

The Pennsylvania State University

The Graduate School

BAYESIAN NONPARAMETRIC APPROACHES FOR FINANCIAL  
OPTION PRICING

A Dissertation in

Statistics

by

Huei-Wen Teng

© 2010 Huei-Wen Teng

Submitted in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

August 2010

The dissertation of Huei-Wen Teng was reviewed and approved\* by the following:

John C. Liechty  
Professor of Marketing and Statistics  
Dissertation Advisor  
Chair of Committee

Bruce G. Lindsay  
Professor of Statistics  
Head of the Department of Statistics

Jingzhi (Jay) Huang  
Associate Professor of Finance

Murali Haran  
Assistant Professor of Statistics

Zhibiao Zhao  
Assistant Professor of Statistics

\*Signatures are on file in the Graduate School.

# Abstract

The price of a financial option equals the discounted expected payoff of the option under the risk-neutral measure, and an option's Greeks are formulas that give the change in an option price with respect to parameters of interest (e.g. the price of the underlying asset). The density that reproduces the observed option price is called the risk-neutral or state price density and is used for a variety of important activities in finance, including providing an arbitrage-free tool for pricing complex and less liquid securities. The importance of understanding this density with respect to asset pricing and risk management has led to a competing number of approaches for making inference about the state price density. Unlike the option prices, Greeks can not be observed in the market and have to be calculated. As Greeks are important for measuring and managing risk as well as executing dynamic trading strategies, developing methods to calculate them efficiently and accurately is of critical importance both in theory and in practice (Broadie and Glasserman, 1996).

We start by proposing a finite-dimensional model for the state price density in a Bayesian framework. This modeling approach can be viewed as a Bayesian Quadrature model, where the locations and weights of support points in the finite-dimensional representation of the risk-neutral density are random variables. This modeling approach allows a 'prior' reference distribution which can be a parametric distribution (e.g. the lognormal density) or which can be uniform and completely non-informative, and it also provides a posterior distribution of the state price density that is consistent with the observed option prices. We asses the performance of the proposed model using simulation studies based on synthetic data and then by contrasting the method with a number of competing methods using S&P 500 index option data.

In contrast to European options, American options can be exercised anytime prior to maturity. We show how our Bayesian Quadrature approach can be ex-

tended to make inference for American options. To tackle this problem, we propose a Bayesian implied random tree model as an extension of the Bayesian Quadrature approach by building a unique binomial tree similar to Rubinstein (1994). The benefits of our approach are demonstrated via simulation study and empirical studies using S&P 100 index option data.

Although finite-difference methods are commonly used to calculate Greeks, these estimates can often be biased and suffer from erratic behavior when the payoff function is discontinuous. We provide new and simple mathematical formulas that overcome these problems and that are applicable to a wide range of complicated options and underlying processes. Moreover, we provide an innovative Bayesian approach to calculate Greeks using observed option prices without any parametric assumptions on the underlying process, so that the proposed method avoids the model misspecification problem. We demonstrate the performance of our methods through simulation studies.

# Table of Contents

<b>List of Figures</b>	viii
<b>List of Tables</b>	xi
<b>Acknowledgments</b>	xii
<b>Chapter 1</b>	
<b>Introduction</b>	1
1.1 The state price density . . . . .	1
1.2 Greeks . . . . .	2
1.3 Organization of dissertation . . . . .	3
<b>Chapter 2</b>	
<b>Literature reviews</b>	5
2.1 Parametric inference methods for the state price density . . . . .	5
2.1.1 Mixture models . . . . .	6
2.1.2 Expansions of functions . . . . .	6
2.1.3 Other diffusion processes . . . . .	6
2.2 Nonparametric inference methods for the state price density . . . . .	7
2.2.1 Kernel-based methods . . . . .	7
2.2.2 Canonical methods . . . . .	8
2.2.3 Curve fitting methods . . . . .	8
2.3 Related methods for estimating option pricing formulas but not state price densities . . . . .	9
2.4 Summary . . . . .	9
<b>Chapter 3</b>	
<b>The Bayesian Quadrature model</b>	10

3.1	The Bayesian Quadrature model . . . . .	11
3.2	Slice Sampling . . . . .	13
3.3	Simulation study one . . . . .	15
3.3.1	Synthetic data sets . . . . .	15
3.3.2	MCMC inference . . . . .	15
3.4	Simulation study two . . . . .	16
3.4.1	Synthetic data sets . . . . .	17
3.4.2	Bayes factors for model selection . . . . .	17
3.4.3	Cross-validation for selecting the number of support points .	18
3.4.4	MCMC inference . . . . .	19
3.5	Empirical study . . . . .	20
3.5.1	Using data set in Aït-Sahalia and Lo (1998) . . . . .	21
3.5.2	Using both call and put options . . . . .	21
3.6	Summary . . . . .	24

## Chapter 4

	<b>The Bayesian implied random tree model</b>	<b>41</b>
4.1	The Bayesian implied random tree model . . . . .	42
4.2	Simulation study . . . . .	48
4.2.1	True model in the simulation study . . . . .	48
4.2.2	Generating synthetic data . . . . .	48
4.2.3	MCMC inference . . . . .	49
4.3	Empirical study . . . . .	51
4.3.1	Descriptions of S&P 100 index options . . . . .	51
4.3.2	Recovery of option prices . . . . .	53
4.3.3	In-sample versus out-of-sample analysis . . . . .	56
4.4	Summary . . . . .	57

## Chapter 5

	<b>Parametric Greeks: The non-Lipschitz pathwise method</b>	<b>77</b>
5.1	Preliminaries . . . . .	83
5.2	Fundamental theorems . . . . .	90
5.3	Greeks of rainbow options . . . . .	96
5.3.1	Greeks of spread options . . . . .	96
5.3.2	Greeks of maximum options . . . . .	97
5.4	Numerical results . . . . .	99
5.5	Summary . . . . .	100

<b>Chapter 6</b>	
<b>Nonparametric Greeks</b>	<b>104</b>
6.1 Our approaches . . . . .	105
6.2 Simulation study . . . . .	105
6.2.1 The Black-Scholes model . . . . .	106
6.2.2 The Heston stochastic volatility model . . . . .	107
6.3 Empirical study . . . . .	111
6.4 Summary . . . . .	115
<b>Chapter 7</b>	
<b>Conclusions and future work</b>	<b>120</b>
7.1 Conclusions . . . . .	120
7.1.1 The Bayesian Quadrature model . . . . .	121
7.1.2 The Bayesian implied random tree model . . . . .	121
7.1.3 Parametric Greeks . . . . .	122
7.1.4 Nonparametric Greeks . . . . .	122
7.2 Future work . . . . .	122
7.2.1 Implied multi-variate state price density . . . . .	123
7.2.2 Implied diffusion processes with an application to Credit Default Swaps . . . . .	124
7.2.3 New tree-based approaches for pricing options . . . . .	128
7.2.4 Do American options bring more information than European options for understanding state price densities? . . . . .	129
7.2.5 Will nonparametric Greeks improve the performance of dynamic hedge? . . . . .	130
<b>Appendix A</b>	
<b>Slice Sampler</b>	<b>132</b>
<b>Appendix B</b>	
<b>Technical proofs of Parametric Greeks</b>	<b>137</b>
B.1 Proof of Theorem ?? . . . . .	137
B.2 Proof of Theorem ?? . . . . .	139
B.3 Proof of Theorem ?? . . . . .	141
B.4 Proof of Theorem ?? . . . . .	142
B.5 Greeks Using the Likelihood Ratio Method . . . . .	145
<b>Bibliography</b>	<b>147</b>

# List of Figures

3.1	Scatter plots for the first simulation study. . . . .	26
3.2	Convergence MCMC analysis for the first simulation study. . . . .	27
3.3	Credible region of residuals plots for the first simulation study. . . . .	28
3.4	Trace plots and convergence test of $\sigma^2$ for the first simulation study.	29
3.5	Trace plots of parameters for the first simulation study. . . . .	30
3.6	True models and scatter plots of option prices for the second simulation study. . . . .	31
3.7	Marginal log-likelihood for the simulation study two. . . . .	32
3.8	Convergence MCMC analysis for the second simulation. . . . .	33
3.9	The 10-fold cross-validation for the second simulation study. . . . .	34
3.10	The 90% credible region of residuals plots for the second simulation study. . . . .	35
3.11	The true and estimated state price density for the second simulation study. . . . .	36
3.12	The 90% credible region of the estimated state price density using call options for the empirical study. . . . .	37
3.13	Scatter plots of the actual and fitted option prices using call options in the empirical study. . . . .	38
3.14	The 90% credible region of the estimated state price density using call and put options. . . . .	39
3.15	Scatter plots of the actual and fitted option prices using call and put options for the empirical study. . . . .	40
4.1	Illustrations of a generalized binomial tree associated with $Q(w, \theta)$ for $N = 3$ . . . . .	44
4.2	Scatter plots of option prices against strike prices of the first data set for the simulation study. . . . .	60
4.3	Scatter plots of option prices against strike prices of the second data set for the simulation study. . . . .	61
4.4	Scatter plots of option prices against strike prices of the third data set for the simulation study. . . . .	62

4.5	Convergence MCMC analysis for the simulation study. . . . .	63
4.6	Trace plots of $\sigma^2$ for the simulation study. . . . .	64
4.7	Trace plots of the locations, $\theta_n$ , and weights, $w_n$ , for $n = 0, \dots, 3$ for the simulation study. . . . .	65
4.8	A time series plot of future prices of the S&P 100 index traded on 11/02/2009. . . . .	66
4.9	Stacked bar plots of the number of transactions against strike prices.	67
4.10	Scatter plots of S&P 100 index options with 19 days maturity. . . .	68
4.11	Scatter plots of S&P 100 index options with 47 days maturity. . . .	69
4.12	Scatter plots of actual and fitted option prices against strike prices and the estimated state price density using European options. . . .	70
4.13	Scatter plots of actual and fitted option prices against strike prices and the estimated state price density using American options. . . .	71
4.14	Scatter plots of actual and fitted option prices against strike prices and the estimated state price density using European and American options. . . . .	72
4.15	Summary of posterior distribution of $R^2$ for the empirical study. . .	73
4.16	Summary of posterior distribution of Root-Mean-Squared-Errors of pricing errors for the empirical study. . . . .	74
4.17	The estimated state price densities using different data sets for the empirical study. . . . .	75
4.18	Trace plots of the log-likelihoods for assessing the convergence of the Markov chain Monte Carlo simulation for the S&P 100 index options data. . . . .	76
6.1	Scatter plots of the actual and fitted values in prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Black-Scholes model. . . .	108
6.2	Scatter plots of the absolute errors in dollar for the prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Black-Scholes model. 109	
6.3	Scatter plots of the absolute relative errors in percentage for the prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Black-Scholes model. . . . .	110
6.4	Scatter plots of actual and fitted prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Heston stochastic volatility model. .	112

6.5	Scatter plots of the absolute errors in price for the prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Heston stochastic volatility model. . . . .	113
6.6	Scatter plots of the absolute relative errors in percentage for the simulation study using the Heston stochastic volatility model. . . . .	114
6.7	The scatter plot of implied volatilities against strike prices for the empirical study. . . . .	116
6.8	The scatter plot of actual and fitted option prices against strike prices for the empirical study. . . . .	117
6.9	The scatter plot of Black-Scholes deltas and nonparametric deltas against strike prices for the empirical study. . . . .	118
6.10	The scatter plot of Black-Scholes gammas and nonparametric gammas against strike prices for the empirical study. . . . .	119

# List of Tables

3.1	Numerical results for the first simulation. . . . .	16
3.2	Marginal log-likelihood for the simulation study two. . . . .	19
3.3	Numerical results for the second simulation study. . . . .	20
3.4	Numerical results using call options for the empirical study. . . . .	22
3.5	Numerical results using call and put options for the empirical study. . . . .	23
4.1	Interior stock prices for the simulation study. . . . .	49
4.2	The probabilities of moving up for the simulation study. . . . .	49
4.3	Acceptance rates for parameters in the MCMC algorithm for the simulation study. . . . .	50
4.4	Numerical results for the simulation study. . . . .	51
4.5	Number of transactions of S&P 100 index options data traded on 11/02/2009. . . . .	53
4.6	Summary statistics of the trading volume of S&P 100 index options data. . . . .	54
4.7	Summary statistics for S&P 100 index options data. . . . .	55
4.8	Root-Mean-Square-Errors of pricing errors for the in-sample versus out-of-sample analysis. . . . .	58
5.1	A sampling of rainbow options . . . . .	79
5.2	Finite-difference formulas for deltas, gammas and cross-gammas . .	81
5.3	Greeks of spread options on two assets . . . . .	102
5.4	Greeks of maximum options on two assets . . . . .	103
6.1	European call option prices, implied volatilities (Imp. Vol.), Black-Scholes deltas, and Black-Scholes gammas for the empirical study using S&P 500 index options. . . . .	115

# Acknowledgments

I gratefully acknowledge my advisor Dr. John Liechty for his crucial contribution. He allows me a great deal of freedom to explore my interests in the field of financial econometrics. With his constant professional guidance, inspiring conversation, optimism, and encouragement, I have learned the most precious lesson in my graduate study.

It is also my pleasure to convey my gratitude to my committee members, Dr. Jay Huang, Dr. Murali Haran, and Dr. Zhibiao Zhao, for their great contributions to this dissertation, their suggestions and kind discussions in different aspects in finance and statistics.

I will also thank the statistics department at Penn State University, where I received high quality education and found my interests in quantitative analysis. I would also like to thank Dr. Bruce Lindsay, Dr. Yu Zhang, Dr. Runze Li, Dr. William Harkness, Dr. Naomi Altman, Dr. James Rosenberger, and Dr. Donald Richards for their great kindness and support in my research and in my life.

In addition, I would like to deliver my special thanks to Prof. Yuh-Dauh Lyuu at National Taiwan University and Prof. Cheng-Der Fuh at Academia Sinica for their continuing support and collaboration on several interesting research projects.

Finally, the list of thanks would not be complete without a mention of my family, to whom I owe a great deal and more. This is for my family.

# **Dedication**

This thesis is dedicated to my dearest family.

Chapter **1**

# Introduction

Options are financial contracts that give the buyer a cash flow contingent on a payoff function that depends on the behavior of an underlying asset. The underlying asset is typically a single stock or an index. Let  $x$  denote the underlying asset's price at maturity, and  $K$  denote the strike price. A call option has the payoff function given by  $\max(x - K, 0)$ , and a put option has the payoff function given by  $\max(K - x, 0)$ . European type options can only be exercised at maturity, whereas American type options can be exercised anytime up to and including maturity. Due to this early exercise feature, American options provide investors more flexibility and are therefore among the most traded options in the market.

## 1.1 The state price density

Standard no-arbitrage, asset pricing theory gives the theoretical option price as the discounted value of the expected payoff function under the risk-neutral measure. The density under this risk-neutral measure is called the risk-neutral density or state price density. Understanding the state price density is critical for risk management and for dynamic hedging purposes. In addition, the state price density can be used to price illiquid or less frequently traded options. As a result, the research of calibrating a state price density using widely traded options remains an important and active area in academia and in industry.

Existing methods to calibrate the state price density can be categorized as either parametric or nonparametric methods. Although a parametric assumption on

the state price density or process is straightforward, misspecification of a parametric assumption may lead to incorrect conclusions and systematic pricing errors. Nonparametric methods are popular because they are free of model misspecification problems. We give a brief overview of the relevant literature in the next chapter and a comprehensive overview on the literature can be found in Jackwerth (1999) and Fengler (2009).

In this dissertation, we first propose a Bayesian Quadrature model to calibrate the state price density using European options. In contrast to parametric methods, our approach avoids the problem of model misspecification. It also provides an option pricing model that is parsimonious and can be easily calculated without extra algebraic efforts and very little numerical efforts. In contrast to nonparametric approaches, the derivation of the full-conditional densities and detailed description of the inference algorithm for our approach is straightforward.

Existing calibration methods focus primarily on using European options, not American options, to calibrate state price density. However, American options are the most traded options and it is natural to explore methods for including American options in the calibration of state price density (Stensoft, 2004). Recently, Alcock and Carmichael (2008) and Alcock and Auerswald (2009) propose nonparametric approaches for pricing American options using extended canonical methods proposed by Stutzer (1996). However, their methods are sensitive to the samples of historic data of the underlying assets. To overcome this problem, we provide a Bayesian implied random tree model to make inference for the state price density using American options.

## 1.2 Greeks

Greeks summarize how option prices change with respect to underlying variables (e.g., underlying asset) and are critically important in asset pricing and risk management. Greeks are the price's sensitivities with respect to certain parameters of interest such as the underlying asset's price, volatility, maturity, and interest rate. Although the price of an option can often be observed in the market, this is not the case for its Greeks; they must be calculated. Because Greeks are important for measuring and managing risk as well as executing dynamic trading strategies, the

problem of how to calculate them efficiently and accurately is critically important both in theory and in practice (Broadie and Glasserman, 1996).

Although finite-difference with resimulation is the standard method for estimating Greeks, it is typically biased and suffers from erratic behavior when the payoff function is discontinuous. Direct methods, such as the pathwise method and the likelihood ratio method, are based on differentiating the prices formulas directly and hence produce unbiased Greeks (Broadie and Glasserman, 1996). The pathwise method differentiates the payoff function, whereas the likelihood ratio method differentiates the densities. Our method differentiates the payoff function, but lifts the Lipschitz continuity requirements on the payoff function. We build a new but simple mathematical formulation so that formulas of Greeks for a broad class of derivative securities can be derived systematically.

Although the Greeks calculated using our proposed mathematics formulation is unbiased, the parametric assumptions in the underlying asset may inevitably lead to systematic errors in the calculation of Greeks if the parametric assumptions on the underlying asset is incorrect. To avoid model misspecification problems, we consider Bayesian nonparametric approaches to calculate the Greeks using the state price density calibrated from the option prices. We demonstrate the performance of our proposed method via simulation studies.

### 1.3 Organization of dissertation

This dissertation is organized as follows. Because the state price density forms a basis of option pricing, we first review existing methods for calibrating state price density in Chapter 2.

We propose a Bayesian Quadrature model in Chapter 3 for calibrating a state price density using European options. For numerical purposes, we provide Gibbs sampling with slice sampling in the Markov chain Monte Carlo algorithm. Simulation studies and empirical studies show that our ability of recovering option prices and that the method is competitive with other popular methods.

We propose a Bayesian implied random tree model in Chapter 4 to calibrate a state price density using American options. We demonstrate the performance of our approach through a simulation study and an empirical study using S&P 100

index options.

We provide the non-Lipschitz pathwise method to calculate unbiased Greeks in Chapter 5. We note that Chapter 5 was done in collaboration (Lyuu and Teng, 2010). In Chapter 6, we provide a new Bayesian nonparametric approach to estimate the Greeks using observed option prices without placing a parametric assumption on the underlying process. The last chapter concludes and discusses potential future research.

# Chapter 2

## Literature reviews

A European call option has a theoretic price

$$e^{-rT} \int \max(x - K) f(x) dx, \max \quad (2.1)$$

where  $r$  is the risk-free interest rate,  $T$  is the time to maturity,  $K$  is the strike price,  $f(x)$  is the state price density. Similarly, a European put option has a theoretic price

$$e^{-rT} \int \max(K - x) f(x) dx. \quad (2.2)$$

Pricing formulas for European options are simply integrals, where the integrands are a product of the payoff function and the state price density.

### 2.1 Parametric inference methods for the state price density

The Black-Scholes model follows from a parametric state price density that is the lognormal density. However, many empirical studies have shown that the Black Scholes lognormal density assumption does not hold (Longstaff, 1995). This leads to more complicated parametric assumptions, which find the best density with respect to the ability to fit a set of empirical option prices. These solutions are attractive as they typically result in closed-form formulas which allows for rapid calculations for pricing other assets or calculating risk positions.

Although parametric methods are straightforward, model-misspecification is a major concern that may lead to systematic errors in asset pricing and incorrect dynamic hedge strategies. Nonparametric methods can be preferable because they are free of model misspecification. For a discussion of the advantage of parametric and nonparametric approaches, we refer to Engle and González-Rivera (1991) and Aït-Sahalia and Lo (1998). Next, we review several popular parametric and nonparametric methods.

### 2.1.1 Mixture models

Mixture models are popular methods for density estimation. Melick and Thomas (1997) and Söderlind and Svensson (1997) consider mixtures of log-normal distributions. Yuan (2009) establishes the mathematics basis using mixtures of log-normal distributions. Recently, Giacomini et al. (2008) proposes mixtures of scaled and shifted  $t$ -distributions. Abadir and Rockinger (2003) develops density functionals based on confluent hypergeometric functions, which cover several popular distributions, such as normal and gammas distributions, and mixtures of these distributions.

### 2.1.2 Expansions of functions

Jarrow and Rudd (1982) develops the Edgeworth expansion to approximate an arbitrary distribution which allows higher moments. Corrado and Su (1996) and Strong and Xu (1999) have applied the Edgeworth expansion for option pricing. Madan and Milne (1994) and Abken et al. (1996) use a system of Hermite polynomials to approximate the state price density. Knight and Satchell (1997) and Jondeau and Rockinger (2001) use Gram-Charlier expansions for the risk-neutral estimation; however, these methods may produce non-positive probabilities.

### 2.1.3 Other diffusion processes

From a time series perspective, more general underlying processes, other than the geometric Brownian motion assumption as in the Black-Scholes model, have been introduced as alternative parametric approaches. These efforts aim at developing

more flexible dynamics of asset prices leading to more accurate option pricing formulae.

For example, Bates (1996a), Bates (1996b), Malz (1996), Madan et al. (1998) propose jump-diffusion models. Hull and White (1987) and Heston (1993) consider stochastic volatility models; Amin and Ng (1993) and Bakshi and Chen (1997) propose stochastic volatility and stochastic interest rates models. Scott (1997), Nandi (1998), Bates (2000), Duffie et al. (2000), Pan (2002), Eraker et al. (2003), and Maheu and McCurdy (2004) consider stochastic volatility models with jumps.

## 2.2 Nonparametric inference methods for the state price density

Although these parametric models have greatly relaxed the Black-Scholes's restrictive assumptions, these models are not derived from comprehensive economic theories. These methods often rely on different assumptions concerning the dynamics of the underlying asset, and need to be simple enough to allow for the derivation of pricing formulas. Therefore, these models can not be expected to capture all the relevant features of the pricing mechanisms. Indeed, there are always limitations on the performance of modeling techniques. Model misspecification is always a major concern as it can lead to erroneous valuations and hedging strategies.

The idea of applying nonparametric methods to option pricing has been suggested recently in a number of papers. A major advantages of nonparametric methods is that they are free of model misspecification. In the following, we review several nonparametric approaches.

### 2.2.1 Kernel-based methods

Breeden and Litzenberger (1978) first presented the following formula connecting option prices and the state price density,

$$f(K) = e^{rT} \frac{\partial^2 C}{\partial K^2}. \quad (2.3)$$

This formula shows that the second derivatives of the call options price with respect

to strike price is equal to the state price density.

Taking advantage of Eq. (2.3), Aït-Sahalia and Lo (1998) proposes a kernel method, by first estimating the functional form of European call option prices using kernel methods, and differentiating this pricing formula twice with respect to the strike price to obtain the state price density. In practice, the kernel methods rely on a large number of observed option prices to achieve reasonable convergence properties. In addition, these methods may produce non-positive density. Therefore, extra modifications to consider the convexity constrain are required. See Aït-Sahalia and Duarte (2003), Härdle and Yatchew (2002), and Birke and Pilz (2009), for example.

Recently, Fan and Mancini (2009) provides a model-guided nonparametric approach to estimate the risk-neutral survival function, which is the cumulative density function of the state price density. This method starts with a parametric assumption on the risk-neutral distribution (but not the state price density), then corrects the estimate risk-neutral distribution using kernel methods for pricing errors.

The kernel-based methods uses Eq. (2.3) for calibrating state price density and are therefore only applicable to European options.

### 2.2.2 Canonical methods

Another category by nonparametric methods is the canonical method or entropy method proposed Stutzer (1996). Stutzer (1996) uses a Bayesian method based on the maximum-entropy principle of Shannon, and proposes the canonical valuation method, which relies on the assumption that the state price density can be approximated by the physical density which satisfies the minimum distance condition represented by the Shannon entropy. As an extension, Kacperczyk et al. (2005) uses a semiparametric scale mixture of betas model for the historic log return data, and convert the predictive physical density to the state price density. However, the estimated state price density depends heavily on the sample of the historic data.

### 2.2.3 Curve fitting methods

Curve fitting methods are a broad class of methods that use a general function to approximate either the implied volatility, which is used to estimate the risk neutral density, or to approximate the state price density directly. In the case of approximating the implied volatility, Shimko (1993) and Malz (1997) fit a quadratic polynomial. The idea of using a polynomial is extended to the use of splines in Brown and Toft (1999) and Campa et al. (1998), among others.

The second group of curve fitting methods approximates the state price density directly. For example, Rubinstein (1994) proposes the implied binomial trees that minimize the distance of discretized probabilities to a lognormal prior distribution, subject to the correct option prices. As an extension, Jackwerth and Rubinstein (1996) maximizes the smoothness of the state price density subject to the correct option prices. Using the same idea of maximizing the smoothness, Mayhew (1995) uses cubic splines for approximating the state price density. In general, it is easier to approximate the implied volatility against strike prices than the state price density Jackwerth (1999).

## 2.3 Related methods for estimating option pricing formulas but not state price densities

There are other methods focusing on estimating the pricing formulas of options but not the state price density. For instance, Hutchinson et al. (1994) proposes neural networks for estimating the pricing formula of an option, and Broadie et al. (2000) provides kernel estimators for call prices and the exercise boundaries using American options.

## 2.4 Summary

Although parametric assumptions often offer straightforward pricing formulas, the major concern with parametric models is model misspecification, which can lead to systematic miss-pricing. Nonparametric methods, such as kernel-based methods, are applicable to only European options. We first propose the Bayesian Quadrature

model in Chapter 3 to calibrate the state price density using European options. To calibrate the state price density using American options, in Chapter 4, we provide the Bayesian implied random tree model as an extension to the Bayesian Quadrature model.

We are also interested in calculating the price sensitivity of these options, which are also known as the Greeks, given a state price density. Unlike option prices, which can be observed in the market, the Greeks can not be observed and need to be calculated based on the pricing model. The Greeks are especially important for risk management and dynamic hedging. We start this investigation by providing the non-Lipschitz pathwise method for deriving formulas for the Greeks under the parametric assumptions in Chapter 5. The concern of model-misspecification motivates us to provide a method to calculate the Greeks without specifying parametric assumptions for the underlying asset. As a result, we study calculating the Greeks nonparametrically based on estimated approach for the state price density using the Bayesian implied random tree model in Chapter 6.

Chapter **3**

## The Bayesian Quadrature model

We have given an overview of existing methods for calibrating state price densities in Chapter 2. In this chapter, we describe a Bayesian Quadrature model for calibrating state price densities and present a slice sampling algorithm within the Markov chain Monte Carlo framework in Section 3.2.

We demonstrate the performance of the method through simulation studies in Sections 3.3 and 3.4. In the first simulation study, we considered a finite and discrete distribution. In the second simulation, study we considered three different continuous distributions, including a lognormal distribution, a mixture of  $t$ -distributions with two components, and a mixture of  $t$ -distributions with three components. The number of components is assumed to be known in the first simulation, and it is selected using a 10-fold cross-validation for the second simulation.

We apply our method to S&P 500 index options data in Section 3.5, where our method performs better than many existing methods and illustrates the need for both call and put data in order to effectively estimate both tails of the state price density using our Bayesian Quadrature approach. We provide a summary of the Bayesian Quadrature model in Section 4.4.

As a remark, the Bayesian Quadrature model is applicable to just European options, and we provide the Bayesian implied random tree model to calibrate the state price density using American options in Chapter 4.

### 3.1 The Bayesian Quadrature model

We define a Quadrature model as a finite distribution with parameters  $w = \{w_0, \dots, w_N\}$  and  $\theta = \{\theta_0, \dots, \theta_N\}$  for some non-negative integer  $N$ , where  $w$  and  $\theta$  are the weights and locations of the support points. Specifically, a Quadrature model has a density function as

$$f(x|w, \theta) = w_0\delta_{\theta_0}(x) + \dots + w_N\delta_{\theta_N}(x),$$

where  $\delta_\varpi(x)$  is the Dirac measure at the point  $\varpi$  such that

$$\delta_\varpi(x) = \begin{cases} 1, & \text{for } x = \varpi, \\ 0, & \text{otherwise.} \end{cases}$$

Let  $N_Q$  denote the number of support points in the Quadrature model,  $N_Q = N+1$ .

Let  $Q(w, \theta)$  denote a Quadrature with parameters  $(w, \theta)$ . In our Quadrature model, the locations  $\theta$  and the weights  $w$  of the support points in the finite-dimensional representation are random variables. When approximating the state price density using a Quadrature, the locations  $\theta$  are constrained to be non-negative and the weights  $w$  are constrained to be nonnegative quantities that sum to one. From a modeling perspective,  $Q(w, \theta)$  is a finite mixture distribution with the point measure as the component densities.

Once a Quadrature model on the state price density is placed, theoretic prices of European options are simply finite number terms of summations. Let  $K$  denote a strike price,  $r$  denote the risk-free interest rate, and  $T$  denote the time to maturity. Given the Quadrature model,  $Q(w, \theta)$ , as the state price density, the price of European call option equals

$$e^{-rT} \sum_{n=0}^N w_n \max(\theta_n - K, 0). \quad (3.1)$$

The price of a European put option equals

$$e^{-rT} \sum_{n=0}^N w_n \max(K - \theta_n, 0). \quad (3.2)$$

Let  $C_{ijk}(w, \theta)$  denote the option price calculated from  $Q(w, \theta)$ . The option has type  $i$ , strike price  $K_j$ , and time to maturity  $T_k$ . We assume that observed option prices are perturbed theoretic option prices generated from a Quadrature model. Specifically, the observed option price  $y_{ijkl}$  follows

$$y_{ijkl} = C_{ijk}(w, \theta) e^{\varepsilon_{ijkl}} \quad (3.3)$$

with identically and independently distributed (i.i.d.) normal errors  $\varepsilon_{ijkl} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ . Let  $y = \{y_{ijkl}\}$  denote the collection of all observed option prices. Therefore, the likelihood is

$$L(y|w, \theta, \sigma) = \prod_{i,j,k,l} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\log y_{ijkl} - \log C_{ijk}(w, \theta))^2}{2\sigma^2}}.$$

For simplicity, we assume a priori that the distribution of  $\sigma^2$  is an inverse-gamma distribution with shape parameter  $\alpha$  and scale parameter  $\beta$ , denoted by  $\sigma^2 \sim IG(\alpha, \beta)$ .

We assume a vague prior for the weights  $w_m$ , such that  $w$  has a Dirichlet distribution with parameter  $\bar{w} = (\bar{w}_0, \dots, \bar{w}_N)$ , denoted by  $w \sim D(\bar{w})$ . Let  $\mathbf{1}_A(\cdot)$  be an indicator function on the support set  $A$ . To avoid zero option prices, we assume a priori that the distribution of the locations  $\theta$  of the support points are uniformly distributed over the support set  $\Theta$ ,

$$\mathcal{A} = \{\theta_n \in \Re^+ : \theta_0 > \theta_1 > \dots > \theta_N, \theta_0 > K_{\max}, \theta_N < K_{\min}\},$$

and  $K_{\max}$  and  $K_{\min}$  are the maximum and the minimum of strike prices respectively. It is important to note that all three of these assumptions for the prior distributions can be changed in cases where appropriate subjective information is available.

Inferences for the parameters of interest is based on the posterior distribution of  $w$ ,  $\theta$ , and  $\sigma^2$ , which is

$$p(w, \theta, \sigma^2 | y) \propto L(y|w, \theta, \sigma^2) p(w|\bar{w}) p(\theta|K_{\max}, K_{\min}) p(\sigma^2|\alpha, \beta).$$

## 3.2 Slice Sampling

Before presenting the main inference algorithm, the Markov chain Monte Carlo (MCMC) algorithm, we describe the idea behind the slice sampling, which we use in part of larger MCMC algorithm. This description builds on the discussion in Damien et al. (1999) and Liu (2001).

If we wish to sample  $x$  and it has a density proportional to a product of  $k$  nonnegative functions  $f_i(x)$ ,

$$p(x) \propto \prod_{k=1}^K f_k(x), \quad (3.4)$$

we can introduce  $K$  auxiliary variables,  $u_1, \dots, u_K$ , as described in Edwards and Sokal (1988), and define a joint distribution for  $(x, u_1, \dots, u_K)$ , which is uniform over the region in which  $0 < u_k < f_k(x)$  for  $k = 1, \dots, K$ .

It is straightforward to show that the marginal distribution of  $x$  is given in Eq. (3.4). We call the set  $\{x : 0 < u_k < f_k(x)\}$  the sub-slice, and the interaction of  $\bigcap_{k=1, \dots, K} \{x : 0 < u_k < f_k(x)\}$  the slice. Samples from the augmented joint distribution can be generated by generating samples from each sub-slice, conditional on the current values of the remaining auxiliary variables and the variable of interest; then samples from the variable of interest  $x$  can be generated by sampling from the slice, conditional on the current value of the auxiliary variables. Sampling from conditional densities, which are created from a joint density that includes auxiliary variables is known as the slice sampling.

To be more explicit, let  $x^0$  denote the current value of  $x$  at each iteration, and let  $U(A)$  denote a uniform distribution over the set  $A$ . The algorithm to sample  $x$  having distribution in Eq. (3.4) is implemented as follows.

1. Start  $x$  randomly.
2. Repeat the following procedures until convergence.
  - (a) Sample  $u_k \sim U([0, f_k(x^0)])$  for  $k = 1, \dots, K$ .
  - (b) Find sub-slices  $S_k = \{x : f_k(x) > u_k\}$  and the slice  $S = \bigcap_{k=1}^K S_k$ .
  - (c) Sample  $x \sim U(S)$ .

To avoid possible floating-point underflow problems, it is typically better to calculate  $g_k(x) = \log(f_k(x))$  instead of  $f_k(x)$ . When a log transformation is used, the auxiliary variable becomes  $z_k = g_k(x^0) - e$ , where  $e$  is independently and identically distributed as an Exponential distribution with mean 1 (denoted by  $e \sim Exp(1)$ ) for  $k = 1, \dots, K$ . The algorithm to sample  $x$  having distribution in Eq. (3.4) is implemented as follows.

1. Start  $x$  randomly.
2. Repeat the following procedures until convergence.
  - (a) Sample  $z_k = g_k(x^0) - e$ , where  $e \sim Exp(1)$ , for  $k = 1, \dots, K$ .
  - (b) Find sub-slices  $S_k = \{x : g_k(x) > z_k\}$  and the slice  $S = \bigcap_{k=1}^K S_k$ .
  - (c) Sample  $x \sim U(S)$ .

We use the slice sampler in the main MCMC algorithm, which is used to generate samples of  $w$ ,  $\theta$  and  $\sigma^2$  that will have a distribution that converges to the posterior distribution. A summary of the MCMC algorithm is given below and a detailed derivation of the proposed slice samplers are given in Appendix A.

1. Start  $w$ ,  $\theta$  and  $\sigma^2$  at random values in the support of their joint distribution.
2. At each iteration, repeat the following procedures until the distribution of the samples from  $w$ ,  $\theta$  and  $\sigma^2$  converge.
  - (a) Sample  $w_m \sim U(T_m)$ , using the slice sampler described in the Appendix, and update  $w_M$  as  $1 - w_1 - \dots - w_{M-1}$ , where the set  $T_m$  is an open interval generated by Eq. (A.6), for  $m = 1, \dots, M - 1$ .
  - (b) Sample  $\theta_m \sim U(S_m)$ , using the slice sampler described in the Appendix, where the set  $S_m$  is an open interval generated by Eq. (A.12), for  $m = 1, \dots, M$ .
  - (c) Sample

$$\sigma^2 \sim IG \left( \alpha + \sum_{i,j,k,l} 1/2, \beta + \sum_{i,j,k,l} (\log y_{ijk} - \log G_{ij}(w, \theta))^2 / 2 \right).$$

### 3.3 Simulation study one

In this simulation study, we consider a finite, discrete distribution with four support points as described in the second column of Table 3.1, and assume that the number of support points is known or that  $N_Q = 4$ . We let  $K$  denote the number of different strike prices used to generate the synthetic data sets.

#### 3.3.1 Synthetic data sets

To investigate how the number of different number of strike prices affects the inference, we generate four data sets having 4, 8, 16, and 32 different strike prices. The strike prices are equally distributed in the interval between 5 and 30. We then calculate theoretical call and put option prices at these strike prices, and generate additional 500 option prices using Eq. (3.3) with  $\sigma^2$  at 0.05. Scatterplots of these four data sets are in Figure 3.1.

#### 3.3.2 MCMC inference

For each MCMC analysis, we begin with random starting values for the parameters, discard the first 100,000 burn-in samples, and make inference using the the following 1000 samples.

The trace plot, autocorrelation function (ACF) plot, and the kernel density estimation (KDE) of the log-likelihood (LL) are used to assess the the convergence of the MCMC algorithm. When the sampling distribution appears to have converged to joint distribution of the simulation is satisfied, we assess model fits using residuals plots and summarize the posterior densities of parameters of interest.

Figure 3.2 shows the trace plots, ACF plots, and KDEs of LL of these four data sets. These plots show that the sampling density appears to have converged.

Figure 3.3 depicts the 90% credible regions of residuals plots for put and call options. These residuals have zero means and constant variances across strike prices, indicating that the proposed method produces a good model fit.

Figure 3.4 shows the trace plots, ACF plots, and KDEs of  $\sigma^2$  of these four data sets, and these plots show that the sampling density of  $\sigma^2$  appears to have converged.

**Table 3.1.** Numerical results for the first simulation.

We generate four data, each with 4, 8, 16, and 32 strike prices, respectively. We report the true parameter values, and the posterior mean and the standard deviation in parenthesis of the parameter. The values of standard deviations below for  $\sigma^2$  and  $w$  are a multiple of 100 and 10, respectively.

Num. Strikes		4	8	16	32
True Value					
$\sigma_Y^2$	<b>0.05</b>	0.05 (0.11)	0.05 (0.08)	0.06 (0.06)	0.05 (0.04)
$w_0$	<b>0.1</b>	0.08 (0.06)	0.10 (0.01)	0.10 (0.00)	0.10 (0.01)
$w_1$	<b>0.2</b>	0.23 (0.08)	0.28 (0.03)	0.34 (0.01)	0.20 (0.02)
$w_2$	<b>0.3</b>	0.30 (0.27)	0.38 (0.02)	0.45 (0.01)	0.30 (0.03)
$w_3$	<b>0.4</b>	0.39 (0.27)	0.24 (0.03)	0.11 (0.02)	0.40 (0.02)
$\theta_0$	<b>5</b>	3.68 (0.60)	4.90 (0.05)	5.07 (0.02)	4.99 (0.01)
$\theta_1$	<b>15</b>	14.76 (0.15)	16.35 (0.06)	16.89 (0.02)	15.01 (0.08)
$\theta_2$	<b>25</b>	25.55 (0.35)	29.31 (0.07)	32.19 (0.02)	25.02 (0.08)
$\theta_3$	<b>35</b>	35.11 (0.44)	37.12 (0.10)	39.70 (0.09)	34.98 (0.01)

Figure 3.5 shows the trace plots of  $w_n$  and  $\theta_n$  for  $n = 0, \dots, 3$ , and Table 3.1 summarizes the posterior means and standard deviations of these parameters.

In summary, Figure 3.5 and Table 3.1 show that the posterior means of these parameters are closer to the true values and the posterior standard deviations decrease as the number of strike prices increases.

### 3.4 Simulation study two

To understand how the proposed method works for a more realistic SPD, we consider the following three distributions,

$$LN(6.7319, 0.0031), \quad (3.5)$$

$$0.4 t_5(800, 100^2) + 0.6 t_5(1100, 60^2), \quad (3.6)$$

$$0.3 t_5(650, 60^2) + 0.5 t_5(850, 100^2) + 0.2 t_5(1150, 70^2). \quad (3.7)$$

### 3.4.1 Synthetic data sets

For each density, we uniformly select 50 strike prices from a proper interval, calculate the theoretical call and put option prices, and perturb the theoretic option prices with normal errors using Eq. (3.3) with  $\sigma^2 = 0.001$  for generating addition 50 option prices as a synthetic data set. Figure 3.6 gives the KDE and scatter plots of the option prices against strike prices for these data sets.

### 3.4.2 Bayes factors for model selection

Selecting the number of support point in the Quadrature model is a model selection problem. Consider selecting one model between two models,  $M_1$  and  $M_2$ , given data  $D$ . Let  $\Theta_i$  be the model parameter (may be of higher dimension) of  $M_i$ . The Bayes factor  $B$  between  $M_1$  and  $M_2$  is defined as the posterior odds in favor of  $M_1$  of the likelihood divided by the prior odds in favor of  $M_1$ , and can be simplified by the Bayes rule as follows.

$$\begin{aligned} B(M_1, M_2) &= \frac{P(M_1|D)}{P(M_2|D)} \div \frac{P(M_1)}{P(M_2)} \\ &= \frac{P(D|M_1)}{P(D|M_2)} \\ &= \frac{\int P(\Theta_1|M_1)P(D|\Theta_1, M_1)d\Theta_1}{\int P(\Theta_2|M_2)P(D|\Theta_2, M_2)d\Theta_2} \end{aligned}$$

where  $P(D|M_i)$  is known as the marginal likelihood for model  $i$ .

Bayesian model comparison is based on Bayes factors. A value of  $B > 1$  means that the data indicate that  $M_1$  is more strongly supported by the data under consideration than  $M_2$ . Therefore, model selection can be solved based on the marginal likelihood for a given model. A model that produces higher marginal likelihood is preferable.

However, the marginal likelihood is a high-dimensional integral, which can not be calculated analytically and has to be approximated by numerical methods. Methods for approximating the marginal likelihood include Laplace approximation, importance sampling, Gaussian quadrature and MCMC simulations. Details about these methods are out of the scope of this chapter. We refer to Kass and Raftery (1995) for references.

We use the harmonic mean of the likelihoods in the MCMC algorithm as an approximation to the marginal likelihood. However, as most log-likelihood has value about -2000, the likelihood is very pretty close to zero. Consequently, it is difficult to calculate the harmonic mean of the likelihood for approximating the marginal likelihood. Therefore, we calculate the harmonic mean of the log-likelihood to approximate the marginal log-likelihood for Bayesian model selection.

Figure 3.7 and Table 3.2 summarize the marginal log-likelihood with respect to the number of support points used in the Quadrature model. Although the marginal log-likelihood is preferred in theory, it performs poorly in practice for this example and does not appear to stable when the number of support points increases. This is possibly because of the slow convergence of approximate marginal log-likelihood using harmonic mean. In the next subsection, we implement cross-validation as an alternative method to select the number of support points used in the Quadrature model.

### 3.4.3 Cross-validation for selecting the number of support points

Cross-validation suggests an optimal model in terms of correctness of out-sample forecasting using the in-sample data set. A 10-fold cross-validation splits the data set randomly into ten sub-data sets and for each sub-data set, we use the remaining nine data sets to calibrate the model and predict the remaining data set. For each MCMC analysis, we begin with random starting values for the parameters, discard the first 10,000 burn-in samples, and make inference using the the following 1000 samples. Figure 3.8 shows the trace plots, ACF plots, and KDEs of LL of these four data sets. These plots show that the sampling density appears to have converged.

Figure 3.9 shows the boxplots of the prediction errors with respect to the number of support points for the three densities (3.5)–(3.7), and Table 3.3 reports the median, the 5-th, and the 95-th quantile of prediction errors.

Based on these the criteria that the prediction error has stabilized and that we want as parsimonious model as possible, we selected  $N_Q = 14$  for the Log Normal data,  $N_Q = 15$  for the mixture of two t-distributions data, and  $N_Q = 17$  for the mixture of three t-distribution data. Figure 3.10 gives the residuals plots,

**Table 3.2.** Marginal log-likelihood for the simulation study two.

We summarize the marginal log-likelihood with respect to the number of support points used in the Quadrature model for Bayesian model selection. The marginal log-likelihood is approximate by the harmonic mean of the log-likelihood in the MCMC algorithm.

$N_Q$	LN	$t_2$	$t_3$
5	-2512.54	-2512.44	-2512.55
6	-2512.34	-2512.31	-2513.01
7	-2510.55	-2510.98	-2510.64
8	-2509.33	-2509.41	-2509.13
9	-2512.46	-2512.46	-2512.75
10	-2510.83	-2510.67	-2510.94
11	-2510.01	-2509.88	-2509.95
12	-2509.78	-2509.73	-2509.94
13	-2510.22	-2510.16	-2510.17
14	-2510.61	-2510.55	-2510.53
15	-2513.26	-2513.2	-2513.21
16	-2509.84	-2509.78	-2509.8
17	-2512.64	-2512.6	-2512.63
18	-2511.79	-2511.47	-2511.63
19	-2510.75	-2510.72	-2510.74
20	-2510.92	-2510.9	-2510.91
21	-2511.78	-2511.75	-2511.78
22	-2512.07	-2512.05	-2512.07
23	-2510.39	-2510.32	-2510.34
24	-2509.3	-2509.33	-2509.3
25	-2510.46	-2510.45	-2510.46

indicating that the method is able to recover the prices across a range of strick prices.

### 3.4.4 MCMC inference

Figure 3.11 gives the posterior estimate of the empirical density of the SPDs, showing that the method is able to recover the essential structure of the densities used to generate the option prices.

**Table 3.3.** Numerical results for the second simulation study.

The 5-th quantile ( $Q_5$ ), median ( $Q_{50}$ ), and 95-th quantile ( $Q_{95}$ ) of the prediction errors in a 10-fold cross-validation for data generated from a lognormal distribution (LN), a mixture of  $t$ -distributions with two components (t2) and a mixture of  $t$ -distributions with three components (t3) are reported with respect to the number of support points, which go from  $N_Q = 5$  to  $N_Q = 20$ . The results for optimal number of support points are underlined, where the optimal number is based on the criteria that the prediction error has stabilized and that we want as parsimonious model as possible.

True Dist. $N_Q$	LN			t2			t3		
	$Q_5$	$Q_{50}$	$Q_{95}$	$Q_5$	$Q_{50}$	$Q_{95}$	$Q_5$	$Q_{50}$	$Q_{95}$
5	1.77	1.98	2.46	1.37	1.52	2.00	1.95	2.35	3.05
6	1.06	1.25	1.41	1.01	1.09	1.74	0.99	1.19	1.68
7	0.80	0.91	2.79	0.63	0.81	2.20	0.78	0.94	1.09
8	0.68	0.76	1.13	0.55	0.84	1.80	0.65	0.78	1.88
9	0.54	0.79	2.45	0.54	0.62	1.00	0.59	0.69	1.19
10	0.59	0.69	2.56	0.49	0.74	1.08	0.56	0.64	0.95
11	0.52	0.59	1.14	0.50	0.61	0.76	0.54	0.63	0.76
12	0.56	0.63	0.72	0.48	0.59	1.02	0.49	0.60	0.65
13	0.47	0.56	0.70	0.48	0.56	0.68	0.48	0.61	0.79
14	<u>0.48</u>	<u>0.57</u>	<u>0.66</u>	0.42	0.53	0.64	0.47	0.52	0.67
15	0.46	0.54	0.75	<u>0.43</u>	<u>0.51</u>	<u>0.57</u>	0.47	0.52	0.59
16	0.46	0.55	0.57	0.43	0.51	0.56	0.45	0.53	0.63
17	0.46	0.52	0.59	0.48	0.53	0.59	<u>0.46</u>	<u>0.54</u>	<u>0.59</u>
18	0.45	0.53	0.57	0.42	0.50	0.61	0.46	0.53	0.57
19	0.45	0.53	0.57	0.42	0.49	0.55	0.44	0.51	0.57
20	0.45	0.52	0.56	0.43	0.49	0.55	0.45	0.51	0.56
21	0.44	0.53	0.87	0.43	0.49	0.54	0.45	0.51	0.56
22	0.44	0.51	0.56	0.42	0.48	0.56	0.45	0.50	0.56
23	0.44	0.52	0.56	0.43	0.48	0.57	0.45	0.50	0.59
24	0.43	0.52	0.55	0.43	0.49	0.54	0.44	0.50	0.59
25	0.43	0.51	0.56	0.42	0.48	0.58	0.44	0.50	0.55

### 3.5 Empirical study

In order to understand how the proposed model would work in practice, we apply it to the four data sets used in Aït-Sahalia and Lo (1998). There are a number of competing methods for estimating the SPD, using this data, making it ideally

suit for comparison.

### 3.5.1 Using data set in Aït-Sahalia and Lo (1998)

We begin with random starting values for the parameters, discard the first 50,000 burn-in samples, and make inference using the the following 1000 samples for each MCMC analysis. Table 3.4 provides a summary of the performance of the proposed Bayesian Quadrature model with a number of competing methods, using an  $R^2$  measure of fit.

We consider a range of different number of support points from five up to ten and two different ways of incorporating the input data. Because of the put-call parity and the assumed symmetry implied with regards to input data, existing methods only use call data as input. We found that for our procedure, as we are explicitly estimating support points based on a payoff function, using only call data and hence call payoff functions, results in considerable uncertainty about the lower tail of the SPD. Please see Figure 3.12.

Even though the lower tail is not as well defined, the model still fits very well and has an extremely tight predictive range as shown in Figures 3.12 to 3.13 .

### 3.5.2 Using both call and put options

We found that the ability to estimate the tail can be improved, along with the ability to recover the option prices, by using the put-call parity to create both call and put input data. This is done by creating a set of put data based on the call data and then calibrating the model using both sets of data. See numerical results in Tables 3.5.

Using this extended input data, the lower tail behaves in a fashion consistent with the upper tail, see Figure 3.14 and the fit actually improves, see Figures 3.14 to 3.15.

Not only is the fit of the proposed method improved using additional put data, the model turns out to be a strong competitor to the existing methods in the literature. Using the  $R^2$  as a measure of this, as is used in earlier studies, the

**Table 3.4.** Numerical results using call options for the empirical study.

We report the 5-th quantile ( $Q_5$ ) and the 95-th quantile ( $Q_{95}$ ) of the posterior distribution of  $R^2$  produced by our Bayesian Quadrature approach with different number of support points  $N_Q = 5$  up to  $N_Q = 10$  for these four data sets used in Aït-Sahalia and Lo (1998), where these data sets vary with respect to the time to maturity in days (T), trading date (Date), and number of observations (#). In addition, we report the  $R^2$  values for the same data sets, as summarized by Giacomini et al. (2008), where "Mixtures of  $t$ " refers to the method used by Giacomini et al. (2008), "Abadir-Rockinger" refers to the functional densities used in Abadir and Rockinger (2003), "Hermite" refers to the method used by Jondeau and Rockinger (2001), "Jumps" refers to the Malz-type jump-diffusion model in Malz (1996), "Mixtures of lognormal" uses the mixtures of lognormal distributions, and "lognormal" uses a lognormal density.

T	46		58		98		176	
Date	05/03/1993		10/20/1993		06/11/1993		09/23/1993	
#	15		20		17		24	
Bayesian Quadrature								
$N_Q$	$Q_5$	$Q_{95}$	$Q_5$	$Q_{95}$	$Q_5$	$Q_{95}$	$Q_5$	$Q_{95}$
5	0.996008	0.997358	0.999718	0.999820	0.999048	0.999303	0.999159	0.999414
6	0.999136	0.999339	0.999774	0.999858	0.999693	0.999757	0.999809	0.999865
7	0.999363	0.999542	0.999748	0.999812	0.999465	0.999592	0.999905	0.999933
8	0.998944	0.999281	0.999869	0.999887	0.999645	0.999755	0.999653	0.999700
9	0.999607	0.999701	0.999931	0.999958	0.997406	0.998663	0.999668	0.999724
10	0.999705	0.999792	0.999960	0.999964	0.999974	0.999989	0.999769	0.999821
Mixtures of $t$	0.99968		0.99982		0.99959		0.99989	
Abadir-Rockinger	0.999922		0.999911		0.999661			
Hermite	0.997214		0.984918		0.993403			
Jumps	0.997926		0.991013		0.995244			
Mixtures of lognormal	0.998267		0.990682		0.996039			
lognormal	0.951508		0.928570		0.980671			

**Table 3.5.** Numerical results using call and put options for the empirical study.

We report the 5-th quantile ( $Q_5$ ) and the 95-th quantile ( $Q_{95}$ ) of the posterior distribution of  $R^2$  produced by our Bayesian Quadrature approach, using the original Call data and Put data built from the original Call data using the put-call parity, with different number of support points  $N_Q = 5$  up to  $N_Q = 10$  for these four data sets used in Aït-Sahalia and Lo (1998), where these data sets vary with respect to the time to maturity in days (T), trading date (Date), and number of observations (#).

T	46	58	98	176
Date	05/03/1993	10/20/1993	06/11/1993	09/23/1993
#	30	40	34	48
Bayesian Quadrature				
$N_Q$	$Q_5$	$Q_{95}$	$Q_5$	$Q_{95}$
5	0.999112	0.999768	0.999545	0.999872
6	0.999376	0.999891	0.999811	0.999932
7	0.999888	0.999969	0.999956	0.999986
8	0.999918	0.999971	0.999985	0.999994
9	0.999479	0.999959	0.999985	0.999996
10	0.999900	0.999977	0.999991	0.999998
	$Q_5$	$Q_{95}$	$Q_5$	$Q_{95}$
	0.999384	0.999663	0.999736	0.999859
	0.999893	0.999943	0.999906	0.999950
	0.999949	0.999976	0.999950	0.999972
	0.999906	0.999947	0.999853	0.999915
	0.999984	0.999995	0.999814	0.999904
	0.999978	0.999994	0.999989	0.999993

Bayesian Quadrature model clearly dominates the lognormal density assumption and it can improve on all of the competing methods depending on the whether the just the call data or call and put data are used and on the number of support points.

One advantage of the Bayesian Quadrature method is that it gives a posterior distribution of the goodness of fit  $R^2$ , or it naturally provides a range of reasonable values. We report the 5-th and 95-th quantile of the posterior distribution of  $R^2$  values in Tables 3.4 and 3.5. One interesting point is that the performance of the Bayesian Quadrature model improves as the number of support points increase.

In the best case that we reported, with 10 support points in the Quadrature model, and using both call and put data, the lower 5-th quantile for the  $R^2$  value is better than the  $R^2$  value for all of the competing methods for all of the data sets considered, except in one case. When there are just 15 observations, the “Abadir-Rockinger” density function approach presented in Abadir and Rockinger (2003) has an  $R^2$  that is slightly higher than the 5-th quantile of best performing Bayesian Quadrature model.

In general, even the Bayesian Quadrature model with just five support points performs at a comparable level with the two strongest performing methods, the “Abadir-Rockinger” density function approach and the mixtures of  $t$  approach as reported in Giacomini et al. (2008) and it dominates the remaining approaches, the “Hermite” approach used in Jondeau and Rockinger (2001), the “Jumps” approach or the Malz-type jump-diffusion model used in Malz (1996), the “Mixtures of lognormal” which uses a mixture of lognormal distributions, and the ”lognormal” which uses a lognormal distributions. The parsimony of this model, just 10 parameters, suggests that there is a limited amount of information in the original option data and it points to the fact that, in practice, the Bayesian Quadrature model is a very computationally efficient tool for pricing options.

### 3.6 Summary

We propose a Bayesian Quadrature model as a nonparametric approach for the state price density estimation, and provide an efficient MCMC algorithm using slice sampling, for making inference about the resulting finite dimensional approx-

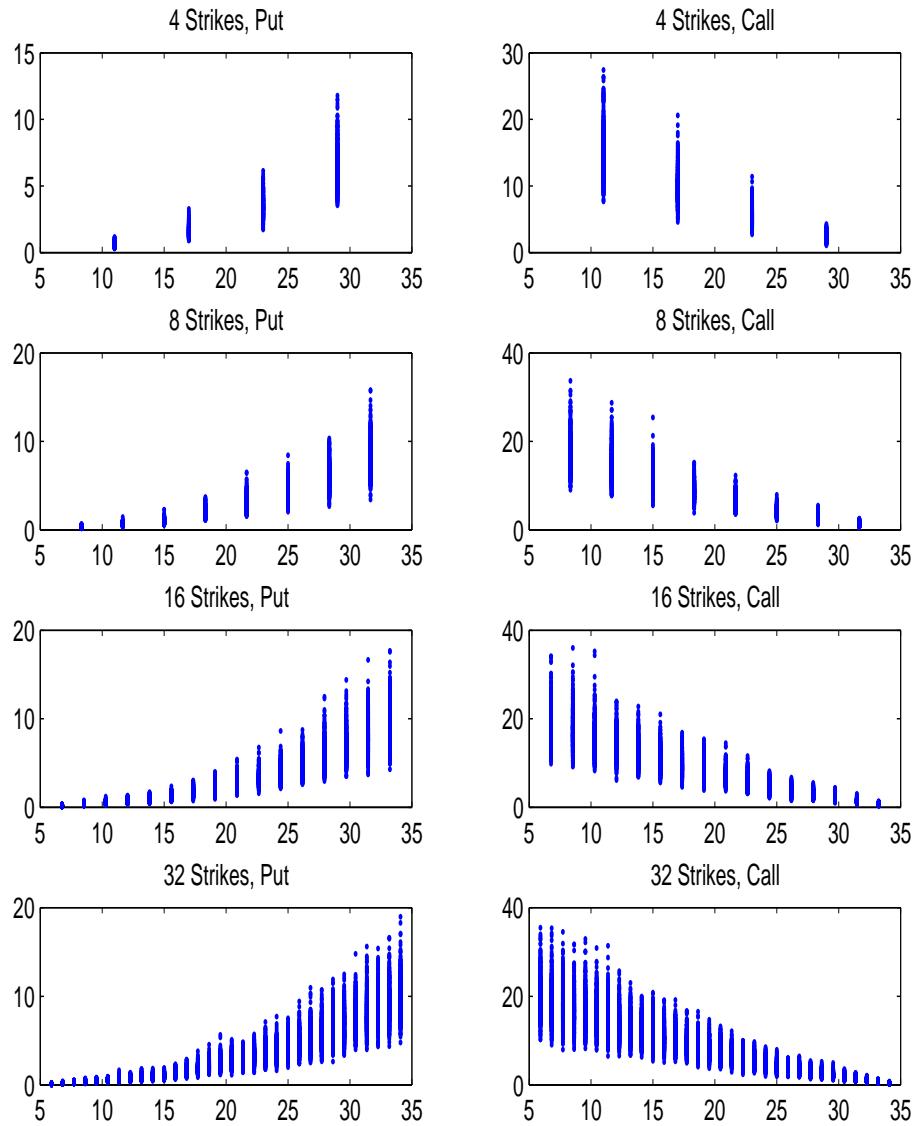
imation to the state price density.

In two simulation studies, we demonstrate that the proposed inference method can correctly recover the option prices and the underlying state price density used to generate these options price, for a variety of different distributions. We provide an empirical study that demonstrates both the ability to recover the underlying option prices and the parsimony of the proposed Bayesian Quadrature method.

The proposed method performs at a comparable level to the best competing methods with as few as five support points in the Bayesian Quadrature and dominate all of the competing methods when the number of support points are increased to ten. One finding from our empirical study is that the Bayesian Quadrature model does better, when it uses both put and call data (were the put data is generated using the put-call parity and using call data).

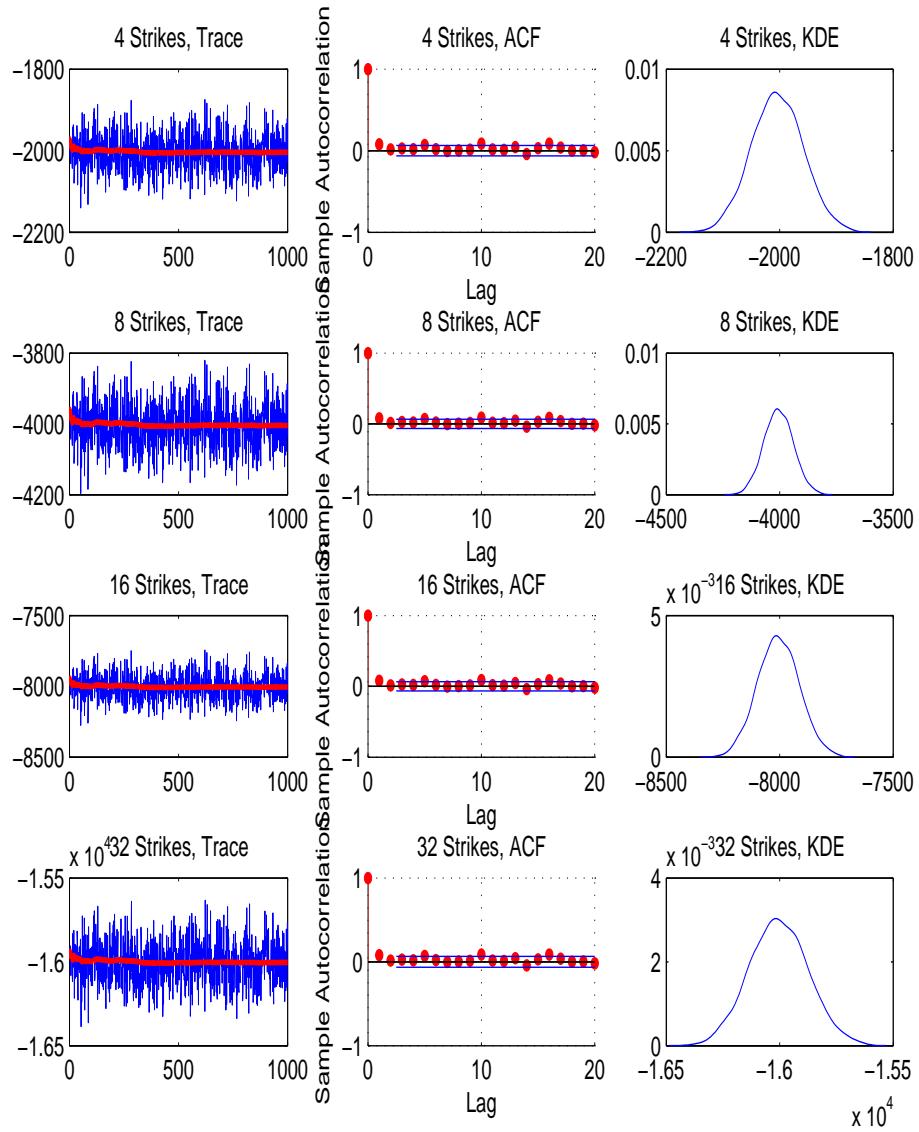
The Bayesian modeling framework, reports a distribution for a range of different summaries of interest. It allows for a predictive distribution (conditional on the observed data) for option prices, it results in a posterior distribution of model fit criteria (e.g.  $R^2$ ) and a posterior distribution of the state price density. It also allows the researcher to include different prior distributions, different from the uninformative distributions that we used in our model description, if the researcher has subjective information that merits more informative priors.

The Bayesian Quadrature model produces good model fit and is easy to implement. However, it calibrates the state price density using only European options. On the other hand, American options can be exercised prior to maturity and hence are very frequently traded. However, to our knowledge, nonparametric methods that can be used to calibrate the state price density using American option prices rarely exist. As result, we provide the Bayesian implied random tree model as a solution in the next chapter. Briefly speaking, the Bayesian implied random tree model begins with the Quadrature model for the state price density at maturity, and builds a unique binomial tree associated with the Quadrature model under certain assumptions, so that American options can be priced based on the binomial tree using backward induction.

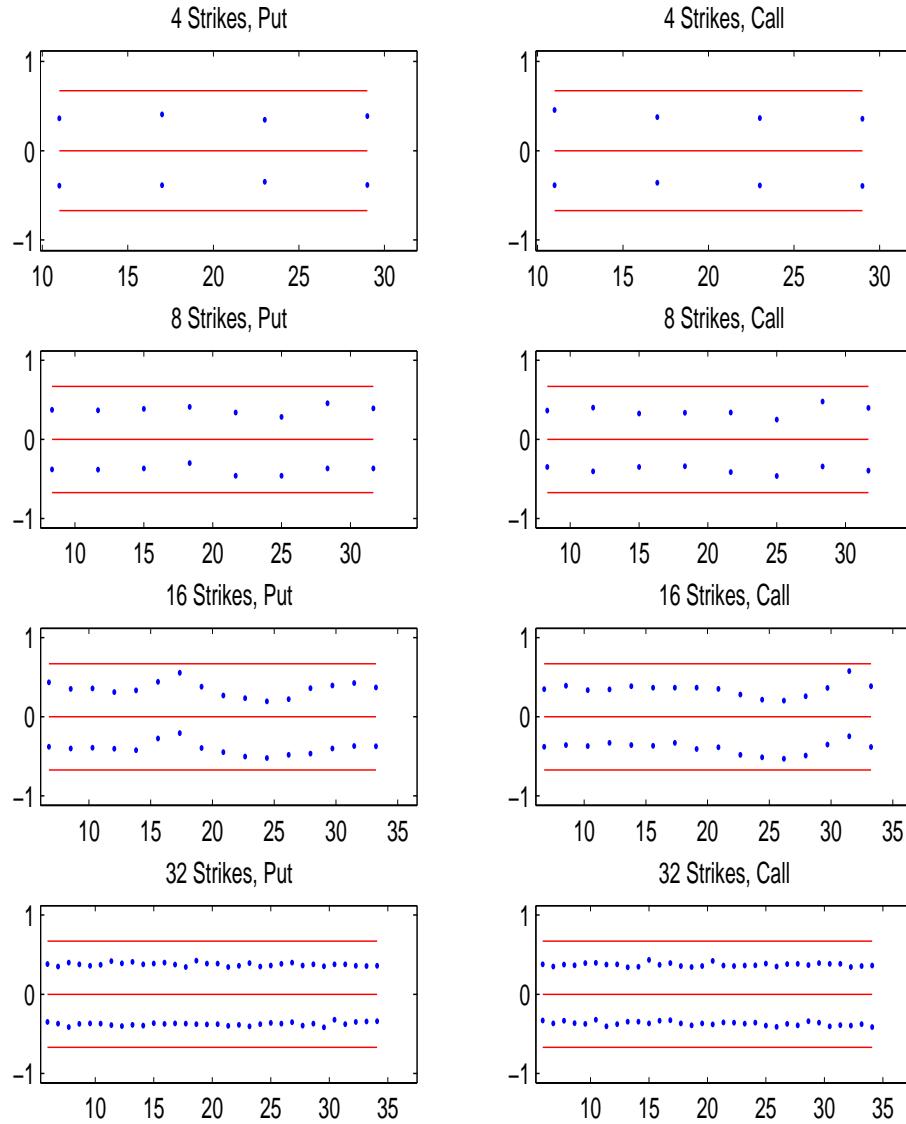


**Figure 3.1.** Scatter plots for the first simulation study.

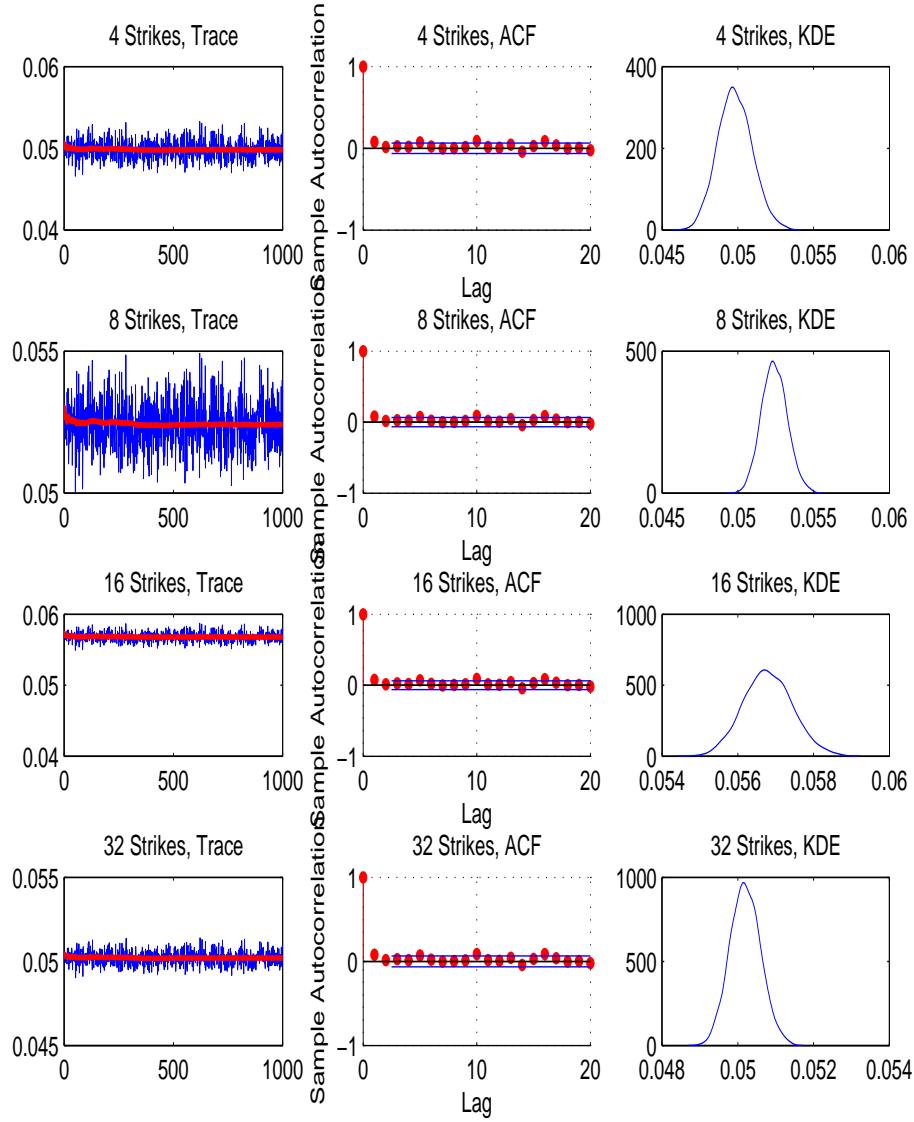
Scatter plots, for the first simulation, of put and call data against strike prices in the left and right panels, respectively, for four data, each with 4, 8, 16, and 32 strike prices, respectively.



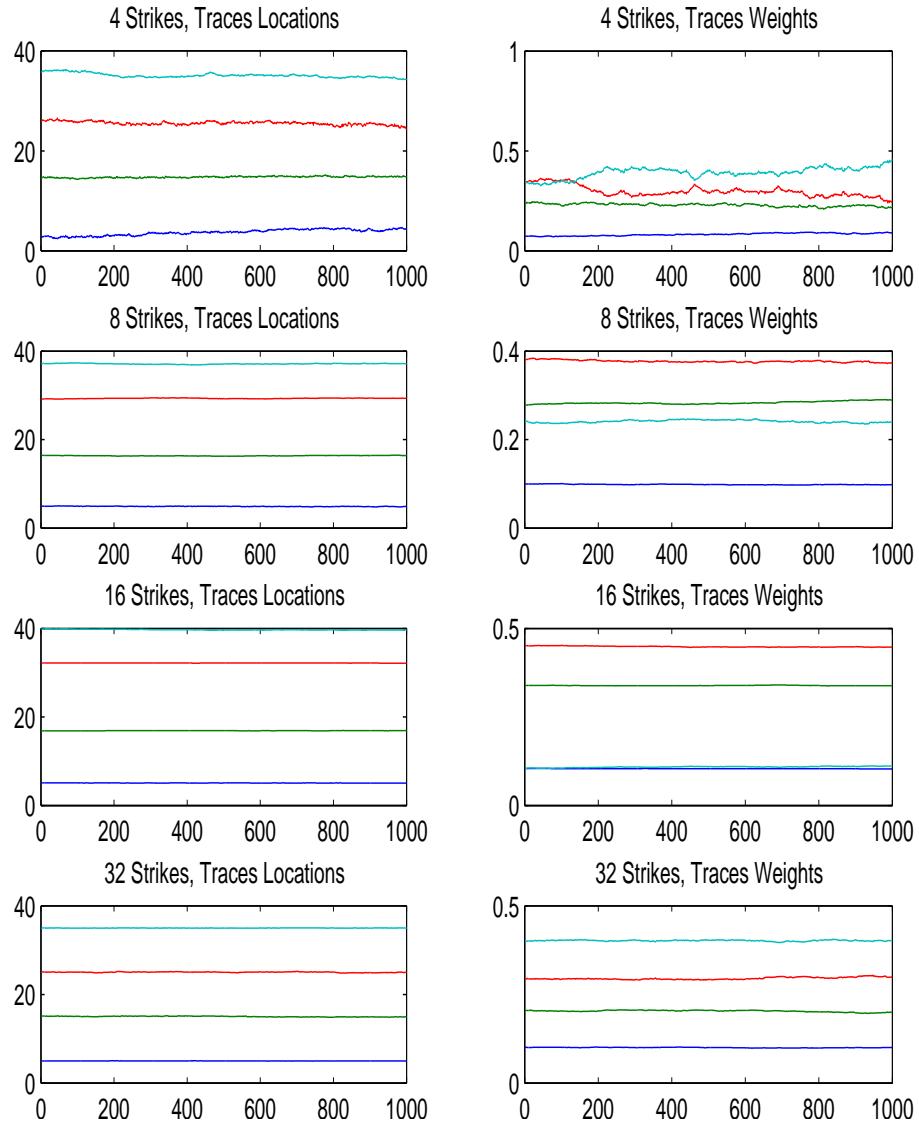
**Figure 3.2.** Convergence MCMC analysis for the first simulation study. The trace plot, autocorrelation function (ACF) plot, and kernel density estimate (KDE) of the log-likelihood for four data sets, each with 4, 8, 16 and 32 strike prices, respectively.



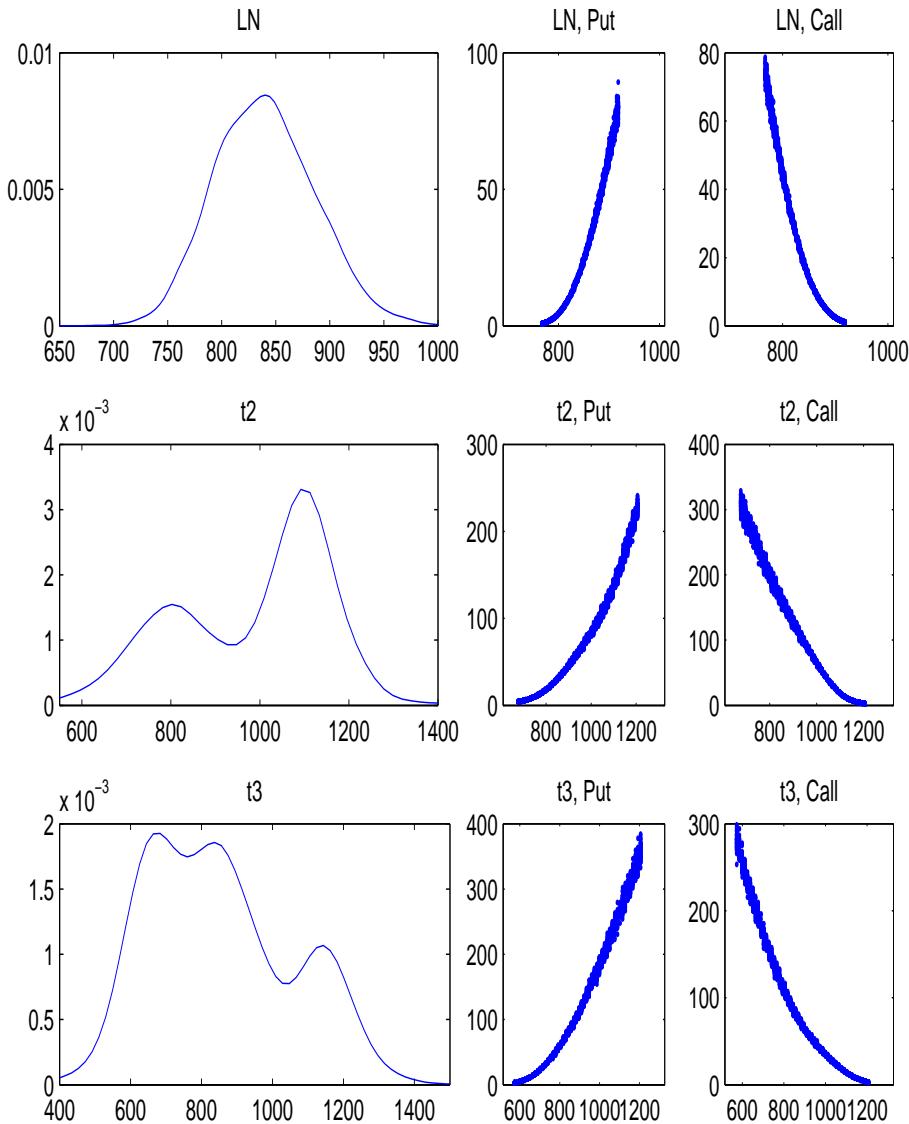
**Figure 3.3.** Credible region of residuals plots for the first simulation study. The 90% credible region of residuals and plots of the residuals against strike prices for put and call in the left and right panels, respectively, for four data sets, each with 4, 8, 16 and 32 strike prices, respectively.



**Figure 3.4.** Trace plots and convergence test of  $\sigma^2$  for the first simulation study. The trace plot, autocorrelation function (ACF) plot, and kernel density estimate (KDE) of  $\sigma^2$  for four data sets, each with 4, 8, 16 and 32 strike prices, respectively.

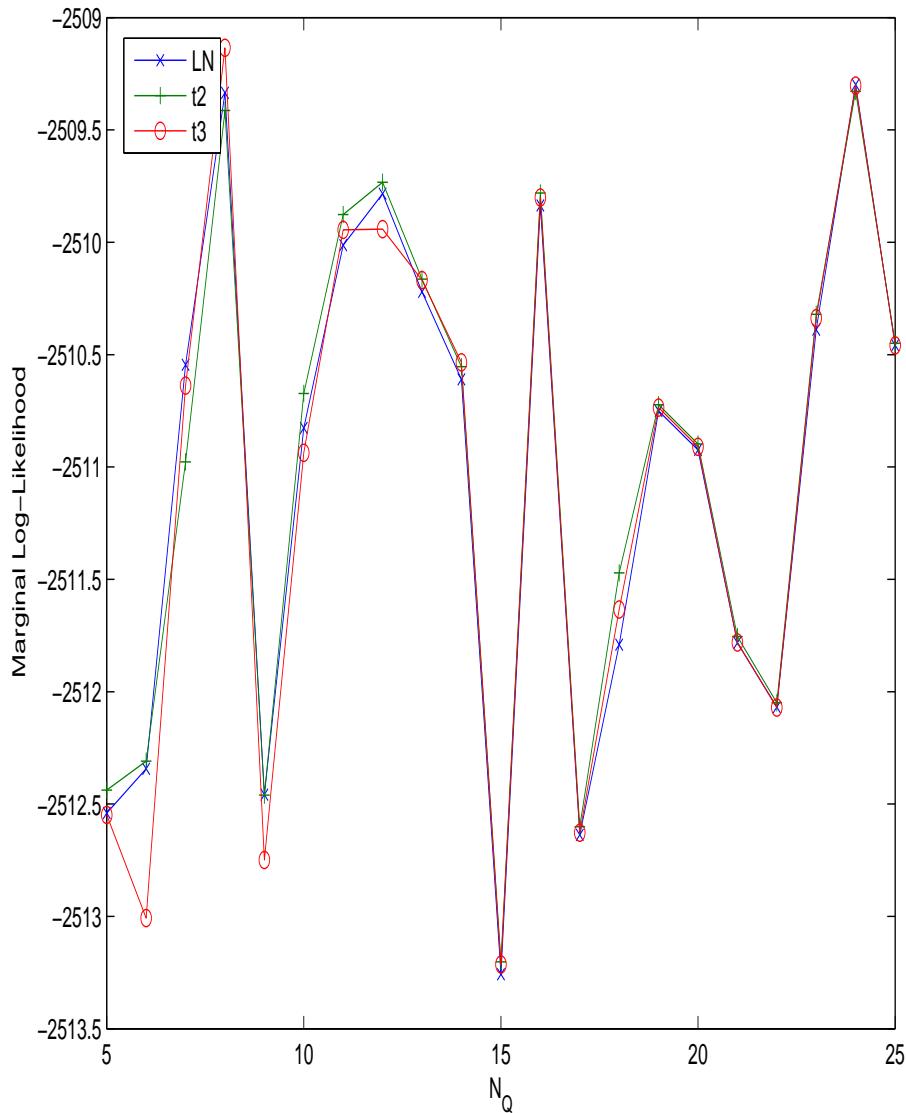


**Figure 3.5.** Trace plots of parameters for the first simulation study.  
Trace plots of the locations,  $\theta_m$ , and weights,  $w_m$ , for  $m = 1, \dots, 4$  for four data sets, each with 4, 8, 16 and 32 strike prices, respectively.



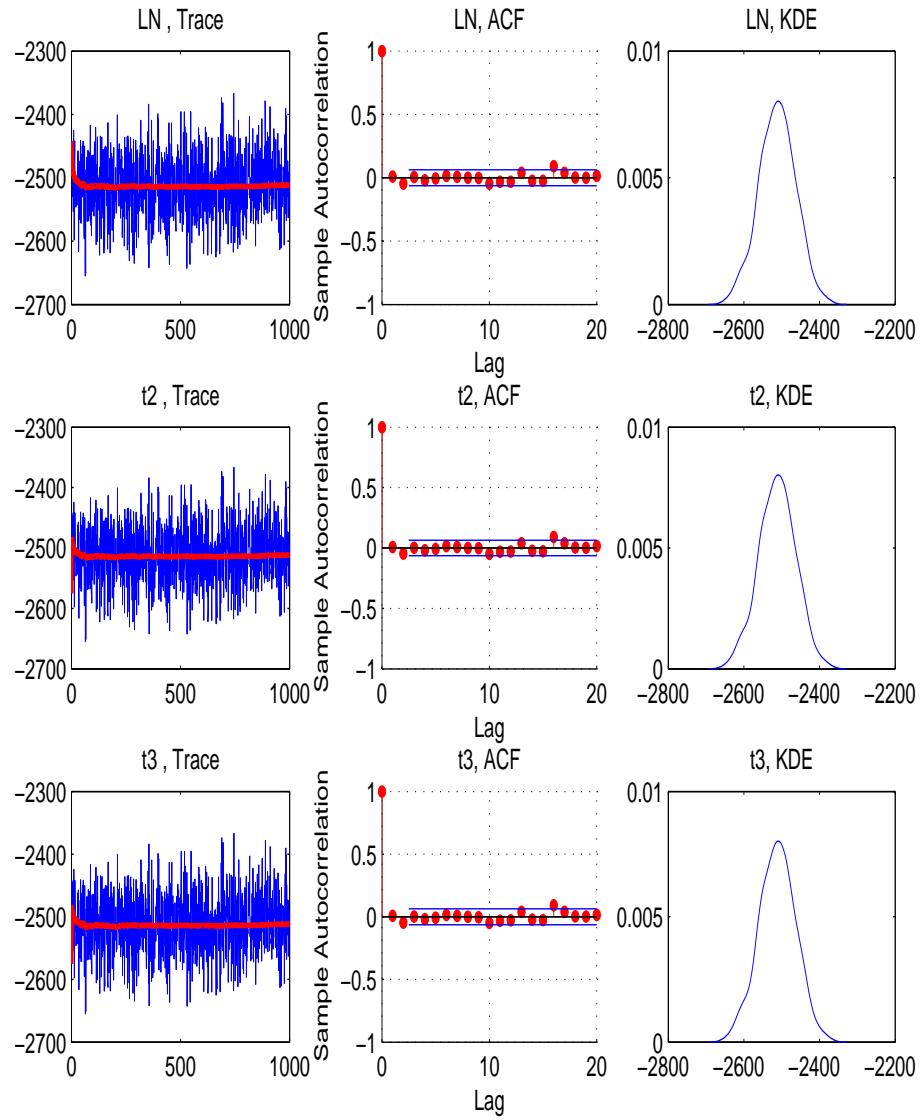
**Figure 3.6.** True models and scatter plots of option prices for the second simulation study.

The true density, scatter plots of put and call data against strike prices for three data, where data is generated from a lognormal distribution (LN), a mixture of  $t$ -distributions with two components (t2), and a mixture of  $t$ -distributions with three components (t3).



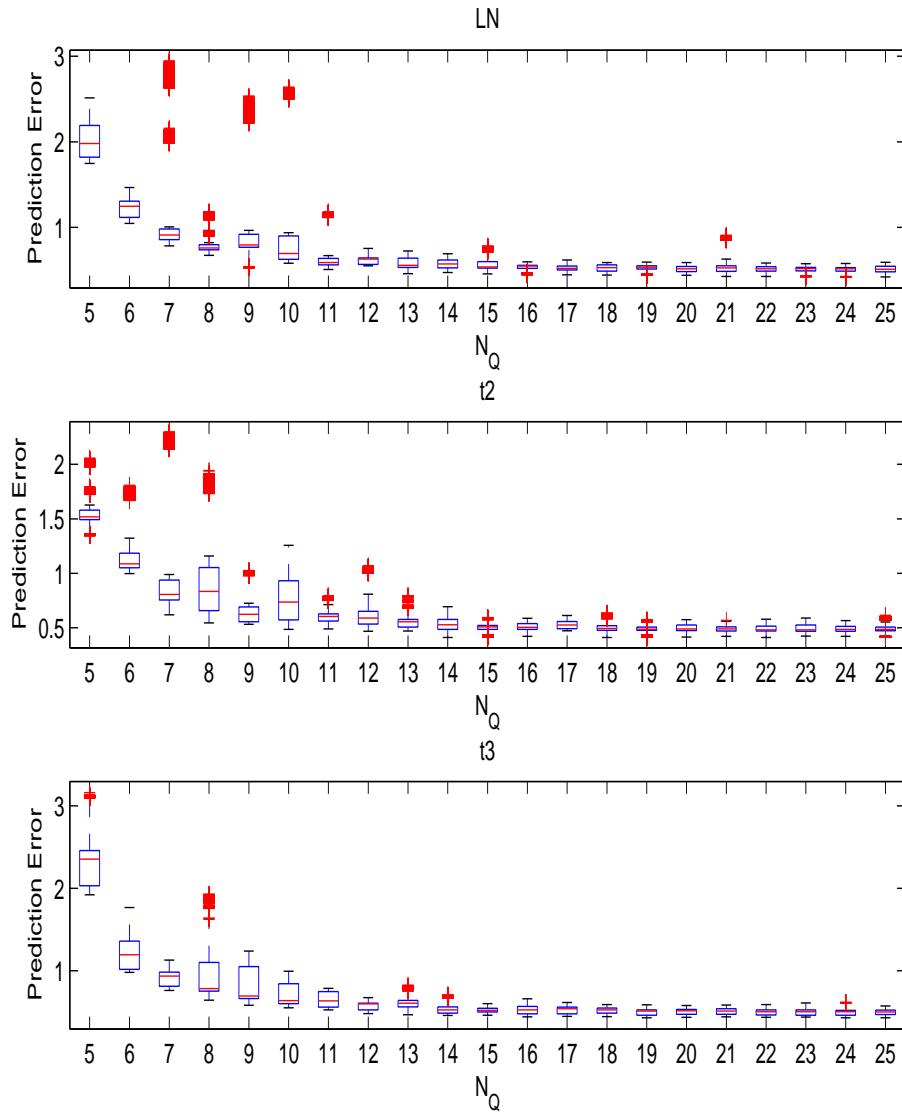
**Figure 3.7.** Marginal log-likelihood for the simulation study two.

We give the scatter plot of marginal log-likelihood with respect to the number of support points in the Quadrature model for three data, where data is generated from a lognormal distribution (LN), a mixture of  $t$ -distributions with two components (t2), and a mixture of  $t$ -distributions with three components (t3). The marginal log-likelihood is approximate using the harmonic mean of the log-likelihood in the MCMC algorithm.



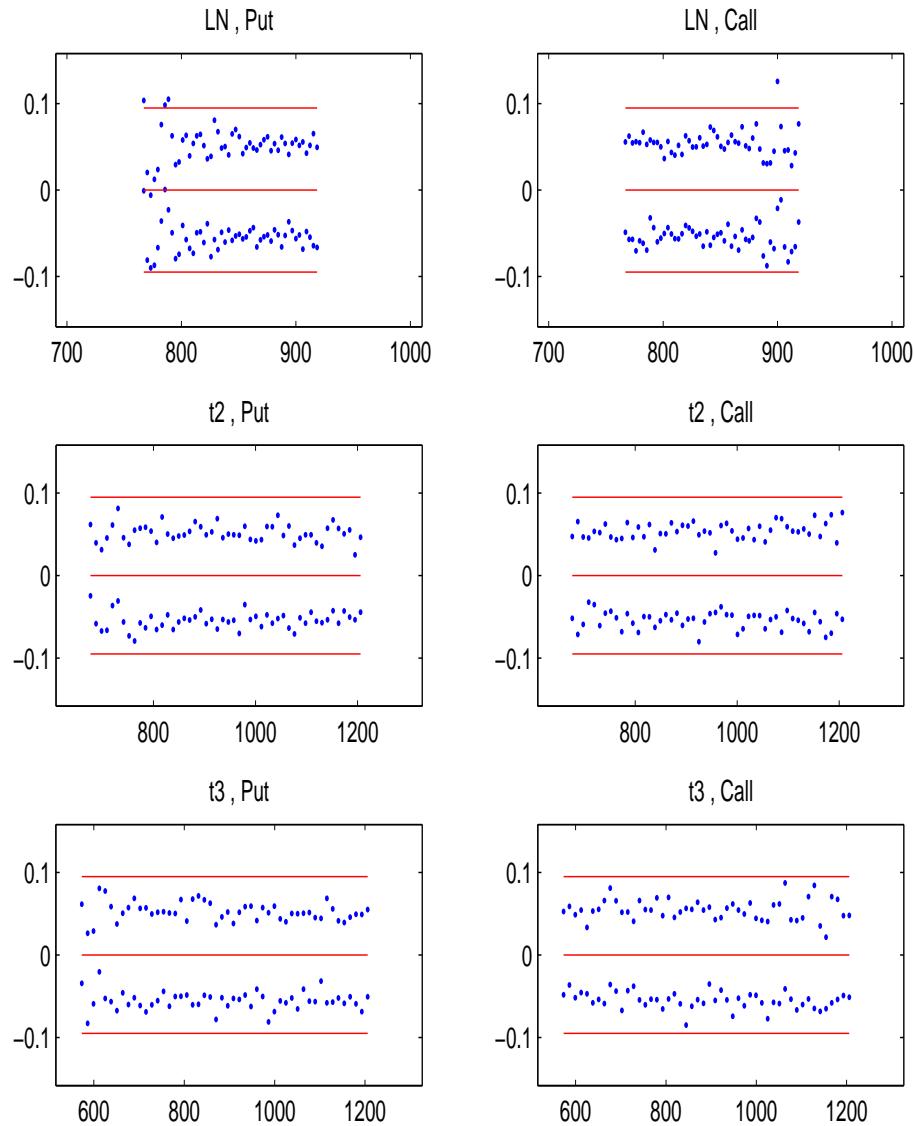
**Figure 3.8.** Convergence MCMC analysis for the second simulation.

The trace plot, autocorrelation function (ACF) plot, and kernel density estimate (KDE) of the log-likelihood for three data, where data is generated from a lognormal distribution (LN), a mixture of  $t$ -distributions with two components (t2), and a mixture of  $t$ -distributions with three components (t3).

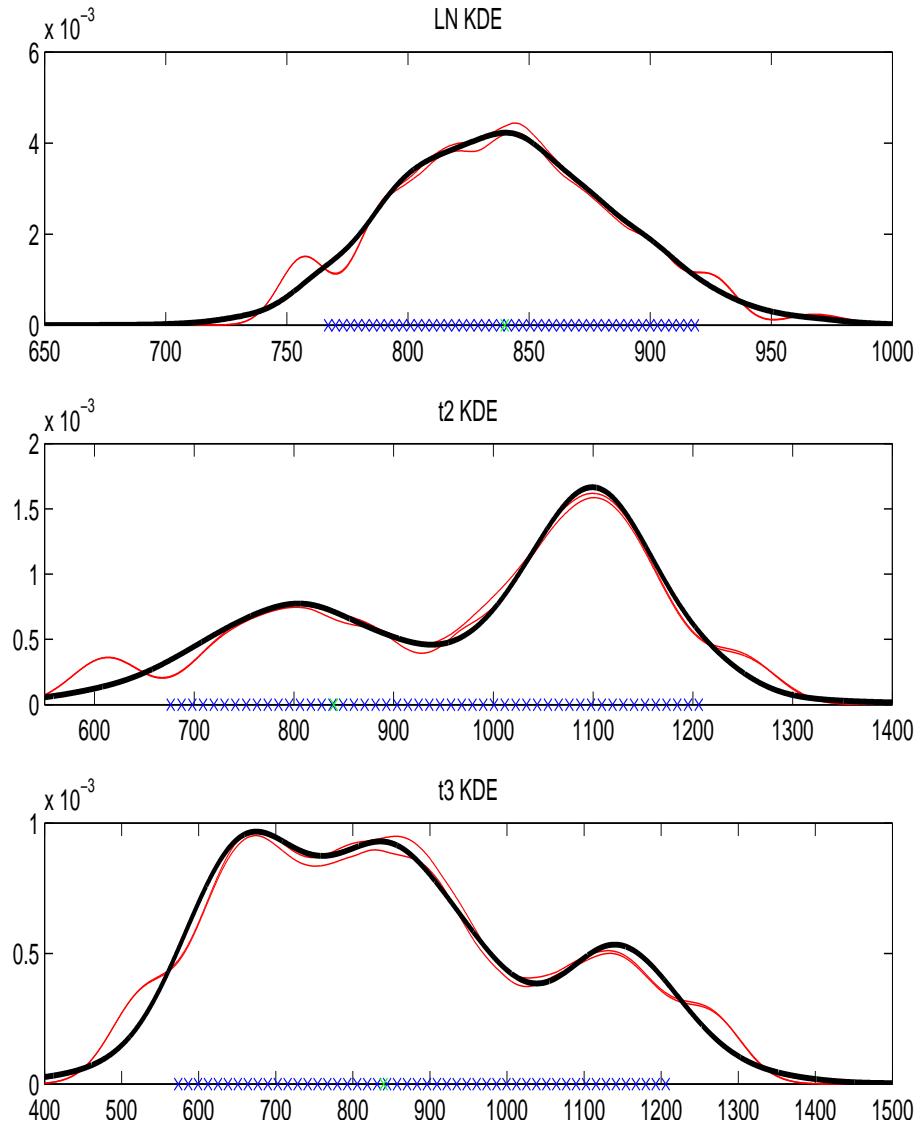


**Figure 3.9.** The 10-fold cross-validation for the second simulation study.

A summary of the 10-fold cross-validation prediction errors using boxplots with respect to the number of support points, from five up to twenty-five, for three data, where data is generated from a lognormal distribution (LN), a mixture of  $t$ -distributions with two components (t2), and a mixture of  $t$ -distributions with three components (t3).

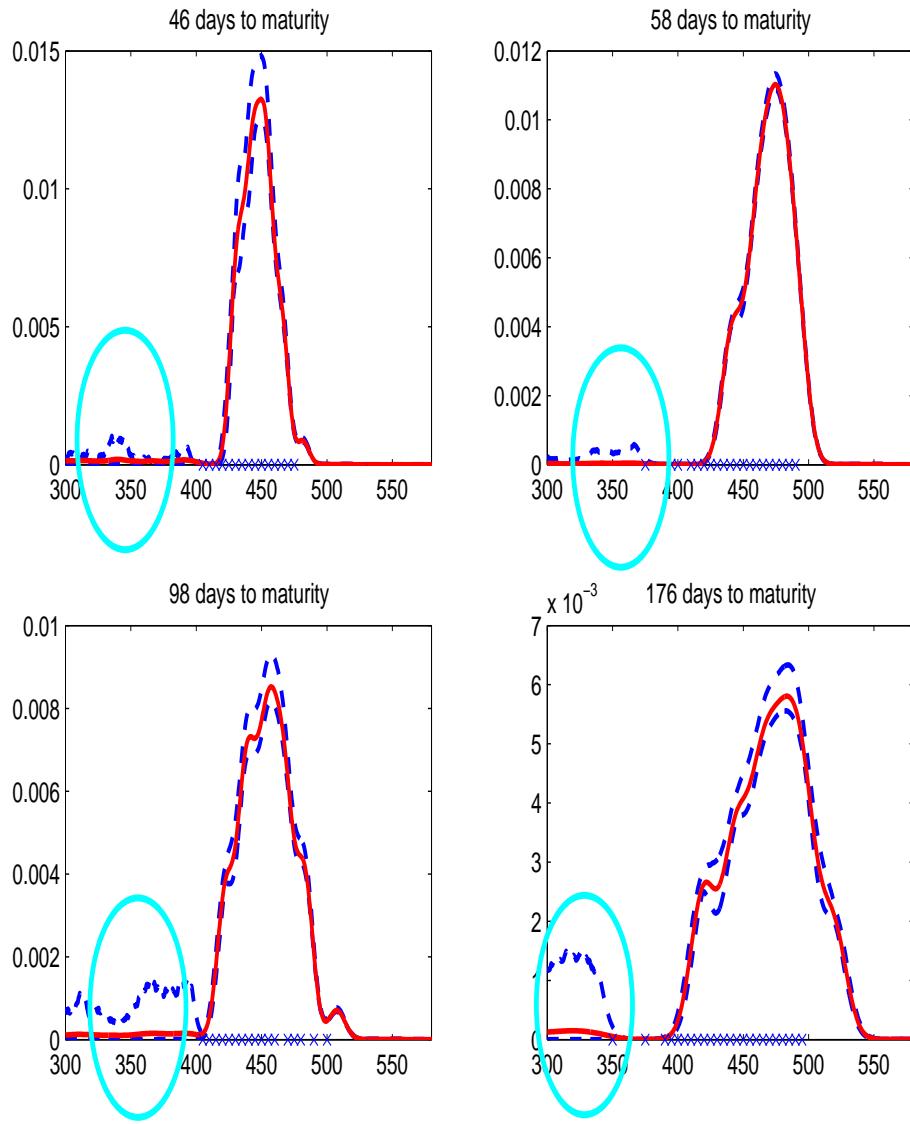


**Figure 3.10.** The 90% credible region of residuals plots for the second simulation study. The 90% credible region of residuals and plots of the residuals against strike prices for put and call in the left and right panels, respectively, for three data, where data is generated from a lognormal distribution (LN), a mixture of  $t$ -distributions with two components (t2), and a mixture of  $t$ -distributions with three components (t3).



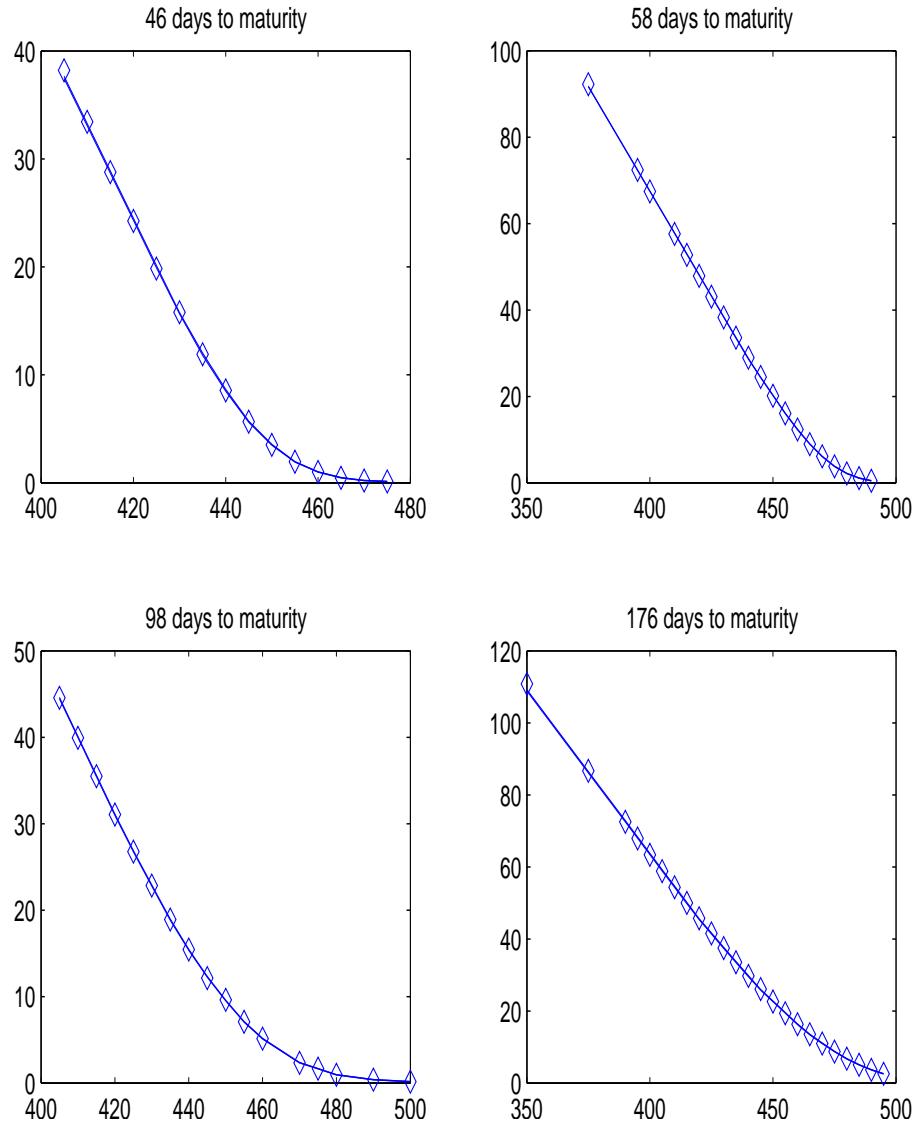
**Figure 3.11.** The true and estimated state price density for the second simulation study.

The true density (in bold) and a 90% credible region of the estimated state price density using the Bayesian Quadrature model for three data, where data is generated from a lognormal distribution (LN), a mixture of  $t$ -distributions with two components (t2), and a mixture of  $t$ -distributions with three components (t3). Strike prices from the original data are represented as crosses on the horizontal axis. The estimated state price density is smoothly visualized using a kernel method.



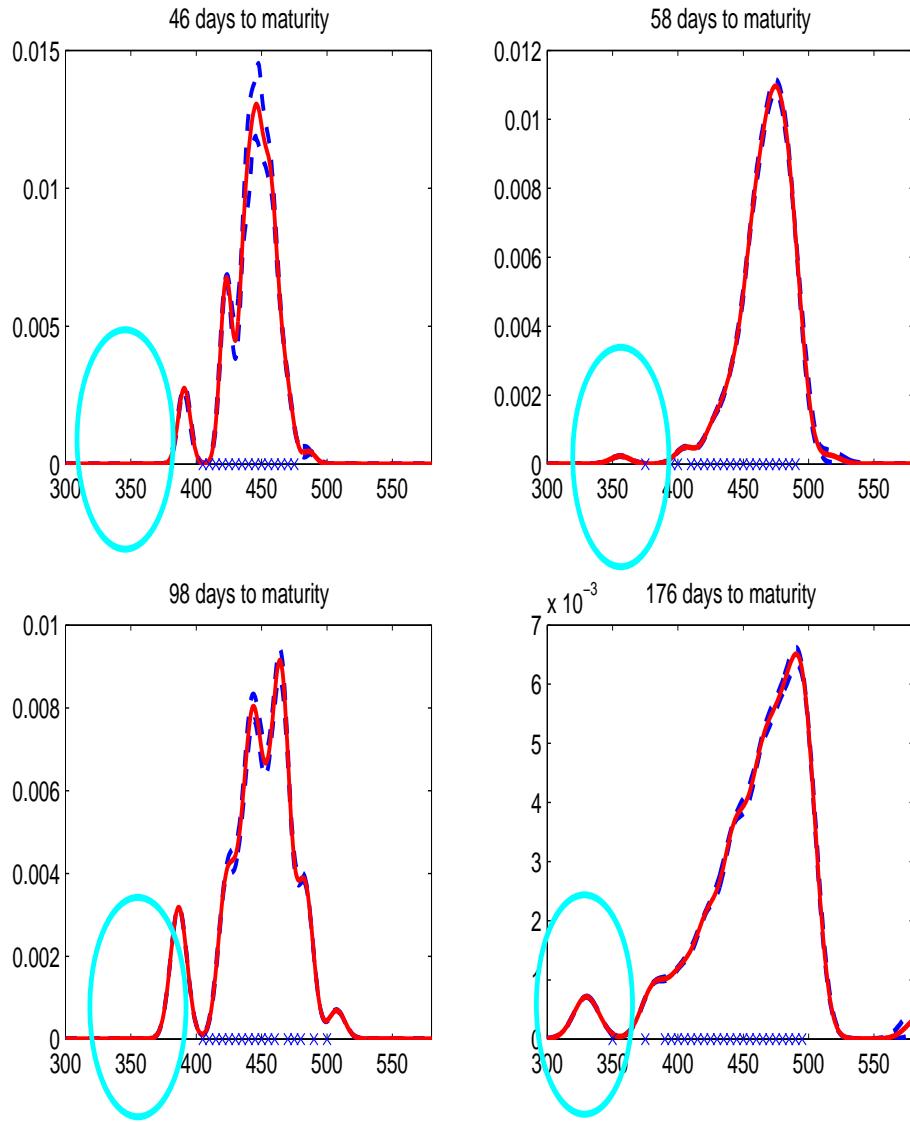
**Figure 3.12.** The 90% credible region of the estimated state price density using call options for the empirical study.

The 90% credible region of the estimated state price density using the Bayesian Quadrature model for four data sets in Aït-Sahalia and Lo (1998). Strike prices from the original data are represented as crosses on the horizontal axis. The estimated state price density is smoothly visualized using a kernel method.



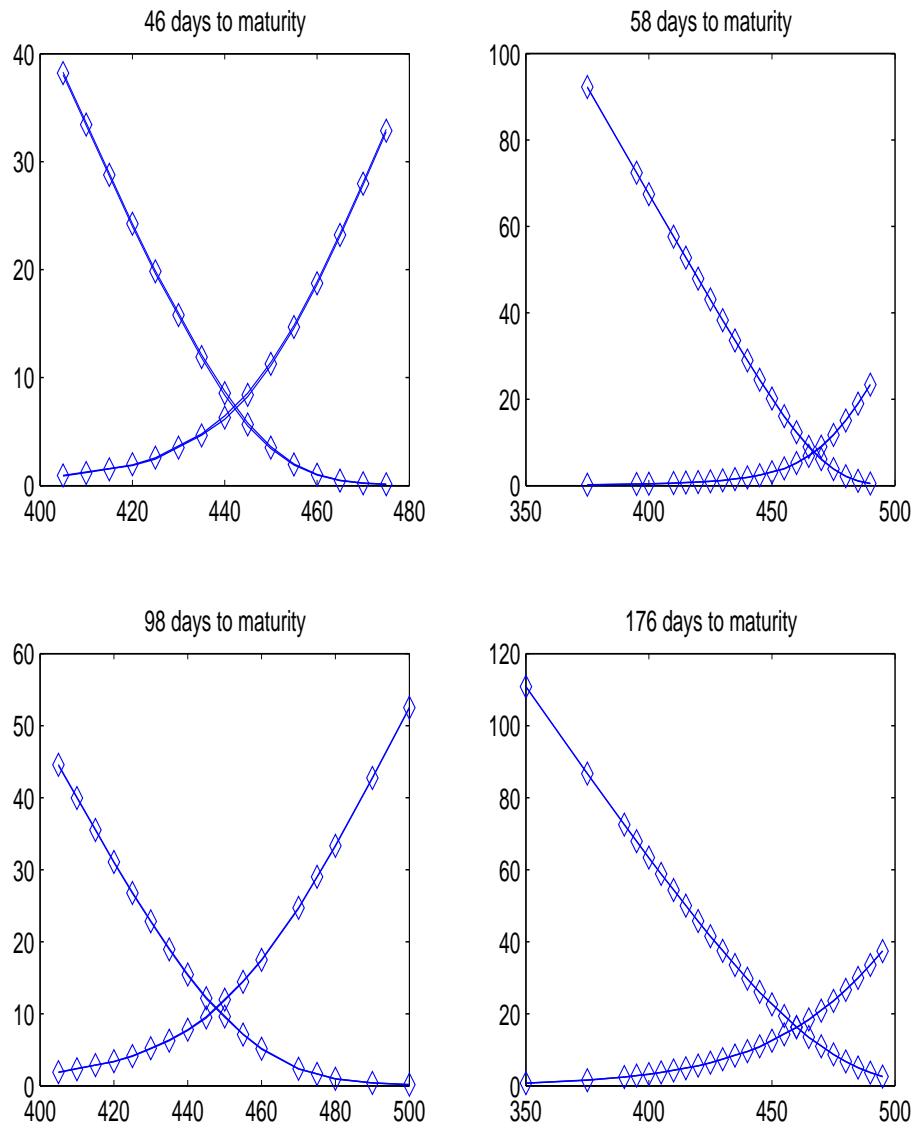
**Figure 3.13.** Scatter plots of the actual and fitted option prices using call options in the empirical study.

Scatter plots of the actual option prices, indicated with a diamond, and the 90% credible region of the fitted option prices, indicated with a line, using ten support points, for the data sets in Aït-Sahalia and Lo (1998).



**Figure 3.14.** The 90% credible region of the estimated state price density using call and put options.

The 90% credible region of the estimated state price density using the Bayesian Quadrature model for four data sets in Aït-Sahalia and Lo (1998), using the original call data and put data built from the original call data using the put-call parity. Strike prices from the original data are represented as crosses on the horizontal axis. The estimated state price density is smoothly visualized using a kernel method.



**Figure 3.15.** Scatter plots of the actual and fitted option prices using call and put options for the empirical study.

Scatter plots of the actual option prices, indicated with a diamond, and the 90% credible region of the fitted option prices, indicated with a line, using ten support points, for four data sets in Aït-Sahalia and Lo (1998), using the original call data and put data built from the original call data using the put-call parity.

Chapter **4**

## The Bayesian implied random tree model

American options are the most commonly traded options on the Chicago Board of Options Exchange and are also among the most challenging to price due to their early exercise feature (Stensoft, 2004). The parametric assumptions of the underlying asset determines the exercise boundary of American options. McKean (1965) formulates the valuation of American options as a free boundary problem. Additional insights about the properties of the optimal exercise boundary are provided by Moerberke (1974) and more recently by Barles et al. (1995), Büttler and Waldvogel (2006), and Kuske and Keller (1998).

Closed-form formulas of American options are not available even under the simplest Black-Scholes assumptions. Therefore, the computations of the optimal exercise boundary and the contract price can be achieved only using numerical methods. One simplified approach for pricing American options is to use Bermudan options, which can be exercised at several pre-defined times before maturity.

A Bermudan option with a sufficiently large set of possible exercise times is treated as an approximation of an American option. However, even Bermudan options are not straightforward to price by Monte Carlo methods. Specifically, at every step on every path, one needs to know the continuation value to make the optimal decision about whether to exercise.

The binomial model by Cox et al. (1979) has been a popular numerical method for pricing American options because it is possible to incorporate the early ex-

ercise features using backward induction. However, the binomial model becomes computationally infeasible for more than a couple of stochastic factors, such as interest rates, dividend, volatility, or multiple underlying asset, due to the curse of dimensionality.

Longstaff and Schwartz (2001) proposes the Least Squares Monte Carlo method. The idea is to estimate the conditional expectation of the payoff from continuing to keep the option alive at each possible exercise point from a cross-sectional least squares regression using the information in the simulated paths. The mathematics foundation about the convergence for the Least Squares simulation is established later by Stensoft (2004). Other numerical methods include lattice methods, methods based on solving partial differential equations, integral equation, and different approximation and extrapolation schemes.

Because there exists no simple closed-form formulas for American option prices, existing nonparametric methods are usually calibrating state price densities using European options. Recently, Alcock and Carmichael (2008) combines the canonical valuation by Stutzer (1996) and the least-square algorithms for pricing American options by Longstaff and Schwartz (2001) in a controlled simulated study for calibrating state price density using American options. However, their method is solely based on a time series of historical prices of the underlying asset, and ignores valuable price-sensitive information embodied within traded option prices. Alcock and Auerswald (2009) extend Alcock and Carmichael (2008), so that conditional information from observed option prices can be incorporated. Liu (2008) proposes an approach called canonical least-squares Monte Carlo to price American options, which turns out to be quite similar to the nonparametric approach of Alcock and Carmichael (2008). However, these methods are extensions of canonical methods by Stutzer (1996) and depend heavily on the historical prices of the underlying asset.

In the following, we present our Bayesian implied random tree models for calibrating the state price density using American options in Section 4.1. We demonstrate our method via a simulation study in Section 4.2 and via an empirical study using S&P 100 index options in Section 4.3, and provide a summary in the final section.

## 4.1 The Bayesian implied random tree model

Recall that we have defined a Quadrature model, denoted as  $Q(w, \theta)$ , as a finite distribution with parameters  $w = \{w_0, \dots, w_N\}$  and  $\theta = \{\theta_0, \dots, \theta_N\}$  for some non-negative integer  $N$ , where  $w$  and  $\theta$  are the weights and locations of the support points. Specifically, a Quadrature model has a density function as

$$f(x|w, \theta) = w_0\delta_{\theta_0}(x) + \dots + w_N\delta_{\theta_N}(x),$$

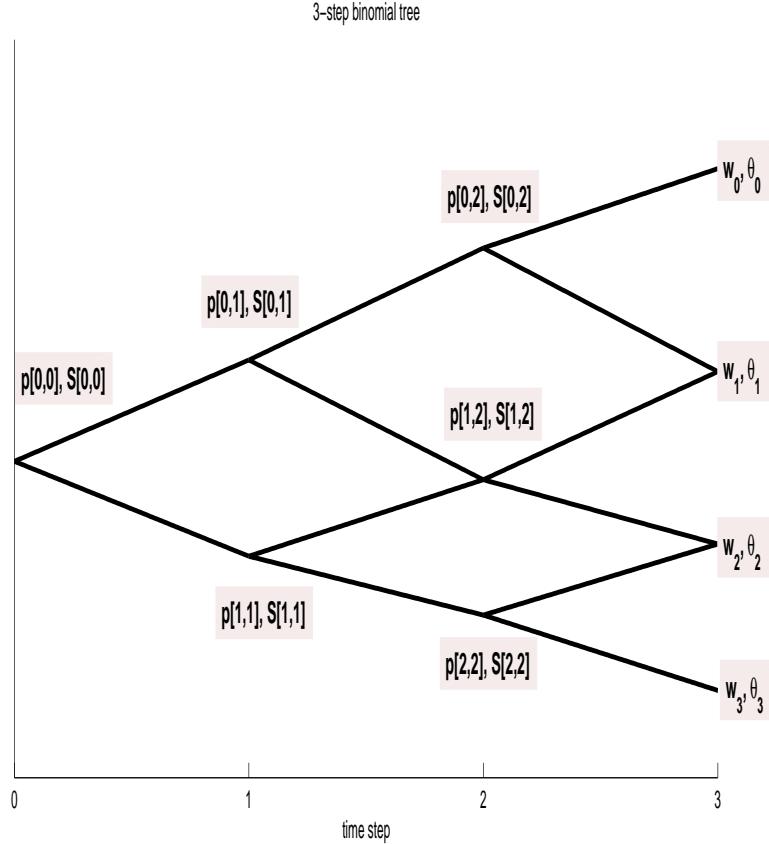
where  $\delta_\omega(x)$  is the Dirac measure at the point  $\omega$  such that

$$\delta_\omega(x) = \begin{cases} 1, & \text{for } x = \omega, \\ 0, & \text{otherwise.} \end{cases}$$

To price path-dependent options, we follow Rubinstein (1994) and build a generalized binomial tree to approximate the underlying process of the risk-neutral measure. Let  $[i, j]$  denote the  $i$ -th node at the  $j$ -th time step for  $i = 0, \dots, j$  and  $j = 0, \dots, N$ . Figure 4.1 shows an example of a generalized binomial tree associated with  $Q(w, \theta)$  for  $N = 3$ . A generalized binomial tree contains the move probability  $p[i, j]$  and the interior stock price  $S[i, j]$  at each node  $[i, j]$ . The move probability  $p[i, j]$  is the probability that a stock price at node  $[i, j]$  will move upward, and the interior stock price  $S[i, j]$  is the (unobservable) stock price at node  $[i, j]$ . Note  $1 - p[i, j]$  is the probability that a stock price at node  $[i, j]$  will move downward.

Figure 4.1 shows that 12 quantities, including move probabilities  $p[i, j]$  for  $i = 0, \dots, j$  and  $j = 0, 1, 2$  and interior stock price  $S[i, j]$  for  $i = 0, \dots, j$  and  $j = 0, 1, 2$ , are needed to determine given  $Q(w, \theta)$ . However, we have seven independent information in  $Q(w, \theta)$ , including four locations and three weights (recall that all weights sum to one). Therefore, there exists an infinite number of solutions for the generalized binomial tree which will produce the Quadrature model as desired. To construct a unique generalized binomial tree associated with  $Q(w, \theta)$ , we make the following assumptions as in Rubinstein (1994).

1. The underlying asset return follows a binomial process.
2. The binomial tree is recombining.



**Figure 4.1.** Illustrations of a generalized binomial tree associated with  $Q(w, \theta)$  for  $N = 3$ .

3. The ending nodal values are ordered from the highest to the lowest.
4. The interest rate is constant.
5. All paths that lead to the same ending node have the same risk-neutral probability.

The solutions of  $p[i, j]$  and  $S[i, j]$  can be calculated using backward induction as follows.

1. Calculate the risk-neutral interest rate between two consecutive steps as

$$\bar{r} = \left( \sum_{i=0}^N w_i \theta_i \right)^{1/N}.$$

2. Calculate the nodal probability  $P[i, j]$  as follows.

$$(a) \ i = 0, \dots, N, \ P[i, N] = w_i / \binom{N}{i}.$$

(b) For  $j = N - 1, \dots, 0$ , for  $i = 0, \dots, j$ ,

$$P[i, j] = P[i, j + 1] + P[i + 1, j + 1].$$

3. Calculate the move probability  $p[i, j]$  as follows.

(a) For  $j = N - 1, \dots, 0$ , for  $i = 0, \dots, j$ ,

$$p[i, j] = P[i, j + 1] / (P[i, j + 1] + P[i + 1, j + 1]).$$

4. Calculate the interior stock price  $S[i, j]$  as follows.

(a) For  $i = 0, \dots, N$ ,  $S[i, N] = \theta_i$ .

(b) For  $j = N - 1, \dots, 0$ , for  $i = 0, \dots, j$ ,

$$S[i, j] = (p[i, j]S[i, j + 1] + (1 - p[i, j])S[i + 1, j + 1]) / \bar{r}.$$

Given such a generalized binomial tree, an American option can be priced by backward induction. For example, American call option equals  $C[0, 0]$ , where  $C[i, j]$  is calculated as follows.

$$C[i, N] = \max(\theta_i - K, 0) \quad \text{for } i = 0, \dots, N,$$

$$C[i, j] = \max(S[i, j] - K, (p[i, j]C[i, j + 1] + (1 - p[i, j])C[i + 1, j + 1]) / \bar{r}) \\ \text{for } j = N - 1, \dots, 0, \quad i = 0, \dots, j.$$

Similarly, American put option price equals  $C[0, 0]$ , where

$$C[i, N] = \max(K - \theta_i, 0) \quad \text{for } i = 0, \dots, N,$$

$$C[i, j] = \max(K - S[i, j], (p[i, j]C[i, j + 1] + (1 - p[i, j])C[i + 1, j + 1]) / \bar{r}) \\ \text{for } j = N - 1, \dots, 0, \quad i = 0, \dots, j.$$

We let  $C_{ijk}(w, \theta)$  denote the option price calculated from  $Q(w, \theta)$  (when the

option type is given by  $i$ , the strike price is given by  $K_j$  and the maturity by  $T_k$ ) and the observed option price  $y_{ijkl}$  follows

$$y_{ijkl} = C_{ijk}(w, \theta) e^{\varepsilon_{ijkl}} \quad (4.1)$$

with i.i.d. normal errors  $\varepsilon_{ijkl} \stackrel{i.i.d.}{\sim} N(0, \sigma^2)$ . Let  $y = \{y_{ijkl}\}$  denote the collection of all observed option prices. Therefore, the likelihood is given by

$$L(y|w, \theta, \sigma^2) = \prod_{i,j,k,l} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_{ijkl} - C_{ijk}(w, \theta))^2}{2\sigma^2}}.$$

Similarly, we assume a priori that the distribution of  $\sigma^2$  is an inverse-gamma distribution with shape parameter  $\alpha$  and scale parameter  $\beta$ , denoted by  $\sigma^2 \sim IG(\alpha, \beta)$ .

We assume a vague prior for the weights  $w_m$ , or  $w$  has a Dirichlet distribution with parameters  $\bar{w} = (\bar{w}_0, \dots, \bar{w}_N)$ , denoted by  $w \sim D(\bar{w})$ .

To avoid zero option prices, we assume a priori that the distribution of the locations  $\theta$  of the support points are uniformly distributed over the support set  $\Theta$ ,

$$\mathcal{A} = \{\theta_n \in \Re^+ : \theta_0 > \theta_1 > \dots > \theta_N, \theta_0 > K_{\max}, \theta_N < K_{\min}\},$$

and  $K_{\max}$  and  $K_{\min}$  are the maximum and the minimum of strike prices respectively.

We again remark that all three of these assumptions for the prior distributions can be changed in cases where appropriate subjective information is available. Inference for the parameters of interest is based on the posterior distribution of  $w$ ,  $\theta$ , and  $\sigma^2$ .

Our inference is based on the posterior distribution of the parameters of interest.

$$p(w, \theta, \sigma^2 | y) \propto L(y|w, \theta, \sigma^2) p(w)p(\theta)p(\sigma^2),$$

where  $p(w)$ ,  $p(\theta)$ , and  $p(\sigma^2)$  are prior densities of parameters  $w$ ,  $\theta$ , and  $\sigma^2$ , respectively.

We denote  $TN(\mu, h, [l, u])$  as a truncated normal distribution restricted to the interval  $[l, u]$  with mean  $\mu$  and variance  $\sigma^2$ . Let  $\Phi(\cdot)$  denote the cumulative distribution function of a standard normal distribution. The MCMC simulation is

implemented as follows.

1. Randomly start the parameters  $w$ ,  $\theta$ , and  $\sigma^2$  in the support of their joint distribution.
2. At each iteration, do the following Metropolis-Hastings update until convergence.
  - (a) Update  $w$  using a random-walk version of the MCMC algorithms.
    - i. Let  $w^c$  denote the current configuration of  $w$ .
    - ii. Propose  $w^p \sim D(Vw^c)$  with a tuning parameter  $V$ .
    - iii. Update  $w$  as  $w^p$  with the acceptance rate
$$\alpha = \min \left\{ 1, \frac{p(w^c|Vw^p)p(w^p, \theta, \sigma^2|y)}{p(w^p|Vw^c)p(w^c, \theta, \sigma^2|y)} \right\}.$$

2. At each iteration, do the following Metropolis-Hastings update until convergence.
  - (a) Update  $w$  using a random-walk version of the MCMC algorithms.
  - (b) Update  $\theta$  using a truncated random-walk version of the MCMC algorithms. We update each variate of  $\theta$ ,  $\theta_n$ , for  $n = 0, \dots, N$ , as follows.

2. At each iteration, do the following Metropolis-Hastings update until convergence.
  - (a) Update  $w$  using a random-walk version of the MCMC algorithms.
  - (b) Update  $\theta$  using a truncated random-walk version of the MCMC algorithms.
    - i. Let  $\theta_n^c$  denote the current configuration of  $\theta_n$ .
    - ii. Propose  $\theta_n^p \sim TN(\theta_n^c, h, [l, u])$  with a tuning parameter  $h$ , where

$$(l, u) = \begin{cases} (K_{\max}, \infty) & \text{for } n = 0, \\ (\theta_{n+1}, \theta_{n-1}) & \text{for } n = 1, \dots, N-1, \\ (0, K_{\min}) & \text{for } n = N; \end{cases}$$

2. At each iteration, do the following Metropolis-Hastings update until convergence.
  - (a) Update  $w$  using a random-walk version of the MCMC algorithms.
  - (b) Update  $\theta$  using a truncated random-walk version of the MCMC algorithms.
    - i. Let  $\theta_n^c$  denote the current configuration of  $\theta_n$ .
    - ii. Propose  $\theta_n^p \sim TN(\theta_n^c, h, [l, u])$  with a tuning parameter  $h$ , where
    - iii. Update  $\theta_n$  as  $\theta^p$  with the acceptance rate

$$\alpha = \min \left\{ 1, \frac{L(y|w, \theta^p, \sigma^2) \left( \Phi \left( \frac{u - \theta_n^p}{\sqrt{h}} \right) - \Phi \left( \frac{l - \theta_n^p}{\sqrt{h}} \right) \right)}{L(y|w, \theta^c, \sigma^2) \left( \Phi \left( \frac{u - \theta_n^c}{\sqrt{h}} \right) - \Phi \left( \frac{l - \theta_n^c}{\sqrt{h}} \right) \right)} \right\}.$$

2. At each iteration, do the following Metropolis-Hastings update until convergence.
  - (a) Update  $w$  using a random-walk version of the MCMC algorithms.
  - (b) Update  $\theta$  using a truncated random-walk version of the MCMC algorithms.
    - i. Let  $\theta_n^c$  denote the current configuration of  $\theta_n$ .
    - ii. Propose  $\theta_n^p \sim TN(\theta_n^c, h, [l, u])$  with a tuning parameter  $h$ , where
    - iii. Update  $\theta_n$  as  $\theta^p$  with the acceptance rate

$$\sigma^2 \sim IG(\alpha + \bar{N}/2, \beta + \sum_{i,j,k,l} (\log y_{ijkl} - \log C_{ijk}(w, \theta))^2/2),$$

where  $\bar{N}$  is the total number of options in the data.

In the generalized binomial tree in Rubinstein (1994), the locations of the support points are pre-specified by the researcher. Usually, these locations are equally

spaced, and the weights are determined using an optimization algorithm. As a result, one disadvantage of the Rubinstein (1994) approach is that an illegitimate density or density with negative weights may appear. In addition, this method usually need a huge number of support points to achieve a good model fit.

In contrast to the method by Rubinstein (1994), our Quadrature model assumes both weights and locations are random variables, which are determined by the data. Furthermore, the prior distributions assumptions placed on the weights and locations ensure that the resulting state price density is a legitimate density.

## 4.2 Simulation study

The simulation study provides an opportunity to test our approach and the subsequent MCMC algorithm. For simplicity, we assume that the risk-free interest rate is constant, and create options with arbitrary strike prices and times to maturity, although we acknowledge that, these assumptions are too simplified and unrealistic in practice.

### 4.2.1 True model in the simulation study

In this simulation study, we begin with a Quadrature model,  $Q(w, \theta)$ , and assume that current stock price is 100. Parameters in  $Q(w, \theta)$  are given as follows.

$$w = \{w_0 = 0.2, w_1 = 0.3, w_2 = 0.4, w_3 = 0.1\}, \quad (4.2)$$

$$\theta = \{\theta_0 = 127.76, \theta_1 = 108.51, \theta_2 = 92.16, \theta_3 = 78.27\}. \quad (4.3)$$

This follows a numerical example given in the appendix of Rubinstein (1994), and a three-step generalized binomial tree can be built from  $Q(w, \theta)$ . The interior stock prices and moving probabilities at node  $[i, j]$  are listed in Tables 4.1 and 4.2.

### 4.2.2 Generating synthetic data

We denote the  $k$ -th time step in the generalized binomial tree by  $T_k$ . To show our ability to recover the model, we consider three data sets, where the first data set generates options with only one maturity ( $T_3$ ), the second data set generates

**Table 4.1.** Interior stock prices for the simulation study.  
 We record  $\theta$  and interior stock prices at node  $[i, j]$  associated with  $Q(w, \theta)$  as defined in Eqs. (4.2) and (4.3) for the simulation study.

$i \setminus j$	0	1	2	3
0	100	109.61	120.23	127.76
1		91.00	98.26	108.51
2			85.42	92.16
3				78.27

**Table 4.2.** The probabilities of moving up for the simulation study.

$i \setminus j$	0	1	2
0	0.533	0.563	0.667
1		0.5000	0.429
2			0.571

options with two different maturities ( $T_2$  and  $T_3$ ), and the third data set generates options with three different maturities ( $T_1$ ,  $T_2$ , and  $T_3$ ). For each data set, we generate both European and American call and put options.

For options with maturity  $T_k$ , we set up ten strike prices equally distributed over an interval. For example, for options with maturity  $T_k$ , we define this interval by  $[S[k, k] S[0, k]]$ , where the end points of the interval are the minimum and the maximum of the interior stock prices at  $T_k$  of the generalized binomial tree. Next, we calculate the theoretic prices of European and American options using the true values as reported in Tables 4.1 and 4.2. Then, we generate a synthetic data set by perturbing the true option price using Eq. (4.1) with  $\sigma^2 = 0.1$ . Scatterplots of these three data sets are given in Figures 4.2 to 4.4.

### 4.2.3 MCMC inference

For each MCMC analysis, we begin with random starting values for the parameters, discard the first 1,000 burn-in samples, and make inference using the the following 1,000 samples. The trace plot, auto correlation function (ACF) plot, and the

**Table 4.3.** Acceptance rates for parameters in the MCMC algorithm for the simulation study.

This table records the acceptance rates of updating parameters,  $\sigma^2$ ,  $w$ ,  $\theta_0$ ,  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ , in the Markov chain Monte Carlo analysis, for the simulated data sets. The first data set includes options with one maturity ( $T_3$ ), the second data set includes options with two different maturities ( $T_2$  and  $T_3$ ), and the third data set includes options with three different maturities ( $T_1$ ,  $T_2$ , and  $T_3$ ).

Acceptance Rate (%)	$T_3$	$T_2, T_3$	$T_1, T_2, T_3$
$\sigma^2$	100	100	100
$w$	76	73	70
$\theta_0$	5	4	3
$\theta_1$	17	10	8
$\theta_2$	11	6	4
$\theta_3$	5	4	41

kenel density estimate (KDE) of the log-likelihood (LL) are used to assess the the convergence of the MCMC algorithm. When the sampling distribution appears to have converged to the joint distribution, we summarize the posterior densities of parameters of interest.

Figure 4.5 shows the trace plots, ACF plots, and KDEs of LL of these three data sets. Figure 4.6 shows the trace plots, ACF plots, and KDEs of  $\sigma^2$  of these three data sets. These plots show that the sampling of the MCMC algorithm appears to have converged.

As a point of interest with respect to the mixing properties of these algorithm, Table 4.3 reports the acceptance rates for the Hastings-Matropolis algorithm. These acceptance rates provide an evidence of good mixing in the Markov chain Monte Carlo simulation.

Figure 4.7 shows the trace plots of  $w_n$  and  $\theta_n$  for  $n = 0, \dots, 3$  of these three data sets. Table 4.4 summarizes the posterior means and standard deviations of these parameters. The simulation results show that the posterior means of these parameters are very close to the true values and that the parameters of the implied random tree can be recovered using an MCMC algorithm.

**Table 4.4.** Numerical results for the simulation study.

We summarize true values, posterior means and standard deviations (in parenthesis) of parameters of the simulation study. The first data set includes options with maturity  $T_3$ , the second data set include options with maturity  $T_2$  and  $T_3$ , and the third data set include maturity  $T_1$ ,  $T_2$ , and  $T_3$ .

Maturities True Values	$T_3$	$T_2, T_3$	$T_1, T_2, T_3$
$\sigma^2$	<b>0.01</b>	0.01 (0.0010)	0.01 (0.0007)
$w_0$	<b>0.2</b>	0.19 (0.0070)	0.20 (0.0043)
$w_1$	<b>0.3</b>	0.29 (0.0196)	0.30 (0.0097)
$w_2$	<b>0.4</b>	0.42 (0.0119)	0.40 (0.0074)
$w_3$	<b>0.1</b>	0.10 (0.0046)	0.11 (0.0038)
$\theta_0$	<b>127.76</b>	127.87 (0.2747)	128.10 (0.2090)
$\theta_1$	<b>108.51</b>	109.83 (0.4022)	108.69 (0.2718)
$\theta_2$	<b>92.16</b>	92.10 (0.3657)	92.39 (0.1856)
$\theta_3$	<b>78.27</b>	78.37 (0.2546)	78.73 (0.2330)

## 4.3 Empirical study

We apply our Bayesian implied random tree model to S&P 100 index options as an empirical study. These options depend on S&P 100 index as an underlying asset, and are traded on the Chicago Board Option Exchange. The S&P 100 index is a capitalization-weighted index of 100 stocks from a broad range of industries. In this empirical study, we focus on S&P100 index options traded on 11/2/2009. The S&P 100 index options contain both European and American options, and as a result they provide a good source for testing our Bayesian implied random tree model.

### 4.3.1 Descriptions of S&P 100 index options

Figure 4.8 gives the time series plot of the S&P 100 index future prices from 9:00 am to 3:00 pm traded on 11/02/1009. The S&P 100 index future prices opened at 488.76 and closed at 485.01. The future price has a mean of 484.57, a median of 485.01, and a standard deviation of 2.60. For simplicity, we assume that the value of the S&P 100 index is fixed at 485.01 for our analysis.

Unlike simulated data, which can be set up arbitrarily, real option prices differ in several important ways. To illustrate these differences, we provide some descriptive summaries. First, we summarize the numbers of transactions for different options and maturities in Table 4.5. We find that options having shorter maturities are traded more frequently, and that American options are traded more frequently than European options.

Second, options are typically only traded for some strikes and in a certain range around the current index level. Figure 4.9 gives the stacked bar plots of the number of transactions against strike prices for different options. This figure shows that American options are more frequently traded than European options, and are traded with a wider range of strike than their European counterparts.

In addition, because options with strikes that are in-the-money are usually more expensive, they are less frequently traded, and as a result their prices may be less reliable. When methods require or can use only one type of options (for example, kernel-based methods only use European call options), they usually convert European put options to European call options using the call-put parity,

$$C + K e^{-rT} = P + S,$$

where  $C$  is the call option price,  $P$  is the put option price,  $S$  is current value of the underlying asset,  $K$  is the strike price,  $r$  is the risk-free interest rate, and  $T$  is the time to maturity. As we will show later, our approach can naturally uses both put and call options for inferences, and avoid the need to convert put options to call options.

Table 4.6 summarizes the 10-th, 50-th, and 90-th percentiles of the trading volume relative to different types of options and times to maturity, and shows that the trading volume varies for different options. We do not incorporate the volume as an information in the model at this stage, but we remark that the volume can be used to weight the importance of an option price for estimating the state price density very easily.

Table 4.7 records the mean, standard deviation, and several percentiles of option prices with respect to different options. Because the most frequently traded options have shorter maturities, we give scatter plots of S&P 100 index option

**Table 4.5.** Number of transactions of S&P 100 index options data traded on 11/02/2009. We tabulate numbers of transactions of S&P 100 index options with respect to different options and maturities (T). This data set shows that options having shorter maturities are more traded than those having longer maturities, and that American options options are more traded than European options.

T	European		American	
	Call	Put	Call	Put
19	75	77	918	681
47	13	20	156	154
75	0	1	34	22
110	0	0	2	3
138	5	0	4	3
775	0	2	0	0

prices, with 19 days and 47 days maturity, against strike prices in Figures 4.10 to 4.11, respectively.

### 4.3.2 Recovery of option prices

We implement our approach for the S&P 100 index options with 19 days to maturity traded on 11/02/2009. To investigate how well our implied random tree model can recover option prices for different types of options, we consider three data sets. The first data set consists of just European options, the second data set consists of just American options, and the third data set consists of both European and American options.

After the MCMC algorithm has converged, we summarize the posterior distribution of  $R^2$  as goodness of fit measures using box plots for the three data sets in Figure 4.15. The  $R^2$  is calculated using

$$1 - \frac{\sum_{i=1}^I (o_i - \hat{o}_i)^2}{\sum_{i=1}^I o_i^2}, \quad (4.4)$$

where  $o_i$  is the actual option price,  $\hat{o}_i$  is the fitted option price, and  $I$  is the total number of the observations. The closer  $R^2$  to one is, the better model fit our model produces. Because the posterior means of  $R^2$  in these three data sets are

**Table 4.6.** Summary statistics of the trading volume of S&P 100 index options data.

We record the 10-th, 50-th, and 90-th percentiles of the volume at each transaction of S&P 100 index options data traded on 11/02/2009 with respect to different options and maturities (T) in days. We write zero for options without transaction.

T	European Call			European Put			American Call			American Put		
	10-th	50-th	90-th	10-th	50-th	90-th	10-th	50-th	90-th	10-th	50-th	90-th
19	1	30	136	2	20	143	1	5	51	1	5	50
47	1	1	35	1	8	23	1	6	50	1	10	100
75	0	0	0	2	2	2	1	2	11	1	1	21
110	0	0	0	0	0	0	1	11	20	5	10	25
138	1	1	2	0	0	0	2	8	100	1	1	1
775	0	0	0	1	5	9	0	0	0	0	0	0

**Table 4.7.** Summary statistics for S&P 100 index options data.

We summarize the mean, standard deviation (Std. Dev.), the minimum, 10-th, 50-th, and 90-th percentiles, and the maximum of the option prices with respect to different options and maturities (T) in days.

Type	T	Mean	Std. Dev.	Min	Quantile			
					10%	50%	90%	Max
European Put	19	6.53	4.66	1.25	1.90	4.70	11.60	27.00
European Call	19	5.94	9.53	0.01	0.05	1.90	11.00	71.10
American Put	19	8.15	6.28	0.10	1.50	7.80	15.17	69.90
American Call	19	7.03	20.34	0.02	0.27	3.50	11.00	228.20
European Put	47	35.56	72.85	2.94	2.95	16.00	84.10	275.70
European Call	47	7.16	5.26	0.53	0.79	6.43	15.07	16.30
American Put	47	28.50	68.84	0.30	1.35	9.50	33.78	320.70
American Call	47	14.69	20.74	0.05	0.88	7.47	41.17	117.84
European Call	75	34.15	0.00	34.15	34.15	34.15	34.15	34.15
American Put	75	31.88	31.58	3.00	9.21	15.70	86.92	88.09
American Call	75	20.06	28.82	0.23	0.26	5.30	38.69	125.80
American Put	110	26.50	2.12	25.00	25.00	26.50	28.00	28.00
American Call	110	83.53	119.67	3.50	3.50	26.00	221.10	221.10
European Put	138	36.90	1.92	35.00	35.00	37.50	39.50	39.50
American Put	138	13.03	11.54	3.00	3.00	9.80	29.50	29.50
American Call	138	29.57	0.61	28.90	28.90	29.70	30.10	30.10
European Call	775	37.20	0.00	37.20	37.20	37.20	37.20	37.20

larger than 99%, we conclude that our approach can recover option prices. We only report  $R^2$  value, as opposed to a marginal log-likelihood value, because we are comparing the performance of the same model on different data sets.

Figure 4.16 summarizes the posterior distribution of Root-Mean-Square-Error (RMSE) of the pricing errors, calculated by

$$\sqrt{\frac{\sum_{i=1}^I (o_i - \hat{o}_i)^2}{I}}. \quad (4.5)$$

These two figures show that our model produces a good model fit as the number of support points become larger than 14. We finally give scatter plots of the actual and fitted options against strike prices, and the estimated state price densities, for these three data sets, in Figures (4.12) to (4.14).

### 4.3.3 In-sample versus out-of-sample analysis

Figure 4.17 gives estimated state price densities using different options with 20 support point in the Quadrature model. It appears that the calibrated state price density using American options and that using European and American options give a very similar shape. However, these seems to be a difference between the state price density calibrated from European options and the other two estimated state price densities.

Although we have shown that our approach has a good in-sample prediction in Section 4.3.2, we are also interested in the out-of-sample prediction performance. In particular, we would like to investigate how well the state price density calibrated from European options can predict prices of American options.

To do this, we implement an in-sample versus out-of-sample analysis as follows. The first data set uses just European options as data, the second data set uses just American options as data, and the third data set uses both European and American options as data. For each data set, we randomly select 90% of the original data as our calibration data, and we use the remaining 10% as hold out data. We repeat this ten times and summarize the average RMSE of the pricing errors as a measure of fit of the in-sample test versus out-of-sample in Table 4.8.

Figure 4.18 gives the trace plots of the log-likelihood of the MCMC algorithm.

The trace plots show that the MCMC algorithm appears to have converged. There are several interesting observations that can be made regarding the in- and out-of-sample performance based on these different data sets.

First, we find that out-of-sample prediction errors are similar to in-sample errors, and conclude that our approach does not have serious over-fitting problems. Second, the prediction errors using American options as calibration data are larger than those using European options as calibration data. This empirical data analysis shows that state price densities calibrated from European options do not do as well as the predictions provided by the American options. In contrast, the state price densities calibrated using American options can predict the prices of European options well. Finally, compared with predictions using just European options and predictions using just American options, the predictions using both European and American options produce smaller prediction errors.

## 4.4 Summary

American option pricing remains an active research area because there are many competing methods that are amenable to improvement. In this chapter, we have extended Bayesian Quadrature model to Bayesian implied random tree model for calibrating the state price density using American options.

We have demonstrated our approaches via simulation and empirical studies, and our Bayesian implied random tree approach appears to be able to recover the model parameters well.

In the empirical study, we focus on the intra-day trading of S&P 100 index options with 19 days maturity traded on 11/02/2009. We implement our methods on three data sets, where the first data set consists of just European options, the second data set consists of just American options, and the third data set consists of both European and American options. Our approach seems to produce very good model fits in terms of high  $R^2$  values. We find that when the number of support points exceed 15, our model produce stable model fit in terms of  $R^2$  and RMSE of pricing errors.

In addition, we find that estimated state price densities using just American options and using both European and American options are very similar in shape.

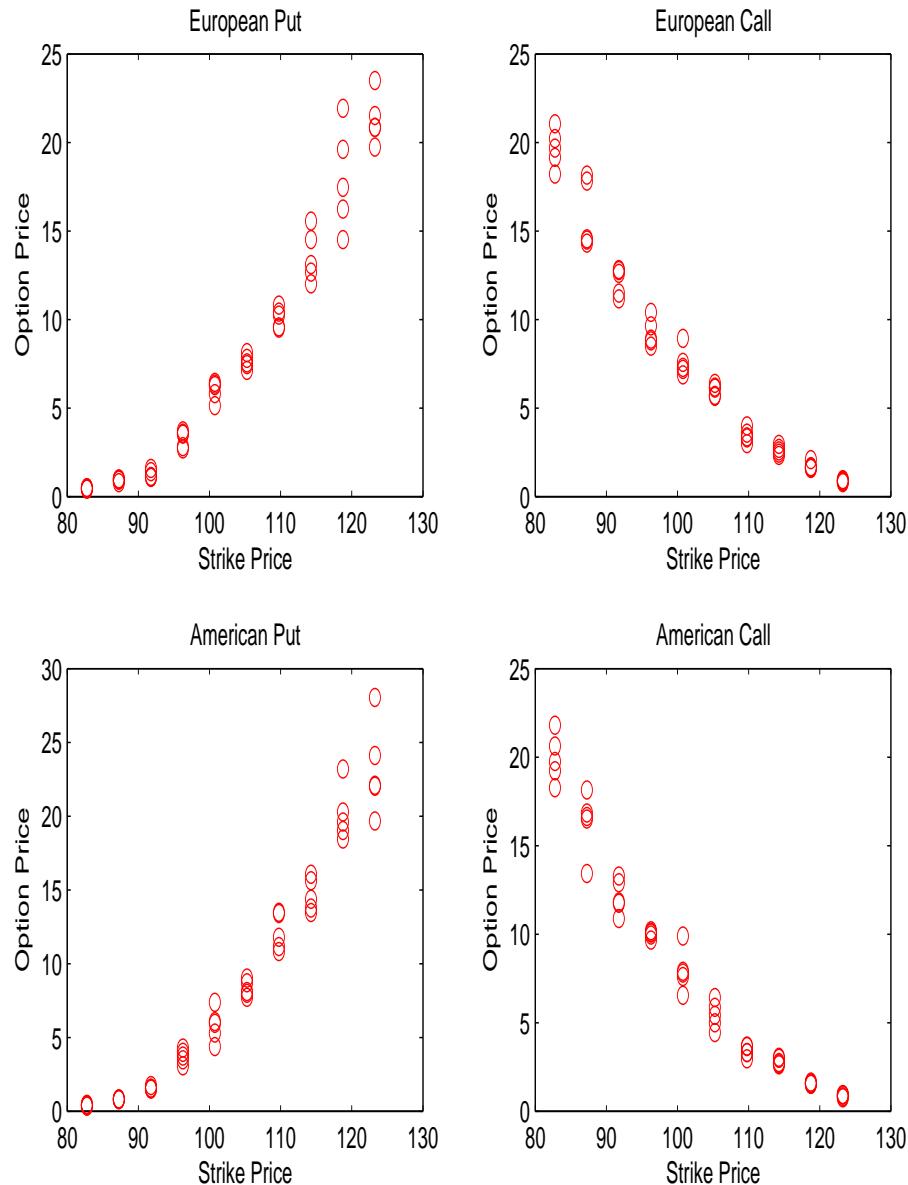
**Table 4.8.** Root-Mean-Square-Errors of pricing errors for the in-sample versus out-of-sample analysis. Calibration Data indicates data used to estimate the state price density, and Prediction Data indicates data that was predicted using the calibrated density. Shaded regions indicate in-sample estimates, and the remaining regions indicate out-of-sample estimates.

Calibration Data	Prediction Data											
	E. Calibration			E. Hold Out			A. Calibration			A. Hold Out		
	5%	50%	95%	5%	50%	95%	5%	50%	95%	5%	50%	95%
European+American	0.96	0.98	1.00	1.00	1.06	1.12	1.11	1.11	1.12	1.13	1.15	1.17
European	0.94	0.98	1.06	0.96	1.14	1.37	1.15	1.21	1.34	1.17	1.24	1.35
American	0.97	0.99	1.02	1.00	1.06	1.12	1.11	1.11	1.12	1.13	1.15	1.17

However, the state price density calibrated using just European options seems to be different from the other two state price densities. In an in-sample versus out-of-sample analysis, we find that the state price density calibrated from American options does a good job of predicting prices of European options, but the state price density calibrated from European options does not do as good of a job predicting American options. One reason may be that the European options are traded with a narrower range of strike price compared with American options.

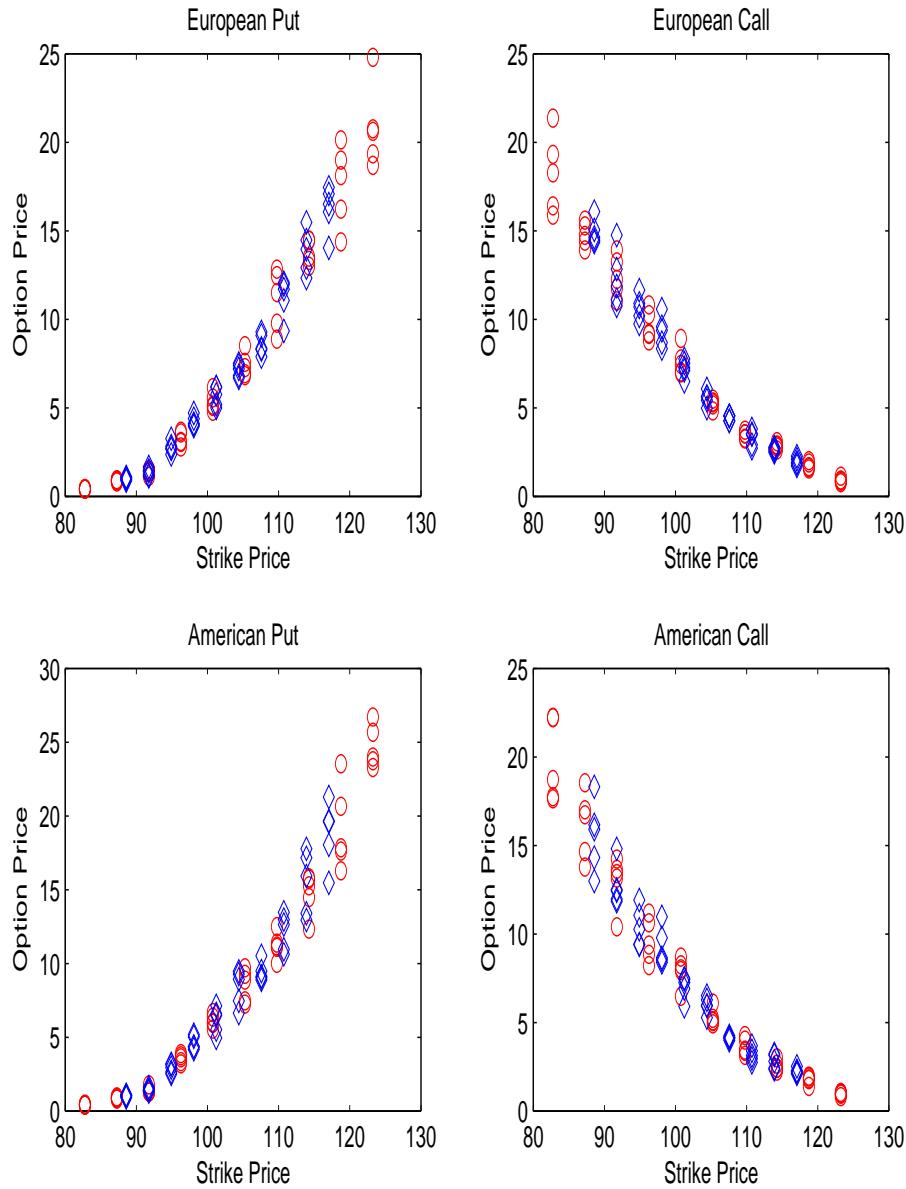
As mentioned before, because most methods focus on European call options, our method has an advantage over other methods, in that our approach can use most valuable information of option prices, e.g., European and American call and put options, at a time, for inferences.

The Bayesian Quadrature model in Chapter 3 and the Bayesian implied random tree model in this chapter are proposed for the state price density calibration. In the following two chapters, we study the price sensitivity of options (also known as the Greeks), which are extremely important for risk management and dynamic hedging. We begin with providing the non-Lipschitz pathwise method to derive the Greeks formulas for options under a parametric assumption on the underlying asset. To avoid model misspecification, we provide a new method in Chapter 6 to calculate the Greeks nonparametrically using finite-difference formulas based on the estimation for the state price density using the Bayesian implied random tree model.



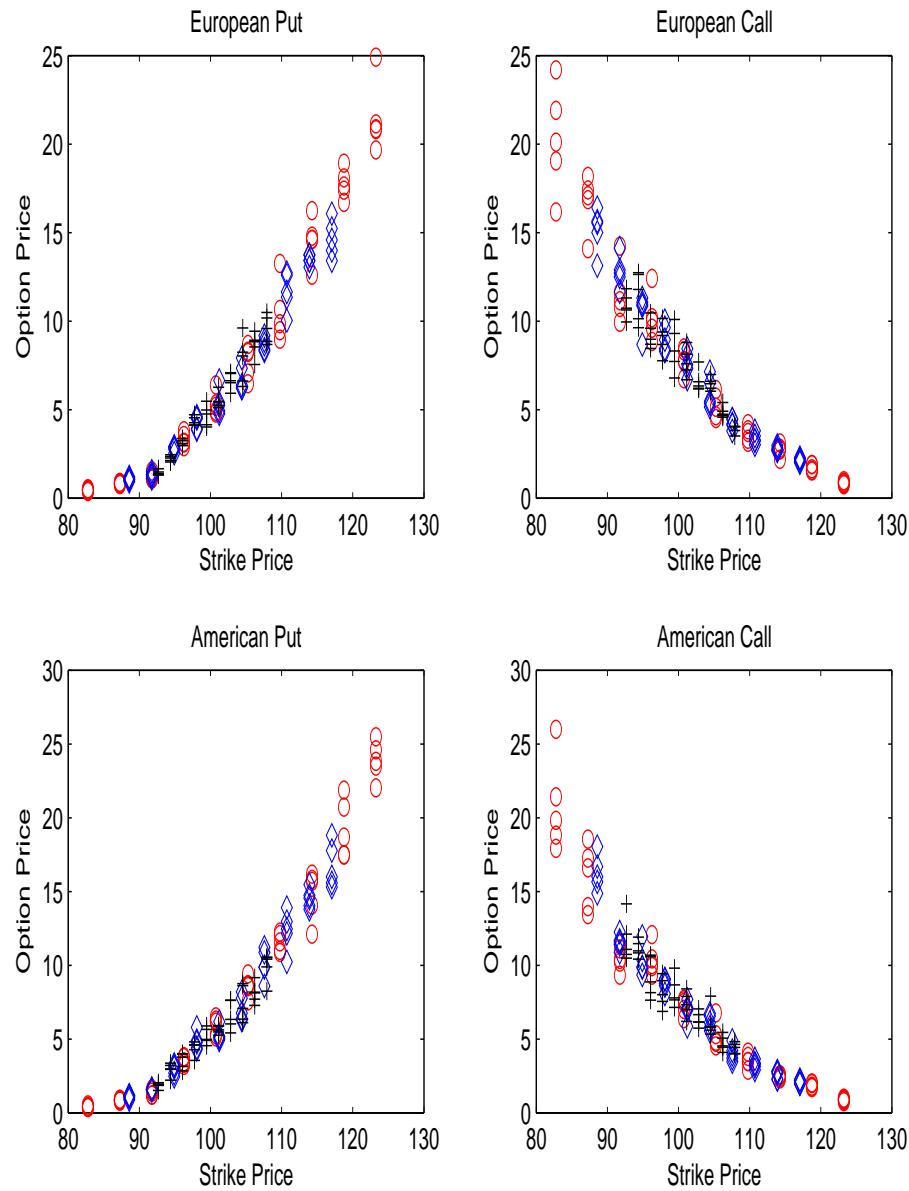
**Figure 4.2.** Scatter plots of option prices against strike prices of the first data set for the simulation study.

These options all have the same maturity,  $T_3$ .



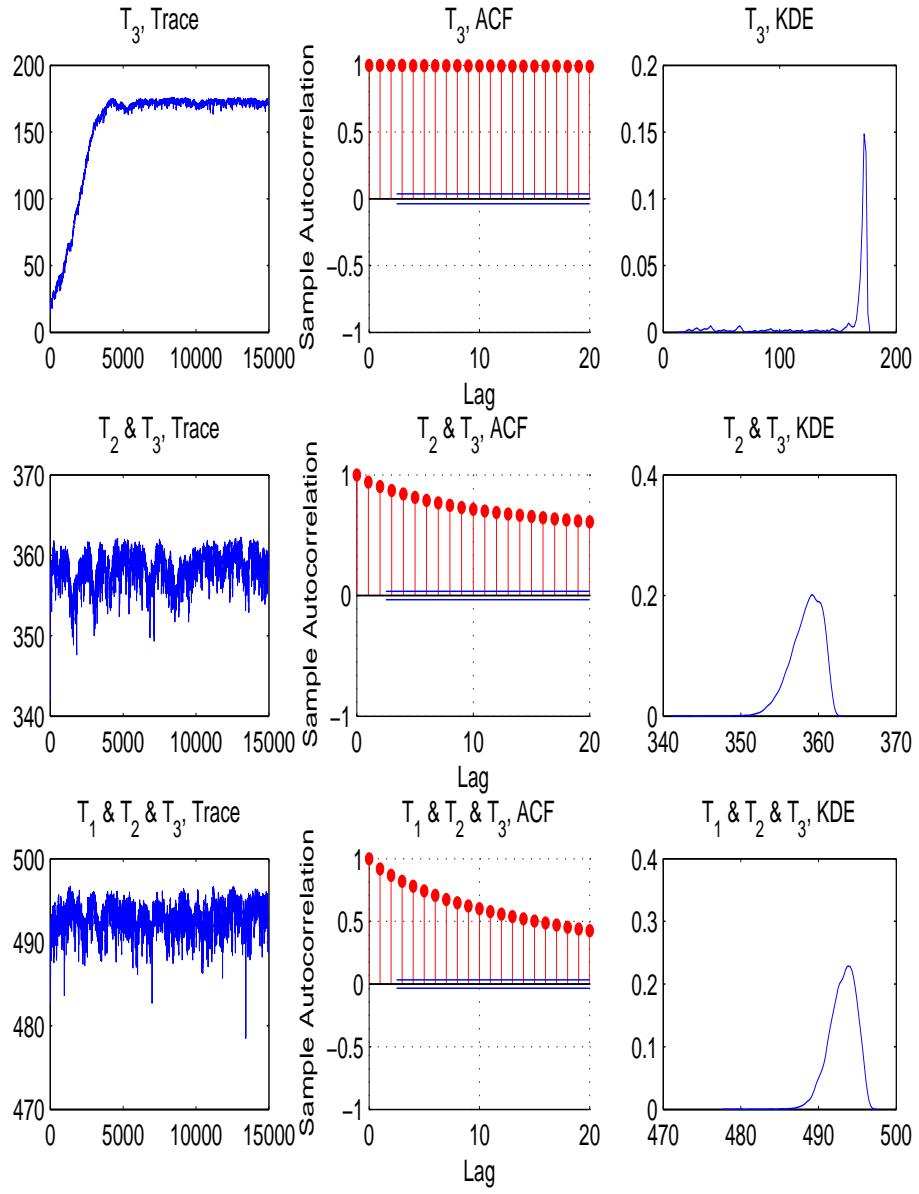
**Figure 4.3.** Scatter plots of option prices against strike prices of the second data set for the simulation study.

The options with maturity  $T_3$  are indicated with a circle, and the options with maturity  $T_2$  are indicated with a diamond.



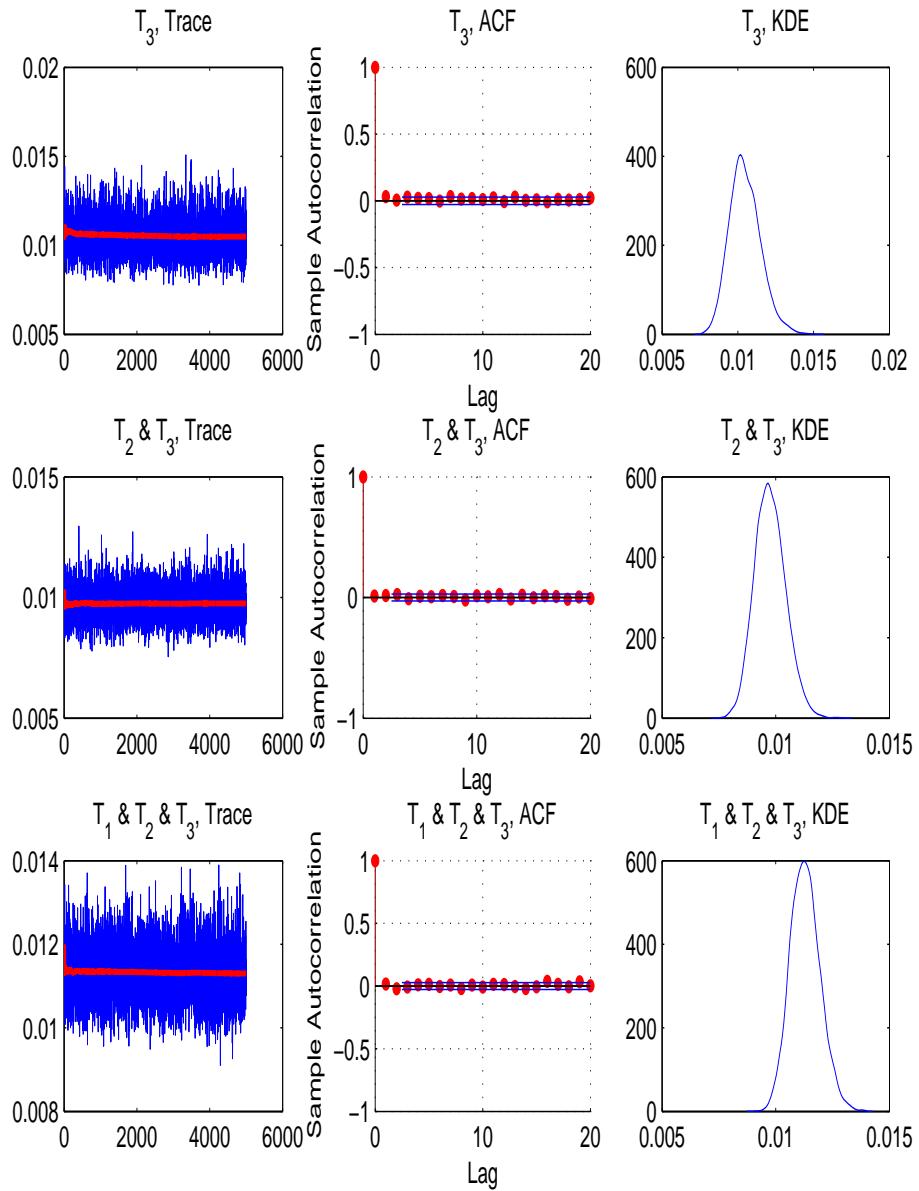
**Figure 4.4.** Scatter plots of option prices against strike prices of the third data set for the simulation study.

The options with maturity  $T_3$  are indicated with a circle, the options with maturity  $T_2$  are indicated with a diamond, and the options with maturity  $T_1$  are indicated with a plus.



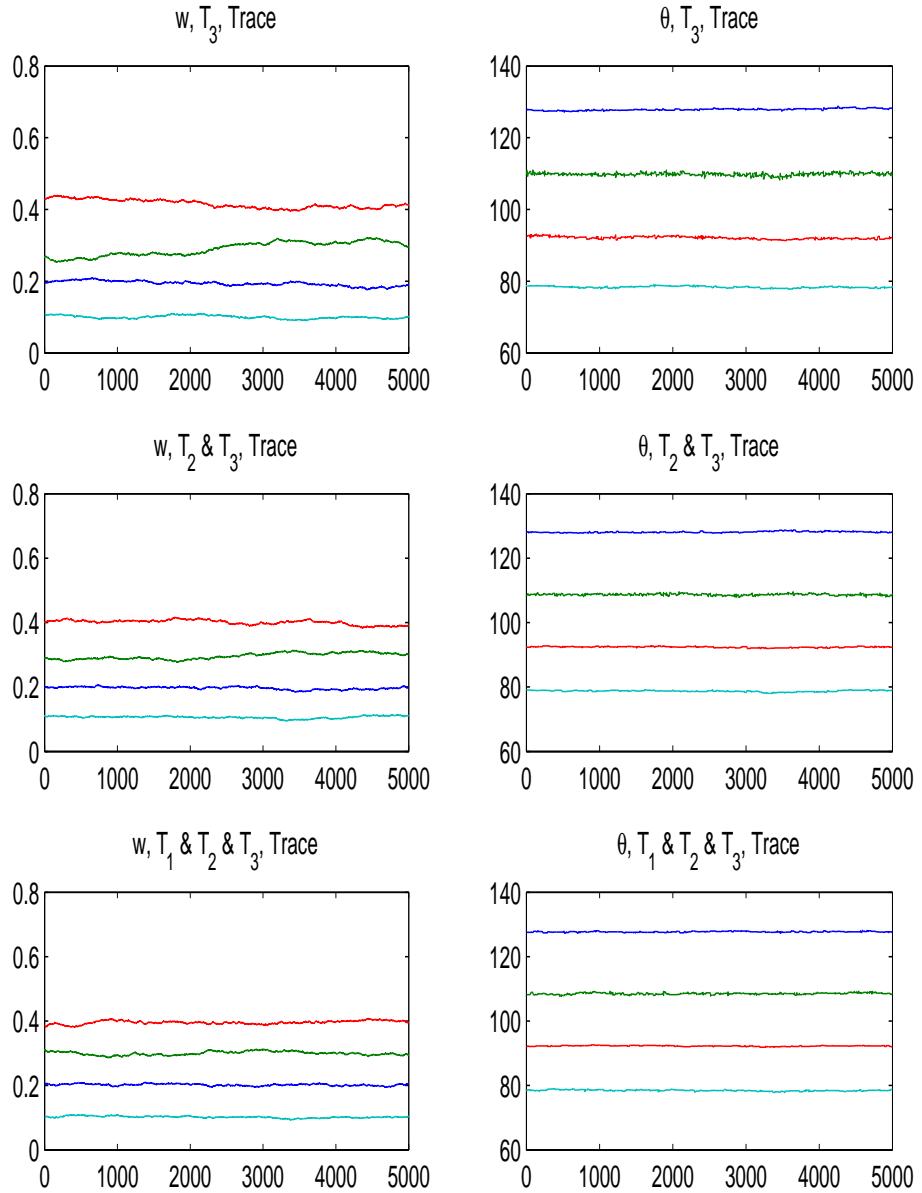
**Figure 4.5.** Convergence MCMC analysis for the simulation study.

The trace plot, autocorrelation function (ACF) plot, and kernel density estimate (KDE) of the log-likelihood are for the simulation data sets. The first data set includes options with one maturity ( $T_3$ ), the second data set includes options with two different maturities ( $T_2$  and  $T_3$ ), and the third data set includes options with three different maturities ( $T_1$ ,  $T_2$ , and  $T_3$ ).



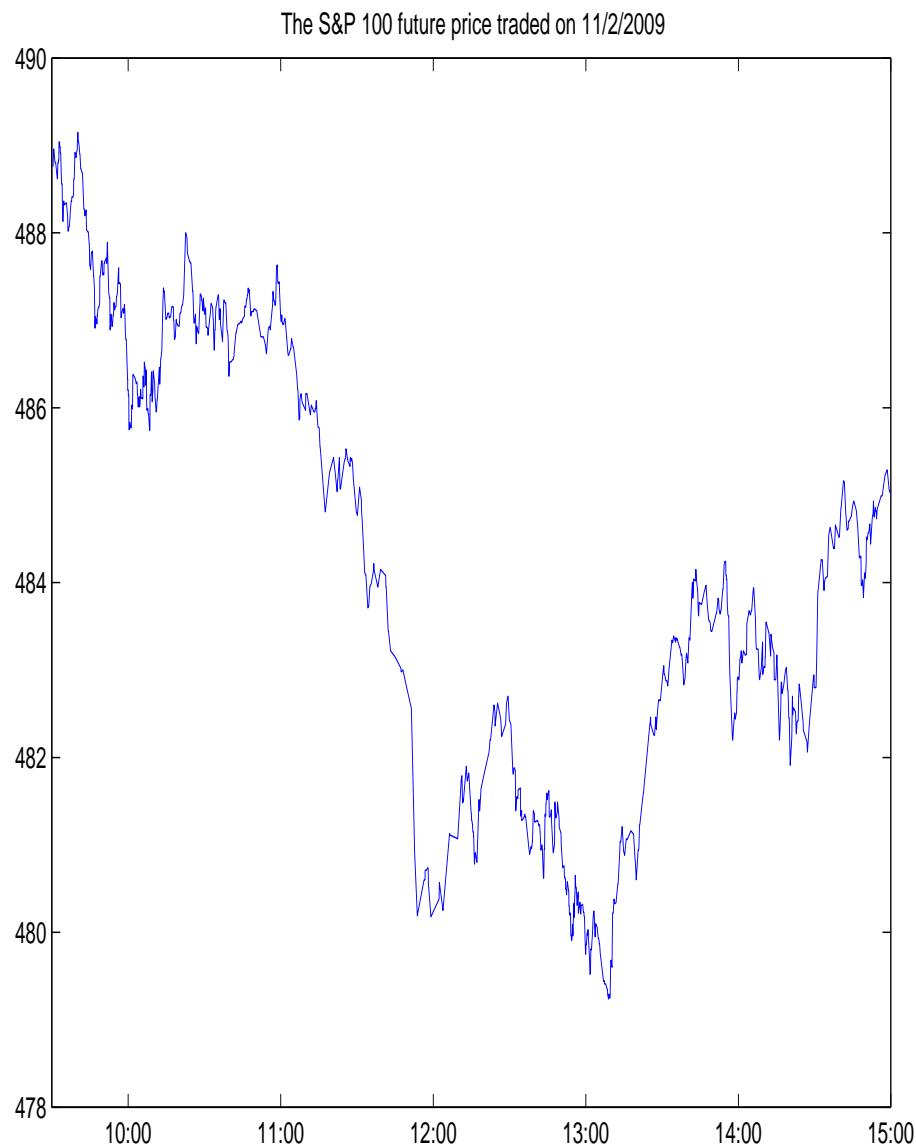
**Figure 4.6.** Trace plots of  $\sigma^2$  for the simulation study.

The trace plot, autocorrelation function (ACF) plot, and kernel density estimate (KDE) of  $\sigma^2$  are depicted for the simulation data sets. The first data set includes options with one maturity ( $T_3$ ), the second data set includes options with two different maturities ( $T_2$  and  $T_3$ ), and the third data set includes options with three different maturities ( $T_1$ ,  $T_2$ , and  $T_3$ ).



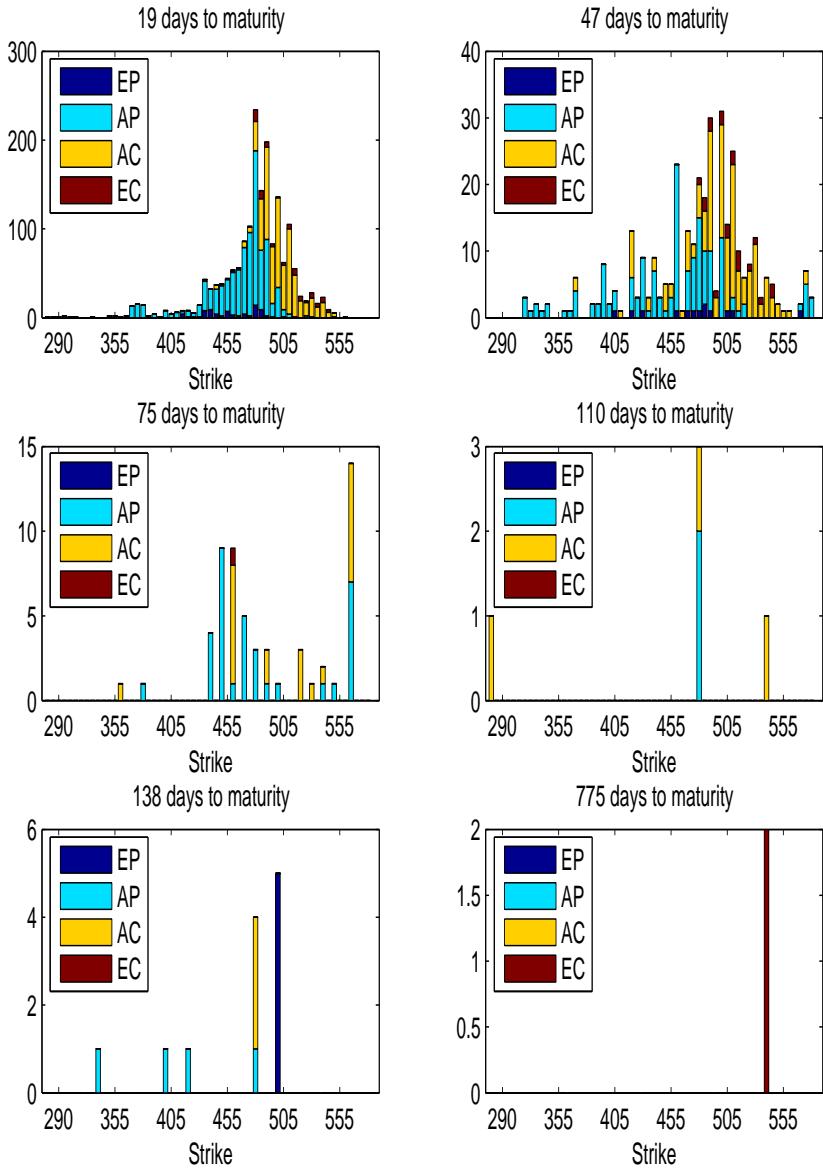
**Figure 4.7.** Trace plots of the locations,  $\theta_n$ , and weights,  $w_n$ , for  $n = 0, \dots, 3$  for the simulation study.

The first data set includes options with one maturity ( $T_3$ ), the second data set includes options with two different maturities ( $T_2$  and  $T_3$ ), and the third data set includes options with three different maturities ( $T_1$ ,  $T_2$ , and  $T_3$ ).

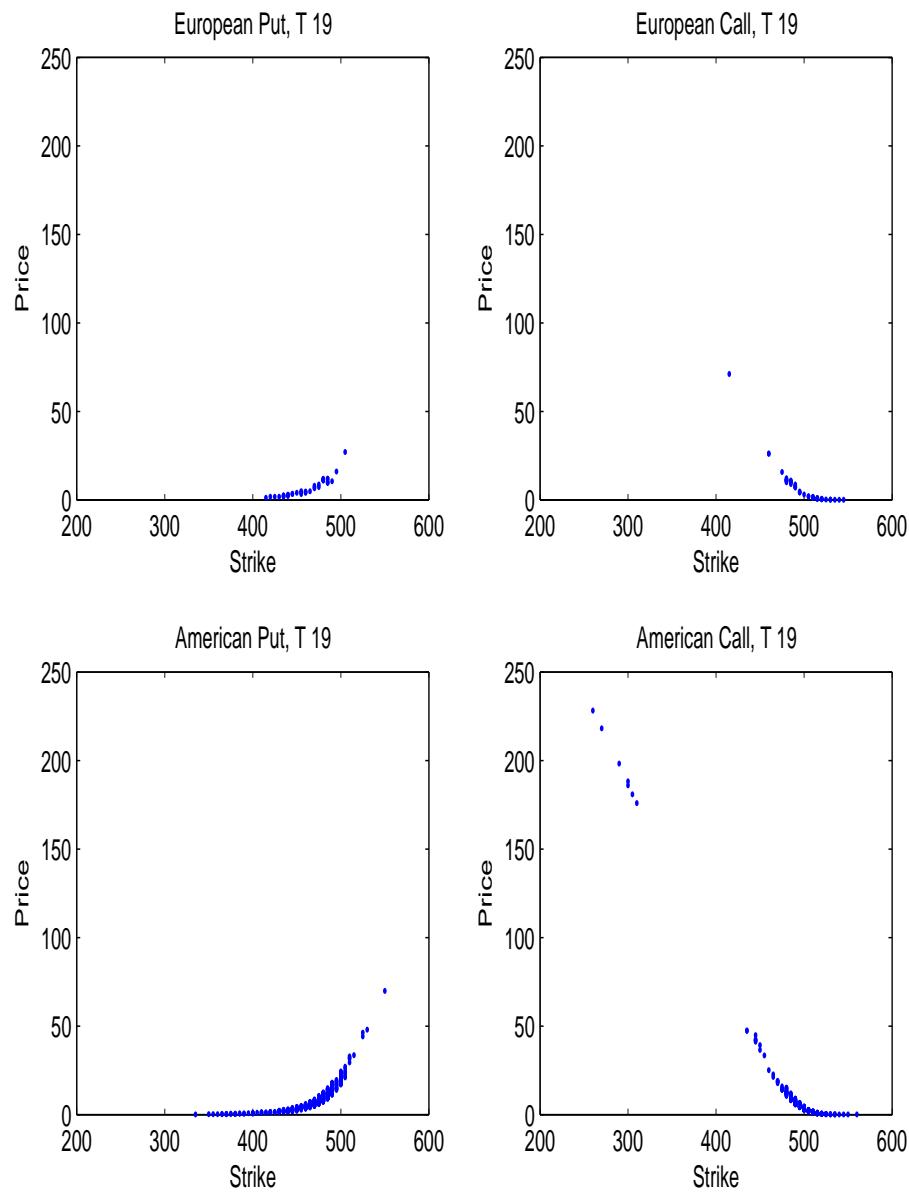


**Figure 4.8.** A time series plot of future prices of the S&P 100 index traded on 11/02/2009.

We depict a time series plot of S&P 100 index future prices from 9:30 am to 3:00 pm on 11/02/2009. This future index opened at 488.76 and closed at 485.01. The traded future price has a mean of 484.57, a median of 485.01, and a standard deviation of 2.60.

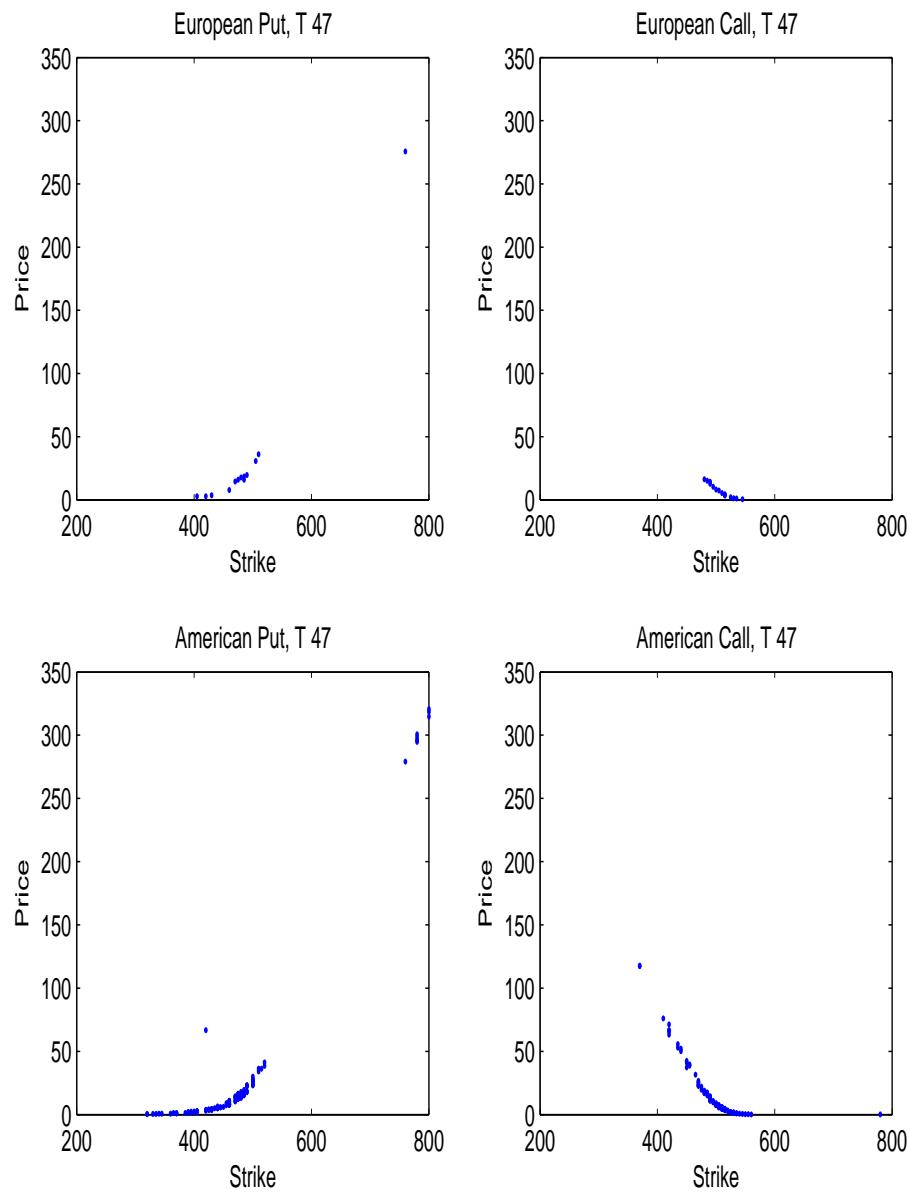


**Figure 4.9.** Stacked bar plots of the number of transactions against strike prices. We present stacked bar plots of the number of transactions against strike prices for options with respect to different maturities. We divide the number of transaction for different options into four groups, including European Put (EP), European Call (EC), American Call (AC), American Put options (AP).



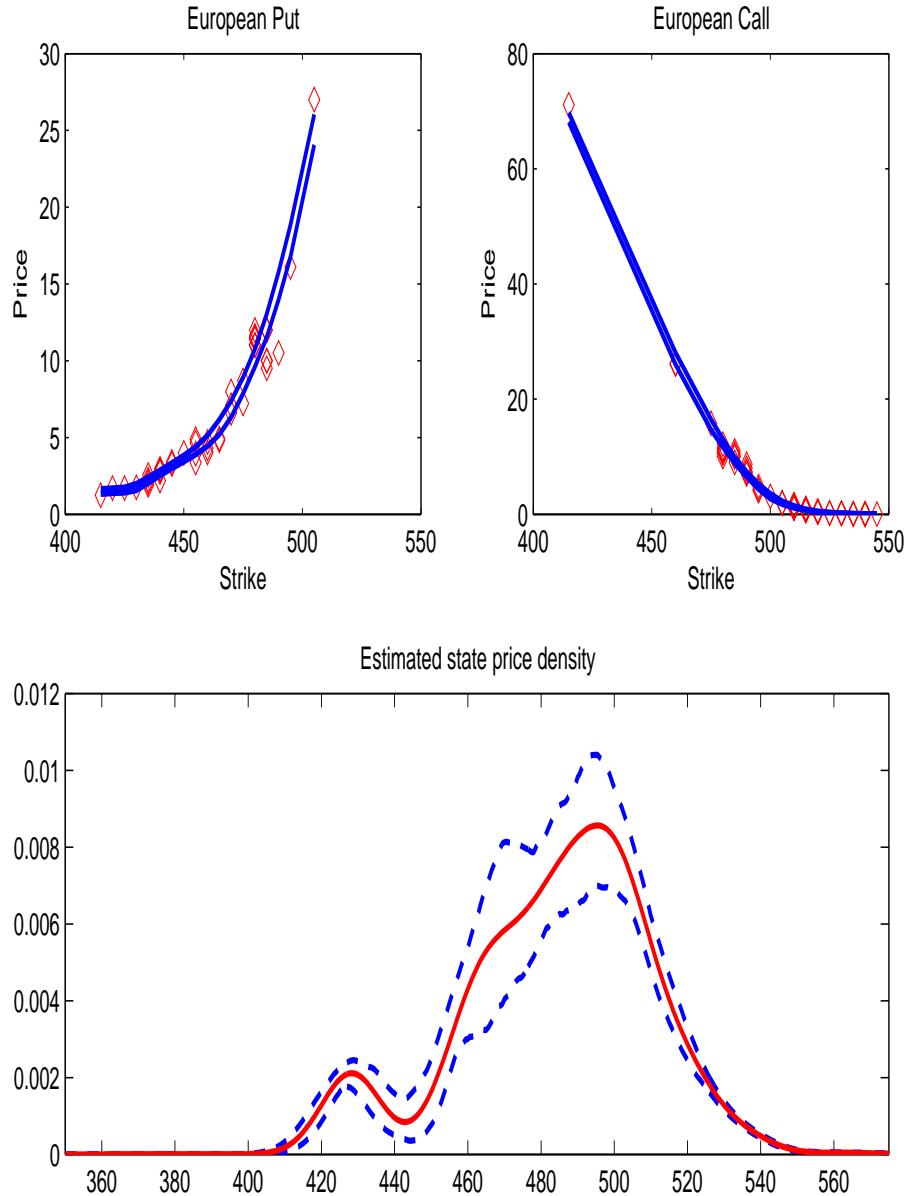
**Figure 4.10.** Scatter plots of S&P 100 index options with 19 days maturity.

We give the scatter plot of S&P 100 index option prices against strike prices for different options into four groups, including European put options, European call options, American put options, and American call options, with 19 days maturity traded on 11/02/2009.



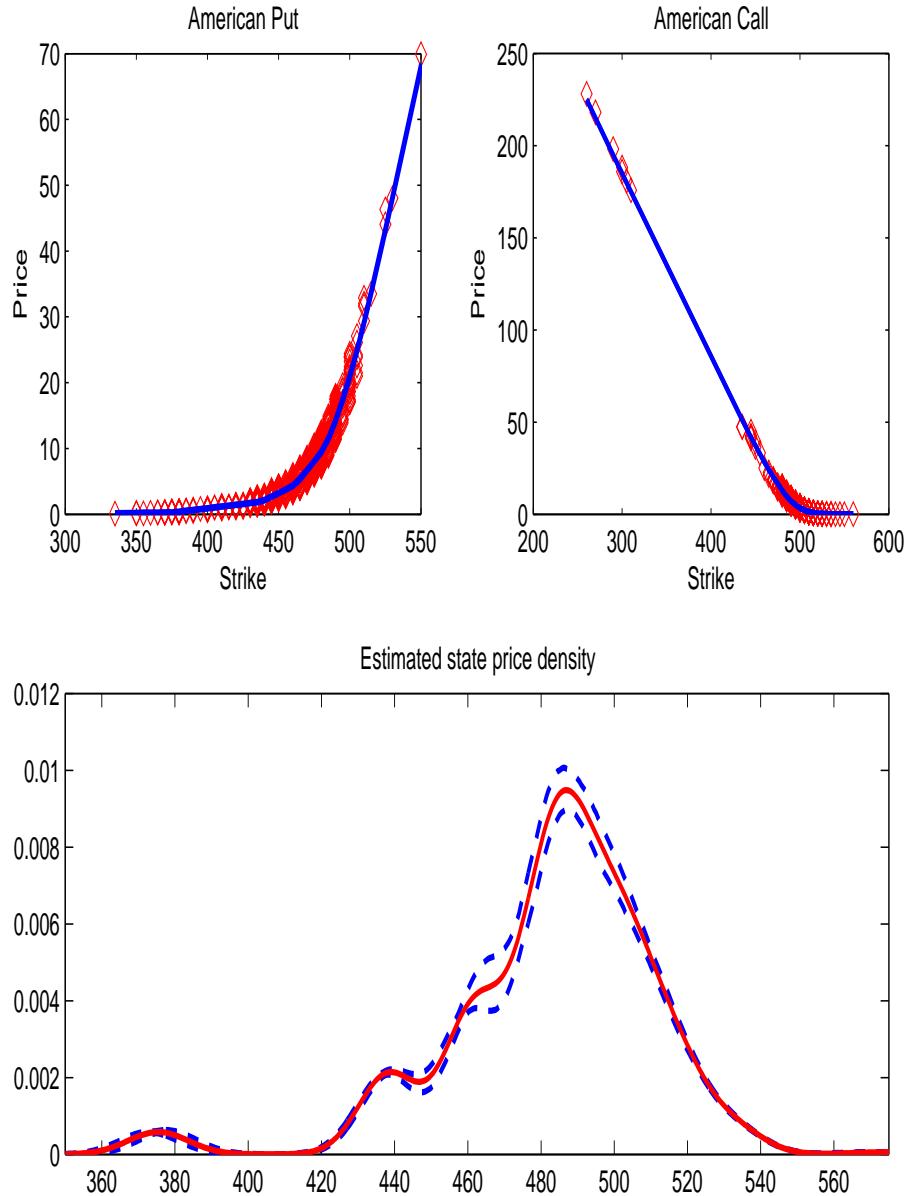
**Figure 4.11.** Scatter plots of S&P 100 index options with 47 days maturity.

We give the scatter plot of S&P 100 index option prices against strike prices for different options into four groups, including European put options, European call options, American put options, and American call options, with 47 days maturity traded on 11/02/2009.



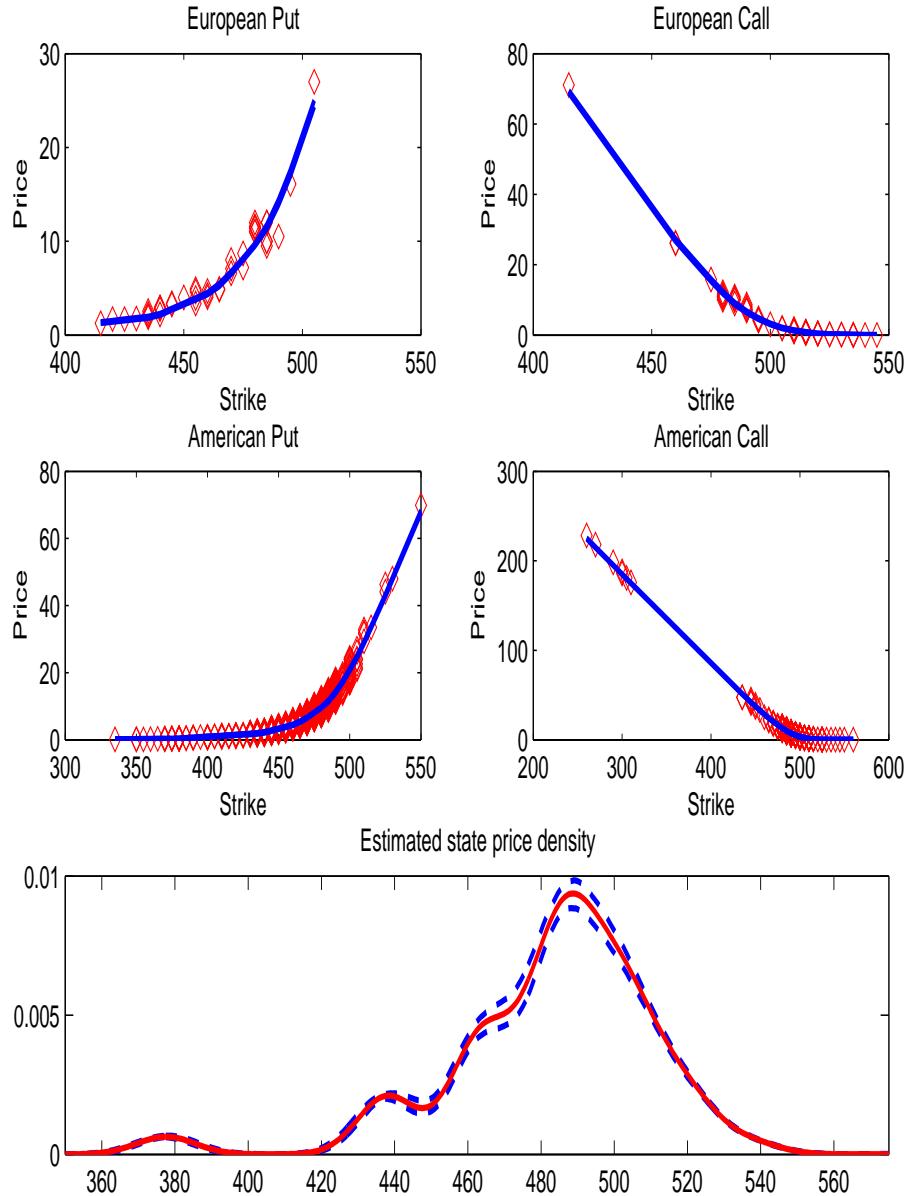
**Figure 4.12.** Scatter plots of actual and fitted option prices against strike prices and the estimated state price density using European options.

The upper panels give scatter plots of actual option prices, against strike prices, indicated with a diamond and the 90% credible region of fitted option prices indicated with a line. In the lowest panel, the posterior mean of the estimated state price density is plotted with a bold line, and the 90% credible region of the estimated state price density indicated is plotted with a dashed line. The estimated state price density is smoothly visualized using a kernel method.



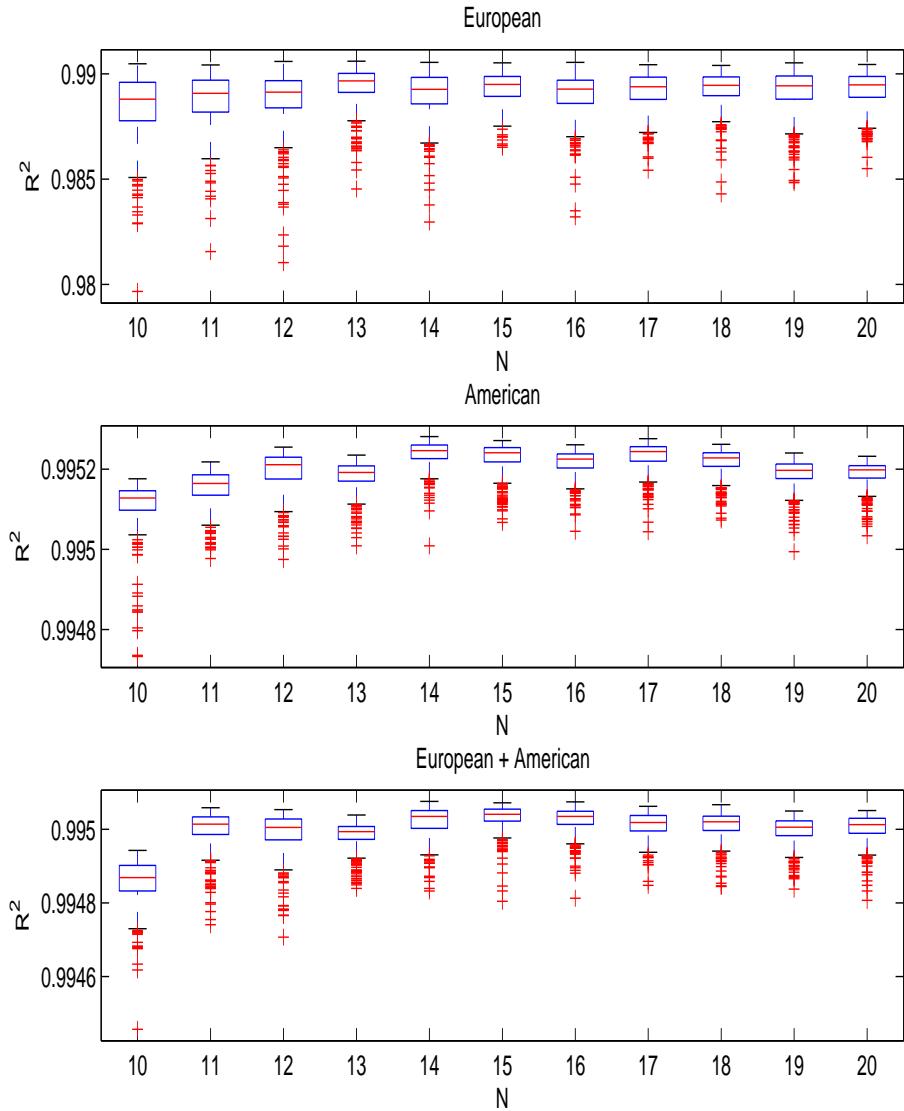
**Figure 4.13.** Scatter plots of actual and fitted option prices against strike prices and the estimated state price density using American options.

The upper panels give scatter plots of actual option prices, against strike prices, indicated with a diamond and the 90% credible region of fitted option prices indicated with a line. In the lowest panel, the posterior mean of the estimated state price density is plotted with a bold line, and the 90% credible region of the estimated state price density indicated is plotted with a dashed line. The estimated state price density is smoothly visualized using a kernel method.

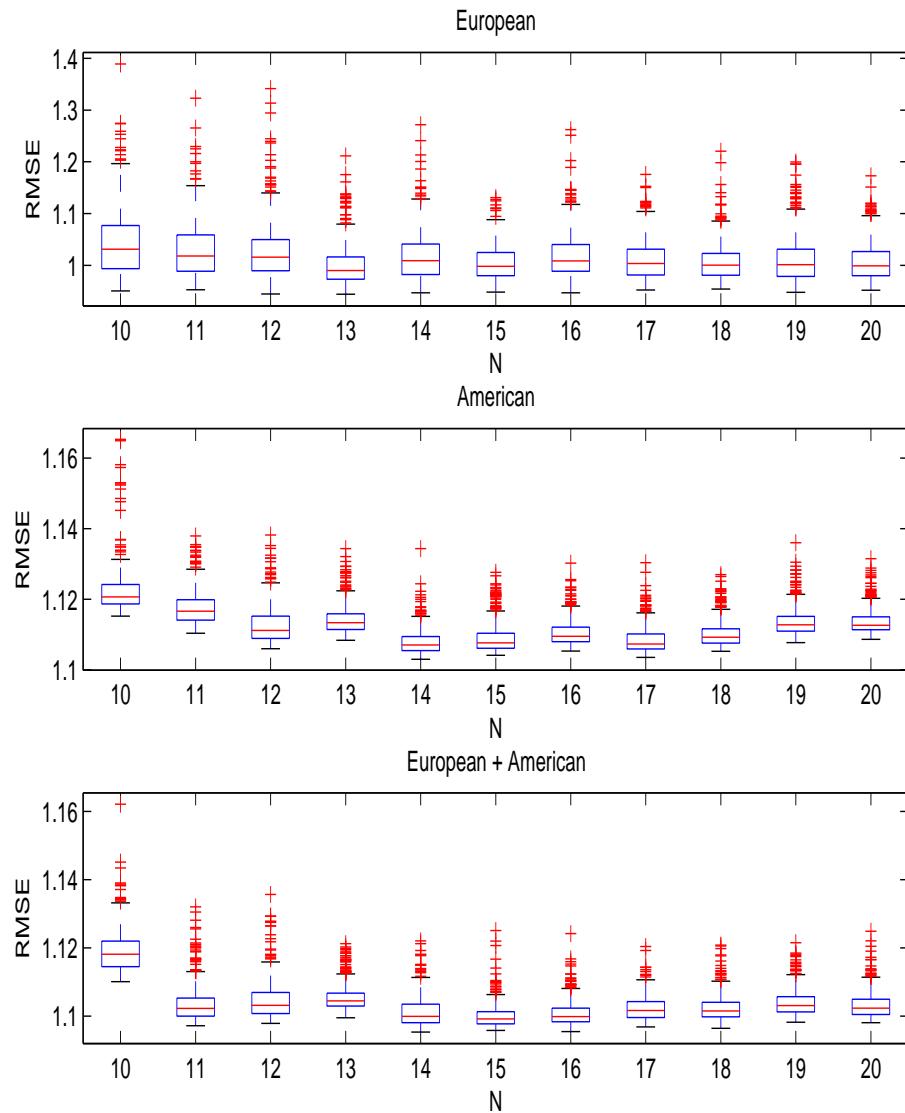


**Figure 4.14.** Scatter plots of actual and fitted option prices against strike prices and the estimated state price density using European and American options.

The upper tow rows of panels give scatter plots of actual option prices, against strike prices, indicated with a diamond and the 90% credible region of fitted option prices indicated with a line. In the lowest panel, the posterior mean of the estimated state price density is plotted with a bold line, and the 90% credible region of the estimated state price density indicated is plotted with a dashed line. The estimated state price density is smoothly visualized using a kernel method.

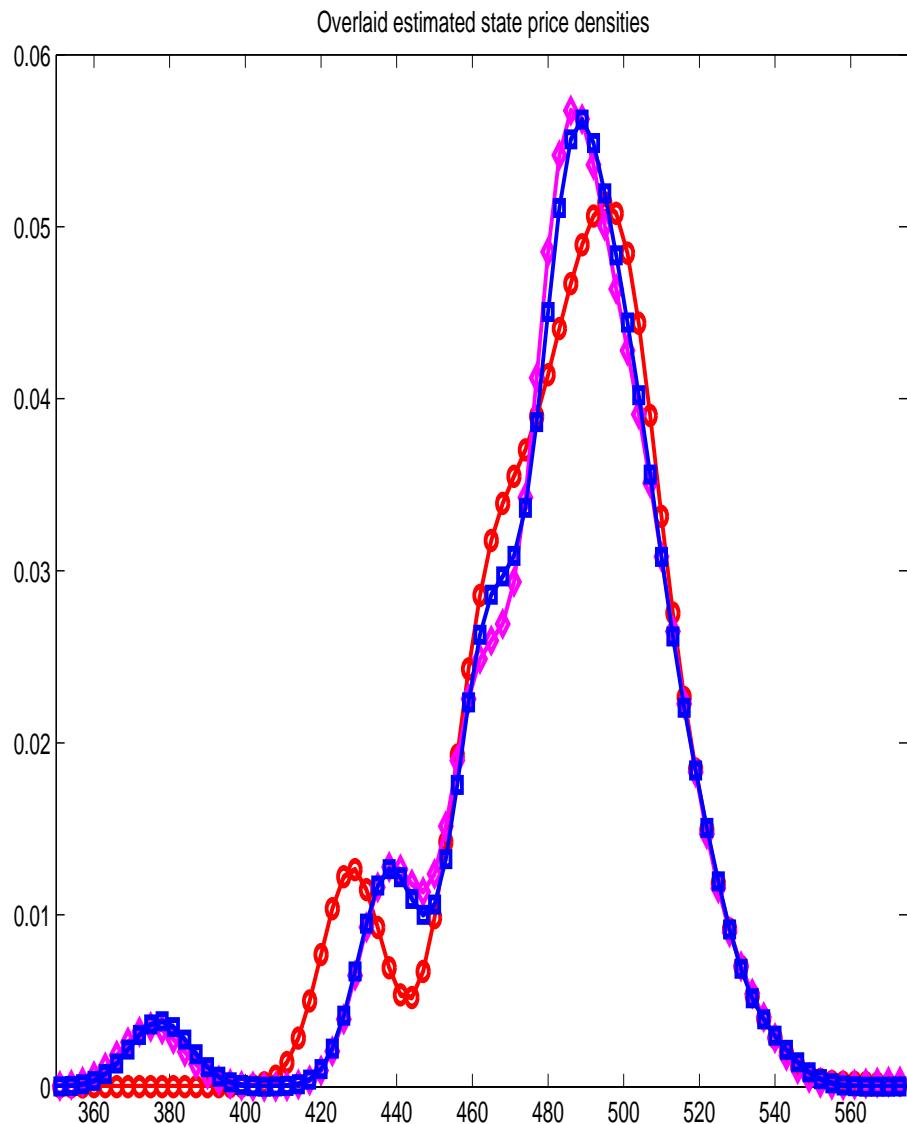


**Figure 4.15.** Summary of posterior distribution of  $R^2$  for the empirical study. A summary of posterior distribution of  $R^2$  as a measure of model fit using box plots with respect to the number of support points, from 10 up to 20, are given for three data sets of S&P 100 index options with 19 days maturity traded on 11/02/2009. The first data set includes just European options, the second data set includes just American options, and the third data set includes both European and American options.



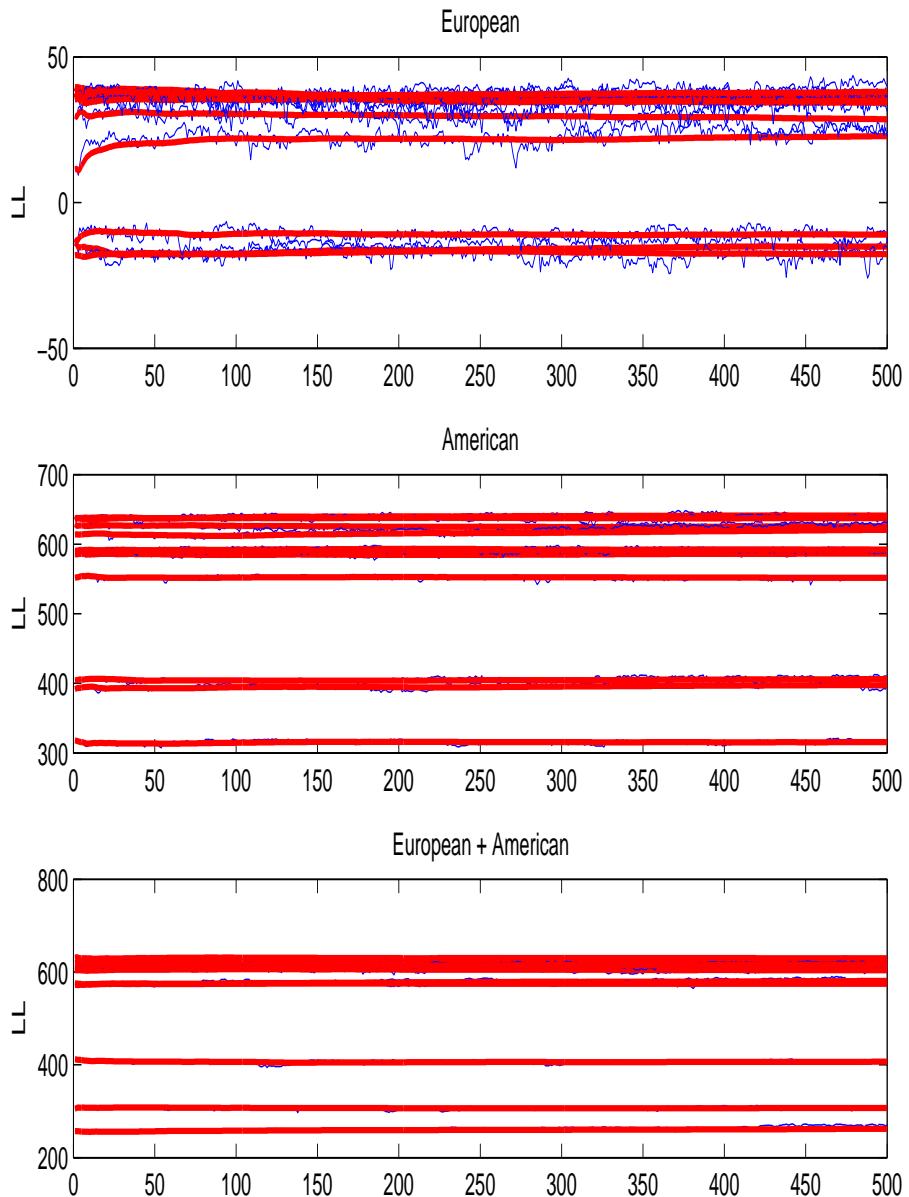
**Figure 4.16.** Summary of posterior distribution of Root-Mean-Squared-Errors of pricing errors for the empirical study.

A summary of posterior distribution of Root-Mean-Square-Error of pricing errors in dollars as a measure of model fit using box plots with respect to the number of support points, from 10 up to 20, are given for three data sets of S&P 100 index options with 19 days maturity traded on 11/02/2009. The first data set includes just European options, the second data set includes just American options, and the third data set includes both European and American options.



**Figure 4.17.** The estimated state price densities using different data sets for the empirical study.

The estimated density using European options data is plotted with a circle-marked line, the estimated density using American options data is plotted with a diamond-marked line, and the estimated density using both European and American options data is plotted with a square-marked line. We find that the former two estimated densities are similar in shape. However, the estimated density using European options data seems to deviate from the other two estimated densities, particularly in the lower tail of the density.



**Figure 4.18.** Trace plots of the log-likelihoods for assessing the convergence of the Markov chain Monte Carlo simulation for the S&P 100 index options data.

We give trace plots of the log-likelihood using three data sets, where the first data set consists of just European options, the second data set consists of just American options, and the third data set consists of both European and American options. For each data, we randomly select 90% of the original data as our calibration data, and use the remaining 10% for predictions; we repeat this ten times, and give ten trace plots of log-likelihoods for the Markov chain Monte Carlo analysis on each of the ten different calibration data sets. These options are S&P 100 index options with 19 days maturity traded on 11/02/2009.

Chapter **5**

## Parametric Greeks: The non-Lipschitz pathwise method

The option price equals the discounted expected payoff of the security under the risk-neutral measure, and Greeks are price sensitivities with respect to parameters of interest. When closed-form formulas do not exist, Monte Carlo simulation has proved invaluable for computing the prices and Greeks of derivative securities. Although finite-difference with resimulation is the standard method for estimating Greeks, it is in general biased and suffers from erratic behavior when the payoff function is discontinuous.

Direct methods, such as the pathwise method and the likelihood ratio method, are proposed to differentiate the prices formulas directly and hence produce unbiased Greeks (Broadie and Glasserman, 1996). The pathwise method differentiates the payoff function, whereas the likelihood ratio method differentiates the densities. When both methods apply, the pathwise method generally enjoys lower variances, but it requires the payoff function to be Lipschitz continuous. Similar to the pathwise method, our method differentiates the payoff function but lifts the Lipschitz continuity requirements on the payoff function. We build a new but simple mathematical formulation, called the non-Lipschitz pathwise method, so that formulas of Greeks for a broad class of derivative securities can be derived systematically. We then present an importance sampling method to estimate the Greeks. Based on the study in this chapter, we note that a separate paper was done in collaboration (Lyuu and Teng, 2010).

Recall that an option is a financial instrument whose payoff is based on other underlying assets such as stocks, indexes, currencies, volatilities, commodities, bonds, mortgages, other derivatives, temperatures, and countless others. Besides an option's price, its Greeks are equally important. Greeks are the price's sensitivities with respect to certain parameters of interest such as the underlying asset's price, volatility, maturity, and interest rate. Although an option's price can often be observed in the market, this is not the case for its Greeks, which must be calculated. As Greeks are important for measuring and managing risk as well as executing dynamic trading strategies, how to calculate them efficiently and accurately is of critical importance both in theory and in practice (Broadie and Glasserman, 1996).

Let us use  $C$  to denote an option's price,  $S$  the underlying asset's spot price,  $\sigma$  the volatility,  $T$  the maturity time, and  $r$  the risk-free interest rate. Common Greeks are delta ( $\Delta = \partial C / \partial S$ ), gamma ( $\Gamma = \partial^2 C / \partial S^2$ ), vega ( $\Lambda = \partial C / \partial \sigma$ ), theta ( $\Theta = \partial C / \partial T$ ), and rho ( $\rho = \partial C / \partial r$ ). For multi-asset options (also called rainbow options), cross-gammas  $\Gamma_{ij} = \partial^2 C / \partial S_i \partial S_j$ , where  $S_i$  and  $S_j$  are different underlying assets, are also important. Note that some Greeks are first-order partial derivatives, whereas others such as gamma and cross-gamma are second-order ones.

Easy-to-calculate closed-form formulas are rare for options with complicated payoff functions; as a result, numerical methods are indispensable for calculating their prices and Greeks. See Table 5.1 for a variety of rainbow options in literature. Deterministic numerical methods for rainbow options such as numerical integration and PDE-related methods suffer from the curse of dimensionality in that the computational complexity grows exponentially with the dimension. They are hence computationally infeasible. A few methods deal with the curse of dimensionality for a certain type of rainbow options, such as the closed-form approximation formulas of Kirk (1995) and Carmona and Durrelman (2003) for spread options and the method of Hörfelt (2008) for options on the  $k$ -th ranked asset. However, these methods are quite restrictive in their domains of applicability and their Greeks are biased in general. As a result, Monte Carlo simulation has proved to be the most important general-purpose numerical scheme for pricing rainbows options. Please refer to Boyle (1977) and Glasserman (2004).

Simulation methods for estimating the Greeks fall into two broad categories:

**Table 5.1.** A sampling of rainbow options

An asterisked rainbow option is one whose Greeks will be derived in the chapter. The Greeks of all these rainbow options listed above except basket and Madonna options can be obtained directly by our method. For basket and Madonna options, our method is applicable after proper changes of variables in their payoff functions. See Nelken (1996) and Wystup (2006) for a more complete list of rainbow options.

Type	Payoff
Margrabe option (Margrabe, 1978)	$\max(S_1(T) - S_2(T), 0)$
Better-off option (Nelken, 1996; Wystup, 2006)	$\max(S_1(T), \dots, S_n(T))$
Worse-off option (Nelken, 1996; Wystup, 2006)	$\min(S_1(T), \dots, S_n(T))$
Binary Maximum option*	$\mathbf{1}_{(\max(S_1(T), \dots, S_n(T)) > K)}$
Maximum option* (Stulz, 1982; Johnson, 1987)	$\max(\max(S_1(T), \dots, S_n(T)) - K, 0)$
Minimum option (Stulz, 1982; Johnson, 1987)	$\max(\min(S_1(T), \dots, S_n(T)) - K, 0)$
Spread option* (Pearson, 1995; Carmona and Durrelman, 2003)	$\max(S_2(T) - S_1(T) - K, 0)$
Basket average option (Hull, 2002)	$\max\left(\frac{S_1(T) + \dots + S_n(T)}{n} - K, 0\right)$
Multi-strike option (Nelken, 1996; Wystup, 2006)	$\max(S_1(T) - K_1, \dots, S_n(T) - K_n, 0)$
Pyramid rainbow option (Nelken, 1996; Wystup, 2006)	$\max( S_1(T) - K_1  + \dots +  S_n(T) - K_n  - K)$
Madonna rainbow option (Nelken, 1996; Wystup, 2006)	$\max(\sqrt{(S_1(T) - K_1)^2 + \dots + (S_n(T) - K_n)^2} - K, 0)$

methods that involve resimulation and those that do not. The first category, based on finite-difference (FD) approximations, is easy to understand and implement. Let  $\psi$  denote the parameter of interest. In the so-called forward FD method, for example, the first-order Greek is approximated by  $[C(\psi + \delta) - C(\psi)]/\delta$ , and the second-order Greek such as gamma by  $[C(\psi + 2h) - 2C(\psi + h) + C(\psi)]/h^2$ . Here,  $\delta$  denotes the perturbed size throughout the chapter and must be suitably small to avoid bias due to higher-order terms. Note that resimulation is required because each of  $C(\psi)$ ,  $C(\psi - h)$ ,  $C(\psi + h)$ , and  $C(\psi + 2h)$  has to be established by simulation. A variant is the more accurate but also more costly central FD method. For this method, the first-order Greek becomes  $[C(\psi + h/2) - C(\psi - h/2)]/h$ , and the second-order Greek becomes  $[C(\psi + h) - 2C(\psi) - C(\psi - h)]/h^2$ . Table 5.2 lists related formulas.

Although FD approximations are straightforward, they have one severe weakness in that deciding on the right  $\delta$  is difficult. If  $\delta$  is too large, the Greek estimate would be biased because of the nonlinearity of  $C(\psi)$ . When  $C(\psi)$  is differentiable, FD approximations should be expected to converge to the true value with  $\delta$  small enough. However, this is not the case numerically for simulation methods. In fact, if  $\delta$  is too small, the variation between the original price  $C(\psi)$  and the perturbed prices  $C(\psi \pm h)$  makes the Greek estimates unstable. Although using common random numbers in resimulation can reduce the estimation error, the above observations remain valid even if variance reduction techniques or stratified sampling are employed (Zazanis and Suri, 1993; Jäckel, 2002). Higher-order Greek estimates are in most cases numerically unstable. That higher-order partial derivatives are estimated at a slower rate of convergence is sometimes referred to as the curse of differentiation.

Methods in the second category produce unbiased estimates but are more elaborate. With the direct method of Broadie and Glasserman (1996), the information from a single simulation is used to estimate multiple Greeks besides the option's price. The direct method does not rely on resimulation in calculating the Greeks. Popular direct methods include the pathwise method and the likelihood ratio method. The pathwise method differentiates each simulated outcome with respect to the parameters of interest. The likelihood ratio method, on the other hand, differentiates the probability density function rather than the outcome.

**Table 5.2.** Finite-difference formulas for deltas, gammas and cross-gammas

$\varphi$  is the option payoff function,  $\varphi(S_i + h)$  denotes the payoff when the initial underlying asset's price  $S_i$  changes to  $S_i + h$ , and so on. The formulas assume the same form if the parameter of interest is not the stock price. For a complete list of formulas with error rates, see Tavella and Randall (2009).

Greeks	<b>Forward FD schemes</b>
$\Delta_i$	$e^{-rT} [E[\varphi(S_i + \delta)] - E[\varphi(S_i)]] / \delta$
$\Gamma_{ii}$	$e^{-rT} [E[\varphi(S_i + 2\delta)] - 2E[\varphi(S_i + \delta)] + E[\varphi(S_i)]] / \delta^2$
$\Gamma_{ij}$	$e^{-rT} [E[(S_i + \delta, S_j + \delta)] - E[(S_i + \delta, S_j)] - E[\varphi(S_i, S_j + \delta)] + E[\varphi(S_i, S_j)]] / \delta^2$
Greeks	<b>Central FD schemes</b>
$\Delta_i$	$e^{-rT} [E[\varphi(S_i + \delta/2)] - E[\varphi(S_i - \delta/2)]] / \delta$
$\Gamma_{ii}$	$e^{-rT} [E[\varphi(S_i + \delta)] - 2E[\varphi(S_i)] + E[\varphi(S_i - \delta)]] / \delta^2$
$\Gamma_{ij}$	$e^{-rT} [E[\varphi(S_i + \delta/2, S_j + \delta/2)] - E[\varphi(S_i + \delta/2, S_j - \delta/2)] - E[\varphi(S_i - \delta/2, S_j + \delta/2)] + E[\varphi(S_i - \delta/2, S_j - \delta/2)]] / \delta^2$

Although the domains of applicability of the pathwise method and the likelihood ratio method overlap, no method dominates the other. When both apply, the pathwise method generally enjoys lower variances. Unfortunately, the applicability of the pathwise method is limited by the requirement of Lipschitz continuity in the payoff function, which is needed for convergence guarantees. Although the likelihood ratio method can differentiate discontinuous payoff functions, it needs the explicit formula of the density function. Fournie et al. (1999) applies integration-by-parts formulas to derive the Greeks using Malliavin calculus. This approach avoids differentiating a discontinuous payoff function and deriving the densities for the underlying securities. Greeks are the expectations of the product of the payoff function and the Malliavin weighting function. Benhamou (2003) further bridges the Malliavin weighting method and the likelihood ratio method. It is shown that the likelihood ratio is the weighting function with the smallest total variance.

Several approaches have been proposed to extend the pathwise method to lift the requirement of Lipschitz continuity. For example, the conditional Monte Carlo method in Fu and Hu (1997) is a very general scheme to smooth the discontinuous integrand by conditioning on some random variables. However, the conditional Monte Carlo method is problem dependent and may be difficult to implement, particularly for higher-order Greeks. In contrast, we provide a new mathematical formulation so that the Greeks can be derived systematically as long as the payoff function belongs in the class  $\mathcal{C}$  defined later. Recently, Liu and Hong (2008) extend the pathwise method to options with discontinuous payoff functions. However, our Greek estimates are unbiased and efficient, but their kernel estimates are biased and may suffer from slow convergence rates.

We now briefly summarize our approach as follows. An option's price equals the discounted expected value of its payoff function under the risk-neutral probability measure (Harrison and Kreps, 1979). The option value is therefore a discounted integral whose integrand is the product of the payoff and a probability density function. Recall that Greeks are the partial derivatives of the price with respect to the parameter of interest. Our method is applicable whenever the payoff function belongs to a class we name  $\mathcal{C}$ .  $\mathcal{C}$  is roughly a family of payoff functions that can be written as a sum of products of differentiable functions and indicator functions with the right kind of support. A formal definition will be given later. For a

payoff function from  $\mathcal{C}$ , we (1) prove that expectation (equivalently, integration) and differentiation can be interchanged, (2) provide the “differentiation” of the indicator functions, and (3) guarantee the validity of the product rule. As a result, as long as an option’s payoff function belongs to  $\mathcal{C}$ , its Greeks can be derived systematically and calculated without bias. The formulas for the Greeks of two rainbow options — the spread option and the maximum option — are given to illustrate the broad applicability of our method. We call our method the non-Lipschitz pathwise method, because our method differentiates the payoff function but lifts the Lipschitz continuity requirements on the payoff function. We will concentrate on deltas and gammas as other Greeks can be treated similarly.

The proposed non-Lipschitz pathwise method has several advantages over other schemes. First, it addresses the requirement of Lipschitz continuity for the payoff function by the pathwise method. This condition has severely restricted the applicability of the pathwise method. For example, digital options cannot be handled by the pathwise method as their payoff function is an indicator function. But they pose no problems for our method. Second, our method gives provably unbiased estimates by doing away with FD. Third, our method is more efficient than FD approximations as resimulation is avoided. This feature is particularly beneficial for higher-order Greeks as we shall see later. Moreover, numerical results show that our method produces Greek estimates with lower standard errors than the likelihood ratio method. Last but not the least, our method is easy to implement and its application is almost mechanical compared with other available methods, such as the conditional Monte Carlo method and the likelihood ratio method. Subsequent sections will establish all these claims.

We organize this chapter as follows. In Section 5.1, the mathematics necessary for the derivation of the Greeks are established. Fundamental theorems to differentiate an integration whose integrand is a product of a differentiable function and several indicator functions are presented in Section 5.2. In Section 5.3, formulas of the Greeks for several popular rainbow options are given. This chapter ends with numerical results and conclusions in Sections 5.4 and 5.5, respectively. The appendices contain proofs for several technical results stated in the main text.

## 5.1 Preliminaries

Let  $\mathbf{x} = (x_1, \dots, x_n)^t$ , where the superscript “ $t$ ” means matrix transpose throughout this dissertation. The  $n$ -variate normal distribution with mean vector  $\mu$  and covariance matrix  $\Sigma$  is expressed as  $N_n(\mu, \Sigma)$ . When  $\Sigma$  is positive definite and  $\mathbf{x} \sim N_n(\mu, \Sigma)$ , the probability density function (pdf) of  $\mathbf{x}$  is

$$f(\mathbf{x}; \mu, \Sigma) = \frac{e^{-(\mathbf{x}-\mu)^t \Sigma^{-1} (\mathbf{x}-\mu)/2}}{\sqrt{(2\pi)^n |\Sigma|}}. \quad (5.1)$$

Above,  $|\Sigma|$  denotes the determinant of  $\Sigma$ . This chapter uses  $n$  to denote the number of underlying assets. Let  $\psi$  denote a parameter of interest throughout this chapter. The indicator function  $\mathbf{1}_A(\mathbf{x})$  is a function from  $\Re^n$  to  $\{1, 0\}$  with a subset  $A \subseteq \Re^n$  such that

$$\mathbf{1}_A(\mathbf{x}) = \begin{cases} 1, & \text{for } \mathbf{x} \in A, \\ 0, & \text{otherwise.} \end{cases}$$

To simplify the presentation, we may simply write down the predicate defining the set  $A$  instead of using the set-theoretical notation as above. The set  $A$  is also called the support of the indicator function.

Let  $\wp$  denote the payoff function of an option. For the Black-Scholes model, the option’s price is given by  $C = e^{-rT} E[\wp]$ , where  $E$  denotes the expectation operator under the risk-neutral measure throughout this dissertation (Harrison and Kreps, 1979). When the expectation  $E[\wp]$  is intractable to calculate deterministically, we estimate the expected value  $e^{-rT} E[\wp]$  by

$$\frac{1}{N} e^{-rT} \sum_{n=1}^N \wp(\omega^{(n)}),$$

where the  $N$  sampled paths  $\omega^{(1)}, \dots, \omega^{(n)}$  are drawn from a proper distribution. This is the standard Monte Carlo method (Glasserman, 2004). The strong law of large numbers guarantees that the estimated number will converge to  $e^{-rT} E[\wp]$  with probability one when the sample size  $N$  is large enough under very loose regularity conditions.

For a rainbow option in the Black-Scholes model, the price of each underlying

asset follows the lognormal diffusion process in the risk-neutral economy,

$$dS_i = rS_i dt + \sigma_i S_i dZ_i, \quad \text{for } i = 1, \dots, n,$$

where

- $S_i$  = the current price of asset  $i$ ,
- $r$  = the risk-free interest rate,
- $\sigma_i$  = the volatility of asset  $i$ ,
- $dZ_i$  = the Wiener process,
- $\rho_{ij}$  = the instantaneous correlation between  $dZ_i$  and  $dZ_j$ .

Let  $S_i(T)$  denote the price of asset  $i$  at maturity and  $\mathbf{S}_T$  denote the asset prices at maturity, i.e.,  $\mathbf{S}_T = (S_1(T), \dots, S_n(T))$ . By the lognormality of  $\mathbf{S}_T$ ,

$$S_i(T) = S_i e^{(r - \sigma_i^2/2)T + \sigma_i \sqrt{T} x_i}, \quad \mathbf{x} \sim N_n(\mathbf{0}, \Sigma), \quad i = 1, \dots, n, \quad (5.2)$$

where the correlation matrix is defined as  $\Sigma = [\rho_{ij}]$  with  $\rho_{ij}$  being the instantaneous correlation between  $dZ_i$  and  $dZ_j$ , and  $\mathbf{x}$  is the underlying randomness driving the transition from the initial stock prices to the stock prices at maturity. Note that  $S_i(T)$  increases monotonically with  $x_i$ . As  $\wp$  depends on  $\mathbf{S}_T$ , which in turn depends on  $\mathbf{x}$ , the rainbow option's price is the following integral:

$$C = e^{-rT} E[\wp(\mathbf{S}_T)] = e^{-rT} \int_{\mathbb{R}^n} \wp(\mathbf{S}_T) f(\mathbf{x}; \mathbf{0}, \Sigma) d\mathbf{x}. \quad (5.3)$$

Note that the payoff function  $\wp$  of a rainbow option depends on  $\mathbf{S}_T$  and recall that  $\psi$  is the parameter of interest. Since  $\mathbf{S}_T$  depends on  $\psi$  and a normally distributed random term  $\mathbf{x}$ , we may use  $\wp(\psi, \mathbf{x})$  for  $\wp$  to make the dependency explicit. Equation (5.3) shows that an option's price can be written as an integral of the form:

$$C = e^{-rT} E[\wp] = e^{-rT} \int_{\mathbb{R}^n} \wp(\psi, \mathbf{x}) f(\mathbf{x}; \mathbf{0}, \Sigma) d\mathbf{x}. \quad (5.4)$$

Note that Eq. (5.4) integrates over the domain of  $\mathbf{x}$ .

A first-order Greek with respect to  $\psi$  equals

$$\frac{\partial}{\partial \psi} e^{-rT} E[\varphi].$$

Our methodology will depend on the validity of the interchange of the order of expectation and differentiation, i.e.,

$$\frac{\partial}{\partial \psi} e^{-rT} E[\varphi] = e^{-rT} E \left[ \frac{\partial \varphi}{\partial \psi} \right]. \quad (5.5)$$

If Eq. (5.5) holds, the right-hand side equals the desired Greek, which is again an expectation and can often be estimated without bias by the standard Monte Carlo method. Equally important, if Eq. (5.5) holds, then there is no need for FD or resimulation. Under the assumption that Eq. (5.5) holds, a second-order Greek with respect to  $\psi$  equals

$$\frac{\partial^2}{\partial \psi^2} e^{-rT} E[\varphi] = \frac{\partial}{\partial \psi} e^{-rT} E \left[ \frac{\partial \varphi}{\partial \psi} \right]. \quad (5.6)$$

If it is again valid to interchange the order of expectation and differentiation in the right-hand side of Eq. (5.6), we have

$$\frac{\partial^2}{\partial \psi^2} e^{-rT} E[\varphi] = e^{-rT} E \left[ \frac{\partial^2 \varphi}{\partial \psi^2} \right]$$

by Eq. (5.5). In this case, a gamma equals the delta of a delta. Exactly the same argument can be repeated for cross-gammas and higher-order Greeks.

Broadie and Glasserman (1996) establish a set of conditions on  $\varphi$  for Eq. (5.5) to hold, which yields the pathwise method. Loosely speaking, the order of differentiation and expectation can be interchanged for Lipschitz continuous payoff functions. The pathwise method, however, may not be applicable when the payoff function is not Lipschitz continuous. Indeed, many payoff functions are not Lipschitz continuous. The payoff function of a digital option, for example, is not Lipschitz continuous. The pathwise method is even less applicable to calculating the Gammas (Broadie and Glasserman, 1996). For example, the European call

option's price is

$$e^{-rT} \int_{\mathbb{R}} (S(T) - K) \mathbf{1}_{\{S(T)>K\}}(x) f(x; 0, \sigma^2) dx,$$

where  $S(T) = S e^{(r-\sigma^2/2)T + \sigma\sqrt{T}x}$  with  $x \sim N(0, \sigma^2)$ . Since the integrand above is Lipschitz continuous with respect to  $S$ , its delta can be derived by the pathwise method as

$$e^{-rT} \int_{\mathbb{R}} e^{(r-\sigma^2/2)T + \sigma\sqrt{T}x} \mathbf{1}_{\{S(T)>K\}}(x) f(x; 0, \sigma^2) dx$$

As the integrand above is no longer Lipschitz continuous, the pathwise method can not be used to derive the gamma of the call option without undergoing some modifications. In this case, Greeks with the pathwise method do not yield to a simple, unified treatment.

One salient feature of an option's payoff is that a positive cash flow occurs only when the underlying asset's prices meet certain conditions, which is the reason it is called a contingent claim. For example, a vanilla call option pays off only when the stock price at maturity is higher than the strike price  $K$ . As a result, its payoff function is

$$\max(S(T) - K, 0) = (S(T) - K) \times \mathbf{1}_{\{S(T)-K>0\}}(x),$$

which is a product of a differentiable function and an indicator function. Recall that  $S(T) = S e^{(r-\sigma^2/2)T + \sigma\sqrt{T}x}$ , where  $x \sim N(0, 1)$ . With  $g(\psi, x) \equiv S(T) - K$ , where  $\psi$  is the parameter of interest, the support  $\{S(T) - K > 0\}$  of the indicator function becomes  $\{g(\psi, x) > 0\}$ . The options studied in this chapter will undergo similar transformations.

We now formalize the above-mentioned transformation. Given any function  $h(\psi, \mathbf{x})$  with pdf  $f(\mathbf{x})$  for  $\mathbf{x}$ , let  $h_\psi(\psi, \mathbf{x})$  denote the partial differentiation of  $h(\psi, \mathbf{x})$  with respect to  $\psi$ . Recall that  $x_k$  is the  $k$ -th component of  $\mathbf{x}$ . We define a class of functions called  $\mathcal{H}_k$  such that  $h(\psi, \mathbf{x}) \in \mathcal{H}_k$  if

1.  $h(\psi, \mathbf{x})$  is  $\psi$ -differentiable;
2.  $\int |h(\psi, \mathbf{x})| f(\mathbf{x}) d\mathbf{x} < \infty$  and  $\int |h_\psi(\psi, \mathbf{x})| f(\mathbf{x}) d\mathbf{x} < \infty$ ;
3.  $h(\psi, \mathbf{x})f(\mathbf{x})$  and  $h_\psi(\psi, \mathbf{x})f(\mathbf{x})$  are uniformly continuous with respect to  $\psi$

and  $x_k$  on a compact set.

When  $n = 1$ , we simply write  $h(\psi, x) \in \mathcal{H}$  as  $\mathbf{x}$  is 1-dimensional.

Let  $\mathbf{x}_k = (x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n)^t$ , which denotes  $\mathbf{x}$  with the  $k$ -th component removed. We now define another class of functions named  $\mathcal{G}_k$ , such that  $g(\psi, \mathbf{x}) \in \mathcal{G}_k$  if the following properties are satisfied:

1.  $g(\psi, \mathbf{x})$  is  $\psi$ -differentiable and  $\partial g(\psi, \mathbf{x})/\partial\psi$  is continuous in  $\psi$ ;
2.  $g(\psi, \mathbf{x})$  is  $x_k$ -differentiable and  $\partial g(\psi, \mathbf{x})/\partial x_k$  is continuous in  $x_k$ ;
3.  $g(\psi, \mathbf{x})$  is strictly monotone in  $x_k$ ;
4. There exists a point for  $x_k$  depending on  $\psi$  and  $\mathbf{x}_k$ , written as  $\chi(\psi, \mathbf{x}_k)$ , such that

$$g(\psi, \mathbf{x})|_{x_k=\chi(\psi, \mathbf{x}_k)} \equiv g(\psi, x_1, \dots, x_{k-1}, \chi(\psi, \mathbf{x}_k), x_{k+1}, \dots, x_n) = 0.$$

When  $n = 1$ , we write  $g(\psi, x) \in \mathcal{G}$  as  $\mathbf{x}$  is 1-dimensional and we replace  $\chi(\psi, \mathbf{x}_k)$  with  $\chi(\psi)$  because  $\mathbf{x}_k$  is 0-dimensional. In general, by “ $g_j(\psi, \mathbf{x}) \in \mathcal{G}_k$  for  $j \in \mathcal{B}$ , where  $\mathcal{B}$  is a finite set of natural numbers” we mean that for each  $j \in \mathcal{B}$ ,  $g_j(\psi, \mathbf{x})$  belongs in  $\mathcal{G}_k$  with corresponding function  $\chi_j(\psi, \mathbf{x}_k)$  such that  $g_j(\psi, \mathbf{x})|_{x_k=\chi_j(\psi, \mathbf{x}_k)} = 0$ . We remark that the fourth assumption above does not require the existence of a closed-form formula for  $\chi(\psi, \mathbf{x}_k)$ . In addition, these four assumptions ensure that the partial differentiation of  $\chi(\psi, \mathbf{x}_k)$  with respect to  $\psi$  can be calculated alternatively as

$$\frac{\partial \chi(\psi, \mathbf{x}_k)}{\partial \psi} = - \left[ \frac{\partial g(\psi, \mathbf{x})/\partial\psi}{\partial g(\psi, \mathbf{x})/\partial x_k} \right]_{x_k=\chi(\psi, \mathbf{x}_k)}$$

by the implicit function theorem Protter and Morrey (1991). As will be shown later in Theorems 1, 2 and 3, the term  $\partial g(\psi, x)/\partial x_k$  in the denominator is nonzero because of the third property that  $g(\psi, \mathbf{x})$  is strictly monotone in  $x_k$ .

We are ready to define the desired class of payoff functions:  $\mathcal{C}$ . In the simplest case, we say  $\varphi(\psi, \mathbf{x}) \in \mathcal{C}$  if  $\varphi(\psi, \mathbf{x})$  can be written as

$$\varphi(\psi, \mathbf{x}) = h(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}), \quad (5.7)$$

where

1.  $h(\psi, \mathbf{x}) \in \mathcal{H}_k$  with pdf  $f(\mathbf{x})$  for  $f(\mathbf{x})$ ;
2.  $g_j(\psi, \mathbf{x}) \in \mathcal{G}_k$  for  $j \in \mathcal{B}$ , a finite set of natural numbers;
3. For  $j \in \mathcal{B}$ ,

$$\int_{\mathbb{R}^n} \left| h(\psi, \mathbf{x}) \frac{\partial g_j(\psi, \mathbf{x}) / \partial \psi}{\partial g_j(\psi, \mathbf{x}) / \partial x_k} \right| f(\mathbf{x}) d\mathbf{x} < \infty.$$

We remove the subscript  $j$  from  $g_j(\psi, \mathbf{x})$  when  $\mathcal{B}$  is a singleton. More generally,  $\wp(\psi, \mathbf{x}) \in \mathcal{C}$  if  $\wp(\psi, \mathbf{x})$  is a summation such that each summand can be written as the right-hand side of Eq. (5.7). Roughly speaking, a payoff function is in  $\mathcal{C}$  if it is a sum of products of a differentiable function and several indicator functions with the right kind of support. Although we may later refer to a payoff function being in  $\mathcal{C}$ , it is merely a loose — albeit convenient — expression because the definition of  $\mathcal{C}$  involves not only the payoff function but also the distribution of  $\mathbf{x}$  as well. For an option with a payoff function  $\wp(\psi, \mathbf{x}) \in \mathcal{C}$ , we will show how to differentiate its price with respect to  $\psi$  in the next section.

We remark that Lipschitz continuity appears to be a subset of  $\mathcal{C}$  under most circumstances of practical importance. More specifically, the domains of Lipschitz continuity and  $\mathcal{C}$  overlap, but do not have the inclusion relation. Lipschitz continuity imposes stronger conditions than  $\mathcal{C}$  in that it requires a function to be continuous. On the other hand,  $\mathcal{C}$  imposes stronger conditions than Lipschitz continuity in that it allows a function to be nondifferentiable on a set containing only a finite number of points, whereas Lipschitz continuity allows a function to be nondifferentiable on a set with measure zero. Recall that a set containing only a finite number of points clearly has measure zero.

Although the focus of this chapter is on Greeks under the Black-Scholes model, our method is applicable to different underlying models. For instance, if the underlying asset price follows a general diffusion process:  $dS_t = \mu(t, S_t) + \sigma(t, S_t)dZ_t$  where  $dZ_t$  is the Wiener process, then  $S_t$  may need to be simulated using discretization scheme, e.g., the Euler scheme,

$$S_{t_{i+1}} = S_{t_i} + \mu(t_i, S_{t_i})\Delta t + \sigma(t_i, S_{t_i})\sqrt{\Delta t}x_{i+1}, \quad (5.8)$$

for  $i = 0, 1, \dots, n - 1$ . Replacing  $\mu(t_i, S_{t_i})\Delta t$  by  $\mu_i$  and  $\sigma(t_i, S_{t_i})\sqrt{\Delta t}$  by  $\sigma_i$  for brevity, we can write Eq. (7.2.2) as

$$S_{t_{i+1}} = S_{t_i} + \mu_i + \sigma_i x_{i+1}.$$

By the recursive formula, we have

$$\begin{aligned} S_{t_n} &= S_{t_{n-1}} + \mu_{n-1} + \sigma_{n-1} x_n \\ &= S_0 + \mu_0 + \dots + \mu_{n-1} + \sigma_0 x_1 + \dots + \sigma_{n-1} x_n \end{aligned}$$

is strictly increasing in  $x_n$ . For illustration purposes, consider a European call option having payoff function,  $\wp(\psi, \mathbf{x}) = h(\psi, \mathbf{x})\mathbf{1}_{\{g(\psi, \mathbf{x})\}}(\mathbf{x})$ , where  $h(\psi, \mathbf{x}) = S(T) - K = S_{t_n} - K$  and  $g(\psi, \mathbf{x}) = h(\psi, \mathbf{x})$  with  $\mathbf{x} = (x_1, \dots, x_n)$  having pdf  $f(\mathbf{x})$ . Clearly,  $g(\psi, \mathbf{x})$  is strictly monotone in  $x_n$ ; hence, as long as (1)  $h(\psi, \mathbf{x})$  belongs in  $\mathcal{H}_n$ , (2)  $g(\psi, \mathbf{x})$  belongs in  $\mathcal{G}_n$ , and (3)

$$\int_{\mathbb{R}^n} \left| h(\psi, \mathbf{x}) \frac{\partial g(\psi, \mathbf{x})/\partial\psi}{\partial g(\psi, \mathbf{x})/\partial x_n} \right| f(\mathbf{x}) d\mathbf{x} < \infty,$$

the payoff function  $\wp(\psi, \mathbf{x})$  belongs in  $\mathcal{C}$ . This example shows that our method is applicable to the general diffusion process, such as the CIR model and the CEV model, as they are special cases of the general diffusion process (Cox et al., 1985; Cox and Ross, 1976; Schroder, 1989). To be sure, our method is also applicable to path-dependent options if their payoff functions belong in  $\mathcal{C}$ .

## 5.2 Fundamental theorems

In Theorem 1, we show how to differentiate an integral whose integrand is a product of a function  $h(\psi, x) \in \mathcal{H}$ , a pdf  $f(x)$  for  $x$ , and indicator functions whose supports are of the form  $\{g_j(\psi, x) > 0\}$ , where  $g_j(\psi, x) \in \mathcal{G}$  for  $j \in \mathcal{B}$ , a finite set of natural numbers.

**Theorem 1.** Suppose  $h(\psi, x) \in \mathcal{H}$  with pdf  $f(x)$  for  $x$  and  $g_j(\psi, x) \in \mathcal{G}$  for  $j \in \mathcal{B}$ ,

a finite set of natural numbers. Then

$$\begin{aligned}
& \frac{\partial}{\partial \psi} \int_{\Re} h(\psi, x) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, x) > 0\}}(x) f(x) dx \\
&= \int_{\Re} h_\psi(\psi, x) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, x) > 0\}}(x) f(x) dx \\
&\quad + \sum_{l \in \mathcal{B}} \left[ h(\psi, x) J_l(\psi, x) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, x) > 0\}}(x) f(x) \right]_{x=\chi_l(\psi)}, \tag{5.9}
\end{aligned}$$

where

$$J_l(\psi, x) = \text{sign} \left( \frac{\partial g_l(\psi, x)}{\partial x_k} \right) \frac{\partial g_l(\psi, x)/\partial \psi}{\partial g_l(\psi, x)/\partial x} \text{ for } l \in \mathcal{B}. \tag{5.10}$$

*Proof.* See Appendix 1.  $\square$

In plain language, Theorem 1 (1) guarantees the validity to interchange the order of differentiation and integration, (2) provides the differentiation of an indicator function with a support of the form  $\{g_j(\psi, x) > 0\}$  when  $g_j(\psi, x) \in \mathcal{G}$ , and (3) establishes the product rule for differentiating a product of functions. Theorem 2 generalizes Theorem 1 to higher dimensions.

**Theorem 2.** Suppose  $h(\psi, \mathbf{x}) \in \mathcal{H}_k$  with pdf  $f(\mathbf{x})$  for  $\mathbf{x}$  and  $g_j(\psi, \mathbf{x}) \in \mathcal{G}_k$  for  $j \in \mathcal{B}$ , a finite set of natural numbers. If

$$\int_{\Re^n} \left| h(\psi, \mathbf{x}) \frac{\partial g_j(\psi, \mathbf{x})/\partial \psi}{\partial g_j(\psi, \mathbf{x})/\partial x_k} \right| f(\mathbf{x}) d\mathbf{x} < \infty \text{ for } j \in \mathcal{B},$$

then

$$\begin{aligned}
& \frac{\partial}{\partial \psi} \int_{\Re^n} h(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \\
&= \int_{\Re^n} h_\psi(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \\
&\quad + \sum_{l \in \mathcal{B}} \int_{\Re^{n-1}} \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}) \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)} d\mathbf{x}_k, \tag{5.11}
\end{aligned}$$

where

$$J_l(\psi, \mathbf{x}) = \text{sign} \left( \frac{\partial g_l(\psi, \mathbf{x})}{\partial x_k} \right) \frac{\partial g_l(\psi, \mathbf{x})/\partial \psi}{\partial g_l(\psi, \mathbf{x})/\partial x_k} \text{ for } l \in \mathcal{B}. \tag{5.12}$$

*Proof.* See Appendix B.2. □

Differentiating an  $n$ -dimensional integral in Theorem 2 yields an  $n$ -dimensional integral and several  $(n - 1)$ -dimensional integrals with the component  $x_k$  removed (see Eq. (5.11)). Since the  $n$ -dimensional integral equals

$$\begin{aligned} & \int_{\Re^n} h_\psi(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x} \\ &= E \left[ h_\psi(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right], \end{aligned} \quad (5.13)$$

in principle, it can be estimated using the Monte Carlo method. Estimating each

$$\int_{\Re^{n-1}} \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}) \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)} d\mathbf{x}_k \quad (5.14)$$

of Eq. (5.11) can be a major challenge for the following reasons, however. The crude Monte Carlo method estimates Eq. (5.14) as follows:

1. Draw  $N$  sample paths  $\mathbf{x}_k^{(1)}, \dots, \mathbf{x}_k^{(N)}$  uniformly over  $\Re^{n-1}$ .
2. Approximate Eq. (5.14) by

$$\frac{1}{N} \sum_{n=1}^N p_1(\mathbf{x}_k^{(n)}),$$

where

$$p_1(\mathbf{x}_k) \equiv \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}) \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)}.$$

Two difficulties arise with the above procedure. For step one, it is difficult to draw  $\mathbf{x}_k$  uniformly over  $\Re^{n-1}$  directly, which is unbounded. For step two, the estimate may be inefficient in that it may have a very large variance, which demands more replications. To overcome these difficulties, we next provide an alternative form of Eq. (5.14) using a distribution for  $\mathbf{x}_k$  with pdf  $q(\mathbf{x}_k)$ . When  $q(\mathbf{x}_k)$  has support

$\Re^{n-1}$ , Eq. (5.14) equals

$$\begin{aligned}
& \int_{\Re^{n-1}} \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \frac{f(\mathbf{x})}{q(\mathbf{x}_k)} \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)} q(\mathbf{x}_k) d\mathbf{x}_k \\
&= E_q \left[ \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \frac{f(\mathbf{x})}{q(\mathbf{x}_k)} \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)} \right] \\
&= E_q \left[ \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)} \tilde{\eta}(\psi, \mathbf{x}_k), \right] \quad (5.15)
\end{aligned}$$

where

$$\tilde{\eta}(\psi, \mathbf{x}_k) = \frac{f(\mathbf{x})|_{x_k=\chi_l(\psi, \mathbf{x}_k)}}{q(\mathbf{x}_k)}$$

and  $\mathbf{x}_k$  has pdf  $q(\mathbf{x})$  under  $E_q$ . Mathematically speaking, the identification of Eqs. (5.14) with (5.15) is a change of measure, and it forms a theoretical basis of the importance sampling method in the following sense: To estimate Eq. (5.14), we sample  $\mathbf{x}_k$  from an easy-to-sample sampling distribution with pdf  $q(\mathbf{x}_k)$ . We now estimate Eq. (5.14) using Eq. (5.15) as follows:

1. Draw  $N$  sample paths  $\mathbf{x}_k^{(1)}, \dots, \mathbf{x}_k^{(N)}$  from the sampling distribution,  $q(\mathbf{x}_k)$ .
2. Calculate the importance weight  $\tilde{\eta}^{(n)}$  for  $n = 1, \dots, N$  by plugging  $\mathbf{x}_k^{(n)}$  into the ratio
$$\frac{f(\mathbf{x})|_{x_k=\chi_l(\psi, \mathbf{x}_k)}}{q(\mathbf{x}_k)}.$$
3. Approximate Eq. (5.14) by

$$\frac{1}{N} \sum_{n=1}^N \tilde{\eta}^{(n)} p_2(\mathbf{x}_k^{(n)}),$$

where

$$p_2(\mathbf{x}_k) \equiv \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)}.$$

As for which  $q(\mathbf{x}_k)$  to use, a good candidate is one that is close to the shape of the integrand of Eq. (5.14). However, finding  $q(\mathbf{x}_k)$  close to the integrand of Eq. (5.14) requires sophisticated analysis and can be extremely difficult, especially in high-dimensional cases (Liu, 2001). For simplicity, we suggest the normal distribution as the sampling distribution. Now, we are ready to present the way to calculate the first-order Greeks under the Black-Scholes model.

**Theorem 3.** *Consider an option under the Black-Scholes model. Suppose the payoff function  $\wp(\psi, \mathbf{x})$  belongs in  $\mathcal{C}$  with  $\mathbf{x} \sim N_n(\mathbf{0}, \Sigma)$ , i.e.,*

$$\wp(\psi, \mathbf{x}) = h(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x})$$

for some functions  $h(\psi, \mathbf{x}) \in \mathcal{H}_k$  and  $g_j(\psi, \mathbf{x}) \in \mathcal{G}_k$  for  $j \in \mathcal{B}$ , a finite set of natural numbers. Moreover, suppose that  $q(\mathbf{x}_k)$  is a pdf for  $\mathbf{x}_k$  with support  $\Re^{n-1}$ . Then

$$\begin{aligned} & \frac{\partial}{\partial \psi} E[\wp(\psi, \mathbf{x})] \\ &= E \left[ h_\psi(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right] \\ &+ \sum_{l \in \mathcal{B}} E_q \left[ \left[ h(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)} \tilde{\eta}_l(\psi, \mathbf{x}_k) \right], \end{aligned} \quad (5.16)$$

where

$$\tilde{\eta}_l(\psi, \mathbf{x}_k) = \frac{f(\mathbf{x}; \mathbf{0}, \Sigma)|_{x_k=\chi_l(\psi, \mathbf{x}_k)}}{q(\mathbf{x}_k)}, \quad (5.17)$$

$J_l(\psi, \mathbf{x})$  equals Eq. (5.12), and  $\mathbf{x}_k$  has pdf  $q(\mathbf{x}_k)$  under  $E_q$ .

*Proof.* Because

$$E[\wp(\psi, \mathbf{x})] = \int_{\Re^n} h(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}; \mathbf{0}, \Sigma) d\mathbf{x},$$

$\partial E[\wp(\psi, \mathbf{x})]/\partial \psi$  equals the right-hand-side of Eq. (5.11) except that  $\mathbf{x}$  has pdf  $f(\mathbf{x}; \mathbf{0}, \Sigma)$  by Theorem 2. Applying Eqs. (5.13) and (5.15), Eq. (5.16) is proved.  $\square$

Now, first-order Greeks can be derived if the payoff function  $\wp$  belongs in  $\mathcal{C}$ . To

derive Greeks like delta, vega, and theta, simply apply its corresponding theorem with  $\psi = S$ ,  $\psi = \sigma$ , and  $\psi = T$ , respectively. To derive rho, apply its corresponding theorem with  $\psi = r$  as follows:

$$\frac{\partial}{\partial r} e^{-rT} E[\varphi] = -Te^{-rT} E[\varphi] + e^{-rT} \frac{\partial}{\partial r} E[\varphi].$$

Alternatively, it is well-known that

$$\Theta + r \sum_{i=1}^n S_i \Delta_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \rho_{ij} \sigma_i \sigma_j S_i S_j \Gamma_{ij} = r\mathcal{C},$$

which provides a numerical approach to deriving theta using deltas, gammas, and cross-gammas. See Cox et al. Cox et al. (1985) and Boyle et al. Boyle et al. (1989) for details. Note that the right-hand side of Eq. (5.11) is a formula for first-order Greeks. When a closed-form formula for  $\chi(\psi, \mathbf{x})$  exists, we simply plug  $\chi(\psi, \mathbf{x}_k)$  into  $x_k$ . If a closed-form formula does not exist, we can plug into  $x_k$  a numerical solution using, say, the Newton-Raphson method. As a result, we only require the existence of  $\chi(\psi, \mathbf{x}_k)$  when calculating first-order Greeks.

The second-order Greek is obtained by applying our theorems again to the corresponding first-order Greek. A closed-form formula for  $\chi_l(\psi, \mathbf{x}_k)$  is required to calculate the second-order Greeks for the most general case. However, Eq. (5.11) can be much simplified before applying differentiation to obtain second-order Greeks when its second term reduces to zero (as in the deltas of spread options and maximum options shown later). When this is the case, we only need to differentiate

$$\int_{R^n} h_\psi(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}, \quad (5.18)$$

of Eq. (5.11) and this reduces to the case of deriving first-order Greeks and is solved by applying Theorem 2 again. In this case, the existence of  $\chi_l(\psi, \mathbf{x}_k)$  is sufficient (i.e., closed-form formulas for  $\chi_l(\psi, \mathbf{x}_k)$  are not needed). As a remark, when differentiating Eq. (5.18) with respect to  $\psi$ , we need to check if (1)  $h_\psi(\psi, \mathbf{x})$  belongs in  $\mathcal{H}_k$ , (2)  $g_j(\psi, \mathbf{x})$  belongs in  $\mathcal{G}_k$  for  $j \in \mathcal{B}$ , and (3)

$$\int_{\mathbb{R}^n} \left| h_\psi(\psi, \mathbf{x}) \frac{\partial g_j(\psi, \mathbf{x}) / \partial \psi}{\partial g_j(\psi, \mathbf{x}) / \partial x_k} \right| f(\mathbf{x}) d\mathbf{x} < \infty \text{ for } j \in \mathcal{B}.$$

Higher-order Greeks can be obtained recursively using Theorem 2.

## 5.3 Greeks of rainbow options

The term “rainbow option” originates from Rubinstein (1991), who describes a rainbow option as a combination of a variety of assets much as rainbow is a combination of a variety of colors. There is an abundance of rainbow options in the literature (see Table 5.1). In the following, we show how to derive the unbiased Greeks of spread and maximum options using our methodology.

### 5.3.1 Greeks of spread options

In this subsection, we derive  $\Delta_1$ ,  $\Gamma_{11}$  and  $\Gamma_{12}$  for the spread option. The payoff function of the spread option is

$$\begin{aligned}\wp &= \max(S_2(T) - S_1(T) - K, 0) \\ &= (S_2(T) - S_1(T) - K)\mathbf{1}_{\{S_2(T) - S_1(T) - K > 0\}}(\mathbf{S}_T).\end{aligned}$$

Spread options are options written on the difference between the values of two stocks or two indexes. They are used in many markets, such as the fixed-income markets, the currency markets, the commodity markets, and especially the energy markets. These options are popular because they are designed to mitigate adverse movements between two market variables. There is extensive literature on pricing the spread options; however, calculating their Greeks is still challenging because no closed-form formulas exist for their prices (Carmona and Durrelman, 2003).

In the fixed income markets and the currency markets, spread options are based on the difference between two interest rates or swap rates. In the commodity markets, spread options are based on the difference between the prices of the same commodity at two different locations or at two different points in time, as well as between the prices of different grades of the same commodity. At New York Mercantile Exchange, spread options are traded on the difference between heating oil and crude oil, as well as between gasoline and crude oil. These spreads are known as crack spreads, hence options on these spreads are known as crack spread options. Another example in the energy market involves options on the price

difference between oil and electricity which are known as the spark spread options (Carmona and Durrelman, 2003).

Using Eq. (5.2) to generate  $\mathbf{S}_T$  and defining

$$\begin{aligned} h(\psi, \mathbf{x}) &= S_2 e^{(r-\sigma_2^2/2)T+\sigma_2\sqrt{T}x_2} - S_1 e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1} - K, \\ g(\psi, \mathbf{x}) &= h(\psi, \mathbf{x}), \end{aligned}$$

we rewrite the payoff function as

$$\wp(\psi, \mathbf{x}) = h(\psi, \mathbf{x}) \mathbf{1}_{\{g(\psi, \mathbf{x}) > 0\}}(\mathbf{x}).$$

We proceed to prove that the payoff function belongs in  $\mathcal{C}$ . Note that  $h(\psi, \mathbf{x}) \in \mathcal{H}_2$  with pdf  $f(\mathbf{x}; \mathbf{0}, \Sigma)$  for  $\mathbf{x}$  and  $g(\psi, \mathbf{x})$  is increasing in  $\psi$  and  $x_2$ . Let

$$\chi(\psi, \mathbf{x}_2) = \frac{\log \frac{S_1 e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1} + K}{S_2} - (r - \sigma_2^2/2)T}{\sigma_2\sqrt{T}}. \quad (5.19)$$

It is easy to verify that  $g(\psi, \mathbf{x})|_{x_2=\chi(\psi, \mathbf{x}_2)} = 0$ ; hence  $g(\psi, \mathbf{x}) \in \mathcal{G}_2$ . We conclude that  $\wp(\psi, \mathbf{x}) \in \mathcal{C}$ .

**Theorem 4.** *The delta, gamma and cross-gamma of the spread option equal*

$$\begin{aligned} \Delta_1 &= e^{-rT} E \left[ -S_1(T) \mathbf{1}_{\{S_2(T) - S_1(T) - K > 0\}}(\mathbf{S}(T)) \right] / S_1, \\ \Gamma_{11} &= e^{-rT} E_q \left[ e^{2((r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1) - ((r-\sigma_2^2/2)T+\sigma_2\sqrt{T}\chi(\psi, \mathbf{x}_2))} \tilde{\eta}(\psi, \mathbf{x}_2) \right] / (S_2\sigma_2\sqrt{T}), \\ \Gamma_{12} &= -e^{-rT} E_q \left[ e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1} \tilde{\eta}(\psi, \mathbf{x}_2) \right] / (S_2\sigma_2\sqrt{T}), \end{aligned}$$

respectively, where  $\chi(\psi, \mathbf{x}_2)$  is defined in Eq. (5.19),  $\tilde{\eta}(\psi, \mathbf{x}_2)$  is defined in Eq. (B.4), and  $\mathbf{x}_2$  has pdf  $q(\mathbf{x}_k) = f(\mathbf{x}_2; \mathbf{0}, 1)$  under  $E_q$ .

*Proof.* See Appendix B.3. □

### 5.3.2 Greeks of maximum options

We now derive the Greeks of the maximum option. Let  $\mathcal{B} = \{1, \dots, n\}$  throughout this section. By conditioning on which asset ends up as the maximum terminal

price, the payoff function can be decomposed into

$$\begin{aligned}\wp &= \max(\max(S_1(T), \dots, S_n(T)) - K, 0) \\ &= \sum_{i \in \mathcal{B}} (S_i(T) - K) \mathbf{1}_{\{S_i(T) > K\}}(\mathbf{S}_T) \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{S_i(T) > S_j(T)\}}(\mathbf{S}_T).\end{aligned}$$

Maximum options depend on the maximum stock price at maturity, and minimum options depend on the minimum stock price. Stulz (1982) derives formulas for maximum and minimum options on two risky assets and Johnson (1987) generalizes formulas to options on several assets. Maximum and minimum options are commonly used in compensation plans, risk-sharing contracts, collateralized loans and secured debts, indexed wages and option bonds Stulz (1982).

Using Eq. (5.2) to generate  $\mathbf{S}_T$  and defining

$$h_i(\psi, \mathbf{x}) = S_i e^{(r - \sigma_i^2/2)T + \sigma_i \sqrt{T}x_i} - K \quad \text{for } i \in \mathcal{B}, \quad (5.20)$$

$$g_{ii}(\psi, \mathbf{x}) = S_i e^{(r - \sigma_i^2/2)T + \sigma_i \sqrt{T}x_i} - K \quad \text{for } i \in \mathcal{B}, \quad (5.21)$$

$$g_{ij}(\psi, \mathbf{x}) = S_i e^{(r - \sigma_i^2/2)T + \sigma_i \sqrt{T}x_i} - S_j e^{(r - \sigma_j^2/2)T + \sigma_i \sqrt{T}x_j} \quad \text{for } i \in \mathcal{B}, j \in \mathcal{B} \setminus i, \quad (5.22)$$

the payoff function can be rewritten as

$$\wp(\psi, \mathbf{x}) = \sum_{i \in \mathcal{B}} h_i(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}). \quad (5.23)$$

We proceed to prove that the payoff function belongs in  $\mathcal{C}$ . Note  $h_i(\psi, \mathbf{x}) \in \mathcal{H}_i$  with pdf  $f(\mathbf{x}; \mathbf{0}, \Sigma)$  for  $\mathbf{x}$  and  $g_{ij}(\psi, \mathbf{x})$  is strictly monotone in  $\psi$  and in  $x_i$ . Let

$$\chi_{ii}(\psi, \mathbf{x}_i) = \frac{\log K/S_i - (r - \sigma_i^2/2)T}{\sigma_i \sqrt{T}} \quad \text{for } i \in \mathcal{B}, \quad (5.24)$$

$$\chi_{ij}(\psi, \mathbf{x}_i) = \frac{\log \frac{S_j e^{(r - \sigma_j^2/2)T + \sigma_j \sqrt{T}x_j}}{S_i} - (r - \sigma_i^2/2)T}{\sigma_i \sqrt{T}} \quad \text{for } i \in \mathcal{B}, j \in \mathcal{B} \setminus i. \quad (5.25)$$

It is easy to verify that  $g_{ij}(\psi, \mathbf{x})|_{x_i=\chi_{ij}(\psi, \mathbf{x}_i)} = 0$ ; hence  $g_{ij}(\psi, \mathbf{x}) \in \mathcal{G}_i$ . We conclude that  $\wp(\psi, \mathbf{x}) \in \mathcal{C}$ .

**Theorem 5.** *The delta, gamma and cross-gamma of the maximum option on n assets are*

$$\Delta_i = e^{-rT} E \left[ e^{(r - \sigma_i^2/2)T + \sigma_i \sqrt{T}x_i} \mathbf{1}_{\{S_i e^{(r - \sigma_i^2/2)T + \sigma_i \sqrt{T}x_i} > K\}}(\mathbf{x}) \right]$$

$$\begin{aligned}
& \times \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{S_i e^{(r-\sigma_i^2/2)T+\sigma_i \sqrt{T}x_i} > S_j e^{(r-\sigma_j^2/2)T+\sigma_j \sqrt{T}x_j}\}}(\mathbf{x}) \Big], \\
\Gamma_{ii} &= e^{-rT} \left\{ E_q \left[ K \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{S_j e^{(r-\sigma_j^2/2)T+\sigma_j \sqrt{T}x_j} < K\}}(\mathbf{x}_i) \tilde{\eta}_{ii}(\psi, \mathbf{x}_i) \right] \right. \\
&+ \sum_{l \in \mathcal{B} \setminus i} E_q \left[ S_l e^{(r-\sigma_l^2/2)T+\sigma_l \sqrt{T}x_l} \mathbf{1}_{\{S_l e^{(r-\sigma_l^2/2)T+\sigma_l \sqrt{T}x_l} > K\}}(\mathbf{x}_i) \right. \\
&\quad \times \left. \prod_{j \in \mathcal{B} \setminus \{l, i\}} \mathbf{1}_{\{S_l e^{(r-\sigma_l^2/2)T+\sigma_l \sqrt{T}x_l} > S_j e^{(r-\sigma_j^2/2)T+\sigma_j \sqrt{T}x_j}\}}(\mathbf{x}_i) \tilde{\eta}_{il}(\psi, \mathbf{x}_i) \right] / \left( S_l^2 \sigma_l \sqrt{T} \right) \Big\}, \\
\Gamma_{ij} &= -e^{-rT} E_q \left[ e^{(r-\sigma_j^2/2)T+\sigma_j \sqrt{T}x_j} \mathbf{1}_{\{S_j e^{(r-\sigma_j^2/2)T+\sigma_j \sqrt{T}x_j} > K\}}(\mathbf{x}_i) \right. \\
&\quad \times \left. \prod_{l \in \mathcal{B} \setminus \{j, i\}} \mathbf{1}_{\{S_j e^{(r-\sigma_j^2/2)T+\sigma_j \sqrt{T}x_j} > S_l e^{(r-\sigma_l^2/2)T+\sigma_l \sqrt{T}x_l}\}}(\mathbf{x}_i) \tilde{\eta}_{ij}(\psi, \mathbf{x}_i) \right] / \left( S_j \sigma_j \sqrt{T} \right),
\end{aligned}$$

where  $\tilde{\eta}_{ij}(\psi, \mathbf{x}_i)$  for  $i, j \in \mathcal{B}$  is defined in Eq. (B.10), and  $\mathbf{x}_i$  has pdf  $f(\mathbf{x}_i; \mathbf{0}, \mathbf{I}_{n-1})$  under  $E_q$  for  $i \in \mathcal{B}$ .

*Proof.* See Appendix B.4.  $\square$

## 5.4 Numerical results

Although the FD approach is intuitive and easy to understand, those advantages are weakened by its shortcomings. A fundamental weakness of the FD approach is that, so far, there are no known hard rules for choosing the right  $\delta$  in calculating the Greeks (Jäckel, 2002). This delicate issue is especially damaging for higher-order Greeks like gammas and cross-gammas. It also highlights the advantage of our method, which does not have to pick  $\delta$ ;  $\delta$  simply does not enter our formulas at all. This section compares our method with various FD methods and the likelihood ratio method. Recall that Table 5.2 lists standard formulas for the forward and central FDs. The formulas of Greeks for rainbow options using the likelihood ratio method are summarized in Appendix B.5.

Tables 5.3 and 5.4 show the results of applying our method to spread options and maximum options, respectively. For the benchmark values, we first compute their option prices then calculate the Greeks by central FD with  $h = 0.1$ . We use numerical integration to calculate the prices of spread and barrier options and

follow Johnson (1987) in calculating the prices of maximum and binary maximum options. These tables contain simulation results with 100,000 paths for deltas, gammas, and cross-gammas so that our method can be compared with forward and central FD methods under various  $\delta$ . We also use common random numbers for the FD estimates.

Several conclusions can be drawn. The deltas given by our method and the FD methods are of similar quality. In particular, our method always gives unbiased deltas, whereas under the FD schemes, deltas may be biased when  $\delta$  is large but quickly converge to the correct value as  $\delta$  approaches zero. But the similarity ends here. Although the desired convergence property continues to be shared by our method in the calculations of gammas and cross-gammas, it is no longer shared by any of the FD methods. For large  $\delta$ , the gammas and cross-gammas calculated by the FD methods are clearly biased. On the other hand, for small  $\delta$ , their gammas and cross-gammas are unstable.

Our method yields Greek formulas and hence avoids these drawbacks arising from FD's being approximate only. Its effectiveness is clearly demonstrated by the numerical superiority to the FD methods in treating gammas and cross-gammas. Both our method and the likelihood ratio method produce unbiased Greeks for rainbow options. However, our method produces estimates with much lower standard errors. It is also not obvious how to apply the likelihood ratio method for path-dependent options.

## 5.5 Summary

Mathematically, Greeks are partial differentiations of an option's price with respect to a parameter of interest. For complex options such as rainbow and path-dependent options, there are usually no analytic formulas for their prices; hence Monte Carlo simulation is often the only available method to estimate them. The FD method is subsequently used to approximate the Greeks. But FD as an approximate differentiation operator can be rather unstable and biased. Two direct methods, the pathwise method and the likelihood ratio method, have been proposed to overcome the drawbacks of FD. However, the Greeks of options whose payoff functions are not Lipschitz continuous can not be obtained using the path-

wise method without modifications if convergence is to be guaranteed.

The major theoretical advantage of our method over the pathwise method is a new mathematical formulation so that the Lipschitz continuity restriction on the payoff function is lifted. We present a rule to interchange the order of integration and differentiation when the integrand can be decomposed into a sum of products of differentiable functions and certain indicator functions. As a result, the Greek formulas of a wide variety of options can be derived systematically. For practical purposes, we propose a useful importance sampling method to estimate these Greeks using the formulas mentioned above. For illustration purposes, formulas and numerical results for the Greeks are given for popular rainbow options. Another key feature of our method is that it is easier to implement and its application is almost mechanical compared with such methods as the likelihood ratio method and the conditional Monte Carlo method. Although the focus of this chapter is on the Greeks of rainbow options under the Black-Scholes model, our method is applicable to alternative models for the underlying assets as long as the payoff function belongs in  $\mathcal{C}$ .

Numerically, our estimators for Greeks are unbiased, whereas estimators obtained from the FD methods are not. Furthermore, the FD methods have to solve the difficult problem of determining the right perturbed size, which is completely eliminated by our method. Also eliminated by our method is resimulation. When the payoff function is not smooth, the FD methods tend to fail, particularly for higher-order Greeks like gammas. In contrast, our method can handle a rich class of payoff functions. These results make our method more generally applicable and useful than FD schemes. Finally, our method enjoys lower variances for rainbow options than the likelihood ratio method.

The non-Lipschitz pathwise method calculates the Greeks under a parametric assumption on the underlying asset. To avoid model misspecification, we investigate how to calculate the Greeks without any parametric assumptions on the underlying asset in the next chapter. The basic idea goes as follows. We use the Bayesian implied random tree model to calibrate the state price density using option prices. Simultaneously, for each sweep of the MCMC algorithm, we calculate the Greeks using a finite-difference approximation based on the implied random tree model. As a result, such Greeks are calculated without parametric

assumptions on the underlying asset.

**Table 5.3.** Greeks of spread options on two assets

Parameters:  $S_1 = S_2 = 40$ ,  $K = 0.5$ ,  $r = 10\%$ ,  $T = 1$  year,  $\sigma_1 = \sigma_2 = 0.3$ , and  $\rho_{12} = \rho_{21} = 0.69$ . For benchmark values, we obtain spread options' prices using numerical integrations and then calculate the Greeks by central FD with  $h = 0.1$ . The standard errors are in parentheses. All simulation results are based on 100,000 trials.

$\delta$	$\Delta_1$	$\Gamma_{11}$	$\Gamma_{12}$	
<b>Benchmark value</b>				
	-0.4340	0.0414		-0.0419
<b>Non-Lipschitz pathwise method</b>				
	-0.4346 (0.0017)	0.0414 (0.0001)	-0.0419 (0.0001)	
<b>Likelihood ratio method</b>				
	-0.4308 (0.0033)	0.0407 (0.0007)	-0.0410 (0.0007)	
<b>Forward FD</b>				
$10^{-1}$	-0.4339 (0.0016)	0.0398 (0.0017)	-0.0422 (0.0017)	
$10^{-2}$	-0.4343 (0.0017)	0.0408 (0.0054)	-0.0306 (0.0044)	
$10^{-3}$	-0.4349 (0.0017)	0.0476 (0.0185)	-0.0215 (0.0107)	
$10^{-4}$	-0.4314 (0.0016)	0.0000 (0.0000)	0.0000 (0.0000)	
$10^{-5}$	-0.4304 (0.0016)	0.0000 (0.0000)	0.0000 (0.0000)	
$10^{-6}$	-0.4355 (0.0017)	-0.0031 (0.0000)	0.0000 (0.0000)	
$10^{-7}$	-0.4335 (0.0016)	-0.0011 (0.0012)	0.0000 (0.0000)	
$10^{-8}$	-0.4312 (0.0016)	-0.0636 (0.1196)	0.0013 (0.0018)	
$10^{-9}$	-0.4338 (0.0016)	-3092.6010 (16.8991)	0.2572 (0.2033)	
<b>Central FD</b>				
$10^{-1}$	-0.4338 (0.0016)	0.0420 (0.0017)	-0.0448 (0.0018)	
$10^{-2}$	-0.4351 (0.0017)	0.0415 (0.0055)	-0.0541 (0.0063)	
$10^{-3}$	-0.4343 (0.0016)	0.0405 (0.0182)	-0.0411 (0.0164)	
$10^{-4}$	-0.4330 (0.0016)	0.0804 (0.0804)	-0.0432 (0.0432)	
$10^{-5}$	-0.4335 (0.0016)	0.0000 (0.0000)	0.0000 (0.0000)	
$10^{-6}$	-0.4344 (0.0016)	0.0000 (0.0000)	0.0000 (0.0000)	
$10^{-7}$	-0.4345 (0.0017)	0.0003 (0.0006)	0.0000 (0.0000)	
$10^{-8}$	-0.4321 (0.0016)	-0.0085 (0.0634)	0.0006 (0.0006)	
$10^{-9}$	-0.4342 (0.0016)	3.4397 (6.3518)	0.0643 (0.1114)	

**Table 5.4.** Greeks of maximum options on two assets

Parameters:  $S_1 = S_2 = 40$ ,  $K = 40$ ,  $r = 10\%$ ,  $T = 1$  year,  $\sigma_1 = \sigma_2 = 0.3$ , and  $\rho_{12} = \rho_{21} = 0.69$ . For benchmark values, we follow Johnson's formula to obtain maximum options's prices (Johnson, 1987). We then calculate the Greeks by central FD with  $h = 0.1$ . The standard errors are in parentheses. Simulation results are based on 100,000 trials.

$\delta$	$\Delta_1$	$\Gamma_{11}$	$\Gamma_{12}$		
<b>Benchmark value</b>					
	0.4322	0.0425		-0.0286	
<b>Non-Lipschitz pathwise method</b>					
	0.4354 (0.0020)	0.0424 (0.0001)		-0.0286 (0.0001)	
<b>Likelihood ratio method</b>					
	0.4393 (0.0069)	0.0437 (0.0013)		-0.0292 (0.0011)	
<b>Forward FD</b>					
$10^{-1}$	0.4338 (0.0019)	0.0411 (0.0017)		-0.0288 (0.0015)	
$10^{-2}$	0.4342 (0.0019)	0.0397 (0.0057)		-0.0289 (0.0047)	
$10^{-3}$	0.4316 (0.0019)	0.0448 (0.0172)		-0.0113 (0.0113)	
$10^{-4}$	0.4324 (0.0019)	0.0233 (0.0166)		0.0000 (0.0000)	
$10^{-5}$	0.4343 (0.0019)	0.0000 (0.0000)		0.0000 (0.0000)	
$10^{-6}$	0.4306 (0.0019)	0.0030 (0.0000)		0.0000 (0.0000)	
$10^{-7}$	0.4325 (0.0019)	0.0006 (0.0014)		0.0000 (0.0000)	
$10^{-8}$	0.4325 (0.0019)	0.1537 (0.1367)		0.0000 (0.0000)	
$10^{-9}$	0.4334 (0.0019)	3083.1500 (19.9935)		0.0000 (0.0000)	
<b>Central FD</b>					
$10^{-1}$	0.4323 (0.0019)	0.0442 (0.0018)		-0.0267 (0.0015)	
$10^{-2}$	0.4336 (0.0019)	0.0440 (0.0056)		-0.0353 (0.0053)	
$10^{-3}$	0.4305 (0.0019)	0.0430 (0.0168)		-0.0358 (0.0217)	
$10^{-4}$	0.4325 (0.0019)	0.0000 (0.0000)		-0.0154 (0.0154)	
$10^{-5}$	0.4298 (0.0019)	0.0000 (0.0000)		0.0000 (0.0000)	
$10^{-6}$	0.4303 (0.0019)	0.0000 (0.0000)		0.0000 (0.0000)	
$10^{-7}$	0.4311 (0.0019)	-0.0011 (0.0007)		0.0000 (0.0000)	
$10^{-8}$	0.4308 (0.0019)	0.0129 (0.0724)		0.0000 (0.0000)	
$10^{-9}$	0.4288 (0.0019)	7.2651 (7.2940)		0.0000 (0.0000)	

Chapter **6**

## Nonparametric Greeks

In Chapter 5, we have proposed a new mathematics formulation so that equations for the Greeks under different parametric assumption on underlying asset can be calculated as long as the payoff function belongs to the class  $\mathcal{C}$ . Briefly speak, a payoff function belongs to  $\mathcal{C}$ , if it is a sum of products of a differentiable function and several indicator functions with the right kind of support. For detailed definition, please refer to Chapter 5. Although our method is also applicable for more general diffusion process, such as the Cox-Ingersoll-Ross process and the Constant Elasticity of Variance process, for illustration purposes, we implement our methods in a Black-Scholes framework, i.e., assuming that underlying asset follows a geometric Brownian motion (Cox et al., 1985; Cox and Ross, 1976; Schroder, 1989).

In practice, Greeks can be calculated in a parametric framework as follows.

1. Determine a parametric modeling, either a generalized assumption on the state price density, or a general diffusion process on the underlying process, that can produce the best fit, and estimate the parameters in the model.
2. Calculate Greeks using the selected model.

However, the model misspecification problem of parametric approaches has motivated us to develop nonparametric methods for calculating Greeks. We describe our method in Section 6.1. We present a simulation study in Section 6.2, an empirical study in Section 6.3, and summarize in the last section.

## 6.1 Our approaches

We calculate Greeks using finite-difference formula based on the Quadrature model and the extended implied random tree model. Recall that  $C[i, j]$  and  $S[i, j]$  denote an option price and an underlying asset's value at the  $j$ -th node in the  $i$ -th time step, respectively. Similar to Rubinstein (1994) and Hull (2002), we calculate delta using a central finite-difference formula as

$$\Delta = (C[0, 1] - C[1, 1])/(S[0, 1] - S[1, 1]), \quad (6.1)$$

and we calculate gamma using a central finite-difference formula as

$$\Gamma = \frac{\Delta[0, 1] - \Delta[1, 1]}{(S[0, 2] - S[2, 2])/2}, \quad (6.2)$$

where

$$\begin{aligned} \Delta[0, 1] &= \frac{C[0, 2] - C[1, 2]}{S[0, 2] - S[1, 2]}, \\ \Delta[1, 1] &= \frac{C[1, 2] - C[2, 2]}{S[1, 2] - S[2, 2]}. \end{aligned}$$

Because Greeks calculated in this way do not require any parametric assumptions on the state price density or the process of the underlying asset, we call these formulas nonparametric Greeks. Delta and gamma calculated using Eqs. (6.1) and (6.2) are called nonparametric delta and nonparametric gamma, respectively.

## 6.2 Simulation study

To investigate the accuracy of our nonparametric Greeks, we consider the Black-Scholes model and a popular Heston stochastic volatility model. For simplicity, we focus on European options in this chapter, although our method can be easily to be extended to calculate Greeks for American options.

### 6.2.1 The Black-Scholes model

We assume that the initial price is  $S_0 = 50$ , the risk-free interest rate is  $r = 0$ , the maturity is  $T = 1/20$  year, and the volatility is  $\sigma = 0.3$ . Because the spirit of our approach is to use a Quadrature model to approximate the state price density, our estimates depend on the amount of available information. Therefore, we would like to investigate how to improve our estimation using different information.

To do this, we consider four data sets each with different ranges of strike prices. The first data set uses 10 strike prices, equally spaced over the range 40 to 60, the second data set uses 20 strike prices, equally spaced over the range 40 to 60, the third data set uses 15 strike prices, equally spaced over the range 35 to 65, and the forth data set uses 30 strike prices, equally spaced over the range 35 to 65. Using Black-Scholes' formulas, we calculate theoretical prices, deltas, and gammas of European call and put options.

We calibrate the Quadrature model using the slice sampler provided in Chapter 3 with 10 and 20 support points. Let  $V_{true}$  denote the true value, and  $V_{fitted}$  denote the fitted value for the true value. In our Bayesian framework, we use the posterior mean of the value of interest in the Markov chain Monte Carlo algorithm as the fitted value. After the MCMC simulation converges, we record the fitted values of the prices, deltas and gammas in Figures 6.1.

To measure the accuracy of an estimate, we calculate the absolute errors in dollar as

$$|V_{true} - V_{fitted}|.$$

We summarize absolute pricing errors in dollar for the prices, deltas and gammas in Figure 6.2. We calculate the absolute relative errors in percentage as

$$100 \times |V_{true} - V_{fitted}| / V_{true}.$$

We summarize absolute relative pricing errors in percentage for the prices, deltas and gammas in Figure 6.3.

Our numerical results show that there are no significant difference in the estimates using the first and the second data sets. However, the estimates of the posterior means for the deltas and gammas using the third data set are largely

improved, in terms of lower absolute errors in dollar and lower absolute relative errors in percentage. This is because that the third data set contains options with a wider range of strike prices, and as a result provides more information to calibrate the state price density.

### 6.2.2 The Heston stochastic volatility model

We compare our approach for the stochastic volatility model in Heston (1993). The Heston stochastic volatility model is

$$\begin{aligned} dS_t &= \alpha S_t dt + \sqrt{v_t} S_t dW_t^{(1)}, \\ dv_t &= \kappa(\vartheta - v_t)dt + \xi \sqrt{v_t} dW_t^{(2)}, \end{aligned}$$

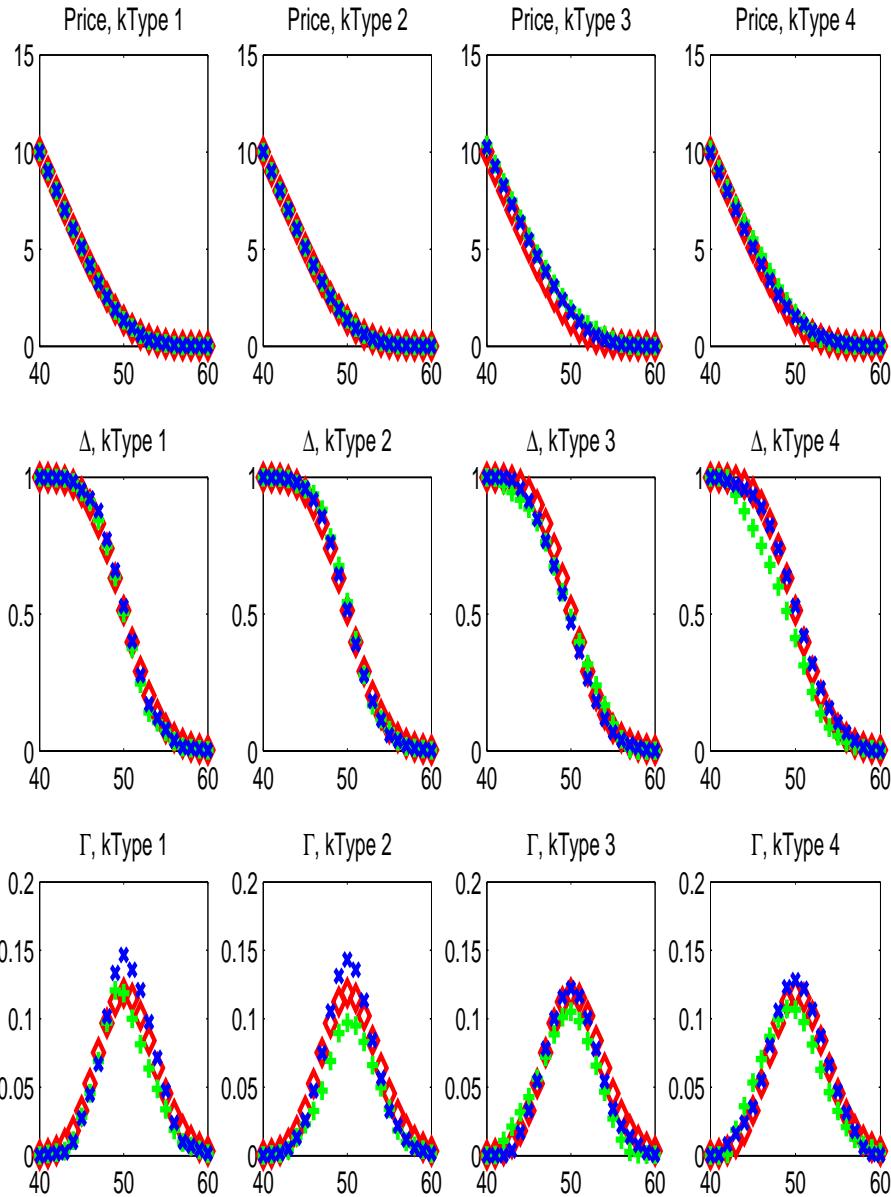
where  $S_t$  is the stock price process, and  $v_t$  is the mean-reverting volatility process with  $\kappa$ ,  $\vartheta$ ,  $\xi$  being the rate of mean reversion, the mean volatility, and the volatility of volatility, respectively.

To generate the synthetic data, we use the parameters  $\kappa = 1.5$ ,  $\vartheta = 0.06$ ,  $\xi = 0.7$ , and  $\rho(\Delta W_1, \Delta W_2) = 0$ . In addition, we assume that the current stock price is  $S_0 = 100$ , the current volatility is  $V_0 = 0.04$ , the risk-free interest rate is  $r = 0.02$ , and the maturity is  $T = 0.25$ . The strike prices are every five units, between 80 and 120.

We use 2000 paths and 250 time steps per path in the Monte Carlo simulation for generating benchmark option prices. Furthermore, the benchmark values of deltas and gammas are calculated using a central finite-difference formula with common random numbers in the resimulation and a perturbed size of 0.1 (Broadie and Glasserman, 1996).

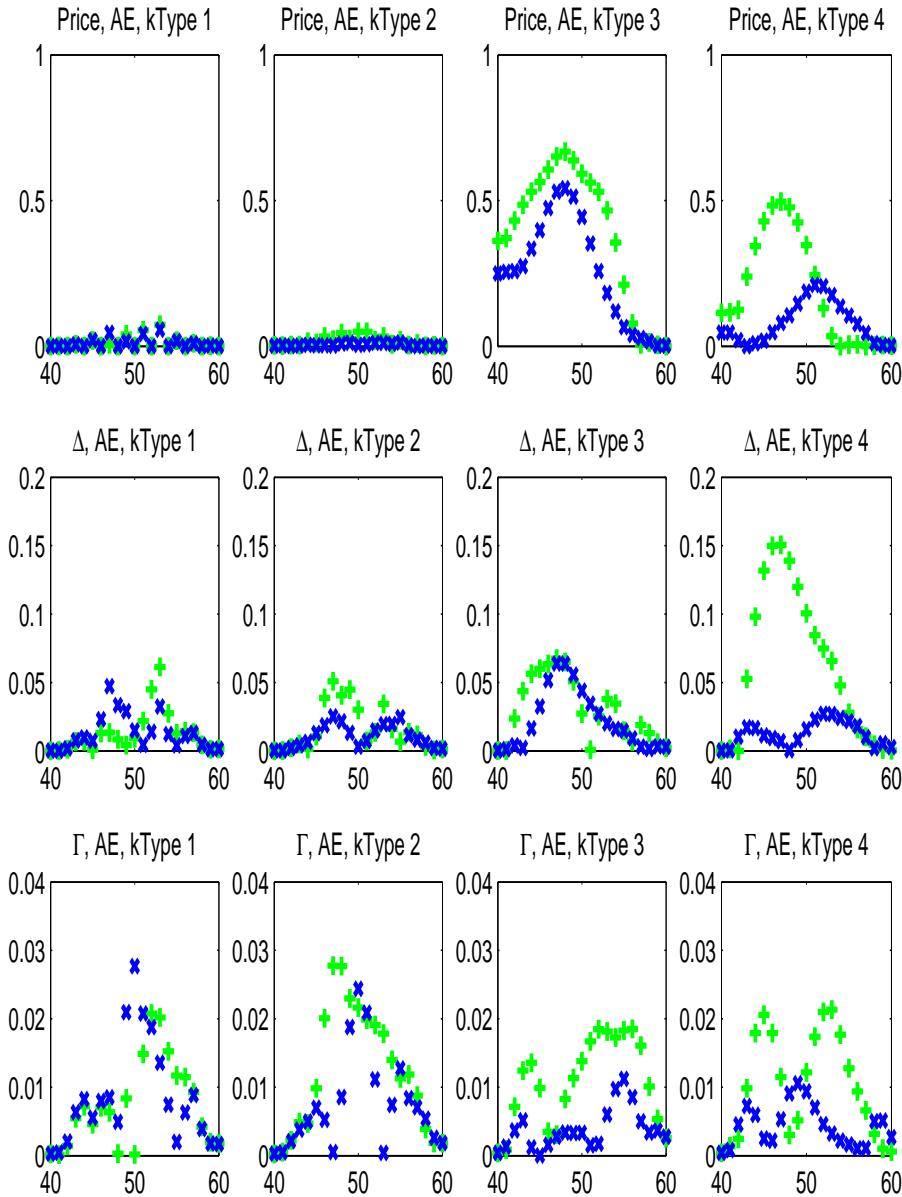
To compare the estimated deltas and gammas for our approach, we implement a Black-Scholes approximation as follows.

1. Calculate an implied volatility  $\sigma_{imp}$  using the Black-Scholes formula for each option at each strike.
2. Calculate the delta by plugging  $\sigma_{imp}$  in the delta formula under the Black-Scholes model.



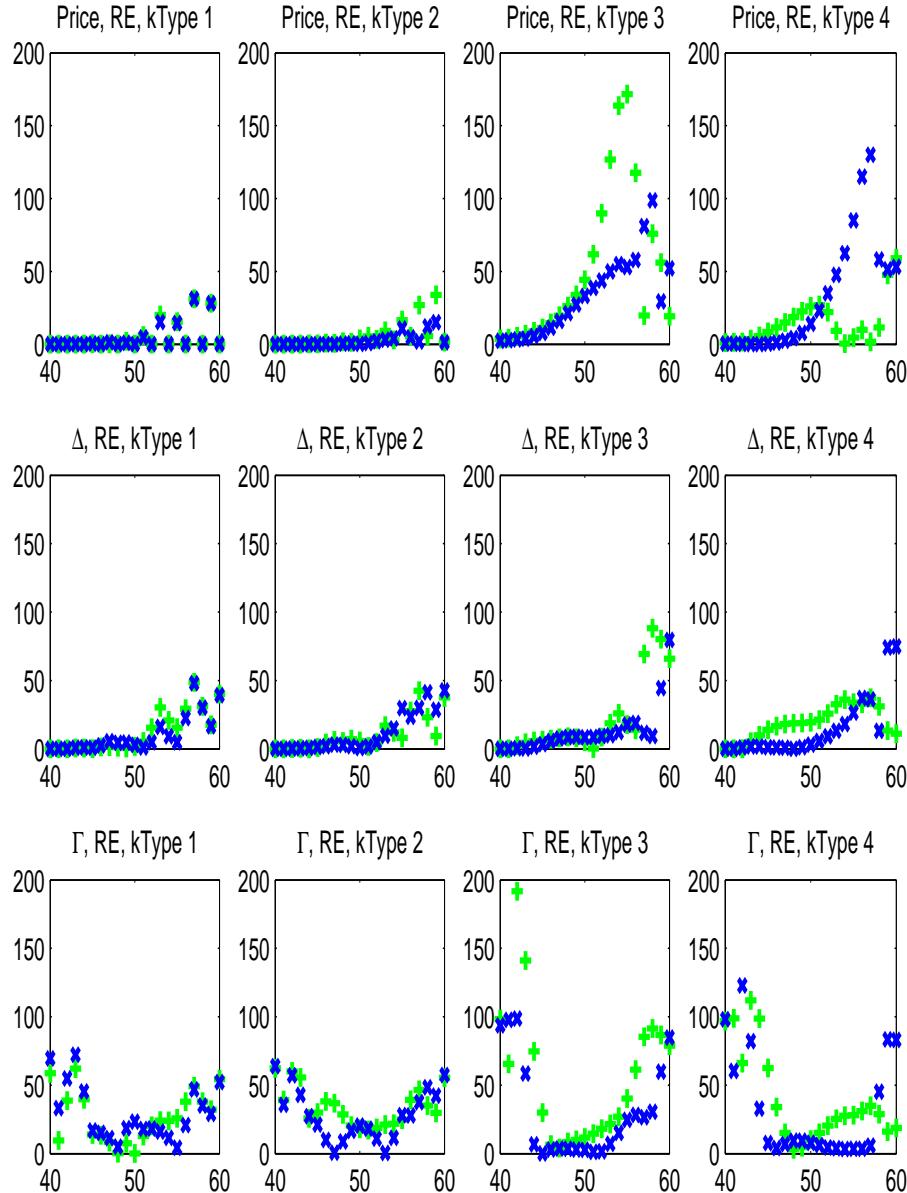
**Figure 6.1.** Scatter plots of the actual and fitted values in prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Black-Scholes model.

Each column gives results using options data with different ranges of strike prices. The first and the second columns use options data having 10 and 20 strike prices equally spaced over the range 40 to 60, respectively. The third and the forth columns use options data having 15 and 30 strike prices equally spaced over the range 35 to 65, respectively. The actual values are plotted with a diamond. The fitted values of our method using 10 and 20 support points are plotted with a plus and a cross, respectively.



**Figure 6.2.** Scatter plots of the absolute errors in dollar for the prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Black-Scholes model.

Each column gives results using options data with different ranges of strike prices. The first and the second columns use options data having 10 and 20 strike prices equally spaced over the range 40 to 60, respectively. The third and the forth columns use options data having 15 and 30 strike prices equally spaced over the range 35 to 65, respectively. The absolute error in dollar is calculated as  $|V_{true} - V_{fitted}|$ , where  $V_{true}$  is the true value and  $V_{fitted}$  is the fitted value. The fitted values of our method using 10 and 20 support points are plotted with a plus and a cross, respectively.



**Figure 6.3.** Scatter plots of the absolute relative errors in percentage for the prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Black-Scholes model.

Each column gives results using options data with different ranges of strike prices. The first and the second columns use options data having 10 and 20 strike prices equally spaced over the range 40 to 60, respectively. The third and the forth columns use options data having 15 and 30 strike prices equally spaced over the range 35 to 65, respectively. The absolute relative error in percentage is calculated as  $100 \times |V_{true} - V_{fitted}| / V_{true}$ , where  $V_{true}$  is the true value and  $V_{fitted}$  is the fitted value. The fitted values of our method using 10 and 20 support points in the Quadrature model are plotted with a plus and a cross, respectively.

3. Calculate the gamma by plugging  $\sigma_{imp}$  as the volatility in the gamma formula under the Black-Scholes model.

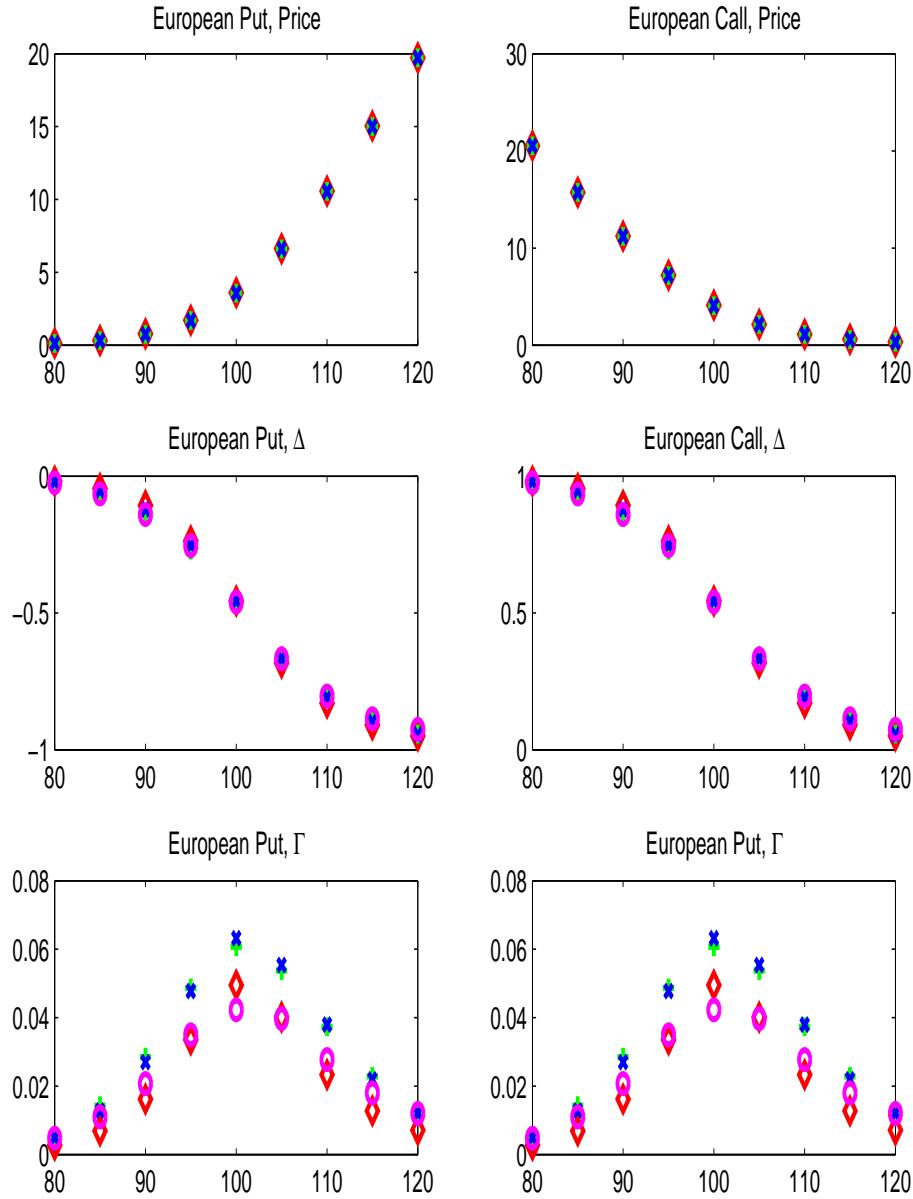
We call these estimates of Greeks the Black-Scholes Greeks. Delta and gamma calculated using the above Black-Scholes approximation are called Black-Scholes delta and Black-Scholes gamma, respectively.

Figure 6.4 gives scatter plots of the actual and fitted values for the prices, deltas, and gammas against strike prices with 10 and 20 support points used in the Quadrature model. Figure 6.5 gives scatter plots of the absolute errors in dollar for the prices, deltas, and gammas against strike prices, and Figure 6.6 gives scatter plots of the absolute relative errors in percentage for the prices, deltas, and gammas against strike prices. These two error plots show that nonparametric deltas are more accurate than Black-Scholes deltas. However, nonparametric gammas do not appear to be more accurate than Black-Scholes gammas. One reason for this may be that the Quadrature model needs more information, for example, by having data from a wider range of strike prices, which is a point of interest for future research.

### 6.3 Empirical study

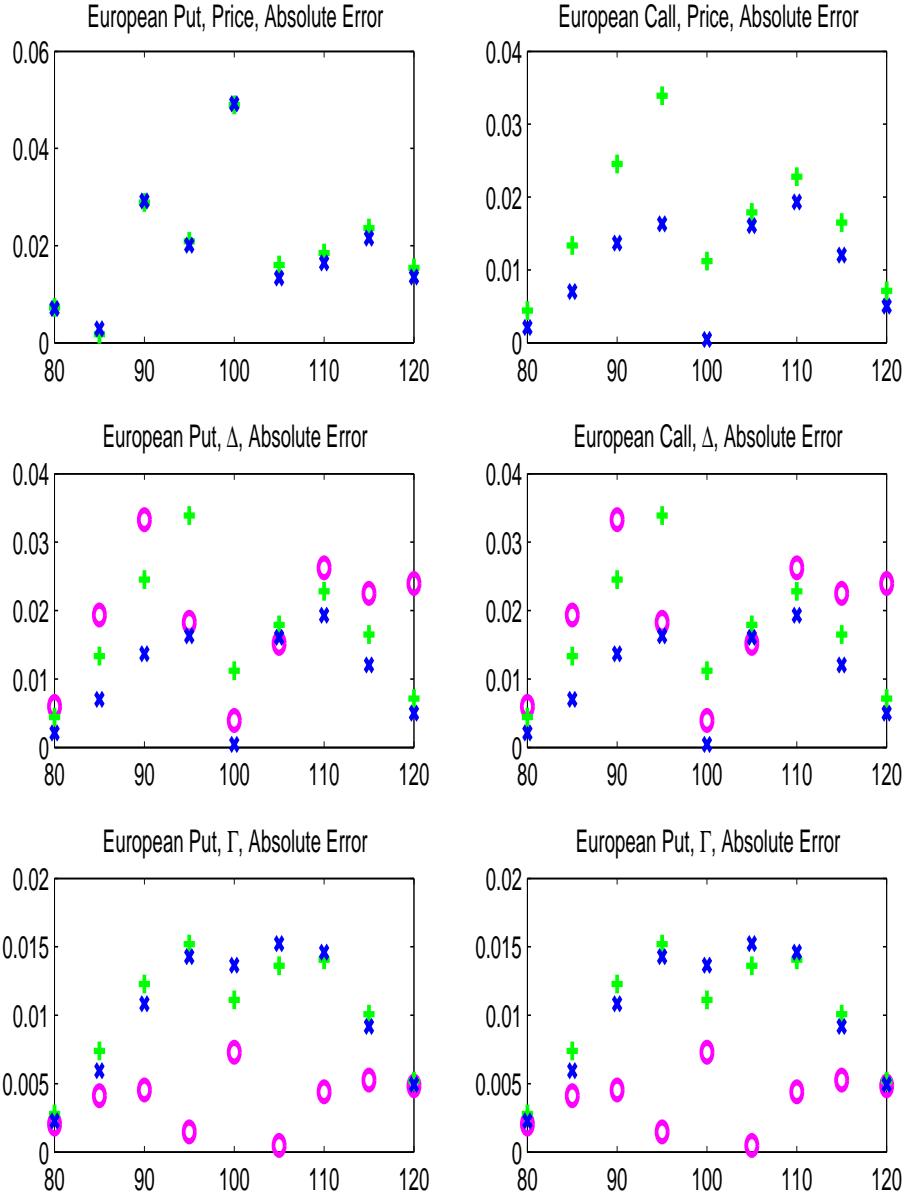
In this section, we implement our method for calculating nonparametric Greeks using a real data set given in Aït-Sahalia and Lo (1998). This is S&P 500 index options data traded on 05/03/1993, where the S&P 500 index level is  $S_0 = 438.25$ . These options have 46 days maturity, and the risk-free interest rate is 4%. We use this data set to compare our nonparametric Greeks and Black-Scholes Greeks. We remark that true values of Greeks are unknown in the empirical study.

Table 6.1 summarizes the option prices, implied volatilities, Black-Scholes deltas, and Black-Scholes gammas. This table shows that the implied volatility is not constant across strike prices. Figure 6.7 gives the scatter plot of implied volatilities against strike prices. This smiling shape of the volatility is called the “volatility smile”, and provides evidence against the constant volatility assumption of Black-Scholes’ model. This plot also shows a larger implied volatility for smaller strike price, and a smaller implied volatility for a larger strike price.



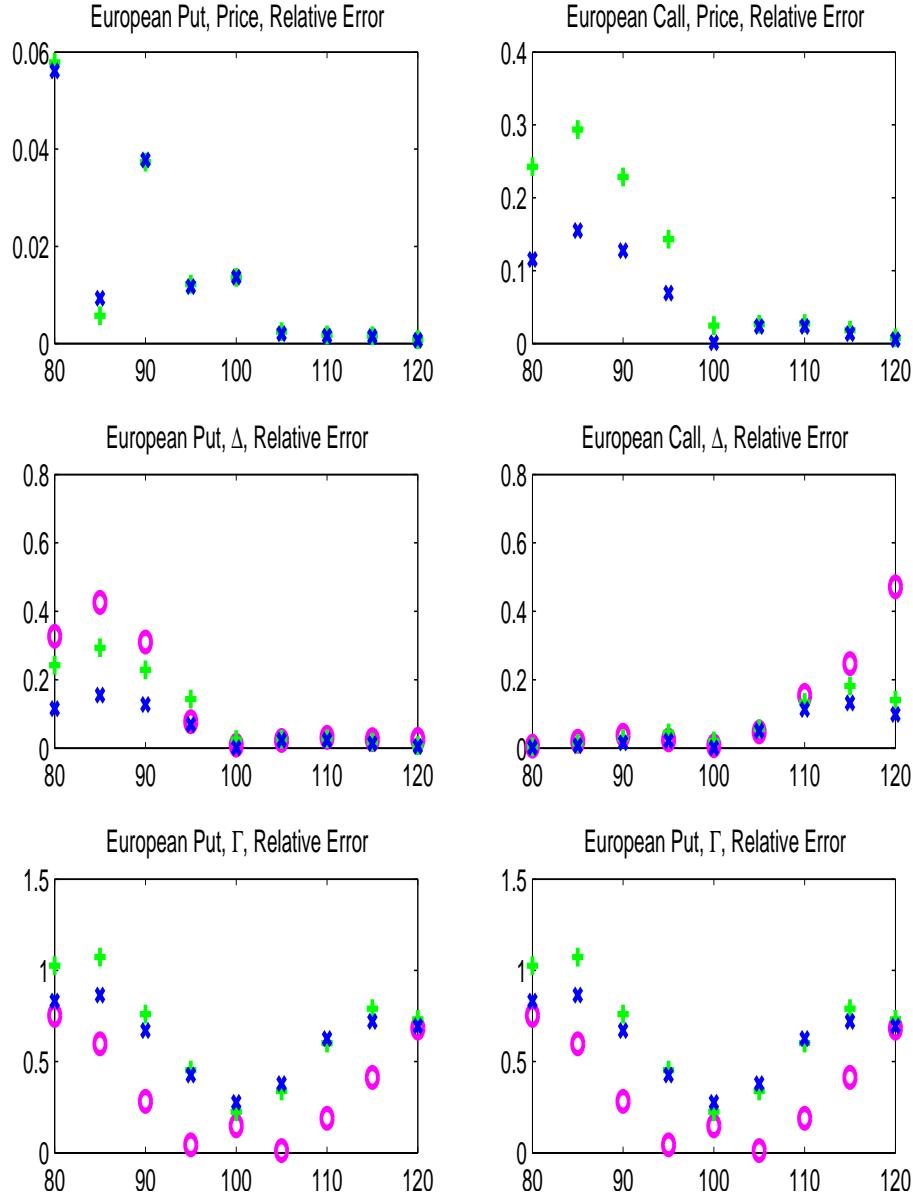
**Figure 6.4.** Scatter plots of actual and fitted prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Heston stochastic volatility model.

The actual values are plotted with a diamond. The fitted Black-Scholes deltas and gammas are plotted with a circle. The fitted nonparametric deltas and gammas with 10 and 20 support points are plotted with a plus and a cross, respectively.



**Figure 6.5.** Scatter plots of the absolute errors in price for the prices, deltas, and gammas, of European put and European call options, against strike prices for the simulation study using the Heston stochastic volatility model.

The absolute error in dollar is calculated as  $|V_{true} - V_{fitted}|$ , where  $V_{true}$  is the true value and  $V_{fitted}$  is the fitted value. The absolute errors in dollar for Black-Scholes deltas and gammas are plotted with a circle, and those for nonparametric deltas and gammas with 10 and 20 support points are plotted with a plus and a cross, respectively.



**Figure 6.6.** Scatter plots of the absolute relative errors in percentage for the simulation study using the Heston stochastic volatility model.

The absolute relative errors in percentage is calculated as  $100 \times |V_{true} - V_{fitted}| / V_{true}$ , where  $V_{true}$  is the true value and  $V_{fitted}$  is fitted value. The absolute relative errors in percentage for Black-Scholes deltas and gammas are plotted with a circle. The absolute relative errors in percentage for nonparametric deltas and gammas with 10 and 20 support points are plotted with a plus and a cross, respectively.

**Table 6.1.** European call option prices, implied volatilities (Imp. Vol.), Black-Scholes deltas, and Black-Scholes gammas for the empirical study using S&P 500 index options. Black-Scholes deltas (B-S  $\Delta$ ) and Black-Scholes gammas (B-S  $\Gamma$ ) are calculated by plugging the implied volatilities as the volatility to the formulas of delta and gamma under the Black-Scholes model, respectively.

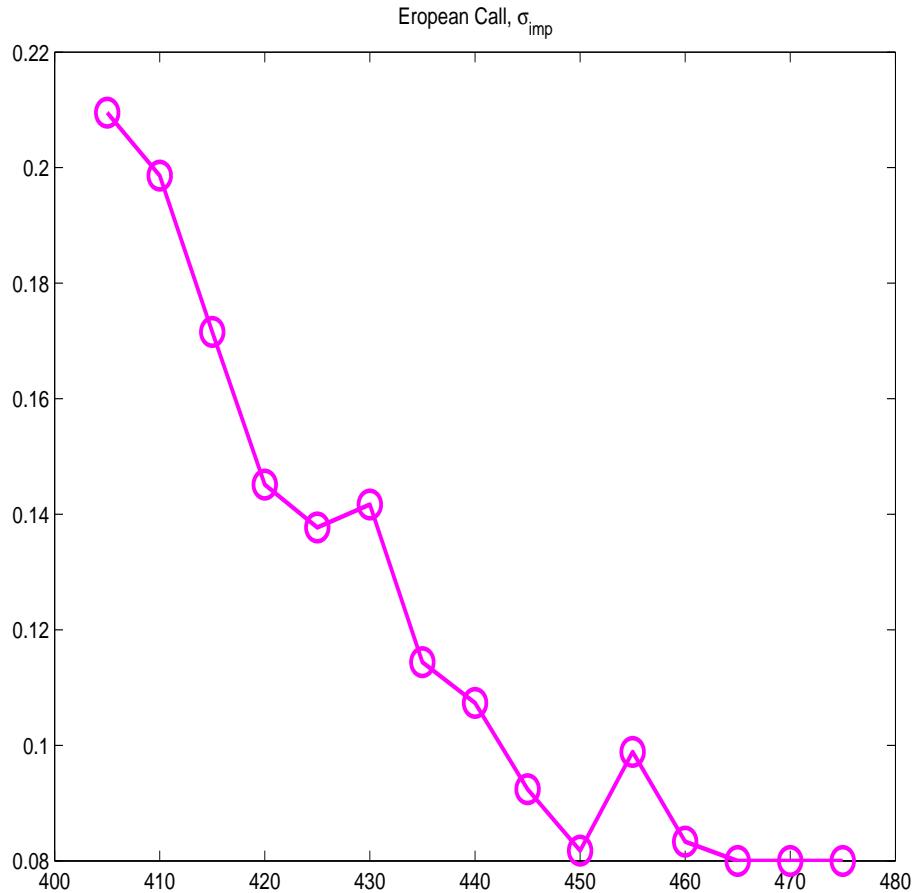
Strike	Price	Imp. Vol.	B-S $\Delta$	B-S $\Gamma$
405	38.21	0.2095	0.82	0.0059
410	33.42	0.1986	0.80	0.0067
415	28.78	0.1716	0.79	0.0079
420	24.24	0.1451	0.78	0.0097
425	19.84	0.1377	0.74	0.0113
430	15.79	0.1417	0.67	0.0121
435	11.91	0.1144	0.63	0.0157
440	8.55	0.1074	0.55	0.0175
445	5.66	0.0924	0.45	0.0204
450	3.52	0.0818	0.34	0.0212
455	1.95	0.0989	0.28	0.0163
460	1.01	0.0833	0.17	0.0143
465	0.47	0.0800	0.10	0.0104
470	0.22	0.0800	0.06	0.0070
475	0.13	0.0800	0.03	0.0043

Figure 6.8 gives the scatter plot of the fitted and actual prices against strike prices. Figure 6.9 gives the scatter plot of Black-Scholes deltas and nonparametric deltas against strike prices, and Figure 6.10 gives the scatter plot of Black-Scholes gammas and nonparametric gammas against strike prices.

We find that there exist apparent differences between our nonparametric Greeks and Black-Scholes Greeks. However, a more comprehensive empirical study would be required to establish the claim that nonparametric Greeks are more accurate Black-Scholes Greeks. We remark that this type of study is outside the current scope of this dissertation, and is worthy of further studies.

## 6.4 Summary

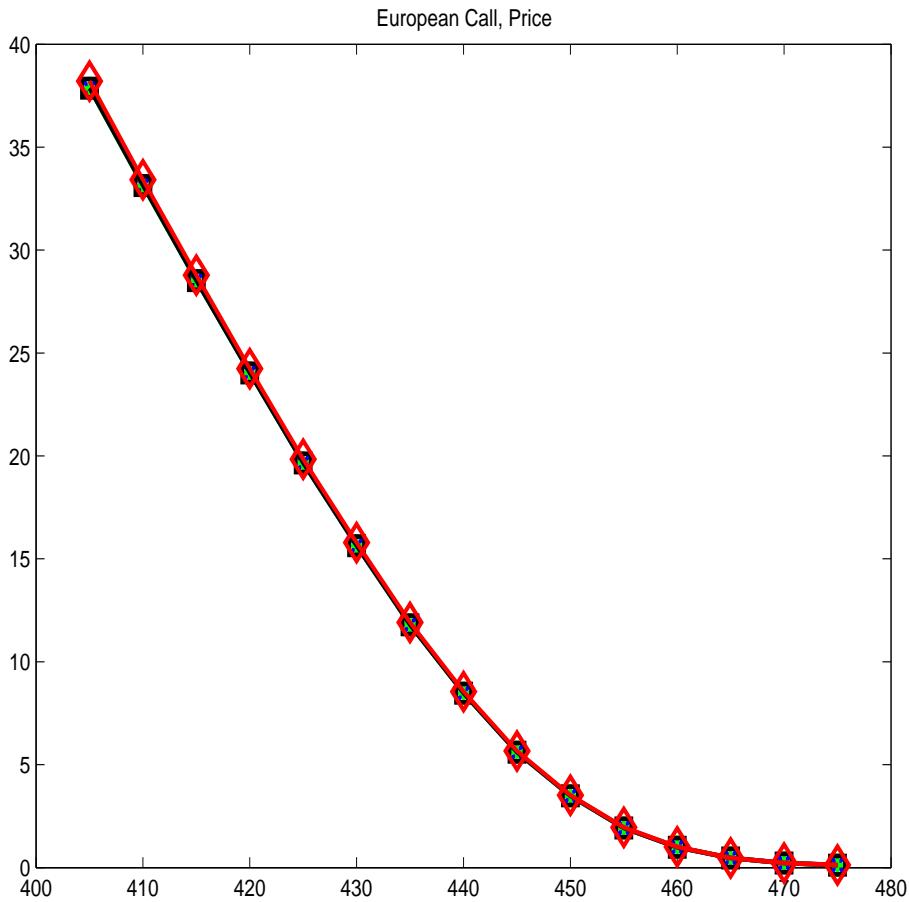
In this chapter, we investigate how to estimate Greeks directly from observed option prices without parametric assumptions on the state price densities or under-



**Figure 6.7.** The scatter plot of implied volatilities against strike prices for the empirical study.

lying process. Our estimated Greeks are calculated by a finite-difference formula based on the Quadrature model and the extended implied random tree model, which allows us to calculate nonparametric Greeks. Hence, we call Greeks using our methods nonparametric Greeks.

In order to study the performance of our proposed method, we consider two simulation studies. The first simulation study uses a Black-Scholes model, and the second simulation study uses the Heston stochastic volatility model. We find that our approach produces good predictions of the price, but the estimated Greeks appear to be biased. There are probably two reasons for the bias. First, the observed information is incomplete in the sense that we can only observe options with several “discrete” strike prices in a “finite range”. This poses a natural limi-

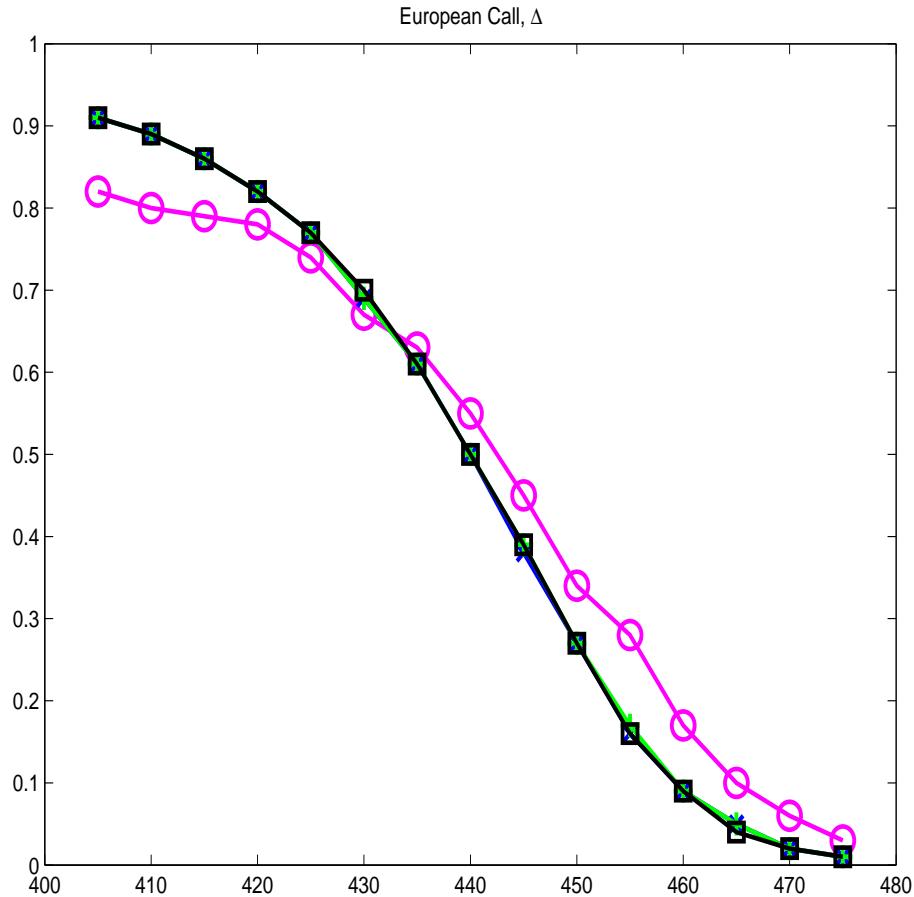


**Figure 6.8.** The scatter plot of actual and fitted option prices against strike prices for the empirical study.

The actual option prices are plotted with a diamond-marked line. The fitted option prices of our methods using 20, 40, and 60 support points are plotted with a plus-, a cross-, and a square-marked line, respectively.

tation on methods, including our method. Second, our methods is approximating a continuous density using a Quadrature model, and approximating a stochastic process using an implied binomial random tree. Therefore, there is a bias in this approximation scheme. Ideally, the bias can be decreases if the number of support points is increases. However, this will also increase the computation burden.

Finally, we apply our approach to a real data set using S&P 500 index options. We find that there are differences between our nonparametric Greeks and Black-Scholes Greeks (which are approximated by plugging the implied volatility to the Greeks formulas under the Black-Scholes model). A comprehensive empiri-

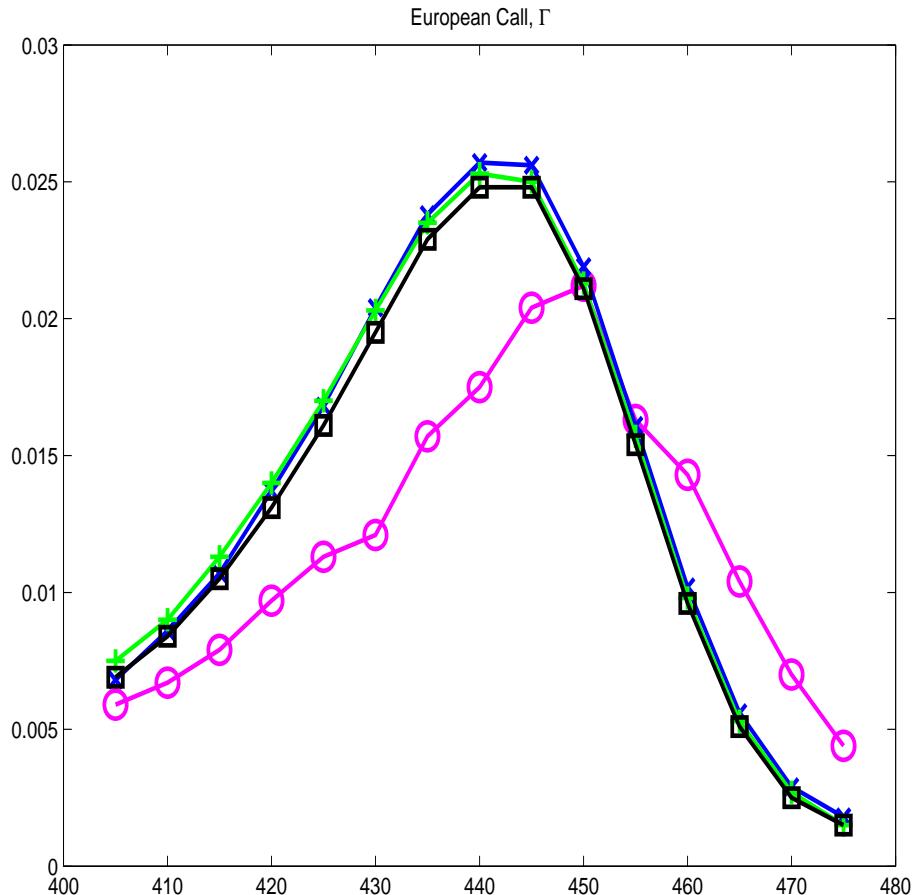


**Figure 6.9.** The scatter plot of Black-Scholes deltas and nonparametric deltas against strike prices for the empirical study.

The Black-Scholes deltas are plotted with a circle-marked line. The nonparametric deltas using 20, 40, and 60 support points are plotted with a plus-, a cross-, and a square-marked line, respectively.

cal studies to establish the superiority of our method is out of the scope of current dissertation, and is worthy of further study.

In summary, our methods are straightforward and easy to implement. In addition, our methods produce very good prediction for the prices, but not always for the Greeks. However, the accuracy of the estimates of the Greeks can be improved as more information on the options is available. Our methods can be used to calculate deltas and gammas directly, and deserve additional empirical investigation for dynamic hedge purposes.



**Figure 6.10.** The scatter plot of Black-Scholes gammas and nonparametric gammas against strike prices for the empirical study.

The Black-Scholes gammas are plotted with a circle-marked line. The nonparametric gammas using 20, 40, and 60 support points are plotted with a plus-, a cross-, and a square-marked line, respectively.

Chapter **7**

# Conclusions and future work

## 7.1 Conclusions

Financial options have dramatically influenced the global finance market. Asset pricing theory states that an option price is equal to the discounted payoff of an underlying asset under the risk-neutral measure. The density under this risk-neutral measure is called the state price density or the risk-neutral density, and is used for a variety of important activities in finance. The importance of understanding this density with respect to asset pricing and risk management has led to a competing number of approaches for inference.

Greeks summarize how option prices change with respect to underlying variables and are critically important in asset pricing and risk management. Greeks are the price's sensitivities with respect to certain parameters of interest such as the underlying asset's price, volatility, maturity, and interest rate. Although the price of an option can often be observed in the market, this is not the case for its Greeks; they must be calculated. Because Greeks are important for measuring and managing risk as well as executing dynamic trading strategies, the ability to calculate them efficiently and accurately is of critical importance both in theory and in practice.

In this dissertation, we aim at providing Bayesian nonparametric approaches to calibrate the state price density and Greeks using observed option prices. Our methods have several advantages over existing methods. First, our approaches do not make any parametric assumptions on the underlying process, hence our

approaches are free of model-misspecification problems. Our approach can be applied directly to the problem of calibrating the state price density using American options, whereas most existing nonparametric methods are inapplicable of using American options in their calibration efforts. In addition, our Bayesian approaches can incorporate prior information based on a researcher's knowledge. We summarize our work as follows.

### 7.1.1 The Bayesian Quadrature model

We first propose a finite-dimensional model for the state price density in a Bayesian framework. This modeling approach can be viewed as a Bayesian Quadrature model, where the locations and weights of support points in the finite-dimensional representation of the state price density are random variables. This modeling approach allows a 'prior' reference distribution which can be a parametric distribution (e.g. the lognormal density) or which can be uniform and completely non-informative, and it also provides a posterior distribution of state price densities that are consistent with the observed option prices. We asses the performance of the proposed model using simulation studies based on synthetic data and then by contrasting the method with a number of competing methods using S&P 500 index option data.

### 7.1.2 The Bayesian implied random tree model

In practice, American options are more popular than European options, mainly because American options can be exercised any time prior to maturity and are hence more flexible. As an extension of the Bayesian Quadrature model, we build a generalized random binomial tree similar to Rubinstein (1994) for realization of the Bayesian Quadrature. We demonstrate the performance of our approach via a simulation study and an empirical study using S&P 100 index options data. In our empirical study, we also find that the state price density calibrated from the European options does not predict the American option prices as well as the state price density calibrated from American options can predict the European options. This is partially because European options are traded with a narrower range of strike prices than American options. This empirical evidence also illustrates the

importance of incorporating most available market information for calibrating the state price density.

### 7.1.3 Parametric Greeks

Greeks summarize how option prices change with respect to underlying variables and are critically important in asset pricing and risk management. Greeks are critical for risk management and dynamic hedge purposes. Unlike option prices, which can be observed in the market, Greeks can not be observed and have to be calculated given a model assumption. Typically, the Greeks are calculated using a partial differentiation of the price formula. However, when options have complicated payoff function, closed-form formulas of option prices are rare. Therefore, numerical methods, such as Monte Carlo simulation and finite-difference methods are the preferred tools. Although the finite-difference method is a straightforward method for calculating Greeks, finite-difference estimates are usually biased. We build a new but simple mathematical formulation so that formulas of Greeks for a broad class of derivative securities can be derived systematically. Our method is applicable to many different types of options and process for the underlying asset.

### 7.1.4 Nonparametric Greeks

We combine the Bayesian Quadrature and implied random tree models with finite-difference formulas in order to provide a general framework for calculating nonparametric estimates of the Greeks. Our approach is applicable to calculate Greeks using European and American options. We demonstrate our approach via simulation and empirical studies.

## 7.2 Future work

The success of our Bayesian nonparametric approaches, for making inference about the state price density using European and American options in Chapters 3 and 4, suggests that the method could be extended to accommodate other exotic options, including path-dependent and rainbow options. One natural extension is to make the locations multidimensional or to make them realizations of diffusion processes.

It is also interesting to investigation if the estimation can be further enhanced using data including exotic options. Both of these extensions, with applications to exotic options and Interest rate options (e.g. Credit Default Spreads) offer additional fruitful areas for future study. Other interesting areas of future research are described below.

### 7.2.1 Implied multi-variate state price density

In this dissertation, we focus on calibrating the state price density of a single underlying asset. However, the need to calibrate multi-variate risk neutral density for underlying assets is a topic that is becoming more important in practice with the rapid development of complicated structured multi-asset options. For example, rainbow options have been created in the finance market to accommodate different speculation or hedge purposes.

One possibility is to extend the Bayesian Quadrature methods by allowing the locations to be multivariate. For example, when we consider a rainbow options depending on  $M$  underlying assets, the Quadrature model can be extended to

$$f(\mathbf{x}|w, \theta) = w_0\delta_{\theta_0}(\mathbf{x}) + \cdots + w_N\delta_{\theta_N}(\mathbf{x}),$$

where  $\mathbf{x} \in \Re^M$ , the weights of the support points  $w = \{w_0, \dots, w_N\}$  are nonnegative and sum to one, and the locations of the support points  $\theta = \{\theta_0, \dots, \theta_N\}$  are of  $M$ -variate, i.e.  $\theta_n = (\theta_{n,1}, \theta_{n,2}, \dots, \theta_{n,M})$  for  $n = 0, \dots, N$ .

Another approach is to construct the dependence among multiple assets using a copula. A copula provides a general framework for formulating a multivariate distribution, so that various general types of dependence can be represented. We refer to Nelsen (1999) for an overview of a copula. A copula is a multivariate joint cumulative distribution function defined on the  $n$ -dimensional unit cube  $[0, 1]^n$ , such that every marginal distribution is uniform on the interval  $[0, 1]$ . Specifically,  $C$  is an  $n$ -dimensional copula if:

1.  $C(\mathbf{u}) = 0$ , whenever  $\mathbf{u} \in [0, 1]^n$  has at least one component equal to 0;
2.  $C(\mathbf{u}) = u_i$ , whenever  $\mathbf{u}$  has all the components equal to 1, except the  $i$ -th one, which is equal to  $u_i$ ;

3.  $C(\mathbf{u})$  is  $n$ -increasing, i.e., for each hyper-rectangle  $B = \times_{i=1}^n [x_i, y_i] \subseteq \mathcal{B}$ , where  $\times$  denotes a Cartesian product.

$$V_C(B) = \sum_{\mathbf{z} \in \times_{i=1}^n [x_i, y_i]} (-1)^{N(\mathbf{z})} C(\mathbf{z}) \geq 0,$$

where  $N(\mathbf{z}) = \text{card}\{k | z_k = x_k\}$ .  $V_C(B)$  is called the  $C$ -volume of  $B$ .

To illustrate the idea of using a copula for constructing a dependence between two random variables  $X$  and  $Y$ , let  $X$  and  $Y$  have continuous cumulative distribution functions  $F_X(x)$  and  $F_Y(\cdot)$ , respectively. Define two random variables  $X' = F_X^{-1}(X)$  and  $Y' = F_Y^{-1}(Y)$  as the inverse transformation of the cumulative distribution functions of  $X$  and  $Y$ , respectively. Using standard probability arguments,  $X'$  and  $Y'$  are both uniform distributions. Since the transforms are invertible, once the dependence of  $X'$  and  $Y'$  is specified using a copula  $C$ , the dependence of the original random variables  $X$  and  $Y$  is also specified. The basic idea of a copula is to simplify the problem by removing consideration of many different marginal distributions by transforming the marginal distributions to uniforms, and then specifying dependence as a multivariate distribution on the uniforms.

Dependence modeling with copula functions is widely used in applications of financial risk assessment and actuarial analysis - for example in the pricing of collateralized debt obligations (Meneguzzo and Vecchiato, 2003). Copulas have also been applied in analyzing a variety of rainbow options (Dong, 2001, 2005). Although the use of a Copula framework offers an interesting avenue for potential future research, the framework of using a copula is limited in terms of the lack of dependence dynamics and the poor representation of extreme events (Lipton and Rennie, 2007).

### 7.2.2 Implied diffusion processes with an application to Credit Default Swaps

We have proposed a Bayesian implied random tree to calibrate the state price density using American options in Chapter 4. Although a tree-based approach can be used to approximate a diffusion process, it usually requires a large number of time-steps and are therefore computationally demanding. An alternative approach

is to consider Monte Carlo based approaches for calibrating the diffusion processes.

To illustrate, we consider the example of calibrating a diffusion process for Credit Default Swaps (CDS). CDS are financial contracts that allow the default risk of a defaultable bond to be transferred in the market. They are like an insurance contract that pays out based on a credit event (e.g. the default of a corporate bond). We assume that the default occurs without warning with an exogenous intensity. The intensity can be calibrated from market prices of various credit sensitive securities, such as liquid debt prices or CDS spreads (Duffie and Huang, 1996). We refer to Duffie and Singleton (2003) for a general introductions to these intensity models.

The link between the intensity process  $X(t)$  and the survival function  $q(t)$  at time  $t$  is described below. Define  $T$  as the time to default for a default-risky bond. The survival function  $q(t)$  is the probability that the default-risky bond defaults no earlier than  $t$ , i.e.,  $q(t) = P(T > t)$ . The intensity,  $X(t)$ , is defined as

$$X(t) = \lim_{dt \rightarrow 0} \frac{P(T \in [t, t + dt] | T > t)}{dt}.$$

Following this definition, we have

$$X(t) = \lim_{dt \rightarrow 0} \frac{q(t) - q(t + dt)}{q(t)dt} = -\frac{q'(t)}{q(t)}.$$

Hence, the survival function is linked by the intensity with the following formula,

$$q(t) = e^{-\int_0^t X(s)ds}.$$

CDS can be priced using intensity process (Duffie and Singleton, 2003). CDS have two potential cash flow streams, a fixed leg and a contingent leg. On the fixed leg side, the buyer of protection makes a series of fixed periodic payments, the CDS premium, until maturity or until the reference credit defaults. On the contingent leg side, the protection seller makes a payment only if the reference credit defaults. The amount of a contingent payment is usually the notional amount multiplied by  $(1 - R)$ , where  $R$  is the recovery rate. The value of the CDS contract to the protection buyer is the difference between the present value of the fixed leg and

that of the contingent leg, i.e.,

$$\begin{aligned} & \text{Value of CDS (to the protection buyer)} \\ = & \text{PV[contingent leg]} - \text{PV[fixed (premium) leg]}. \end{aligned}$$

If we define  $y$  as the annual Credit Default Swaps premium,  $d_i$  as accrual days (expressed in a fraction of one year between payment days),  $q(t_i)$  as survival probability at time  $t_i$ , and  $D(t_i)$  as the discount factor for the particular date  $t_i$ , then the present value of the fixed leg is

$$\sum_{i=1}^N D(t_i)q(t_i)y d_i + \sum_{i=1}^N D(t_i)\{q(t_{i-1}) - q(t_i)\}y \frac{d_i}{2},$$

and the present value of the contingent leg is

$$(1 - R) \sum_{i=1}^N D(t_i)\{q(t_{i-1}) - q(t_i)\}.$$

The CDS spread is set so that the swap transaction is zero, which results in

$$y = \frac{(1 - R) \sum_{i=1}^N D(t_i)\{q(t_{i-1}) - q(t_i)\}}{\sum_{i=1}^N D(t_i)q(t_i)d_i + \sum_{i=1}^N D(t_i)\{q(t_{i-1}) - q(t_i)\}\frac{d_i}{2}}.$$

If we assume that  $X(t)$  is stochastic, the CDS spread becomes the expectation over the probability of default, or

$$y = \frac{E\left[(1 - R) \sum_{i=1}^N D(t_i)\{q(t_{i-1}) - q(t_i)\}\right]}{E\left[\sum_{i=1}^N D(t_i)q(t_i)d_i + \sum_{i=1}^N D(t_i)\{q(t_{i-1}) - q(t_i)\}\frac{d_i}{2}\right]}. \quad (7.1)$$

To build a likelihood, we assume observed CDS spreads are theoretic CDS values contaminated with some random noises. For example, let  $y$  denote an observed CDS spread, which follows

$$y = E[g_1(X)]/E[g_2(X)] + \varepsilon, \quad (7.2)$$

where  $E[g_1(X)]/E[g_2(X)]$  equals the right-hand-side of Eq (7.1) with an error

$$\varepsilon \sim N(0, \tau).$$

One common parametric modeling for intensity process is the Cox-Ingersoll-Ross (CIR) model. The intensity model follows

$$dX_t = \kappa(\vartheta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad (7.3)$$

for parameters  $\kappa$ ,  $\vartheta$ , and  $\sigma$ , and  $W_t$  is the standard Brownian motion. Let  $X_0$  be the initial value of the intensity process. We generate  $N$  realizations of the intensity process at discrete dates  $t_0, t_1, \dots, t_m$  with  $\Delta t = t_j - t_{j-1}$  for  $j = 1, \dots, m$ , to approximate the theoretic CDS spread as in Eq. (7.1) using a Monte Carlo estimate of the integral. Let  $X_{t_j}^n$  denote the intensity at the  $j$ -th time point for the  $n$ -th realization for  $j = 1, \dots, m$  and  $n = 1, \dots, N$ .  $X_{t_j}^n$  is simulated using

$$X_{t_j}^n - X_{t_{j-1}}^n = \kappa(\vartheta - X_{t_{j-1}}^n)\Delta t + \sigma\sqrt{X_{t_{j-1}}^n\Delta t} Z, \quad Z \sim N(0, 1) \quad (7.4)$$

for  $j = 1, \dots, m$  and  $n = 1, \dots, N$ . In this parametric framework, the parameters of interest include  $X_0$ ,  $\kappa$ ,  $\vartheta$ , and  $\sigma$ .

An extension to this framework that would help to avoid model misspecification uses the following semi-parametric model,

$$dX_t = \kappa(\vartheta - X_t)dt + \sigma\sqrt{X_t}d\tilde{W}_t,$$

where the definition of  $d\tilde{W}_t$  can be modeled by the Bayesian Quadrature model,  $Q(w, \theta)$ , without a parametric assumption. We generate a sample of intensity processes using

$$X_{t_j}^n - X_{t_{j-1}}^n = \kappa(\vartheta - X_{t_{j-1}}^n)\Delta t + \sigma\sqrt{X_{t_{j-1}}^n}\tilde{Z}, \quad \tilde{Z} \sim Q(w, \theta). \quad (7.5)$$

The parameters of interests in this semi-parametric framework include  $X_0$ ,  $\kappa$ ,  $\vartheta$ ,  $\sigma$ ,  $w$ , and  $\theta$ .

Another alternative approach is to use the CIR model as a prior distribution for

the intensity process. Let  $\Lambda$  denote the  $N$  realizations of the intensity processes,

$$\Lambda = \{X_{t_j}^n\} = \begin{bmatrix} X_{t_0}^1 & X_{t_1}^1 & \cdots & X_{t_m}^1 \\ X_{t_0}^2 & X_{t_1}^2 & \cdots & X_{t_m}^2 \\ \vdots & \vdots & \ddots & \vdots \\ X_{t_0}^N & X_{t_1}^N & \cdots & X_{t_m}^N \end{bmatrix}.$$

Let  $\Delta X_{t_j}^n = X_{t_j}^n - X_{t_{j-1}}^n$  for  $j = 1, \dots, m$  and  $n = 1, \dots, N$ . The prior distribution of  $\Lambda$  given  $\kappa$ ,  $\vartheta$ , and  $\sigma$ , is

$$p(\Lambda|X_0, \kappa, \vartheta, \sigma) = \prod_{n=1}^N \prod_{j=1}^m \frac{1}{\sqrt{2\pi\sigma^2 X_{t_{j-1}}^n \Delta t}} e^{-\frac{((\Delta X_{t_j}^n) - \kappa(\vartheta - X_{t_{j-1}}^n)\Delta t)^2}{2\sigma^2 X_{t_{j-1}}^n \Delta t}}$$

The inference focuses on the joint posterior distribution,

$$p(\kappa, \vartheta, \sigma, X_0, \Lambda|y) \propto L(y|\Lambda)p(\Lambda|X_0, \kappa, \vartheta, \sigma, \kappa_0)p(\kappa)p(\vartheta)p(\sigma)p(X_0).$$

The joint distribution can be sampled using a Markov chain Monte Carlo algorithm. It would be interesting to investigate the performance of using these three competing models.

### 7.2.3 New tree-based approaches for pricing options

Our proposed Bayesian implied random trees calibrate the state price density using observed option prices. On the other aspect, it is challenging to build an innovative tree for pricing purposes. In other words, it is challenging to price different options given an arbitrary parametric assumptions on the underlying process.

A classic tree-based approach is the binomial model by Cox et al. (1979). The Cox-Ross-Rubinstein binomial trees are popular numerical methods for pricing options. However, the binomial model is not computationally feasible to handle more than a couple of stochastic factors, such as interest rates, dividend, volatility, or multiple underlying asset, due to the curse of dimensionality. Therefore, it would be interesting to build an innovative tree which requires less time-steps to attain the same accuracy.

In addition, many challenges occur when building tree-based model to incorporate (1) different parametric assumptions on the underlying asset (2) multivariate factors, including the interest rates, dividend, and so on, and (3) multivariate underlying assets for pricing rainbow options.

When the underlying process is complicated, the CRR binomial tree can not perform very well. Therefore, many tree-based numerical methods have been proposed for improvement. For example, in the GARCH model, Rubinstein's (1998) Edgeworth tree combines the analytical formulas for moments of the cumulative return under GARCH developed in Duan et al. (1999, 2002) in order to yield a simple recombining binomial tree for option valuation in the GARCH context.

Tree-based approaches are popular and useful for pricing American options. However, it is in general difficult to build a binomial tree to price a basket of American options. Therefore, creating a tree-based approach to incorporate multivariate underlying assets is worthy of future studies.

#### **7.2.4 Do American options bring more information than European options for understanding state price densities?**

In Chapter 3, we have shown the state price density calibrated using American options can predict European option prices for one data set using S&P 100 index options. However, the state price density calibrated using European options can not predict American option prices as well as vice versa.

One reason is that more American options are traded, and they are traded with a wider range of strike prices. Therefore, they provide more information the tail of the state price densities. As a result, the state price density calibrated from American options seems to be more reliable.

This fact highlights one advantage of our approach, as our approach can calibrate state price densities using both American and European options at the same time, whereas most nonparametric methods only use European (call) options, such as kernel-based methods. This raises the interesting question as to whether state price densities calibrated from European options and from American options are the same.

### 7.2.5 Will nonparametric Greeks improve the performance of dynamic hedge?

In an empirical study in Chapter 6, we find that there is an apparent difference in our nonparametric Greeks and the Black-Scholes Greeks (which are approximated by plugging the implied volatility to the Greeks formulas under the Black-Scholes model). Therefore, it is interesting to compare the dynamic hedge effectiveness using our nonparametric Greeks and Greeks calculated using other methods.

Dynamic hedging is a professional procedure widely used by options dealers to hedge delta, gamma, or vega exposures. To illustrate dynamic hedging, we focus on how to perform a dynamic delta hedge. This hedging procedure involves adjusting the hedged portfolio as the underlying process moves; this dynamic updating results in its name.

Assume that the hedging or rebalancing can be implemented at discrete times,  $t_0, t_1, \dots, t_m$ , in the hedging time horizon  $[0, \tau]$  with  $0 \leq \tau \leq T$ , where  $T$  is the maturity of an option. Without loss of generality, we assume that every period between two consecutive rebalancing times is the same, i.e.,  $t_j - t_{j-1} = \Delta t$  for  $j = 1, \dots, m$ . For dynamic delta hedging, we construct the delta hedged portfolio  $\Pi(t)$  at time  $t$  as

$$\Pi(t) = C(t) + \Delta(t)S(t) + B(t),$$

where  $C(t)$  is the value of the hedged option,  $\Delta(t)$  is the option's delta,  $S(t)$  is the value of the underlying asset, and  $B(t)$  is the value of the money market account, at time  $t$ , respectively.

At the beginning of the hedge horizon, we own one option,  $\Delta$  shares of the underlying asset, and  $-(C(0) + \Delta(0)S(0))$  dollars in the money market account. Therefore, the initial value of the hedged portfolio is zero. At each rebalancing time  $t_j$ ,  $\Delta(t_j)$  is re-calculated and the money market account is adjusted by

$$B(t_j) = e^{r\Delta t}B(t_{j-1}) - S(t_j)(\Delta(t_j) - \Delta(t_{j-1})),$$

where  $r$  is the risk-free interest rate. Thus, the hedged portfolio is self-financing, which means that this portfolio does not have extra infusion or withdrawal of money in the hedge period. To measure the dynamic hedge effectiveness, we cal-

culate the dynamic delta hedge error, which is defined as the absolute value of the hedged portfolio at the end of the hedge horizon of the option,  $|\Pi(\tau)|$ . For a more detailed description on the dynamic hedge, we refer to Hull (2002).

Appendix **A**

## Slice Sampler

To begin with, we define

$$g_{ij}(w, \theta) = - \sum_{k=1}^{N_{ij}} v_{ijk} (\log y_{ijk} - \log G_{ij}(w, \theta))^2 / 2\sigma_\varepsilon^2. \quad (\text{A.1})$$

The logarithm of the posterior distribution can be written as

$$\left( \sum_{i=1,2} \sum_{j=1,\dots,N_i} g_{ij}(w, \theta) + \log(p(\sigma_\varepsilon^2 | \alpha_\varepsilon, \beta_\varepsilon)) \right) \mathbf{1}_{\mathcal{W}}(x) \mathbf{1}_\Theta(\theta) + J,$$

for some constant  $J$ . Given an arbitrary value  $z$ , we define the set

$$B_{ij}(z) = \{(w, \theta) \in \Re^{M \times M} : g_{ij}(w, \theta) > z\}.$$

To simplify the set  $B_{ij}(z)$ , we calculate

$$\begin{aligned} a &= \sum_{k=1}^{N_{ij}} v_{ijk}, \\ b &= -2 \sum_{k=1}^{N_{ij}} v_{ijk} \log y_{ijk}, \\ c &= \sum_{k=1}^{N_{ij}} v_{ijk} (\log y_{ijk})^2 + 2\sigma_\varepsilon^2 z. \end{aligned}$$

Note that  $a$  and  $b$  only depend on the observed options prices and volume, but  $c$  also depends on  $z$ . Now,  $B_{ij}(z)$  equals

$$\begin{aligned} & \left\{ (w, \theta) \in \Re^{M \times M} : a(\log G_{ij}(w, \theta))^2 + b \log G_{ij}(w, \theta) + c < 0 \right\} \\ = & \left\{ (w, \theta) \in \Re^{M \times M} : G_{ij}(w, \theta) \in \left( e^{\frac{-b-\sqrt{b^2-4ac}}{2a}}, e^{\frac{-b+\sqrt{b^2-4ac}}{2a}} \right) \right\} \\ = & \left\{ (w, \theta) \in \Re^{M \times M} : \sum_{m=1}^M w_m \wp_{ij}(\theta_m) \in (l_{ij}, u_{ij}) \right\}, \end{aligned}$$

where

$$l_{ij} = e^{rT + \frac{-b-\sqrt{b^2-4ac}}{2a}}, \quad (\text{A.2})$$

$$u_{ij} = e^{rT + \frac{-b+\sqrt{b^2-4ac}}{2a}}. \quad (\text{A.3})$$

The slice  $T_m$  to update  $w_m$  is the intersections of sub-slices  $T_{m0}$  and  $T_{mij}$  defined as follows. Because  $w$  has sum to one, let  $T_{m0}$  be an open interval

$$T_{m0} = (0, p_{-mM}), \quad (\text{A.4})$$

where

$$p_{-mM} = 1 - w_1 - \cdots - w_{m-1} - w_{m+1} - \cdots - w_{M-1}.$$

Let  $w^0 = (w_1^0, \dots, w_M^0)$  and  $\theta^0 = (\theta_1^0, \dots, \theta_M^0)$  being the current values of  $w$  and  $\theta$  at each iteration in the MCMC algorithm, respectively. We sample the auxiliary variable  $z_{ij} = g_{ij}(w^0, \theta^0) - e$  with  $e \sim \text{Exp}(1)$ , and define the sub-slice  $T_{mij}$  as

$$T_{mij} = \{w_m \in R : g_{ij}(w, \theta) > z_{ij}\}.$$

Standard algebra gives

$$e^{rT} G_{ij}(w, \theta) = \alpha_{ij} w_m + \beta_{ij},$$

where

$$\begin{aligned}\alpha_{ij} &= \wp_{ij}(\theta_m) - \wp_{ij}(\theta_M), \\ \beta_{ij} &= \gamma_{ij-mM} + p_{-mM} \wp_{ij}(\theta_M), \\ \gamma_{ij-mM} &= \sum_{\substack{m'=1, \dots, M \\ m' \neq m, M}} w_m \wp_{ij}(\theta_m).\end{aligned}$$

Recall  $l_{ij}$  and  $u_{ij}$  are given in Eqs. (A.2)-(A.3) given a constant  $z$ . Now,  $T_{mij}$  equals

$$\begin{aligned}T_{mij} &= \{w_m : g_{ij}(w, \theta) > z_{mij}\} \\ &= \{w_m : \sum_{m=1}^M w_m \wp_{ij}(\theta_m) \in (l_{ij}, u_{ij})\} \\ &= \{w_m : \alpha_{ij} w_m + \beta_{ij} \in (l_{ij}, u_{ij})\},\end{aligned}$$

and is determined by  $\alpha_{ij}$ ,

$$T_{mij} = \begin{cases} ((l_{ij} - \beta_{ij})/\alpha_{ij}, (u_{ij} - \beta_{ij})/\alpha_{ij}) & \text{for } \alpha_{ij} > 0, \\ ((u_{ij} - \beta_{ij})/\alpha_{ij}, (l_{ij} - \beta_{ij})/\alpha_{ij}) & \text{for } \alpha_{ij} < 0, \\ (-\infty, \infty) & \text{for } \alpha_{ij} = 0. \end{cases} \quad (\text{A.5})$$

Because  $T_{m0}$  and each  $T_{mij}$  are open intervals containing  $w_m^0$ , the slice  $T_m$  is again an interval containing  $w_m^0$ . In summary, we have

$$T_m = T_{m0} \bigcap_{\substack{i=1,2 \\ j=1, \dots, N_i}} T_{mij}, \quad (\text{A.6})$$

where  $T_{m0}$  and  $T_{mij}$  are in Eqs. (A.4) and (A.5).

Similarly, the slice  $S_m$  for  $\theta_m$  is the intersection of the sub-slices  $S_{m0}$  and  $S_{mij}$  are defined as follows. To ensure each stock prices non-negative, we first define

$$S_{m01} = \{\theta_m \in \Re^M : \theta_m > 0\}. \quad (\text{A.7})$$

Let  $\theta_{-m}$  denote the vector  $\theta$  with  $\theta_m$  removed,  $\theta_{-m} = \{\theta_1, \dots, \theta_{m-1}, \theta_{m+1}, \dots, \theta_M\}$ .

To ensure that the minimum of  $\theta$  is less than  $c_{\min}$ , we define

$$S_{m02} = \begin{cases} (0, \infty) & \text{for } \min(\theta_{-m}) < c_{\min}, \\ (0, c_{\min}) & \text{for } \min(\theta_{-m}) > c_{\min}. \end{cases} \quad (\text{A.8})$$

Similarly, to ensure that the maximum of  $\theta$  is larger than  $c_{\max}$ , we define

$$S_{m03} = \begin{cases} (0, \infty) & \text{for } \max(\theta_{-m}) > c_{\max}, \\ (c_{\max}, \infty) & \text{for } \max(\theta_{-m}) < c_{\max}. \end{cases} \quad (\text{A.9})$$

We define

$$S_{m0} = S_{m01} \cap S_{m02} \cap S_{m03}, \quad (\text{A.10})$$

where the sets  $S_{m01}$ ,  $S_{m02}$  and  $S_{m03}$  are given in Eqs. (A.7)-(A.9). Furthermore, we sample the auxiliary variable  $z_{ij} = g_{ij}(w^0, \theta^0) - e$  with  $e \sim \text{Exp}(1)$ , and define the sub-slice  $S_{mij}$  as

$$S_{mij} = \{\theta_m \in \Re : g_{ij}(w, \theta) > z_{ij}\}.$$

Recall the definitions of  $l_{ij}$  and  $u_{ij}$  in Eqs. (A.2)-(A.3) given a constant  $z$ .  $S_{mij}$  equals

$$\begin{aligned} & \left\{ \theta_m \in \Re : \sum_{m=1}^M w_m \wp_{ij}(\theta_m) \in (l_{ij}, u_{ij}) \right\} \\ &= \left\{ \theta_m \in \Re : w_m \wp_{ij}(\theta_m) \in (l_{ij} - \gamma_{ij-m}, u_{ij} - \gamma_{ij-m}) \right\}, \end{aligned}$$

where

$$\gamma_{ij-m} = \sum_{\substack{m'=1, \dots, M \\ m' \neq m}} w_{m'} \wp_{ij}(\theta_{m'}).$$

Because  $w_m \wp_{ij}(\theta_m)$  is non-negative by definition,  $S_{mij}$  can be simplified as

$$S_{mij} = \begin{cases} \left( c_j - \tilde{u}_{m1j}/w_m, \quad c_j - \tilde{l}_{m1j}/w_m \right) & \text{for } i = 1, \\ \left( c_j + \tilde{l}_{m2j}/w_m, \quad c_j + \tilde{u}_{m2j}/w_m \right) & \text{for } i = 2. \end{cases} \quad (\text{A.11})$$

where

$$\begin{aligned}\tilde{l}_{mij} &= \begin{cases} l_{ij} - \gamma_{ij-m} & \text{for } l_{ij} - \gamma_{ij-m} \geq 0, \\ -\infty & \text{for } l_{ij} - \gamma_{ij-m} < 0, \end{cases} \\ \tilde{u}_{mij} &= u_{ij} - \gamma_{ij-m}.\end{aligned}$$

Because  $S_{m0}$  and  $S_{mij}$  are open intervals containing  $\theta_m^0$ ,  $S_m$  is again an open interval containing  $\theta_m^0$ . In summary,

$$S_m = S_{m0} \bigcap_{\substack{i=1,2 \\ j=1,\dots,N_i}} S_{mij}, \quad (\text{A.12})$$

where  $S_{m0}$  and  $S_{mij}$  are given in Eqs. (A.10)-(A.11).

Appendix **B**

## Technical proofs of Parametric Greeks

### B.1 Proof of Theorem 1

*Proof.* Let  $f(\psi, x) = h(\psi, x)f(x)$ . Define  $I_j = \{x \in \Re : g_j(\psi, x) > 0\}$  for  $j \in \mathcal{B}$  and  $I = \bigcap_{j \in \mathcal{B}} I_j$ . The third assumption for  $g_j(\psi, x) \in \mathcal{G}_k$  requires  $g_j(\psi, x)$  to be strictly monotone in  $x$ . Since  $g_j(\psi, x)$  is strictly increasing in  $x$  if  $\text{sign}(\partial g_j(\psi, x)/\partial x) = 1$  and  $g_j(\psi, x)$  is strictly decreasing in  $x$  if  $\text{sign}(\partial g_j(\psi, x)/\partial x) = -1$ , we split the set  $\mathcal{B}$  into two disjoint sets  $\mathcal{B}_R$  and  $\mathcal{B}_L$  defined as follows:

$$\begin{aligned}\mathcal{B}_R &= \{j \in \mathcal{B} : \text{sign}(\partial g_j(\psi, x)/\partial x) = 1\}, \\ \mathcal{B}_L &= \{j \in \mathcal{B} : \text{sign}(\partial g_j(\psi, x)/\partial x) = -1\}.\end{aligned}$$

As a result, we have  $I_j = (\chi_j(\psi), \infty)$  for  $j \in \mathcal{B}_R$  and  $I_j = (-\infty, \chi_j(\psi))$  for  $j \in \mathcal{B}_L$ . When  $\mathcal{B}_R$  is nonempty, let  $R$  denote the index such that  $\chi_R(\psi)$  is the maximum of  $\chi_j(\psi)$  for  $j \in \mathcal{B}_R$ . When  $\mathcal{B}_L$  is nonempty, let  $L$  denote the index such that  $\chi_L(\psi)$  is the minimum of  $\chi_j(\psi)$  for  $j \in \mathcal{B}_L$ . As a result,  $I$  equals

$$\left\{ \begin{array}{ll} \text{case 1} & : I_R \cap I_L, \text{ when } \mathcal{B}_L \text{ and } \mathcal{B}_R \text{ are both nonempty and } \chi_L(\psi) > \chi_R(\psi); \\ \text{case 2} & : \phi, \text{ when } \mathcal{B}_L \text{ and } \mathcal{B}_R \text{ are both nonempty and } \chi_L(\psi) \leq \chi_R(\psi); \\ \text{case 3} & : I_R, \text{ when } \mathcal{B}_L \text{ is empty;} \\ \text{case 4} & : I_L, \text{ when } \mathcal{B}_R \text{ is empty.} \end{array} \right.$$

In case 1,  $I = I_R \cap I_L = (\chi_R(\psi), \chi_L(\psi))$  is a finite open interval. When  $l$  is neither  $R$  nor  $L$ , it is clear that  $\prod_{j \in \mathcal{B} \setminus j} \{x : g_j(\psi, x) > 0\}$  equals  $I_R \cap I_L$ , and hence

$$\left[ \prod_{j \in \mathcal{B} \setminus j} \mathbf{1}_{\{g_j(\psi, x) > 0\}} \right]_{x=\chi_l(\psi)} = 0.$$

On the other hand, when  $l$  is  $R$  or  $L$ , it is clear that  $\prod_{j \in \mathcal{B} \setminus j} \{x : g_j(\psi, x) > 0\}$  dominates  $I_R \cap I_L$ , and hence

$$\left[ \prod_{j \in \mathcal{B} \setminus j} \mathbf{1}_{\{g_j(\psi, x) > 0\}} \right]_{x=\chi_l(\psi)} = 1.$$

As a result, we only need to prove the reduced formula for Eq. (5.9),

$$\begin{aligned} & \frac{\partial}{\partial \psi} \int_{\mathfrak{R}} f(\psi, x) \mathbf{1}_{\{g_R(\psi, x) > 0\}}(x) \mathbf{1}_{\{g_L(\psi, x) > 0\}}(x) dx \\ &= \int_{\mathfrak{R}} f_\psi(\psi, x) \mathbf{1}_{\{g_R(\psi, x) > 0\}}(x) \mathbf{1}_{\{g_L(\psi, x) > 0\}}(x) dx \\ & \quad + [f(\psi, x) J_L(\psi, x)]_{x=\chi_L(\psi)} + [f(\psi, x) J_R(\psi, x)]_{x=\chi_R(\psi)}. \end{aligned} \quad (\text{B.1})$$

Note that the left-hand side of Eq. (B.1) equals

$$\begin{aligned} & \frac{\partial}{\partial \psi} \int_{\chi_R(\psi)}^{\chi_L(\psi)} f(\psi, x) dx \\ &= \int_{\chi_R(\psi)}^{\chi_L(\psi)} f_\psi(\psi, x) dx + f(\psi, \chi_L(\psi)) \frac{\partial \chi_L(\psi)}{\partial \psi} - f(\psi, \chi_R(\psi)) \frac{\partial \chi_R(\psi)}{\partial \psi} \end{aligned} \quad (\text{B.2})$$

by the Leibniz rule, where the partial differentiation of  $\chi_l(\psi)$  with respect to  $\psi$  can be calculated alternatively as

$$\frac{\partial \chi_l(\psi)}{\partial \psi} = - \left[ \frac{\partial g_j(\psi, x)/\partial \psi}{\partial g_j(\psi, x)/\partial x} \right]_{x=\chi_l(\psi)}, \quad l \in \mathcal{B}$$

by the implicit function theorem. Recall that Eq. (5.10) defines

$$J_l(\psi, x) = \text{sign} \left( \frac{\partial g_l(\psi, x)}{\partial x} \right) \frac{\partial g_l(\psi, x)/\partial \psi}{\partial g_l(\psi, x)/\partial x}, \quad l \in \mathcal{B}.$$

Now, it is clear that Eq. (B.2) equals

$$\begin{aligned}
& \int_{\chi_R(\psi)}^{\chi_L(\psi)} f_\psi(\psi, x) dx - f(\psi, \chi_L(\psi)) \frac{\partial g_L(\psi, x)/\partial\psi}{\partial g_L(\psi, x)/\partial x} + f(\psi, \chi_R(\psi)) \frac{\partial g_R(\psi, x)/\partial\psi}{\partial g_R(\psi, x)/\partial x} \\
= & \int_{\mathfrak{R}} f_\psi(\psi, x) \mathbf{1}_{\{g_R(\psi, x) > 0\}}(x) \mathbf{1}_{\{g_L(\psi, x) > 0\}}(x) dx - f(\psi, \chi_L(\psi)) \frac{\partial g_L(\psi, x)/\partial\psi}{\partial g_L(\psi, x)/\partial x} \\
& + f(\psi, \chi_R(\psi)) \frac{\partial g_R(\psi, x)/\partial\psi}{\partial g_R(\psi, x)/\partial x} \\
= & \int_{\mathfrak{R}} f_\psi(\psi, x) \mathbf{1}_{\{g_R(\psi, x) > 0\}}(x) \mathbf{1}_{\{g_L(\psi, x) > 0\}}(x) dx + [f(\psi, x) J_L(\psi, x)]_{x=\chi_L(\psi)} \\
& + [f(\psi, x) J_R(\psi, x)]_{x=\chi_R(\psi)}
\end{aligned}$$

which is exactly the right-hand side of Eq. (B.1). Hence, the proof of Eq. (5.9) for case 1 is done. The proofs for cases 2, 3, and 4 are similar and are therefore omitted.

□

## B.2 Proof of Theorem 2

*Proof.* Let  $f(\psi, \mathbf{x}) = h(\psi, \mathbf{x})f(\mathbf{x})$ . With  $\mathbf{x}_k$  fixed, Theorem 1 implies

$$\begin{aligned}
& \frac{\partial}{\partial\psi} \int_{\mathfrak{R}} f(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) d\mathbf{x}_k \\
= & \int_{\mathfrak{R}} f_\psi(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) d\mathbf{x}_k + \sum_{l \in \mathcal{B}} \left[ f(\psi, \mathbf{x}) J_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right]_{x_k=\chi_l(\psi, \mathbf{x}_k)} \quad (\text{B.3})
\end{aligned}$$

Because  $h(\psi, \mathbf{x}) \in \mathcal{H}_k$ ,  $\int_{\mathfrak{R}^n} \left| f(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right| d\mathbf{x}$  is finite, and we can exchange the order of integrals by the Fubini Theorem:

$$\begin{aligned}
& \frac{\partial}{\partial\psi} \int_{\mathfrak{R}^n} f(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) d\mathbf{x} \\
= & \frac{\partial}{\partial\psi} \int_{\mathfrak{R}^{n-1}} \left[ \int_{\mathfrak{R}} f(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) d\mathbf{x}_k \right] d\mathbf{x}_k \\
= & \lim_{h \rightarrow 0} \int_{\mathfrak{R}^{n-1}} \left[ \int_{\mathfrak{R}} \frac{f(\psi + h, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi+h, \mathbf{x}) > 0\}}(\mathbf{x}) - f(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x})}{h} d\mathbf{x}_k \right] d\mathbf{x}_k.
\end{aligned}$$

Let  $\{h_n\}$  be a sequence of numbers such that  $h_n \rightarrow 0$  as  $n \rightarrow \infty$ . Define a sequence of functions  $\{q_n(\mathbf{x}_k)\}$  by

$$q_n(\mathbf{x}_k) = \int_{\mathfrak{R}} \frac{f(\psi + h_n, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi + h_n, \mathbf{x}) > 0\}}(\mathbf{x}) - f(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x})}{h_n} dx_k,$$

and define  $q(\mathbf{x}_k)$  as the right-hand side of Eq. (B.3). Since Eq. (B.3) holds,  $q_n(\mathbf{x}_k) \rightarrow q(\mathbf{x}_k)$  pointwise almost everywhere. We only need to show there exists an integrable function,  $p(\mathbf{x}_k)$ , such that  $|q_n(\mathbf{x}_k)| < p(\mathbf{x}_k)$  for all  $n$ , then by Lebesgue's dominated convergence theorem,

$$\lim_{n \rightarrow \infty} \int_{\mathfrak{R}^{n-1}} q_n(\mathbf{x}_k) d\mathbf{x}_k = \int_{\mathfrak{R}^{n-1}} q(\mathbf{x}_k) d\mathbf{x}_k.$$

Since  $\{h_n\}$  is an arbitrary sequence with  $h_n \rightarrow 0$  as  $n \rightarrow \infty$ ,

$$\begin{aligned} & \text{LHS of Eq. (5.11)} \\ &= \lim_{h \rightarrow 0} \int_{\mathfrak{R}^{n-1}} \left[ \int_{\mathfrak{R}} \frac{f(\psi + h, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi + h, \mathbf{x}) > 0\}}(\mathbf{x}) - f(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_j(\psi, \mathbf{x}) > 0\}}(\mathbf{x})}{h} dx_k \right] d\mathbf{x}_k \\ &= \lim_{n \rightarrow \infty} \int_{\mathfrak{R}^{n-1}} q_n(\mathbf{x}_k) d\mathbf{x}_k \\ &= \int_{\mathfrak{R}^{n-1}} q(\mathbf{x}_k) d\mathbf{x}_k \\ &= \text{RHS of Eq. (5.11)}. \end{aligned}$$

We now prove that there exists an integrable function  $p(\mathbf{x}_k)$  such that  $|q_n(\mathbf{x}_k)| < p(\mathbf{x}_k)$  for all  $n$ . We only need to prove the case when  $\mathcal{B}$  contains only one member; the proofs for other cases are similar. Without loss of generality, assume  $\partial g(\psi, \mathbf{x}) / \partial \psi > 0$  and  $\partial g(\psi, \mathbf{x}) / \partial x_k > 0$ . Also assume that  $h(\psi, \mathbf{x})$  and  $h_\psi(\psi, \mathbf{x})$  are increasing in  $\psi$  and  $x_k$ . Let  $a = \chi(\psi, \mathbf{x}_k)$  and  $b_n = \chi(\psi + h_n, \mathbf{x}_k)$ . Since  $h_n \rightarrow 1$ ,  $|h_n| < 1$  for  $n$  large enough. Note that  $f(\psi, \mathbf{x})$  and  $f_\psi(\psi, \mathbf{x})$  are nondecreasing in  $\psi$  and we may assume that  $f(\psi, \mathbf{x})$  is nondecreasing in  $x_k$  for  $x_k \in [b_n, a]$  without loss of generality. For sufficiently large  $n$ ,

$$\begin{aligned} |q_n(\mathbf{x}_k)| &= \left| \int_{b_n}^a \frac{f(\psi + h_n, \mathbf{x})}{h_n} dx_k + \int_a^\infty \frac{f(\psi + h_n, \mathbf{x}) - f(\psi, \mathbf{x})}{h_n} dx_k \right| \\ &\leq \left| \int_{b_n}^a \frac{f(\psi + h_n, \mathbf{x})}{h_n} dx_k \right| + \left| \int_a^\infty \frac{f(\psi + h_n, \mathbf{x}) - f(\psi, \mathbf{x})}{h_n} dx_k \right| \\ &\leq \left| f(\psi + 1, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n) \left( \frac{a - b_n}{h_n} \right) \right| + \left| \int_a^\infty f_\psi(\psi + 1, \mathbf{x}) dx_k \right|. \end{aligned}$$

Taylor expansion gives

$$a - b_n = \frac{g_\psi(\psi, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n)}{g_{x_k}(\psi, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n)} h_n + O(h_n^2).$$

As a result, as  $n$  is large enough,

$$\left| \frac{a - b_n}{h_n} \right| \leq \left| \frac{g_\psi(\psi, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n)}{g_{x_k}(\psi, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n)} \right| + 1.$$

To bound  $q_n(\mathbf{x}_k)$ , we choose  $p(\mathbf{x}_k)$  as

$$\begin{aligned} & p(\mathbf{x}_k) \\ = & |f(\psi + 1, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n)| \left( \left| \frac{g_\psi(\psi, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n)}{g_{x_k}(\psi, x_1, \dots, x_{k-1}, a, x_{k+1}, \dots, x_n)} \right| + 1 \right) \\ & + \left| \int_{\chi(\psi, \mathbf{x}_k)}^{\infty} f_\psi(\psi + 1, \mathbf{x}) dx_k \right|, \end{aligned}$$

which is integrable by the assumptions on  $f(\psi, \mathbf{x})$  and  $g(\psi, \mathbf{x})$ .  $\square$

### B.3 Proof of Theorem 4

*Proof.* Applying Theorem 3 and noticing that  $h(\psi, \mathbf{x})|_{x_1=\chi_1(\psi, \mathbf{x}_1)} = 0$ , we obtain

$$\Delta_1 = e^{-rT} E \left[ -e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1} \mathbf{1}_{\{g(\psi, \mathbf{x})>0\}}(\mathbf{x}) \right],$$

or equivalently,  $\Delta_1 = e^{-rT} E \left[ -S_1(T) \mathbf{1}_{\{S_2(T)-S_1(T)-K>0\}}(\mathbf{S}_T) \right] / S_1$ .

To derive  $\Gamma_{11}$ , note that  $-e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1} \in \mathcal{H}_2$  with pdf  $f(\mathbf{x}; \mathbf{0}, \Sigma)$  and  $g(\psi, \mathbf{x}) \in \mathcal{G}_2$  with  $\chi(\psi, \mathbf{x}_2)$  defined in Eq. (5.19). As a result, we conclude that  $-e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1} \mathbf{1}_{\{g(\psi, \mathbf{x})>0\}} \in \mathcal{C}$ . Note that

$$\text{sign} \left( \frac{\partial g(\psi, \mathbf{x})}{\partial x_2} \right) \frac{\partial g(\psi, \mathbf{x})/\partial S_1}{\partial g(\psi, \mathbf{x})/\partial x_2} = - \frac{e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1}}{S_2 \sigma_2 \sqrt{T} e^{(r-\sigma_2^2/2)T+\sigma_2\sqrt{T}x_2}},$$

and let

$$\tilde{\eta}(\psi, \mathbf{x}_2) = \frac{f(\mathbf{x}; \mathbf{0}, \Sigma)}{f(\mathbf{x}_2; \mathbf{0}, \mathbf{I}_{n-1})} \Big|_{x_2=\chi(\psi, \mathbf{x}_2)}. \quad (\text{B.4})$$

where  $\chi(\psi, \mathbf{x}_2)$  is defined in Eq. (5.19). Now

$$\frac{e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1}}{S_2 \sigma_2 \sqrt{T} e^{(r-\sigma_2^2/2)T+\sigma_2\sqrt{T}x_2}} \Big|_{x_2=\chi(\psi, \mathbf{x}_2)} = \frac{e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1}}{S_2 \sigma_2 \sqrt{T} e^{(r-\sigma_2^2/2)T+\sigma_2\sqrt{T}\chi(\psi, \mathbf{x}_2)}}.$$

Apply Theorem 3 to  $\Delta_1$  to obtain

$$\Gamma_{11} = e^{-rT} E_q \left[ e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1} \frac{e^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}x_1}}{S_2 \sigma_2 \sqrt{T} e^{(r-\sigma_2^2/2)T+\sigma_2\sqrt{T}\chi(\psi, \mathbf{x}_2)}} \tilde{\eta}(\psi, \mathbf{x}_2) \right],$$

where  $\mathbf{x}_2 \sim N(0, 1)$  under  $E_q$ . To derive  $\Gamma_{12}$ , note that

$$\text{sign} \left( \frac{\partial g(\psi, \mathbf{x})}{\partial x_2} \right) \frac{\partial g(\psi, \mathbf{x})/\partial S_2}{\partial g(\psi, \mathbf{x})/\partial x_2} = \frac{1}{S_2 \sigma_2 \sqrt{T}}.$$

Hence

$$\Gamma_{12} = e^{-rT} E_q \left[ -e^{(r-\sigma_1^2/2)T+\sigma_1 \sqrt{T}x_1} \frac{1}{S_2 \sigma_2 \sqrt{T}} \tilde{\eta}(\psi, \mathbf{x}_2) \right].$$

□

## B.4 Proof of Theorem 5

*Proof.* To derive  $\Delta_i$ , rewrite  $\varphi(\psi, \mathbf{x})$  as

$$\varphi(\psi, \mathbf{x}) = h_i(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) + \sum_{l \in \mathcal{B} \setminus i} h_l(\psi, \mathbf{x}) \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}). \quad (\text{B.5})$$

Note that  $g_{ij}(\psi, \mathbf{x}) \in \mathcal{G}_i$ ,  $g_{li}(\psi, \mathbf{x}) \in \mathcal{G}_i$ ,

$$\text{sign} \left( \frac{\partial g_{ij}(\psi, \mathbf{x})}{\partial x_i} \right) \frac{\partial g_{ij}(\psi, \mathbf{x})/\partial S_i}{\partial g_{ij}(\psi, \mathbf{x})/\partial x_i} = \frac{1}{S_i \sigma_i \sqrt{T}} \quad \text{for } i, j \in \mathcal{B}, \quad (\text{B.6})$$

$$\text{sign} \left( \frac{\partial g_{li}(\psi, \mathbf{x})}{\partial x_i} \right) \frac{\partial g_{li}(\psi, \mathbf{x})/\partial S_i}{\partial g_{li}(\psi, \mathbf{x})/\partial x_i} = -\frac{1}{S_i \sigma_i \sqrt{T}} \quad \text{for } l \in \mathcal{B} \setminus i. \quad (\text{B.7})$$

Let

$$\eta_i(\psi, \mathbf{x}) = f(\mathbf{x}; \mathbf{0}, \boldsymbol{\Sigma})/f(\mathbf{x}_i; \mathbf{0}, \mathbf{I}_{n-1}) \quad \text{for } i \in \mathcal{B}. \quad (\text{B.8})$$

Apply Theorem 3 to obtain

$$\begin{aligned} \Delta_i = & e^{-rT} \left\{ E \left[ e^{(r-\sigma_i^2/2)T+\sigma_i \sqrt{T}x_i} \prod_{j \in \mathcal{B}} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right] \right. \\ & + E_q \left[ h_i(\psi, \mathbf{x}) \frac{1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{ii}(\psi, \mathbf{x}_i)} \\ & + \sum_{l \in \mathcal{B} \setminus i} E_q \left[ h_i(\psi, \mathbf{x}) \frac{1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \\ & \left. + \sum_{l \in \mathcal{B} \setminus i} E_q \left[ h_l(\psi, \mathbf{x}) \frac{-1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \right\} \end{aligned}$$

$$= \mathrm{e}^{-rT} \{(\mathbf{I}) + (\mathbf{II}) + (\mathbf{III}) + (\mathbf{IV})\},$$

where **(I)**, **(II)** and **(III)** result from differentiating the first term in Eq. (B.5) and **(IV)** results from differentiating the second term in Eq. (B.5). Notice that  $h_i(\psi, \mathbf{x})|_{x_1=\chi_{ii}(\psi, \mathbf{x}_1)} = 0$ , so **(II)** = 0. For **(III)**,

$$h_i(\psi, \mathbf{x})|_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} = S \mathrm{e}^{(r-\sigma_l^2/2)T+\sigma_l\sqrt{T}x_l} - K,$$

and

$$\begin{aligned} & \left. \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right|_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \\ &= \left( \prod_{j \in \mathcal{B} \setminus \{l, i\}} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \Big|_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \right) \left( \mathbf{1}_{\{g_{ii}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \Big|_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \right) \\ &= \left( \prod_{j \in \mathcal{B} \setminus \{l, i\}} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right) \mathbf{1}_{\{g_{ll}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \\ &= \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}). \end{aligned} \quad (\text{B.9})$$

Therefore, for each term in **(III)**,

$$\begin{aligned} & E_q \left[ h_i(\psi, \mathbf{x}) \frac{1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \\ &= E_q \left[ \left( S_l \mathrm{e}^{(r-\sigma_l^2/2)T+\sigma_l\sqrt{T}x_l} - K \right) \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \tilde{\eta}_{il}(\psi, \mathbf{x}_i) \right] / \left( S_i \sigma_i \sqrt{T} \right), \end{aligned}$$

where

$$\tilde{\eta}_{il}(\psi, \mathbf{x}_i) = \eta_i(\psi, \mathbf{x})|_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \quad \text{for } i, l \in \mathcal{B}. \quad (\text{B.10})$$

Similarly, for each term in **(IV)**,

$$\begin{aligned} & E_q \left[ h_l(\psi, \mathbf{x}) \frac{-1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{ii}(\psi, \mathbf{x}_i)} \\ &= -E_q \left[ \left( S_l \mathrm{e}^{(r-\sigma_l^2/2)T+\sigma_l\sqrt{T}x_l} - K \right) \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \tilde{\eta}_{il}(\psi, \mathbf{x}_i) \right] / \left( S_i \sigma_i \sqrt{T} \right). \end{aligned}$$

As a result, **(III)** and **(IV)** sum to zero. Apply Theorem 3 to  $\Delta_i$  to obtain

$$\Gamma_{ii}$$

$$\begin{aligned}
&= e^{-rT} \left\{ E_q \left[ e^{(r-\sigma_i^2/2)T+\sigma_i\sqrt{T}x_i} \frac{1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{ii}(\psi, \mathbf{x}_i)} \right. \\
&\quad \left. + \sum_{l \in \mathcal{B} \setminus i} E_q \left[ e^{(r-\sigma_i^2/2)T+\sigma_i\sqrt{T}x_i} \frac{1}{S_l \sigma_l \sqrt{T}} \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} \right\}.
\end{aligned}$$

Note that  $g_{ij}(\psi, \mathbf{x})|_{x_i=\chi_{ii}(\psi, \mathbf{x}_i)} = -g_{jj}(\psi, \mathbf{x})$ ,  $g_{ij}(\psi, \mathbf{x})|_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} = g_{lj}(\psi, \mathbf{x})$ , and

$$\begin{aligned}
\left. \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right|_{x_i=\chi_{ii}(\psi, \mathbf{x}_i)} &= \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{-g_{jj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}), \\
\left. \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \right|_{x_i=\chi_{il}(\psi, \mathbf{x}_i)} &= \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}).
\end{aligned} \tag{B.11}$$

As a result,

$$\begin{aligned}
&\Gamma_{ii} \\
&= e^{-rT} \left\{ E_q \left[ \frac{K}{S_i} \frac{1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus i} \mathbf{1}_{\{-g_{jj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \tilde{\eta}_{ii}(\psi, \mathbf{x}_i) \right] \right. \\
&\quad \left. + \sum_{l \in \mathcal{B} \setminus i} E_q \left[ \frac{S_l e^{(r-\sigma_l^2/2)T+\sigma_l\sqrt{T}x_l}}{S_i} \frac{1}{S_i \sigma_i \sqrt{T}} \prod_{j \in \mathcal{B} \setminus l} \mathbf{1}_{\{g_{lj}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \tilde{\eta}_{il}(\psi, \mathbf{x}_i) \right] \right\}.
\end{aligned}$$

To differentiate  $\Delta_i$  with respect to  $S_j$ , note that only  $\mathbf{1}_{\{g_{ij}(\psi, \mathbf{x}) > 0\}}(\psi, \mathbf{x})$  depends on  $S_j$ . Since  $g_{ij}(\psi, \mathbf{x}) \in \mathcal{G}_i$  and

$$\text{sign} \left( \frac{\partial g_{ij}(\psi, \mathbf{x})}{\partial x_i} \right) \frac{\partial g_{ij}(\psi, \mathbf{x}) / \partial S_j}{\partial g_{ij}(\psi, \mathbf{x}) / \partial x_i} = \frac{-e^{(r-\sigma_j^2/2)T+\sigma_j\sqrt{T}x_j}}{S_i \sigma_i \sqrt{T} e^{(r-\sigma_i^2/2)T+\sigma_i\sqrt{T}x_i}},$$

apply Theorem 3 to obtain

$$\begin{aligned}
&\Gamma_{ij} \\
&= e^{-rT} E_q \left[ e^{(r-\sigma_i^2/2)T+\sigma_i\sqrt{T}x_i} \frac{-e^{(r-\sigma_j^2/2)T+\sigma_j\sqrt{T}x_j}}{S_i \sigma_i \sqrt{T} e^{(r-\sigma_i^2/2)T+\sigma_i\sqrt{T}x_i}} \prod_{l \in \mathcal{B} \setminus j} \mathbf{1}_{\{g_{il}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \eta_i(\psi, \mathbf{x}) \right]_{x_i=\chi_{ij}(\psi, \mathbf{x}_j)} \\
&= e^{-rT} E_q \left[ \frac{-e^{(r-\sigma_j^2/2)T+\sigma_j\sqrt{T}x_j}}{S_i \sigma_i \sqrt{T}} \prod_{l \in \mathcal{B} \setminus j} \mathbf{1}_{\{g_{jl}(\psi, \mathbf{x}) > 0\}}(\mathbf{x}) \tilde{\eta}_{ij}(\psi, \mathbf{x}_i) \right].
\end{aligned}$$

□

## B.5 Greeks Using the Likelihood Ratio Method

The relationship between  $\mathbf{S}_T$  and  $\mathbf{x}$  are in Eq. (5.2). Let  $\mathbf{s} = (s_1, \dots, s_n)^t$  and  $f_{\mathbf{S}_T}(\mathbf{s})$  be the probability density function of  $\mathbf{S}_T$ . Define

$$d_i(s_i) = \frac{\log(s_i/S_i) - (r - \sigma_i^2/2)T}{\sigma_i \sqrt{T}} \quad \text{for } i = 1, \dots, n,$$

and  $\mathbf{d}(\mathbf{s}) = (d_1(s_1), \dots, d_n(s_n))^t$ . Then  $\frac{\partial d_i(s_i)}{\partial S_i} = -1/(S_i \sigma_i \sqrt{T})$ , and

$$f_{\mathbf{S}_T}(\mathbf{s}) = f(\mathbf{d}(\mathbf{s}); \mathbf{0}, \boldsymbol{\Sigma}) \left( \prod_{i=1}^n \frac{1}{s_i \sigma_i \sqrt{T}} \right). \quad (\text{B.12})$$

Recall that  $\mathbf{v}_i$  is an  $n \times 1$  vector having 1 in the  $i$ -th component and 0 elsewhere. By straightforward calculations,

$$\begin{aligned} \frac{\partial}{\partial S_i} f_{\mathbf{S}_T}(\mathbf{s}) &= f_{\mathbf{S}_T}(\mathbf{s}) \left[ \frac{\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{s})}{S_i \sigma_i \sqrt{T}} \right], \\ \frac{\partial^2}{\partial S_i^2} f_{\mathbf{S}_T}(\mathbf{s}) &= f_{\mathbf{S}_T}(\mathbf{s}) \left[ \frac{(\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{s}))^2 - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{v}_i - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{s}) \sigma_i \sqrt{T}}{S_i^2 \sigma_i^2 T} \right], \\ \frac{\partial^2}{\partial S_i \partial S_j} f_{\mathbf{S}_T}(\mathbf{s}) &= f_{\mathbf{S}_T}(\mathbf{s}) \left[ \frac{(\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{s})) (\mathbf{v}_j^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{s})) - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{v}_j}{S_i S_j \sigma_i \sigma_j T} \right]. \end{aligned}$$

Greeks using the likelihood ratio method are (Broadie and Glasserman, 1996)

$$\begin{aligned} \Delta_i &= e^{-rT} E \left[ \wp(\mathbf{S}_T) \frac{\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{S}_T)}{S_i \sigma_i \sqrt{T}} \right], \\ \Gamma_{ii} &= e^{-rT} E \left[ \wp(\mathbf{S}_T) \frac{(\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{S}_T))^2 - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{v}_i - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{S}_T) \sigma_i \sqrt{T}}{S_i^2 \sigma_i^2 T} \right], \\ \Gamma_{ij} &= e^{-rT} E \left[ \wp(\mathbf{S}_T) \frac{(\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{S}_T)) (\mathbf{v}_j^t \boldsymbol{\Sigma}^{-1} \mathbf{d}(\mathbf{S}_T)) - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{v}_j}{S_i S_j \sigma_i \sigma_j T} \right], \quad i \neq j, \end{aligned}$$

where the expectation is taken over the random variable  $\mathbf{S}_T$  whose probability density function is in Eq. (B.12). Or equivalently after change of variables,

$$\begin{aligned} \Delta_i &= e^{-rT} E \left[ \wp(\mathbf{S}_T) \frac{\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{x}}{S_i \sigma_i \sqrt{T}} \right], \\ \Gamma_{ii} &= e^{-rT} E \left[ \wp(\mathbf{S}_T) \frac{(\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{x})^2 - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{v}_i - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{x} \sigma_i \sqrt{T}}{S_i^2 \sigma_i^2 T} \right], \\ \Gamma_{ij} &= e^{-rT} E \left[ \wp(\mathbf{S}_T) \frac{(\mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{x}) (\mathbf{v}_j^t \boldsymbol{\Sigma}^{-1} \mathbf{x}) - \mathbf{v}_i^t \boldsymbol{\Sigma}^{-1} \mathbf{v}_j}{S_i S_j \sigma_i \sigma_j T} \right], \quad i \neq j, \end{aligned}$$

where the expectation is taken over the random variable  $\mathbf{x}$ . Recall that the formula connecting  $\mathbf{S}_{\mathbf{T}}$  and  $\mathbf{x}$  is in Eq. (5.2).

# Bibliography

- Abadir, K. and M. Rockinger (2003). Density functionals, with an option-pricing application. *Econometric Theory* 19, 778–811.
- Abken, P., D. Madan, and S. Ramamurtie (1996). Estimation of risk-neutral and statistical densities by Hermite polynomial approximation.
- Aït-Sahalia, Y. and J. Duarte (2003). Nonparametric option pricing under shape restrictions. *J. Econometrics* 116, 9–47.
- Aït-Sahalia, Y. and A. Lo (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance* 53, 499–547.
- Alcock, J. and D. Auerswald (2009). Empirical tests of canonical nonparametric American option pricing methods. *Journal of Futures Markets* 30, 509–532.
- Alcock, J. and T. Carmichael (2008). Nonparametric American option pricing. *Journal of Futures Markets* 28, 717–748.
- Amin, K. and V. Ng (1993). Option valuation with systematic stochastic volatility. *Journal of Finance* 48, 881–910.
- Bakshi, G. and Z. Chen (1997). An alternative valuation model for contingent claims. *Journal of Financial Economics* 44, 123–165.
- Barles, G., J. Burdeau, M. Romano, and N. Samsoen (1995). Critical stock price near expiration. *Mathematical Finance* 5, 77–95.
- Bates, D. (1996a). Dollar jump fears, 1984-1992: Distributional abnormalities implicit in currency futures options. *Journal of International Money and Finance* 15, 65–93.
- Bates, D. (1996b). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies* 9, 69–107.

- Bates, D. (2000). Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics* 94, 181–238.
- Benhamou, E. (2003). Optimal malliavin weighting function for the computation of the Greeks. *Mathematical Finance* 13, 37–53.
- Birke, M. and K. Pilz (2009). Nonparametric option pricing with no-arbitrage constraints. *Journal of Financial Econometrics* 7, 53–76.
- Boyle, P. (1977). Options: A Monte Carlo approach. *Journal of Financial Economics* 4, 323–338.
- Boyle, P., J. Evnine, and S. Gibbs (1989). Numerical evaluation of multivariate contingent claims. *Review of Financial Studies* 2, 241–250.
- Breeden, D. and R. Litzenberger (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business* 51, 621–651.
- Broadie, M., J. Detemple, E. Ghysels, and O. Torres (2000). Nonparametric estimation of American options' exercise boundaries and call prices. *Journal of Economic Dynamics and Control* 24, 1829–1857.
- Broadie, M. and P. Glasserman (1996). Estimating security price derivatives using simulation. *Management Science* 42, 269–285.
- Brown, G. and K. Toft (1999). Constructing binomial trees from multiple implied probability distributions. *Journal of derivatives* 7, 83–100.
- Büttler, H.-J. and J. Waldvogel (2006). Pricing callable bonds by means of Greens function. *Mathematical Finance* 6, 53–88.
- Campa, J., P. Chang, and R. Reider (1998). Implied exchange rate distributions: Evidence from OTC option markets. *Journal of International Money and Finance* 17, 117–160.
- Carmona, R. and V. Durrleman (2003). Pricing and hedging spread options. *SIAM Review* 45, 627–685.
- Corrado, C. and T. Su (1996). S&P 500 index option tests of Jarrow and Rudd's approximate option valuation formula. *Journal of Futures Markets* 6, 611–629.
- Cox, J., J. Ingersoll, and S. Ross (1985). An intertemporal general equilibrium model of asset prices. *Econometrica* 53, 363–384.
- Cox, J. and S. Ross (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics* 3, 145–166.

- Cox, J., S. Ross, and M. Rubinstein (1979). Option pricing: A simplified approach. *Journal of Financial Economics* 7, 229–263.
- Damien, P., J. Wakefield, and S. Walker (1999). Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 61, 331–344.
- Dong, Q. (2001). Basket implied volatility surface. Derivatives Week 4 June.
- Dong, Q. (2005). Pricing basket options with skew. Wilmott Magazine July.
- Duffie, D. and M. Huang (1996). Swap rates and credit quality. *Journal of Finance* 51, 921–949.
- Duffie, D., J. Pan, and K. Singleton (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68, 1343–1376.
- Duffie, D. and K. Singleton (2003). *Credit risk: pricing, measurement, and management* (1st ed.). Princeton, NJ: Princeton University Press.
- Edwards, R. and A. Sokal (1988). Generalization of the Fortuin-Kasteleyn-Swendsen-Wang representation and Monte Carlo algorithm. *Physical Review D* 38, 2009–2012.
- Engle, R. and G. González-Rivera (1991). Semiparametric ARCH models. *Journal of Business and Economic Statistics* 9, 345–359.
- Eraker, B., M. Johannes, and N. Polson (2003). The impact of jumps in volatility and returns. *Journal of Finance* 58, 1269–1300.
- Fan, J. and L. Mancini (2009). Option pricing with model-guided nonparametric methods. *Journal of American Statistical Association* 104, 1351–1372.
- Fengler, M. (2009). Arbitrage-free smoothing of the implied volatility surface. *Quantitative Finance* 9, 417–428.
- Fournie, E., J. Lasry, J. Lebuchoux, P. Lions, and N. Touzi (1999). Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics* 3, 391–412.
- Fu, M. and J.-Q. Hu (1997). *Conditional Monte Carlo: Gradient Estimation and Optimization Applications*. Norwell, MA: Kluwer.
- Giacomini, R., A. Gottschling, C. Haefke, and H. White (2008). Mixtures of  $t$ -distributions for finance and forecasting. *Journal of Econometrics* 144, 175–192.
- Glasserman, P. (2004). *Monte Carlo Methods in Financial Engineering*. NY: Springer-Verlag.

- Härdle, W. and A. Yatchew (2002). Dynamic nonparametric state price density estimation using constrained least squares and the bootstrap. Discussion paper 16, Quantification and Simulation of Economics Processes, Humboldt-Universitaet zu Berlin.
- Harrison, J. and D. Kreps (1979). Martingale and arbitrage in multiperiods securities markets. *Journal of Economic Theory* 20, 381–408.
- Heston, S. (1993). A closed-form solutions for options with stochastic volatiltiy with applications to bond and currency options. *Review of Financial Studies* 6, 327–343.
- Hörfelt, P. (2008). A short cut to the rainbow. *Risk* 21, 90–93.
- Hull, J. (2002). *Options, Futures, and Other Derivatives*. Upper Saddle River, NJ: Prentice Hall.
- Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42, 281–300.
- Hutchinson, J., A. Lo, and T. Poggio (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance* 49, 771–818.
- Jäckel, P. (2002). *Monte Carlo Methods in Finance*. West Sussex, UK: Wiley.
- Jackwerth, J. (1999). Option-implied risk-neutral distributions and implied binomial trees: A literature review. *Journal of Derivatives* 7, 66–82.
- Jackwerth, J. and M. Rubinstein (1996). Recovering probability distributions from option prices. *Journal of Finance* 51, 1611–1631.
- Jarrow, R. and A. Rudd (1982). Approximate option valuation for arbitrary stochastic process. *Journal of Financial Economics* 10, 347–369.
- Johnson, H. (1987). Options on the maximum or the minimun of several assets. *Journal of Financial and Quantitative Analysis* 22, 277–283.
- Jondeau, E. and M. Rockinger (2001). Gram-Charlier densities. *Journal of Economic Dynamics and Control* 25, 1457–1483.
- Kacperczyk, M., P. Damien, and S. Walker (2005). A new class of Bayesian semiparametric models with applications to option pricing. McCombs Research Paper Series No. IROM-08-05.
- Kass, R. and A. Raftery (1995). Bayesian factors. *Journal of the American Statistical Association* 90, 773–795.

- Kirk, E. (1995). Correlation in the energy markets. In R. Jameson (Ed.), *Managing Energy Price Risk*, pp. 71–78. London: Risk Publications and Enron.
- Knight, J. and S. Satchell (1997). Pricing derivatives written on assets with arbitrary skewness and kurtosis. Mimeo, Trinity College, University of Cambridge.
- Kuske, R. and J. B. Keller (1998). Optimal exercise boundary for an American put option. *Applied Mathematical Finance* 5, 107–116.
- Lipton, A. and A. Rennie (2007). Credit correlation: Life after copulas (ed.). World Scientific.
- Liu, G. and L. Hong (2008). Pathwise estimation of the greeks of financial options. Working Paper.
- Liu, J. (2001). *Monte Carlo Strategies in Sciетific Computing*. NY: Springer-Verlag.
- Liu, Q. (2008). Pricing American options by canonical least-squares Monte Carlo. Available at SSRN: <http://ssrn.com/abstract=1145331>.
- Longstaff, F. (1995). Option pricing and the martingale restriction. *Review of Financial Studies* 8, 1091–1124.
- Longstaff, F. and E. Schwartz (2001). Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies* 14, 113–147.
- Lyuu, Y.-D. and H.-W. Teng (2010). Unbiased and efficient Greeks of financial options. *Finance and Stochastics*, to appear.
- Madan, D., P. Carr, and E. Chang (1998). The variance gamma process and option pricing. *European Finance Review* 2, 79–105.
- Madan, D. and F. Milne (1994). Contingent claims valued and hedged by pricing and investing in a basis. *Mathematical Finance* 4, 223–245.
- Maheu, J. and T. McCurdy (2004). News arrival, jumps dynamics, and volatility components for individual stock returns. *Journal of Finance* 59, 755–793.
- Malz, A. (1996). Using option prices to estimate realignment probabilities in the European Monetary System: The case of sterling-mark. *Journal of International Money and Finance* 15, 717–748.
- Malz, A. (1997). Estimating the probability distribution of the future exchange rate from option prices. *Journal of Derivatives* 2, 18–36.
- Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance* 33, 177–186.

- Mayhew, S. (1995). On estimating the risk-neutral probability distribution implied by option prices. Working Paper, Purdue University.
- McKean, H. (1965). Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics. *Industrial Management Review* 6, 32–39.
- Melick, W. and C. Thomas (1997). Recovering an asset's implied PDF from option prices: An application to crude oil during the Gulf crisis. *Journal of Financial and Quantitative Analysis* 32, 91–115.
- Meneguzzo, D. and W. Vecchiato (2003). Copula sensitivity in collateralized debt obligations and basket default swaps. *Journal of Futures Markets* 24, 37–70.**
- Moerberke, P. (1974). On optimal stopping and free boundary problems. *Journal of Mathematics* 4, 539–578.
- Nandi, S. (1998). How important is the correlation between returns and volatility in a stochastic volatility model? Empirical evidence from pricing and hedging in the S&P 500 index options market. *Journal of Banking and Finance* 22, 589–610.
- Nelken, I. (1996). *The Handbook of Exotic Options: Instruments, Analysis, and Application*. NY: McGraw-Hill.
- Nelsen, R. (1999). *An Introduction to Copulas*. New York: Springer.
- Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63, 3–50.
- Pearson, N. (1995). An efficient approach for pricing spread options. *Journal of Derivatives* 3, 76–91.
- Protter, M. and C. Morrey (1991). *A First Course in Real Analysis*. New York: Springer-Verlag.
- Rubinstein, M. (1991). Somewhere over the rainbow. *Risk* 4, 63–66.
- Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance* 69, 771–818.
- Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *Journal of Finance* 44, 211–219.
- Scott, L. (1997). Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Application of Fourier inversion methods. *Mathematical Finance* 7, 413–426.

- Shimko, D. (1993). Bounds on probability. *RISK* 6, 33–37.
- Söderlind, P. and L. Svensson (1997). New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics* 40, 383–429.
- Stensoft, L. (2004). Convergence of the least squares Monte-Carlo approach to American option valuation. *Management Science* 50, 1193–1203.
- Strong, N. and X. Xu (1999). Do S & P 500 index options violate the martingale restriction? *Journal of Futures Markets* 19, 499–521.
- Stulz, R. (1982). Options on the minimum or the maximum of two risky assets. *Journal of Financial Economics* 10, 161–185.
- Stutzer, M. (1996). A simple nonparametric approach to derivative security valuation. *Journal of Finance* 51, 1633–1652.
- Tavella, D. and C. Randall (2009). *Pricing Financial Instruments: The Finite Difference Method*. NY: John Wiley & Sons.
- Wystup, U. (2006). *Foreign Exchange Options and Structured Products*. Hoboken, NJ: Wiley.
- Yuan, M. (2009). State price density estimation via nonparametric mixtures. *Annals of Applied Statistics* 3, 963–984.
- Zazanis, M. and R. Suri (1993). Convergence rates of finite-difference sensitivities estimates for stochastic systems. *Operations Research* 41, 694–703.

# Vita

Huei-Wen Teng

## RESEARCH INTERESTS

- Statistics: Bayesian analysis, modeling, nonparametric methods, Markov chain Monte Carlo simulation
- Finance: Asset pricing, risk management, credit derivatives, the Greeks

## EDUCATION

- Pennsylvania State University, University Park, PA  
Ph.D. in Statistics Aug 2010  
M.S. in Statistics May 2009
- National Taiwan University, Taiwan  
M.B.A. in Finance (Major in Financial Engineering) Jun 2004  
B.S. in Mathematics Jun 2001
- Linz University, Austria, Exchange Student Sep 1999–Jan 2000

## PUBLICATIONS

1. Y.-D. Lyuu and H.-W. Teng (2008). Unbiased and Efficient Greeks of Rainbow Options, *Finance and Stochastics*, to appear.
2. C.-D. Fuh, H.-W. Teng and R.-H. Wang (2008). On-Line VWAP Trading Strategies, *Sequential Analysis*, to appear.
3. T.-C. Yang, H.-W. Teng, and M. Haran (2009). The Impacts of Social Capital on Infant Mortality in The U.S.: A Spatial Investigation, *Applied Spatial Analysis and Policy*, 2, 3, 211–227.

## PRESENTATIONS

1. ENAR Spring Meetings, New Orleans, LA, 2010.
2. The Midwest Finance Association Meeting, Las Vegas, NV, 2010.
3. Joint Statistical Meetings, Washington DC, 2009.
4. Conference on Statistical Models and Methods in Quantitative Finance and Related Topics, Academia Sinica, Taiwan, 2009.
5. The Eighth International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, Montreal, Canada, 2008.
6. The Midwest Finance Association Meeting, San Antonio, TX, 2008.

## TEACHING EXPERIENCES

- Math/Stat 418, Probability (evaluation: 5.85 out of 7) Summer 2008
- Math/Stat 418, Probability (evaluation: 5.25 out of 7) Spring 2008