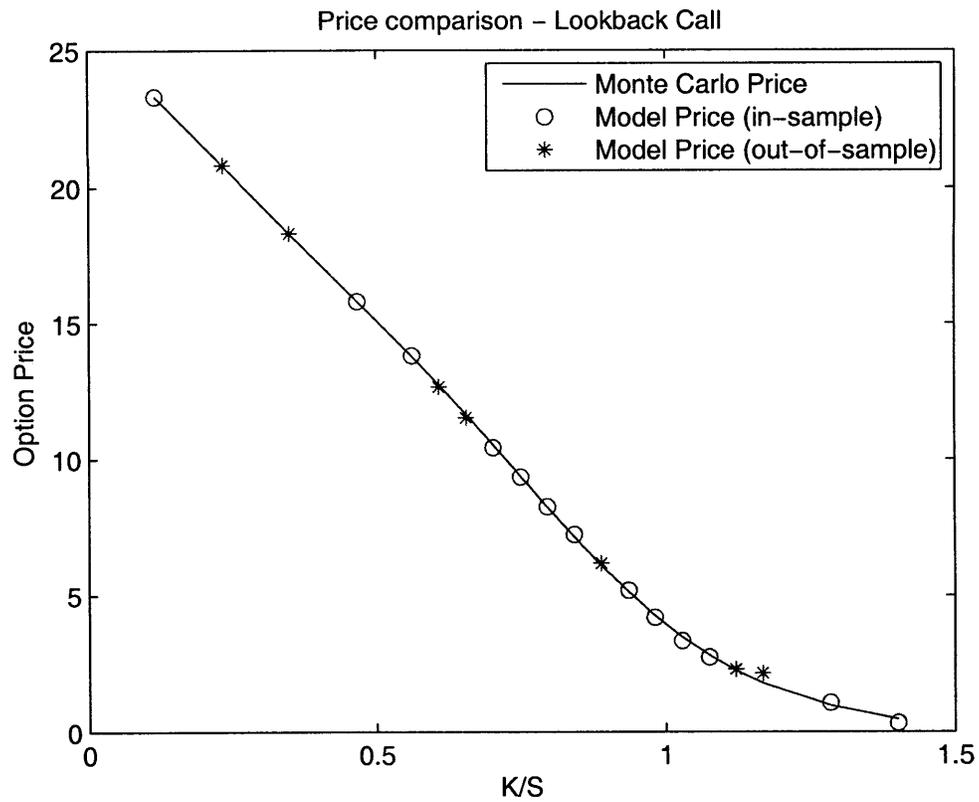


Figure 7.6: Comparison of Model Price and Monte Carlo Price for a Lookback Call option.

We tabulate the results in Table 7.3:

No.	T	K/S	$\Gamma_{implied}$	Monte Carlo Price	Model Price	Error	Sampling
1	18	0.117	1.74	23.313	23.311	-0.003	in
2	18	0.234	1.74	20.813	20.811	-0.003	out
3	18	0.35	1.74	18.313	18.311	-0.003	out
4	18	0.467	1.74	15.812	15.811	-0.002	in
5	18	0.561	1.73	13.791	13.811	0.019	in
6	18	0.607	1.71	12.744	12.674	-0.07	out
7	18	0.654	1.68	11.658	11.538	-0.12	out
8	18	0.701	1.63	10.534	10.43	-0.104	in
9	18	0.748	1.58	9.382	9.348	-0.034	in
10	18	0.794	1.53	8.251	8.267	0.017	in
11	18	0.841	1.49	7.153	7.24	0.088	in
12	18	0.888	1.47	6.097	6.187	0.089	out
13	18	0.935	1.49	5.137	5.187	0.049	in
14	18	0.981	1.53	4.254	4.187	-0.068	in
15	18	1.028	1.57	3.471	3.333	-0.138	in
16	18	1.075	1.56	2.813	2.732	-0.081	in
17	18	1.121	1.4	2.247	2.291	0.043	out
18	18	1.168	1.44	1.785	2.146	0.36	out
19	18	1.285	1.62	0.967	1.064	0.097	in
20	18	1.402	1.9	0.469	0.325	-0.144	in

Table 7.3: Finding the Implied Gamma, the Quadratic Relationship, and Error for a Lookback Call option.

7.4 Experiment 4 : Comparison with market prices for American put option

In this section, we compare our results to the market prices of American style put options. The underlying security is Microsoft stock and the option expiry date is 25 weeks ahead. Similarly as previous experiments, we find $\Gamma_{implied}$ that best matches our model price to the market price. We observe from Figure 7.7 a nice quadratic relationship between the $\Gamma_{implied}$ and the strike price K. For in-the-money and out-of-the-money situations, modeler's have larger risk aversion as implied by bigger magnitude of Γ . From Figure 7.8, we can see our model price matches market price well.

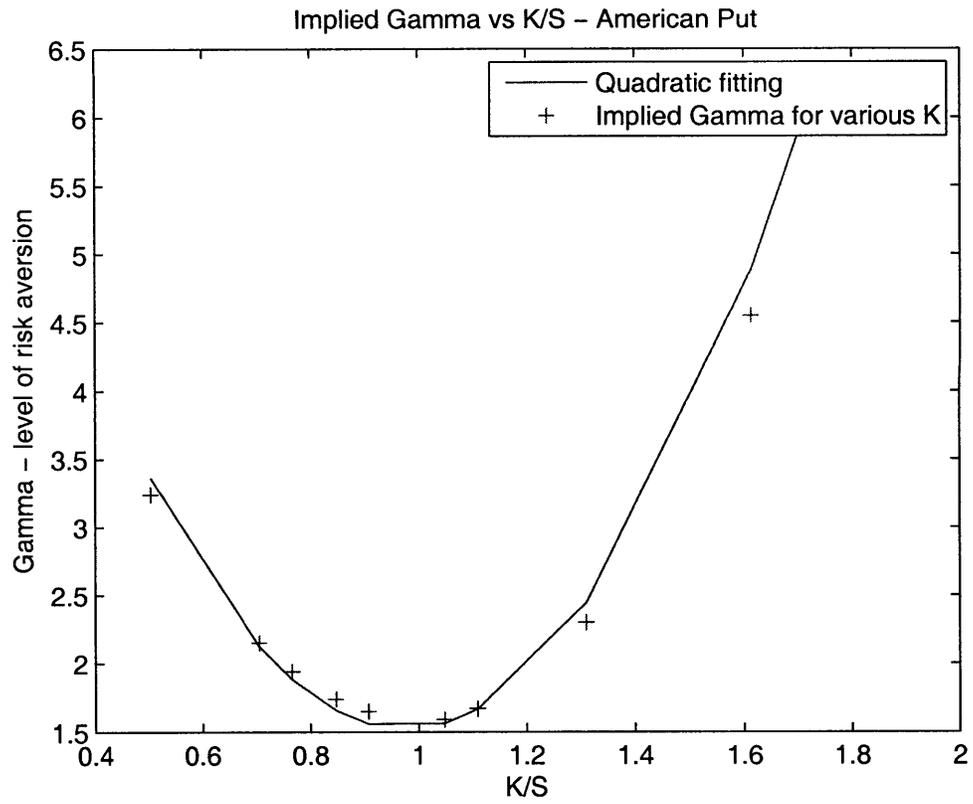


Figure 7.7: $\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for an American Put option.

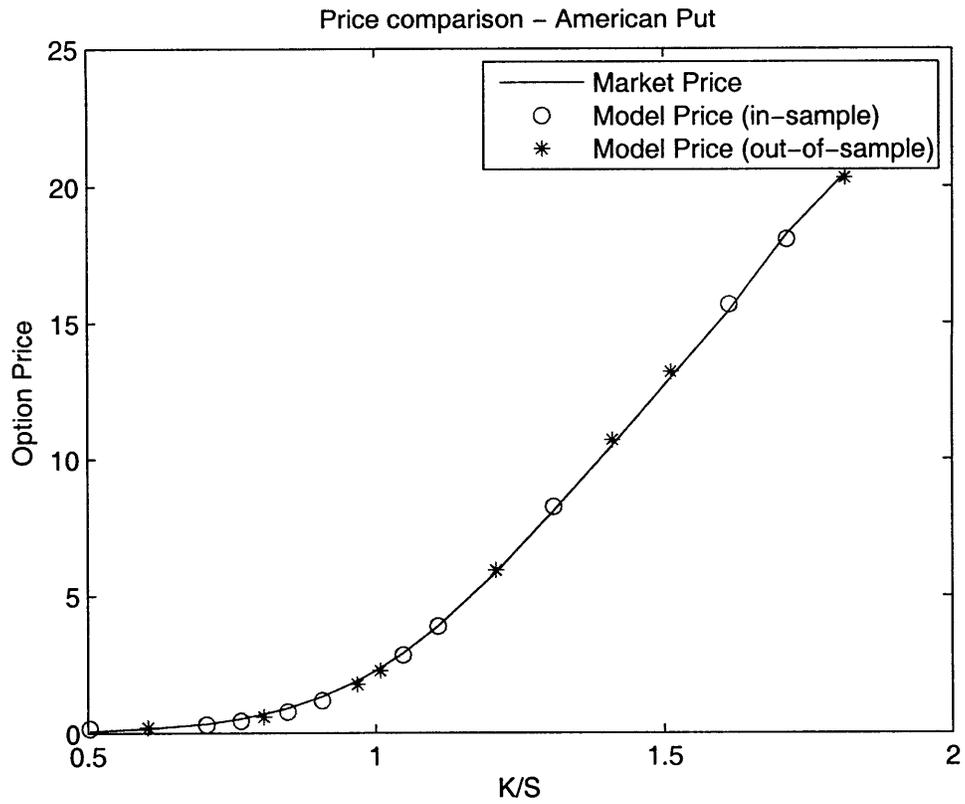


Figure 7.8: Comparison of Model Price and Market Price for an American Put option.

We tabulate the results in Table 7.4:

No.	T	K/S	$\Gamma_{implied}$	Market Price	Model Price	Error	Sampling
1	25	0.504	3.24	0.065	0.17	0.105	in
2	25	0.605	2.62	0.17	0.201	0.031	out
3	25	0.706	2.15	0.34	0.305	-0.035	in
4	25	0.766	1.94	0.525	0.442	-0.083	in
5	25	0.806	1.83	0.695	0.589	-0.106	out
6	25	0.847	1.74	0.905	0.778	-0.127	in
7	25	0.907	1.65	1.325	1.182	-0.143	in
8	25	0.968	1.6	1.895	1.764	-0.132	out
9	25	1.008	1.59	2.365	2.266	-0.099	out
10	25	1.048	1.59	2.91	2.847	-0.063	in
11	25	1.109	1.67	3.9	3.895	-0.005	in
12	25	1.21	1.9	5.85	5.939	0.089	out
13	25	1.31	2.3	8.1	8.258	0.158	in
14	25	1.411	2.87	10.5	10.703	0.203	out
15	25	1.512	3.63	12.975	13.2	0.225	out
16	25	1.613	4.55	15.45	15.679	0.229	in
17	25	1.714	6.35	18.25	18.068	-0.182	in
18	25	1.815	7.7	20.45	20.303	-0.147	out

Table 7.4: Finding the Implied Gamma, the Quadratic Relationship, and Error for an American Put option.

7.5 Experiment 5 : Comparison with market prices for European Style Index option

In this section, we compare our results to the market prices of options based on Index. The index we picked is 1/100 Dow Jones Industrial Average (DJIA) which uses price information of 30 major US listed stocks. DJIA is calculated by adding the prices of the 30 stocks and divide by a common divisor while the market capitalization of stocks does not matter in its calculation. Thus all these 30 stocks are considered equal weighted.

Similar as previous experiments, we find $\Gamma_{implied}$ that best matches our model price to the market price for these index options. We observe from Figure 7.9 a nice quadratic relationship between the $\Gamma_{implied}$ and the strike price K. For in-the-money and out-of-the-money situations, modeler's have larger risk aversion as implied by bigger magnitude of Γ . From Figure 7.10, we observe a match between our price using appropriate Γ and market price.

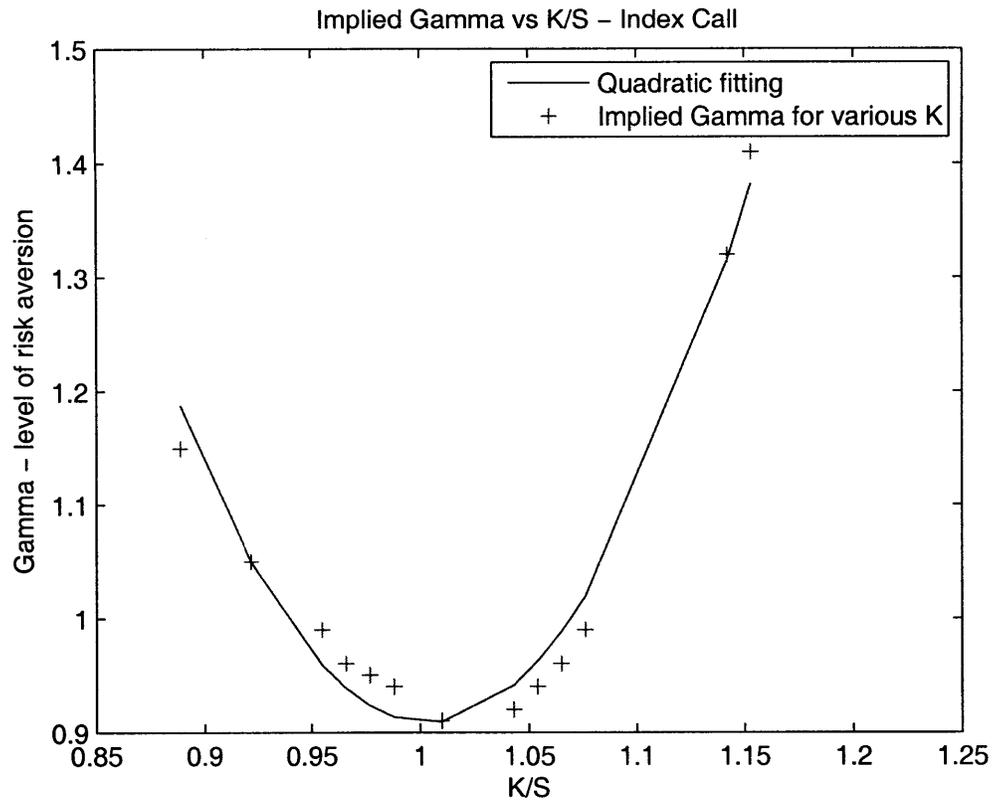


Figure 7.9: $\Gamma_{implied}$ as a function of $\frac{K}{S_0}$ for an Index Call option.

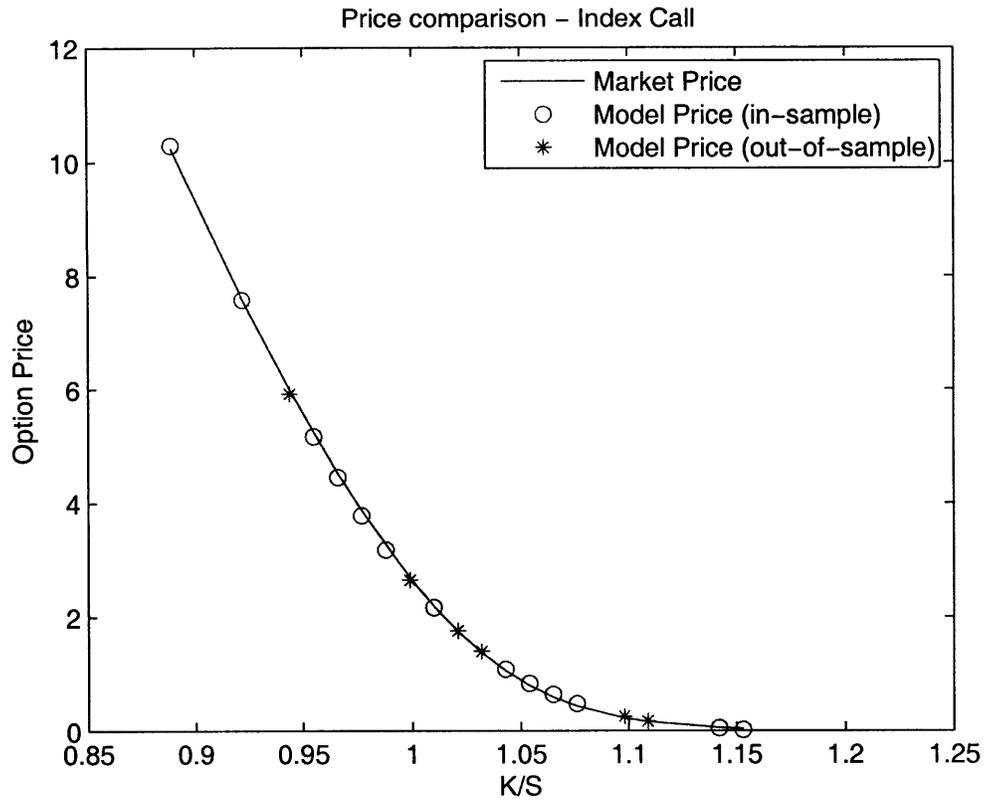


Figure 7.10: Comparison of Model Price and Market Price for an Index Call option.

We tabulate the results in Table 7.5:

No.	T	K/S	$\Gamma_{implied}$	Market Price	Model Price	Error	Sampling
1	8	0.889	1.15	10.25	10.295	0.045	in
2	8	0.922	1.05	7.6	7.582	-0.018	in
3	8	0.944	1.01	6	5.918	-0.082	out
4	8	0.955	0.99	5.25	5.166	-0.084	in
5	8	0.966	0.96	4.525	4.452	-0.073	in
6	8	0.977	0.95	3.875	3.778	-0.097	in
7	8	0.988	0.94	3.275	3.175	-0.1	in
8	8	0.999	0.92	2.69	2.647	-0.043	out
9	8	1.01	0.91	2.19	2.167	-0.023	in
10	8	1.021	0.91	1.75	1.762	0.012	out
11	8	1.032	0.92	1.38	1.402	0.022	out
12	8	1.043	0.92	1.06	1.086	0.026	in
13	8	1.054	0.94	0.8	0.835	0.035	in
14	8	1.065	0.96	0.595	0.641	0.046	in
15	8	1.076	0.99	0.435	0.476	0.041	in
16	8	1.098	1.07	0.225	0.254	0.029	out
17	8	1.109	1.13	0.16	0.17	0.01	out
18	8	1.142	1.32	0.055	0.047	-0.008	in
19	8	1.153	1.41	0.04	0.018	-0.022	in

Table 7.5: Finding the Implied Gamma, the Quadratic Relationship, and Error for an Index Call option.

7.6 Remarks

Table 7.6 summarizes the errors between our model prices and market prices (or simulation prices when market data are not obtainable) from Experiment 1-5. We can make the following observations:

1. In all experiments the errors we report in out-of-sample data are between $[0, 0.626]$. Moreover, except for Asian option, the errors of in-sample and out-of-sample data are of the same order. The reason why the error is large for Asian option is explained in Section 7.2.
2. The errors for the single asset European option, and European style Index option are smaller than other types of options.
3. The average absolute error for American options (Experiment 4) is larger.

Option type	European	Asian	Lookback	American	Index
max absolute error (in-sample)	0.017	0.087	0.144	0.229	0.1
max absolute error (out-of-sample)	0.055	0.626	0.36	0.225	0.082
average absolute error (in-sample)	0.0096	0.0419	0.0649	0.113	0.0475
average absolute error (out-of-sample)	0.0088	0.4544	0.0983	0.129	0.033

Table 7.6: Analysis on errors of our model prices compared with market prices (or simulation prices)

Chapter 8

Conclusion

In this thesis, we combine ϵ -arbitrage and robust optimization to solve, via linear optimization methods, option pricing models that can model transaction costs, high-dimensional options, high-dimensional price dynamics, etc. The main advantage of our model is that unlike current SDP and simulation methods, our approach scales in a polynomial way (as opposed to exponential way) with the dimension of the original pricing problem. We illustrate our method and report results for a variety of options (European, Asian, Lookback, American, Index) using empirical data, which show that our approach produces prices that are close to those observed in options market.

Bibliography

- [1] Black, F., M. Scholes. 1973. Pricing of options and corporate liabilities, *J. of Political Economy*, 81, 637-654.
- [2] Ben-Tal, A., A. Nemirovski, 1998, Robust convex optimization, *Mathematics of Operations Research*, 23, 769-805.
- [3] Bertsimas, D., L. Kogan and A.W. Lo, 2001, Hedging derivative securities and incomplete markets: an ϵ -arbitrage approach, *Operations Research*, 49, 3, 372-397.
- [4] Bertsimas, D., M. Sim, 2004, The price of robustness, *Operations Research*, 52, 1, 35-53.
- [5] Bertsimas, D., D. Pachamanova, 2008, Robust multiperiod portfolio management in the presence of transaction costs, *Computers and Operations Research*, 35, 3-17.
- [6] Bertsimas, D., D. Pachamanova and M. Sim, 2004, Robust Linear Optimization under General Norms, *Operations Research Letters*, 32, 510-516.
- [7] Cox, J.C, S.A. Ross, 1976, The valuation for alternative stochastic processes, *Journal of Financial Economics*, 3, 145-166.
- [8] El-Ghaoui, L., H. Lebret. 1997. Robust solutions to least-square problems to uncertain data matrices. *SIAM Journal of Matrix Analysis and Applications*. 18. 1035-1064.
- [9] Hull, J., A. White, 1987, The pricing of options on assets with stochastic volatility, *Journal of Finance*, 42, 281-300.
- [10] Madan, D.B, P.P. Carr and E.C. Chang, 1998, The variance gamma process and option pricing, *European Finance Review*, 2, 79-105.

- [11] Merton, R.C., 1971. Optimum consumption and portfolio rules in a continuous-time model. *J. Economy Theory*, 3, 373-413.
- [12] Merton, R.C., 1976, Option pricing when underlying stock returns are discontinuous, *Journal of Financial Economics*, 3, 125-144.
- [13] Pachamanova, D.A, 2002, A robust optimization approach to finance, PhD thesis Massachusetts Institute of Technology.
- [14] Rubinsten, M., 1994 Non-parametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE option classes from August 23,1976 through August 31, 1978, *Journal of Finance*, 40, 455-480.
- [15] Soyster, A. L. 1973, Convex programming with set-inclusive constraints and applications to inexact linear programming, *Operations Research*, 21,1154-1157.
- [16] Taylor, S.J., 1994, Modelling stochastic volatility: A review and comparative study, *Mathematical Finance*, 4(2), 183-204.

Appendice A

Proof of Theorem 3.1

Proof : there exists a primal-dual pair:

(Primal)

$$\begin{aligned} \max_{\mathbf{A}} \quad & (\mathbf{a})' \hat{\mathbf{x}} \\ \text{s.t.} \quad & \mathbf{G} \cdot \text{vec}(\mathbf{A}) \leq \mathbf{d} \end{aligned}$$

(Dual)

$$\begin{aligned} \min_{\mathbf{p}} \quad & (\mathbf{p})' \mathbf{d} \\ \text{s.t.} \quad & (\mathbf{p})' \mathbf{G} = (\hat{\mathbf{x}})' \\ & \mathbf{p} \geq 0 \end{aligned}$$

Suppose that $\hat{\mathbf{x}}$ satisfies $(\mathbf{a})' \hat{\mathbf{x}} \leq b$ for all $\tilde{\mathbf{A}} \in P^A$. It follows that $\max_{\mathbf{A}} (\mathbf{a})' \hat{\mathbf{x}} \leq b$. It means that primal problem is feasible and bounded, and so its dual. Hence, there exists a vector $\mathbf{p} \in \mathfrak{R}^{l \times 1}$ that satisfy the dual constraints. Further, by strong duality, the optimal objective function value of the dual equals $\max_{\mathbf{A}} (\mathbf{a})' \hat{\mathbf{x}} \leq b$ and thus $(\mathbf{p})' \mathbf{d} \leq b$.

The reverse direction. Since $P^A \neq \emptyset$, the primal problem is feasible. Suppose there exists a vector $\mathbf{p} \in \mathfrak{R}^{l \times 1}$ that satisfy the dual constraints, then it implies that the dual problem is also feasible. Since both primal and dual are feasible, there must be bounded and their optimal objective function values must be equal. Furthermore, $\min_{\mathbf{p}} (\mathbf{p})' \mathbf{d} \leq (\mathbf{p})' \mathbf{d} \leq b$. Therefore, by strong duality $\max_{\mathbf{A}} (\mathbf{a})' \hat{\mathbf{x}} = \min_{\mathbf{p}} (\mathbf{p})' \mathbf{d} \leq b$, hence $\hat{\mathbf{x}}$ satisfies $(\mathbf{a})' \hat{\mathbf{x}} \leq b$ for all $\tilde{\mathbf{A}} \in P^A$. $\{$