Financial Methods for Online Advertising

Bowei Chen

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Computer Science

University College London

16th January 2015

I, Bowei Chen, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

"Whenever a theory appears to you as the only possible one, take this as a sign that you have neither understood the theory nor the problem which it was intended to solve."

– Sir Karl Raimund Popper

Objective Knowledge: An Evolutionary Approach (1972)

Abstract

Online advertising, a form of advertising that reaches consumers through the World Wide Web, has become a multi-billion dollar industry. Using the state of the art computing technologies, online auctions have become an important sales mechanism for automating transactions in online advertising markets, where advertisement (shortly ad) inventories, such as impressions or clicks, are able to be auctioned off in milliseconds after they are generated by online users. However, with providing non-guaranteed deliveries, the current auction mechanisms have a number of limitations including: the uncertainty in the winning payment prices for buyers; the volatility in the seller's revenue; and the weak loyalty between buyer and seller. To address these issues, this thesis explores the methods and techniques from finance to evaluate and allocate ad inventories over time and to design new sales models. Finance, as a sub-field of microeconomics, studies how individuals and organisations make decisions regarding the allocation of resources over time as well as the handling of risk. Therefore, we believe that financial methods can be used to provide novel solutions to the non-guaranteed delivery problem in online advertising. This thesis has three major contributions. We first study an optimal dynamic model for unifying programmatic guarantee and real-time bidding in display advertising. This study solves the problem of algorithmic pricing and allocation of guaranteed contracts. We then propose a multi-keyword multi-click ad option. This work discusses a flexible way of guaranteed deliveries in the sponsored search context, and it's evaluation is under the no arbitrage principle and is based on the assumption that the underlying winning payment prices of candidate keywords for specific positions follow a geometric Brownian motion. However, according to our data analysis and other previous research, the same underlying assumption is not valid empirically for display ads. We therefore study a lattice framework to price an ad option based on a stochastic volatility underlying model. This research extends the usage of ad options to display advertising in a more general situation.

Acknowledgements

I have been very lucky to learn from a number of outstanding researchers at University College London (UCL) during my PhD study. First of all, I wish to express my greatest thanks to my first supervisor, Jun Wang, whose excellent guidance has been absolutely essential for the completion of this thesis. I would like to thank him for introducing me into the field of online advertising, and his consistent encouragement and support. Over the years, I have learned many things from him: high quality research, effective communication skills and hard-working attitude. I can not ask for more and the high standards he has imparted to me will benefit my research efforts my whole life long.

I am grateful to Ingemar Cox for being my second supervisor and his constructive technical guidance on my research. I would like to thank Mohan Kankanhalli for many enjoyable discussions during his visit at UCL, and his support in my research and job search. I would like to thank Sofia Ceppi for supervising me at Microsoft Research Cambridge and her advice on algorithmic game theory and mechanism design. I would like to thank Philip Treleven and Dell Zhang for being my thesis examiners and their useful comments. I also appreciate Philip's guidance and support on my internship and job search. I would like to thank Jinqiu Hua for recommending me to UCL and his kind advice on talks, research, and everything in between. I am grateful to all the faculty, staff, and students at UCL who have helped me in ways large and small. I appreciate the funding supports from the Faculty of Engineering (Dean's Prize), the UCL Advances (PhD Enterprise Scholarship) and the UCL BITE (Secondment Exchange Award).

I would like to express my deep gratitude to my family and friends, for their continuous love, support, and encouragement. Special thanks go to my wife Xin Wang, who has always been there right beside me through the most difficult periods of my PhD journey. Without her many sacrifices, I would not have made it this far. Finally, I would love to dedicate this thesis to my baby boy Yanchen Chen, who has brought so much happiness and joy to our lives.

Contents

List of Figures

List of Tables

Notation

The following mathematical notations are used throughout the thesis. For the reader's convenience, in addition to their definitions here, they are also described in their first occurrences in each chapter. Other notations used in Chapters 3-5 are further described in Tables 3.1, 4.1 and 5.1, respectively.

Notation	Description
$\mathbf{BIN}(n, p)$	Binomial distribution for n trials with probability $p \in [0, 1]$, $n \in \mathbb{Z}_*$
$\mathcal{A},\ \mathcal{F},\ \mathcal{M}$	Collection of sets or σ-fields
$\text{cov}[X, Y],\ \mathbf{\Sigma}$	Covariance, correlation matrix
$f(x)$ or $f_X(x)$	Density function
$\lvert\mathbf{\Sigma}\rvert$	Determinant of $\mathbf{\Sigma}$
$\emptyset$	Empty set
$\mathbb{E}[\cdot],\ \mathbb{E}[\cdot \mid \mathcal{F}]$	Expectation, conditional expectation on $\mathcal{F}$
$\exp\{\cdot\}$	Exponential function
$f_{X,Y}(x, y)$	Joint density function
$\mathbf{LN}(\mu, \sigma^2)$	Log-normal distribution with mean μ and variance σ^2, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}$
$\max\{\cdot\}$	Maximum function
$\ln\{\cdot\}$	Natural logarithm function
$\mathbf{N}(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}$
$\mathcal{N}(x)$	Cumulative distribution function of a standard normal, $x \in \mathbb{R}$
$\mathbb{P},\ \mathbb{Q}$	Probability measure
$X,\ \mathbf{X}$	Random variables, vector or matrix
$\mathbb{Z},\ \mathbb{Z}_*$	Set of integers $\{\ldots, -1, 0, 1, \ldots\}$, non-negatives integers $\{0, 1, \ldots\}$
$\mathbb{R},\ \mathbb{R}_+$	Set of real numbers $(-\infty, \infty)$, non-negative real numbers $[0, \infty)$
Ω	Sample space
$W(t),\ \mathbf{W}(t)$	Standard Brownian motion
$\sigma[\cdot],\ \text{var}[\cdot]$	Standard deviation, variance
$\Sigma,\ \Pi$	Sum, product
$\mathbf{s.t.}$	Subject to
$\mathbf{X^T}$	Transpose of $\mathbf{X}$
$\mathbf{U}[a, b]$	Uniform distribution on closed interval $[a, b]$, $a \in \mathbb{R}, b \in \mathbb{R}$

Chapter 1

Introduction

The rapid growth of the Internet and the World Wide Web has transformed the way that information is being accessed and used. They have also transformed the business of advertising. Traditional advertising reaches consumers through televisions, radios, newspapers, magazines, etc. Today, many marketing messages are delivered through the Internet and the World Wide Web to online users [Evans, 2009].

By most accounts, the very first online ad was an email sent by Gary Thuerk on 3rd May 1978, who was a marketing manager at the Digital Equipment Corporation (DEC) and also known as the father of spam [Templeton, 2008]. The recipient list was about 400 ARPANET[1] users on the west coast of the United States. This email was an invitation to users to the demonstrations of the DEC's new product. Although some users were happy about the notification, the majority felt annoyed. Despite the generally negative reactions at the beginning, online advertising grew rapidly.

Online advertising has now become a multi-billion dollar industry and a significant source of revenue for many Web based businesses, such as Google, Facebook, Yahoo! and AppNexus. According to the Interactive Advertising Bureau (IAB)[2], the US-only online advertising annual revenues for 2013 totalled \$42.8 billion, \$6.2 billion (or 17.0%) higher than in 2012. In Europe, the region's online ads spending rose by 11.9% in 2013 to total €27.3 billion (equivalent to \$36.4 billion) and the major Western European countries were with 5%-18% annual growth rates respectively[3]. The

[1] The Advanced Research Projects Agency Network (ARPANET) was one of the world's first operational packet switching networks, the first network to implement TCP/IP, was then decommissioned on 28th February 1990.

[2] http://www.iab.net

[3] http://www.emarketer.com

UK has the biggest digital market size in Western Europe[4], whose online advertising market grew 15.6% in 2013, to £6.3 billion (equivalent to $9.23 billion), and has been estimated to account for more than 1/3 of the total revenue of Western Europe in 2014. The above numbers reveal the fact that online advertising has become one of most fast-growing industries, which also means there is emerging demand of the high-quality research in this field.

In fact, online advertising are becoming a new scientific sub-discipline in computer science, bringing the gap among the areas of information systems, data mining, artificial intelligence, machine learning, economics, marketing science, operations research, etc. It arises many interdisciplinary challenges, such as advertising auction mechanism design, re-targeting models, programmatic bidding strategies and risk-aware advertising technologies. Good and timely solutions to those challenges can feed back into a better product or market design that will generate more economic value and produce more benefits to people and businesses associated with it.

1.1 Major Types of Online Ads

It is difficult to use the Internet without seeing online ads. They can be found in almost all types of Web pages, ranging from an online newspaper, to a search engine results page (SERP), to a Facebook homepage, etc.

Display advertising is one of the most popular types of online advertising. It is graphical information that appears next to content on Web pages, pop-up videos, emails, etc. These ads, often referred to as banners, come in standardized ad sizes, and can include text, logos, pictures, or more recently, rich media [IAB, 2013, Jansen, 2011]. Figure 1.1 presents an example of display ads. On the Yahoo! Cars Web page, there are three ad slots (shown in dotted blue line boxes) and Yahoo! can display these three ads together at a same time if a user visits the site. Therefore, display ads are usually sold on the basis of 1000 views of display. Each display is called an *impression* and the cost for 1000 impressions is called the *cost per mille (CPM)*.

Sponsored search, also called search advertising, is another popular type of advertising. As the name implies, the advertising is triggered by a search behaviour [Jansen, 2011]. In Figure 1.2, within the search box is one term that a user submitted to the

[4] http://www.iabuk .net

search engine. The term is collectively known as the *query*. Along with the click on the search button, the query term is what triggers the results on the SERP to appear. The shown SERP (see Figure 1.1) has two types of result listings in response to the submitted query: organic results and paid results. *Organic search results* are the Web page listings that most closely match the user's search query based on relevance. *Paid results* are basically ads – the websites have paid to have their Web pages display for certain keywords, so these listings show up when someone runs a search query containing those keywords. In sponsored search, the value of an ad is not measured by impressions, instead, it is measured if the user clicks on it. The cost of each valid click is called the *cost per click (CPC)*.

Many ad inventories, such as display impressions and clicks, are auctioned off in real time. In display advertising, auctions are normally run in ad exchanges and each auction can target to a single impression from a specific group of users, called the *real-time bidding (RTB)* [Google, 2011, Yuan et al., 2013]. In sponsored search, auctions are mainly run by search engines. These auctions are slightly different to RTB because ad inventories are keyword-based (therefore, called *keyword auctions*) and the search engine also needs to consider the position effects on the probability of clicks [Edelman et al., 2007, Varian, 2007, Lahaie and Pennock, 2007, Börgers et al., 2013]. These topics are not discussed here, instead, a detailed review on advertising sales mechanisms are provided in Section 2.1.1.

1.2 Market Participants

Participants in online advertising markets can be divided into three groups: users, sellers and buyers. They are also called *interactive entities* [Evans, 2009, Jansen, 2011].

Online users can be anyone who use the Internet and the World Wide Web. Normally, a *user* issues ad-hoc topics to express his information needs, such as searching on Google or surfing Yahoo! Cars.

Sellers in online advertising include publishers, search engines, and supply-side platforms (SSPs). A *publisher* is an individual or organisation that prepares, issues, and disseminates content for public distribution [IAB, 2013]. Simply, publishers are those who have the space for ads to be displayed. A *search engine* is a company that indexes documents and then attempts to match documents by relevancy to the users'

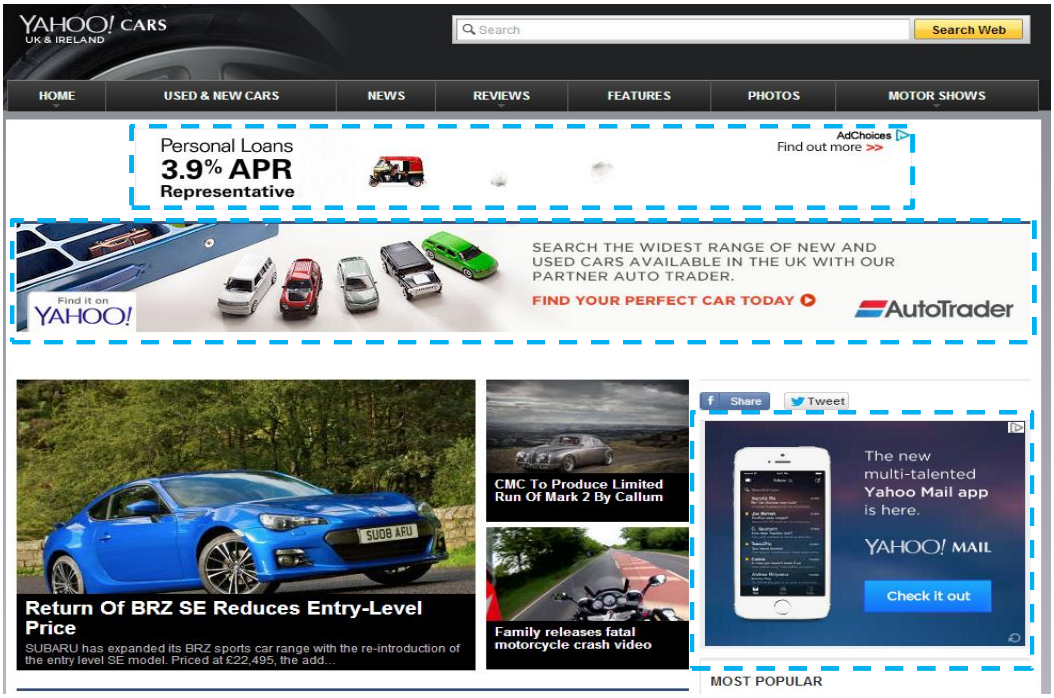

Figure 1.1: Display ads (in dotted blue line box) on the Yahoo! Cars Web page.

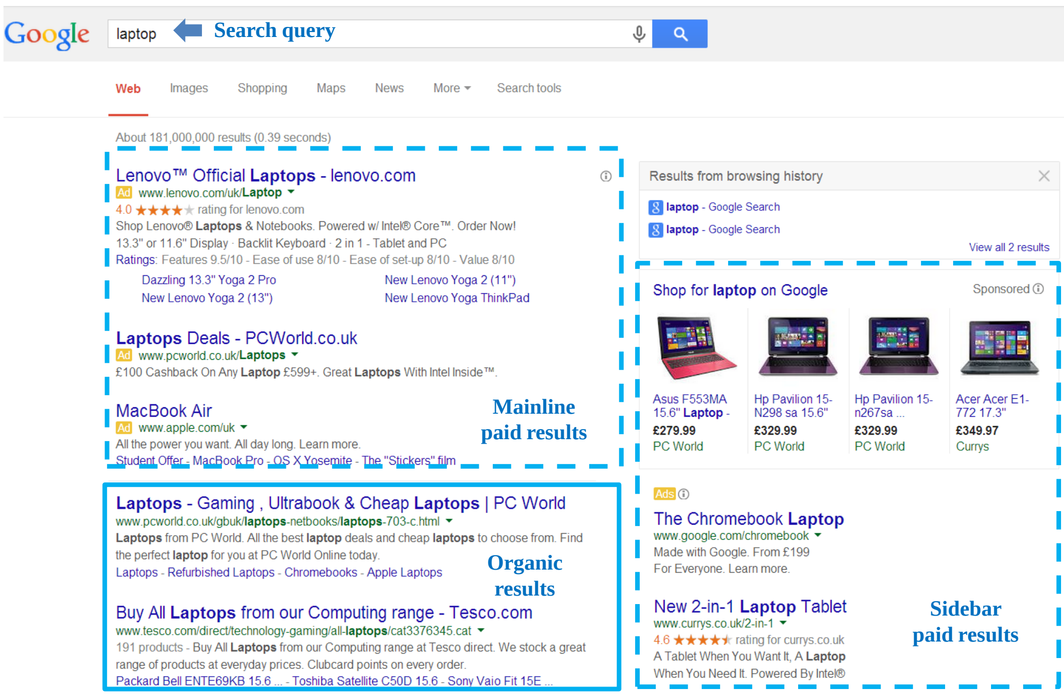

Figure 1.2: Organic search results (in blue line box) and paid search results (in dotted blue line box) on the Google Sponsored search results page (SERP).

search queries. A *supply-side platform (SSP)* is an intermediary company with the single mission of enabling publishers to manage their display impressions and maximise revenue. In the display advertising market, sellers refer to publishers or SSPs; in sponsored search, sellers normally refer to search engines.

Buyers include advertisers and demand-side platforms (DSPs). An *advertiser* is the one who requires places or slots to place marketing messages–online ads–in order to attract the attention of particular online users [Jansen, 2011]. In display advertising, advertisers sometimes join an automated bidding platform to get better or more impressions at low cost and such a platform is called the *demand-side platform (DSP)*.

1.3 Fundamental Challenges

Despite the strong growth of online advertising, the following fundamental challenges are worth of investigation:

1. In either keyword auction or RTB, the highly volatile bid prices make it difficult for online buyers to predict their campaign costs and returns, and thus further complicate their budget planning for advertising. This can be regarded as the price risk inherited from the spot market because bid prices reflect the changes of supply and demand of ad inventories, and the price risk cannot be avoid in an auction mechanism.

2. Guaranteed contracts are a solution to reducing the price risk because they specify the ad inventories and their delivery prices in advance for online buyers. However, they are mostly still sold through private human negotiations (particularly in display advertising) and therefore only a small number of inventories can be sold. According to eMarketer [2013], only 20% display ad inventories in 2012 were sold in terms guaranteed contracts, even through which generated 75% of publishers' revenue. This shows the current problem while great potential of selling guaranteed inventories programmatically.

3. The evolution of online advertising sales mechanisms (see Section 2.1.1) shows that online buyers always require more certainty and more control of their invested ad inventories. "They want more data to inform their bids, exposure to inventory not currently available to them, and preferred access to inven-

tory" [OpenX, 2013]. The current auction mechanisms and guaranteed contracts can partially fulfil the needs on the buyer side. For example, RTB allows advertisers to bid for their targeted users; guaranteed contracts allows advertisers to reserve the inventories that will be created in the future period. However, in general, online buyers have less flexibility and control of advertising in the process, which motivates us to design new ad products that can provide flexible guaranteed deliveries for advertisers.

4. Creating a loyal buyers' base is one important task that online sellers face [OpenX, 2013]. The "pay-as-you-go" attribute of auction mechanisms allows buyers to switch from one seller to another in the next bidding without paying any cost. If the sellers' ad inventories are similar to each other (also called *substitutable* in economics [Mankiw, 2006]), their revenues are difficult to be stable over time and to be optimised. This is because buyers rationally choose to advertise on the lower cost advertising placement and dynamically adjust their strategies across different sellers. This will be more obvious when the programmatic trading becomes more popular as algorithms can give the fastest feedback to buyers to choose the ad inventories with better effectiveness and profit conversion. Hence, sellers must take multiple issues into account in order to maintain their buyers on the long run and one possible solution is to establish contractual relationships with buyers.

The first and fourth statements explain why we study the non-guaranteed problem; the second and third statements describe how challenging the problem is. The first statement tells there is price risk inherent from auction mechanisms which makes online advertising difficult for advertisers. That is one motivation because we want to make advertisers satisfied. The fourth statement describes our second motivation and which is from the seller's perspective. Guaranteed deliveries allow a seller to expand and maintain his loyal advertisers so that his long-term revenues can be increased or stabilised. The second statement reveals the fact that most of profitable guaranteed contracts nowadays are not programmatically sold; therefore, there is much room for improvement on automating the selling of guaranteed inventories. The third statement indicates, apart from automation, offering more flexibility to advertisers is another challenge in providing guaranteed deliveries.

1.4 Proposed Solutions Using Financial Methods

The research carried out through this thesis focuses on using financial models to provide solutions to the non-guaranteed delivery problem in online advertising. Specifically, we consider how can online sellers, such as publishers, search engines and SSPs, develop novel computational algorithms or automated systems to provide guaranteed advertising deliveries for online buyers such as advertisers and DSPs.

By its very nature, advertising is a prominent feature of economic life [Bagwell, 2001]. Economic models contribute to online advertising on many aspects. Marketing researchers use the social choice theory to explore the effectiveness of advertising types and users' engagement. Computer scientists and economists employ the applied game-theoretical models to design online auctions, such as the Generalised Second-Price (GSP) auction [Edelman et al., 2007]. However, finance, as a sub-field of microeconomics, has received surprisingly little attention in online advertising. Two features can distinguish financial studies from other economic resource allocation decisions. First, financial decisions are spread out over time. Second, they are usually not known with certainty in advance by either the decision makers or anybody else [Bodie et al., 2009]. These two features make financial methods particularly suitable and interesting to solve the non-guaranteed delivery problem in online advertising.

This thesis contributes to the field of online advertising both methodologically and algorithmically: the former is supported by mathematical models and statistical analysis; and the latter is validated by empirical experiments. There are three major contributions of this thesis, which are discussed in Chapters 3-5:

1. We study an optimal dynamic model for a publisher or SSP who engages in RTB and wants to provide the guaranteed delivery of display impressions. The model mimics the advanced booking system in the airline industry, and considers both allocation and pricing of estimated future impressions. This work will be introduced in Chapter 3 [Chen et al., 2014b].

2. We propose a multi-keyword multi-click ad option for sponsored search. The proposed option allows an advertiser to: (i) target a set of ad keywords for a cer-

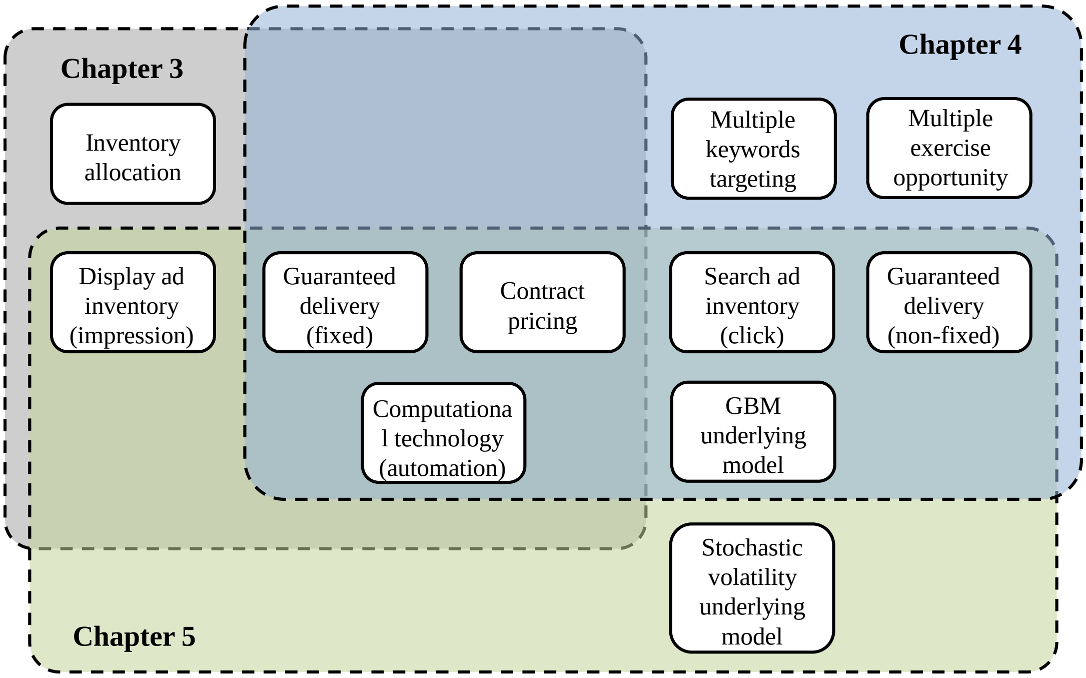

Figure 1.3: Differences and relationships of the contributions in Chapters 3-5.

tain number of clicks; (ii) multiply exercise the option at any time prior to or on the contract expiration date; and (iii) switch the guaranteed delivery request from one keyword to another without paying any cost. We discuss the option evaluation models based on the assumption that the underlying winning payment prices of targeted keywords follow an univariate (or multivariate) geometric Brownian motion (GBM) [Samuelson, 1965a, Wilmott, 2006]. This work will be described in Chapter 4 [Chen et al., 2014a].

3. We discuss a display ad option for display advertising and study a lattice framework for evaluation. The proposed option allows an advertiser to pay a fixed CPM or CPC for an impression or click that is same or different to its underlying measurement model from real-time auctions. The display ad option can be priced for those situations where the GBM assumption is not valid empirically. We use the stochastic volatility (SV) model to describe the underlying price movement and construct a censored binomial lattice to approximate the underlying SV model. This work will be discussed in Chapter 5 [Wang and Chen, 2012, Chen and Wang, 2014].

Figure 1.3 shows the structural relationships of studies in Chapters 3-5. Overall, our three studies provide novel solutions to the non-guaranteed delivery problem in online advertising via contract mechanisms (i.e., guaranteed contracts in Chapter 3 and option contracts in Chapters 4-5). Therefore, each chapter discusses the contract pricing models depending on different advertising environments and the pricing can algorithmically adapt to the changes of supply and demand of inventories in auction mechanisms.

Chapters 4-5 are also uniquely different to each other. Chapter 3 focuses on bringing automation into selling premium impressions in display advertising apart from RTB and discusses both optimal allocation and pricing. The guaranteed contracts are as same as those sold through human direct negotiations: advertisers are guaranteed with needed deliveries at the pre-specified price in the pre-specified future period, therefore, they need to pay the full amount of reservation in advance. Chapter 4 moves several steps further and introduces a flexible guaranteed contract (i.e, ad option) into sponsored search. The flexibility contains non-obligatory exercise right, multiple keywords targeting, multiple clicks exercising, etc. The contract pricing is based on the GBM underlying model as it is suitable for the search ad inventories according to our data analysis and other previous research. Chapter 5 extends the option idea into display advertising and proposes an option contract which allows its buyer to have different payment schemes to the underlying ad format. In addition, as the GBM underlying model is not suitable for display advertising, we study the SV underlying model for more general situations.

1.5 Structure of the Thesis

The rest of the thesis is organized as follows:

Chapter 2 reviews the related literature in online advertising and financial theory. Section 2.1 gives a chronological review of the development of online advertising sales mechanisms and discusses the related work of guaranteed advertising deliveries. Section 2.2 introduces the basic financial concepts that will be used for the research of this thesis, reviews the related work of revenue management (with special focus on dynamic pricing models in airline industry), and discusses financial options and their evaluation methods.

Chapter 3 proposes an optimal dynamic model for unifying programmatic guarantee (PG) and real-time bidding (RTB) in display advertising. Section 3.1 introduces the background and the overview of this study. Section 3.2 formulates the problem, discusses our assumptions and provides a solution. Section 3.3 presents the results of our experimental evaluation and Section 3.4 summarises the chapter.

Chapter 4 discusses an ad option tailored for the unique environment of sponsored search, where multiple ad keywords and certain number of required clicks are considered. Chapter 4.1 introduces the background and indicates the problem. Section 4.2 introduces option structure, and the process of buying, selling and exercising. Section 4.3 discusses the option pricing methods. Section 4.4 analyses the effects on search engine's revenue. Section 4.5 presents our experimental evaluation and Section 4.6 summarises the chapter. The mathematical results used throughout the chapter are provided in Section 4.7.

Chapter 5 studies an ad option for display advertising. Section 5.1 introduces the background and indicates the problem. Section 5.2 investigate several lattice methods for pricing a display ad option with the GBM underlying model. Section 5.3 discusses our lattice method for pricing a display ad option with the SV underlying model. Section 5.4 presents our experimental results. Section 5.5 summarises the chapter. The mathematical results used throughout the chapter are provided in Section 5.6.

Chapter 6 sums up what has been learned in this endeavor and suggests the directions for future research.

Appendix A provides a brief summary of the technical terms that are used throughout the thesis. Appendix B lists the publications and submissions that have been completed during my PhD study at University College London.

Chapter 2

Background

This chapter reviews the literature on online advertising and introduces the preliminaries of modern financial theory. The mathematical content of the chapter is kept to the minimum necessary to achieve a proper understanding of important concepts and models in these two fields. Detailed mathematical modelling and data analysis will be discussed in Chapters 3-5.

2.1 Literature Review on Online Advertising

The following is a chronological review of the development of online advertising sales mechanisms, and then a review of the research related to guaranteed advertising deliveries.

2.1.1 Evolution of Online Advertising Sales Mechanisms

The first banner ads were introduced on 27th October 1994, when HotWired (today Wired News, part of Lycos) signed fourteen banner ads with AT&T, Club Med and Coor'z Zima [Bruner, 2005, Evans, 2009]. These banner ads were largely sold on the number of impressions – individuals who saw the ads – which was the model used by most traditional media for brand advertising [Evans, 2009]. Many online ads were subsequently sold based on 1000 viewers per ad. This is also referred to as the *cost-per-mille* (CPM) measurement (or payment) model. In early online advertising, advertisers paid flat fees to show their ads a fixed number of times. Advertisements were negotiated on a case-by-case basis, minimum contracts for purchases were large and entry was slow [Edelman et al., 2007].

In 1994, search engine InfoSeek introduced the concept of targeting ads to keyword search queries, albeit against display banners not text ads [Bruner, 2005]. However, paying by the number of viewers remained the norm until Procter & Gamble negotiated a deal with Yahoo! in 1996. Procter & Gamble was allowed to pay for ads only on the *cost-per-click* (CPC) basis – Yahoo! was paid only when an online user clicked on the ad [Lahaie and Pennock, 2007, Evans, 2009]. This was the Web version of paying for direct response commonly used by advertisers for things such as mail and telephone solicitations [Bruner, 2005]. Search engine OpenText first tried to put together targeted search queries with paid listings in 1996, but it met with considerable outcry from users [Bruner, 2005, Jansen, 2011].

On 21st February 1998, GoTo.com (then Overture Services, now owned by Yahoo!) launched a sponsored search business model in which the search engines ranked the Web sites based on how much the sites are willing to pay to be placed at the top of the search results under a real-time competitive bidding process [Edelman et al., 2007, Lahaie and Pennock, 2007, Jansen, 2011].

In the original design of GoTo.com's auction, each advertiser submitted a bid reporting his willingness to pay on a per-click basis for a particular search keyword. Advertisers can: (i) target their ads instead of paying for a banner ad that would be shown to everyone visiting a Web site; (ii) specify which keywords were relevant to their products and how much each of those keywords was worth to them based on the users' clicks. Also, ads were sold on the CPC basis. Every time a user clicked on a sponsored link, an advertiser's account was automatically billed the amount of the advertiser's most recent bid. The sponsored links to advertisers were arranged in descending order of bids, making highest bids the most prominent. The GoTo's auction is actually a *generalised first-price (GFP) auction* [Edelman et al., 2007, Jansen, 2011]. The ease of use, the very low entry costs, and the transparency of the mechanism quickly led to the success of GoTo's paid search platform. Yahoo! and MSN soon adopted the GoTo's concept and implemented the GFP auction model on their advertising platforms. However, the underlying auction mechanism itself was far away from perfect. Under the GFP auction framework, the advertiser who can react to competitors' moves fastest had a substantial advantage. The mechanism therefore encouraged inefficient investments in gaming the system, causing volatile prices and allocative inefficiencies [Edelman et al., 2007].

Google provided its solution to these problems by launching its own sponsored search platform, Google AdWords, in February 2002 [Edelman et al., 2007]. Google AdWords adopted many of GoTo.com's concepts but introduced some significant changes. First, Google continued the sales-by-impression model in parallel before finally dropping it altogether in favour of the CPC measurement model [Jansen, 2011]. Second, Google changed the auction model from the GFP auction to a more stable *generalised second-price (GSP)* auction. In the simplest GSP auction, suppose there are n ad positions, an advertiser in position i pays a CPC equal to the bid of an advertiser in position $i + 1$ plus a minimum increment. This second-price structure makes the market more user friendly and less susceptible to gaming [Varian, 2007, Edelman et al., 2007]. Third, Google also changed the standard allocation rule. Instead of ranking ads by bid price alone, the platform computed a quality score derived from the bid amount and *click-through rate (CTR)*. CTR measures the rate at which the searchers click on an ad's link. These factors were later enhanced with other factors such as keyword relevancy and landing-page quality. Google's approach ensured that no advertiser can just buy their way to the top search results while getting no clicks. Recognizing these advantages, Yahoo!, Microsoft and other major search engines subsequently switched their sponsored search platforms from GFP auctions to GSP auctions.

It is worth mentioning that the GSP auction is similar but different to another famous second-price auction mechanism, called the *Vickrey-Clark-Groves (VCG) auction* [Vickrey, 1961, Clarke, 1971, Groves, 1973], which was recently implemented by Facebook [Varian and Harris, 2013]. The VCG auction has the same allocation rule as the GSP auction. In its simplest version, ad slots are allocated to advertisers by ranking their bids in decreasing order. However, winning advertisers are charged differently to the GSP auction. For the position i, the winning advertiser pays for the externalities that he imposes on others rather than the highest bid next to him.

From a mechanism design perspective, the GSP and VCG auctions have their own advantages and disadvantages. Here we simply compare the two auction models along the following four dimensions: incentive compatibility, solution equilibrium, revenue maximisation and allocative efficiency. *Incentive compatibility* (also called *truth-telling*) essentially refers to offering the right incentives that make advertisers reveal their value truthfully [Narahari et al., 2009]. Truth-telling is a dominant strategy under VCG auctions but not under GSP auctions. In a GSP auction, advertisers may shade their bids to maximise their expected utilities. Varian [2007] and Edelman et al. [2007] proved there is a special equilibrium under GSP auctions, called the *symmetric Nash equilibria (SNE)* (or the *locally Envy-free equilibrium*). In the SNE, the expected revenue of a GSP auction is at least as same as a VCG auction. However, the SNE does not always hold in front of random bids. *Allocative efficiency* is achieved when the social utility (i.e., the sum of utilities) of all winning advertisers is maximised. Allocative efficiency ensures that the inventories are allocated to the advertisers who value them most. VCG auctions satisfy allocative efficiency while GSP auctions do not satisfy this property. Even though VCG auctions offer more economic properties, most of the current advertising platforms still use GSP auctions. Edelman et al. [2007] pointed out several possible reasons. First, in many situations, GSP auctions are more profitable for a publisher or search engine. Second, switching from GSP auctions to VCG auctions may generate substantial transition costs. Third, the payment rules of GSP auctions are simple and easy to explain to advertisers.

The introduction of ad networks was another milestone in online advertising. The original ad networks were set up in 1997 to address the problem for advertisers who want to advertise across many different websites [OpenX, 2010]. By aggregating inventory across multiple sites, ad networks offered advertisers the ability to reach the size of audience that they had come to expect from traditional channels like televisions. However, there are several limitations of ad networks. First, there are many intermediaries in the value chain between publishers and advertisers, each taking a slice of "profit cake". For example, an ad network who cannot sell some particular inventories may offer them at a cheaper price to another ad network. Second, advertisers may spend much time and effort on exploring which network is the best one to purchase inventories. Third, to maximise revenue, publishers may spend much time and effort on the allocation of inventories among ad networks.

Ad exchanges came up to improve the limitations of ad networks, which are the technology platforms that facilitate the buying and selling of ad inventories from multiple ad networks [Muthukrishnan, 2009, OpenX, 2010]. Three major exchanges were acquired in 2007: Yahoo! bought Right Media in April, Google bought DoubleClick in May and Microsoft bought AdECN in August. Each company quickly made vast pools of ad inventories, which greatly improved the experience for many participants to transact centrally.

The arrival of ad exchanges and ad networks brought another innovative sale mechanism to online advertising (mainly for display advertising) – real-time bidding (RTB) – a programmatic trading technique designed to help advertisers take advantage of increased data and inventory liquidity [PubMatic, 2010]. Before RTB, buying from multiple exchanges was time-consuming and inefficient for advertisers. They had to use a different system to access each exchange. And since a typical campaign would pull inventories from more than one exchange, there was no easy way to achieve de-duplicated reach or to cap the number of impressions that audiences would receive from any given campaign [Google, 2011]. RTB therefore was originally conceived as an advertiser-focused solution and many DSPs provided services based on it.

2.1.2 Guaranteed Advertising Deliveries

In the following discussion, the research related to guaranteed advertising deliveries is reviewed. As described in Section 2.1.1, guaranteed contracts appeared in the early stages of online advertising but were negotiated by advertisers and publishers privately [Edelman et al., 2007]. Each negotiation contains an amount of needed inventories over a certain period of time and a pre-specified guaranteed price. Hence, in discussing the guaranteed delivery, the following issues must be considered: allocation and pricing. Many studies discussed the two issues separately. Allocation models will be explored first, and then pricing models.

Feldman et al. [2009] studied an ads selection algorithm for a publisher whose objective is not only to fulfil the guaranteed contracts but also to deliver the well-targeted display impressions to advertisers. This research was more relevant to a service matching problem. The allocation of impressions between the guaranteed and non-guaranteed channels was first discussed by Ghosh et al. [2009], where a publisher was considered to act as a bidder who bids for guaranteed contracts. This modelling setting was reasonably good as the publisher acts as a bidder who would allocate impressions to online auctions only when other winning bids are high enough. Balseiro et al. [2011] investigated the same allocation problem but used some stochastic control models. Simply, they considered, for a given price of an impression, the publisher can decide whether to send it to ad exchanges or assign it to an advertiser with a fixed reserve price. The decision making process aims to maximise the expected total revenue. Roels and Fridgeirsdottir [2009] proposed a similar allocation framework to Balseiro et al. [2011], where the publisher can dynamically select which guaranteed buy requests to accept and to deliver the guaranteed impressions accordingly. However, compared to Balseiro et al. [2011], the uncertainty in advertisers' buy requests and the traffic of website were explicitly modelled under the revenue maximisation objective. Recently, a lightweight allocation framework was proposed by Bharadwaj et al. [2012]. They used a simple greedy algorithm to simplify the computations of revenue maximisation.

Bharadwaj et al. [2010] discussed two algorithms for pricing the guaranteed display contracts. Each contract contains a bulk of impressions and the proposed algorithms solved the revenue optimisation problem for the given number of users' visits (i.e., the demand level). However, their work did not consider the auction effects on the contract pricing, and the developed algorithms were purely based on the statistics of users' visits.

Consider if the online advertising market is bulling (i.e., the winning payment prices of specific ad inventories from online auctions increase) and the non-guaranteed selling looks more profitable for publishers, they may want to cancel the sold guaranteed contracts before the time that the targeted inventories will be created. Online auctions with cancellations were recently discussed by Babaioff et al. [2009] and Constantin et al. [2009]. They both considered a design that a publisher can cancel the sold guaranteed contracts but needs to pay a penalty to advertisers. The proposed auctions with cancellations enjoyed some economic properties, such as allocative efficiency and equilibrium solution. However, there may exist speculators who pursue the cancellation penalty only. In fact, the discussed cancellation penalty is very similar to the over-selling booking of flight tickets. Several over-selling booking models were discussed by Talluri and van Ryzin [2005].

Up to this point, the reviewed guaranteed contracts are all for display advertising. Salomatin et al. [2012] studied a framework of guaranteed deliveries for sponsored search, under which advertisers are able to send their guaranteed requests to a search engine. Each guaranteed request includes the needed number of clicks and the ad budget. The search engine then decides the guaranteed delivery according to search queries and available positions. Since the allocation decision is based on the joint revenue maximisation from guaranteed deliveries and keyword auctions, some advertisers may not receive all their demanded clicks. In such cases, the search engine pays a penalty. However, advertisers still have less control of the ad exposure time and the position of the ad. In addition, with the number of guaranteed advertisers increasing, it is less likely that advertisers will meet their business needs in such a mechanism.

The ad option concept was initially introduced by Moon and Kwon [2010] (even though Meinl and Blau [2009] discussed the possibility of Web service derivatives, their proposal was not intended for online advertising). The ad option buyer can be guaranteed the right to choose the minimum payment between CPM and CPC once CTR is realized. This option contract was similar to a *paying the worst and cash option* [Zhang, 1998]. Moon and Kwon [2010] suggested an evaluation of the option under the framework of a Nash bargaining game. Simply, they considered two utility functions: one for advertiser and one for publisher. The objective function is the product of these two utilities and each utility function is restricted by a negotiation power. The option price is the optimal solution which maximises the negotiated joint utility. The work of Moon and Kwon [2010] motivates the research in Chapters 4-5. However, in this study, the proposed ad options differ from theirs in contract structure and evaluation methods.

2.2 Preliminaries on Financial Methods

Finance mainly studies how people make decisions regarding the allocation of resources over time and the handling of risk [Mankiw, 2006]. In this section, an incomplete sketch of financial models is presented. First, some basic concepts, such as uncertainty, risk and time value of money are introduced. Then, revenue management models for airline industry are discussed, which lays the foundation for the research in Chapter 3. Financial options and their evaluation techniques are explored, which form the prelude for the research in Chapters 4-5.

2.2.1 Uncertainty, Risk and Time Value of Money

Uncertainty and risk are manifestations of the same underlying force – randomness [Schmid, 2012]. They are closely related but slightly different concepts. Uncertainty is lack of certainty. Therefore, an uncertain environment is one in which the individual decision maker is not absolutely sure of the consequences of any particular actions. Individuals make decisions according to some rules but the outcomes of those decisions are not known with certainty at the time the decision is taken. Therefore, *uncertainty* is a non-quantifiable form of randomness. Its application to real-world situations is not well-charted. *Risk* is randomness in which events have measurable probabilities [Knight, 1921, Schmid, 2012]. Probabilities may be attained either by deduction using theoretical models or induction using the observed frequency of events. This notion implies that a choice sometimes has an influence on the outcome. We may simply distinguish between the two concepts as follows: "uncertainty exists whenever one does not know for sure what will occur in the future. Risk is the uncertainty that 'matters' because it affects people's welfare. Thus, uncertainty is a necessary but not a sufficient condition for risk. " [Bodie et al., 2009].

There is a time dimension in an uncertain world, since decisions and outcomes are separated in time. Time value of money refers to the fact that money in hand today is worth more than the expectation of the same amount to be received in the future. There are three reasons why this is true [Sundaresan, 2000, Bodie et al., 2009]: the first is that one can invest it, earn interest, and end up with more in the future; the second is that the purchasing power of money can change over time because of inflation; the third is that the receipt of money expected in the future is, in general, uncertain.

2.2.2 Revenue Management

Revenue management (RM) investigates specific resource allocation problems within finance. The following three types of decisions faced by a seller of products or services are considered to be the RM research [Talluri and van Ryzin, 2005]: structural decisions, price decisions and quantity decisions. *Structural decisions* refer to the sales mechanism design. For example, that method a seller uses to deliver his products or services to consumers, such as posted prices, private negotiations, auctions and so on. *Price decisions* are means of evaluating a seller's products or services under a specific market structure. This includes, how to set posted prices, how to set reserve prices in auctions, how to price products across different categories, and how to adjust price over time. *Quantity decisions* refer to the capacity problem. For example, whether to accept or reject a purchase request for a given stock level, or how to allocate products to different types of consumers.

In the following discussion, several RM models that were used in the airline industry are briefly reviewed as the situation of selling a given stock of flight tickets by a deadline is similar to selling future ad inventories in advance. Talluri and van Ryzin [2005] introduced the basic mathematical settings of both price-based RM models and capacity-based RM models for pricing flight tickets. Price-based RM models aim at maximizing the seller's expected revenue by setting the optimal dynamic prices over time; capacity-based RM models aim at maximizing the seller's expected revenue by allocating the remaining tickets effectively. In fact, most of these two types of models consider both pricing and allocation problems for the estimated level of supply and demand. The main difference between them is which variable (i.e., price or capacity) is the control variable in optimisation. Anjos et al. [2004] studied a dynamic model which finds the optimal price to charge for flight tickets under one-way pricing. Their model was based on two underlying assumptions of the consumer's behaviour. The first assumption stated that consumers are price sensitive. If the ticket price increases, consumers are less willing to buy. The second assumption stated that consumers are time sensitive. Their needs to purchase the tickets increase when approaching to the flight departure time. Anjos et al. [2005] then discussed a general framework and examined several pricing policies under various formulations of consumer behaviour. Malighetti et al. [2009] analysed the pricing policy adopted by Ryanair, a main low-cost carrier in Europe. They examined several preference choice functions for one-way pricing using a wide range of actual prices on all of Ryanairs routes, thus validating assumptions made by Anjos et al. [2004, 2005]. Chapter 3 considers the sale of identical future ad inventories in advance, which are based on the economic settings discussed by Anjos et al. [2004, 2005] and Malighetti et al. [2009]. However, our work is more sophisticated because the salvage value of ad inventories on the delivery date is not zero, which is determined by the level of competitions in future real-time auctions.

2.2.3 Options and Option Pricing Methods

Options have been widely used in many fields: financial options are an important derivative to speculate profits as well as to hedge risk [Wilmott, 2006]; real options are an effective decision-making tool to evaluate business projects and corporate risk management [Boer, 2002]. The research carried our throughout Chapters 4-5 is closely related to financial options, whose evaluation is regarded as one of the most important application areas of mathematics today. Sundaresan [2000], Constantinides and Malliaris [2001] and Hobson [2004] provided good surveys on financial options and option pricing models. In the following discussion, the basic option concepts are introduced and the studies related to the research in Chapters 4-5 are reviewed.

2.2.3.1 Standard Options

A *standard option* (or *vanilla option*) is a contract in which the seller grants the buyer the right, but not the obligation, to enter into a transaction with the seller to either buy or sell the underlying asset at a fixed price on or prior to a fixed date [Wilmott, 2006]. The underlying asset can be a stock, bond, foreign currency, or index such as S&P-100, FTSE-100 etc. The market price of the underlying asset is called the *underlying price*; the fixed price is called the *strike price*; and the fixed date is called the *expiration date* (or *maturity date*). The seller grants this right in exchange for a certain amount of money, called the *option price*.

An option is called the *call option* if its buyer has the right to buy the underlying asset in the future. Another case is called the *put option* where the option buyer has the right to sell. The simplest standard option is the *European option* [Wilmott, 2006], which can be exercised only on the expiration date. This differs from an *American option* [Wilmott, 2006], which can be exercised at any time during the contract period. Both European and American options are *standard options*.

2.2.3.2 Exotic Options

In the beginning of the 1980s, standard options became more widely understood and their trading volume exploded. Financial institutions began to search for alternative forms of options, known as *exotic options* [Zhang, 1998], to meet their new business needs. Among them, two types of options, multi-asset options and multi-exercise options, are particularly relevant to our research.

Multi-asset options are options written on at least two underlying assets [Zhang, 1998]. These underlying assets can be stocks, bonds, currencies and indices in either the same category or different markets. Several types of multi-asset options are worth mentioning, such as basket options, dual-strike options, rainbow options, paying the

Table 2.1: Comparison between the n-keyword 1-click ad option (see Chapter 4) and other relevant options: $C_i(t)$ is the price of the ith underlying asset at time $t \in [0, T]$ and T is the contract expatriation date (if there is only one underlying asset its price is denoted by $C(t)$); F_i is the strike price (i.e., the fixed payment price) of the ith underlying asset (if there is only one strike price we denote it by F); ω_i is the weight of ith asset in a basket-type option.

Option contract	Payoff function	Underlying variable	Exercise opportunity	Early exercise	Strike price	Application area
n-keyword 1-click ad option (keyword exact or broad match)	$\max\{C_1(t) - F_1, \ldots, C_n(t) - F_n, 0\}$	Multiple	Single	Yes	Multiple	Keywords
n-keyword 1-click ad option (keyword broad match)	$\max\left\{\sum_{i=1}^{k_1} \omega_{1i} C_{1i}(t) - F_1, \cdots, \sum_{i=1}^{k_n} \omega_{ni} C_{ni}(t) - F_n, 0\right\}$	Multiple	Single	Yes	Multiple	Keywords
European standard call option [Wilmott, 2006]	$\max\{C(T) - F, 0\}$	Single	Single	No	Single	Equity stock, or index
American standard call option [Wilmott, 2006]	$\max\{C(t) - F, 0\}$	Single	Single	Yes	Single	Equity stock, or index
European basket call option [Krekel et al., 2006]	$\max\left\{\sum_{i=1}^{n} \omega_i C_i(T) - F, 0\right\}$	Multiple	Single	No	Single	Index of equity stocks, bonds or foreign, currencies
European dual-strike call option [Zhang, 1998]	$\max\{C_1(T) - F_1, C_2(T) - F_2, 0\}$	Double	Single	No	Double	Equity stocks, or indexes of equity stocks, or bonds, foreign currencies
European rainbow call on max option [Ouwehand and West, 2006]	$\max\{\max\{C_1(T), \ldots, C_n(T)\} - F, 0\}$	Multiple	Single	No	Single	
European paying the best and cash option [Johnson, 1987]	$\max\{C_1(T), C_2(T), F\}$	Double	Single	No	Single	
European quotient call option [Zhang, 1998]	$\max\{C_1(T)/C_2(T) - F, 0\}$	Double	Single	No	Single	

best and cash options, and quotient options. Table 2.1 provides a brief summary of these multi-asset options, and compares them to standard options and our proposed multi-keyword multi-click ad options (see Chapter 4) along the following seven dimensions: payoff function, underlying variable, exercise opportunity, early exercise opportunity, strike price and application area. This comparison indicates that our proposed ad options are more complex than others and are, therefore, more difficult to evaluate.

In Table 2.1, it is worth emphasising basket options and dual-strike options. *Basket options* are those options whose payoff is determined by the weighted sum of underlying asset prices [Wilmott, 2006]. This structure can be extended to the keyword broad match scenario, where the weights are the probabilities that sub-phrases occur in search queries. *Dual-strike options* are the options with two different strike prices for two different underlying assets [Zhang, 1998]. The simplest version of our proposed ad options is a dual-strike call option, which allows an advertiser to switch his targeted two ad keywords during the contract lifetime. However, in sponsored search, the number of candidate ad keywords to choose from is usually more than two, so the two keywords are extended to higher dimensions (see Chapter 4). In addition, as an advertiser usually needs more than a single click for guaranteed delivery, the dual-strike call option is extended to a multi-exercise option.

Multi-exercise options are a generalisation of American options, which provide a buyer with more than one exercise right and sometimes control over one or more other variables [Villinski, 2004], such as the amount of the underlying asset exercised in certain time periods. Multi-exercise options have become more prevalent over the past decade, particularly, in the energy industry, such as electricity swing options and water options. Contributors to the multi-exercise options include Deng [2000], Deng and Oren [2006], Clewlow and Strickland [2000], Villinski [2004], Weron [2006], Marshall et al. [2011] and Marshall [2012]. Their work is not discussed further here as our proposed ad options in Chapter 4 are simple examples of multi-exercise options. Compared to the energy industry, the multi-exercise opportunity in sponsored search is more flexible. Ad options are proposed that can allow advertisers to exercise options at any time in the option lifetime, i.e. the exercise time is not pre-specified, and no minimum number of clicks is required for each exercise. Therefore, there is no penalty fee if the advertiser does not exercise the minimum clicks. In addition, there is no transaction fee for the ad option in sponsored search.

2.2.3.3 Option Pricing Methods

Motivated by an attempt to model the fluctuations of asset prices, Brownian motion (i.e., the continuous-time random walk process [Shreve, 2004]) was first introduced by Bachelier [1900] to price an option. However, the impact of his work was not recognised by financial community for many years. Sixty five years later, Samuelson [1965b] replaced Bachelier's assumptions on asset price with a geometric form, called the *geometric Brownian motion (GBM)*. In the GBM model, the proportional price changes are exponentially generated by a Brownian motion, thereby solving the problem of negative asset prices in option pricing. While the GBM model is not appropriate for all financial assets in all market conditions, it remains the reference model against which any alternative dynamics are judged.

The research of Samuelson highly affected Black and Scholes [1973] and Merton [1973], who then examined the option pricing based on the GBM underlying model. They constructed a portfolio from risky and risk-less underlying assets to replicate the value of a European option. Risky assets can be stocks, foreign currencies, indices, and so on; risk-less assets can be bonds. Once the value of the replicated portfolio is estimated, the option value can be obtained accordingly. The pricing methods proposed by Black and Scholes [1973] and Merton [1973] were based on the assumption that investors on the market cannot obtain arbitrage. Therefore, the replicated portfolio is treated as a self-adjusting process whose least expectation of returns increase at the same speed as the constant bank interest rate. If considering the constant bank interest rate as a discount factor, the discounted value of the replicated portfolio would be a Martingale [Björk, 2009], whose probability measure is called the *risk-neutral probability measure*. Since a closed-form pricing formula can be obtained from the settings of Black and Scholes [1973] and Merton [1973], we normally call their work as the *Black-Scholes-Merton (BSM) option pricing formula*.

The BSM option pricing formula spurred research in option evaluation. Various numerical procedures have appeared in this field, including lattice methods, finite difference methods, Monte Carlo simulations, and so on. These numerical procedures are capable of evaluating more complex options when the closed-form solution does not exist. In the following discussion, lattice methods for pricing options with a parametric underlying process are reviewed. This provides literature for the research in Chapter 5.

Sharpe [1978] initiated the concept of pricing a call option written on an asset with simple up and down two-state price changes. We call it the *one-step binomial lattice method* and use it as a pedagogical framework to explain the continuous-time option pricing model without reference to stochastic calculus. Cox et al. [1979] then developed a multi-step binomial framework, called the *Cox-Ross-Rubinstein (CRR) model*, which can converge with the BSM pricing formula if the length of the time step is sufficiently small. Boyle [1986] proposed a trinomial lattice, whereby the asset price can either move upwards, downwards, or stay unchanged in a given time period. Boyle also discussed the pricing for an option with two underlying assets via two correlated trinomial lattices [Boyle, 1988] and investigated how a multinomial lattice can be developed to price an option with a single asset [Boyle et al., 1989]. Other contributors to lattice methods include Kamrad and Ritchken [1991], Tian [1993] and Haahtela [2010]. The mathematical results of these lattice methods are presented in Table 2.2, where the movement scale is the ratio of the price in the next state to the current one, and the transition probability is the risk-neutral probability that the asset price moves from the current state to the next one.

As discussed above, lattice methods adopt Samuelson's GBM assumption for the underlying asset price. However, the GBM assumption may not always be empirically valid. This motivates a general Ornstein-Uhlenbeck (OU) diffusion process for option pricing. Nelson and Ramaswamy [1990] discussed the conditions under which a sequence of binomial processes converges weakly to an OU diffusion process and investigated its application to pricing an option written on an asset with constant volatility. Primbsa et al. [2007] then proposed a pentanomial lattice method that incorporates the skewness and kurtosis of the underlying asset price and found that the limiting distribution is compounded Poisson. Nelson and Ramaswamy [1990] and Primbsa et al. [2007] only solved the lattice pricing for the non-GBM underlyings which have a constant volatility. Florescu and Viens [2008] proposed a lattice method that deals with the stochastic volatility (SV) underlying model. However, their method is not very practical in terms of computational efficiency as the transition probabilities are restricted by

Table 2.2: Summary of the relevant lattice methods for pricing a call option with the GBM underlying model: Δt is the length of each time step; r is the continuous-time compounding risk-less interest rate; σ_1, σ_2 are the volatilities for the underlying asset price and ρ is the correlation; u, m, d are the state transition scale of underlying asset price in upward, unchanged and downward movements; $q_1, q_2, \dots, q_n$ are the risk-neutral transition probabilities from upward movement to downward movement.

Model	Movement scales u, d (or u, m, d)	Transition probabilities $q_1, q_2, \cdots q_k$
Binomial lattice (one factor)		
CRR	$u = e^{\sigma\sqrt{\Delta t}}$, $d = 1/u$.	$q_1 = \frac{e^{r\Delta t}-d}{u-d}$, $q_2 = 1 - q_1$.
Tian-BIN	$u = \frac{\gamma\zeta}{2}(\zeta+1+\sqrt{\zeta^2+2\zeta-3})$, $d = \frac{\gamma\zeta}{2}(\zeta+1-\sqrt{\zeta^2+2\zeta-3})$, $\gamma = e^{r\Delta t}$, $\zeta = e^{\sigma^2\Delta t}$.	$q_1 = \frac{e^{r\Delta t}-d}{u-d}$, $q_2 = 1 - q_1$.
Haahtela-BIN	$u = e^{\sqrt{e^{\sigma^2\Delta t}-1}+r\Delta t}$, $d = e^{-\sqrt{e^{\sigma^2\Delta t}-1}+r\Delta t}$.	$q_1 = \frac{e^{r\Delta t}-d}{u-d}$, $q_2 = 1 - q_1$.
Trinomial lattice (one factor)		
Boyle-TRIN	$u = e^{\lambda\sigma\sqrt{\Delta t}}$, $m = 1$, $d = e^{-\lambda\sigma\sqrt{\Delta t}}$,	$q_1 = \frac{(\zeta+\gamma^2-\gamma)u-(\gamma-1)}{(u-1)(u^2-1)}$, $\gamma = e^{r\Delta t}$, $q_2 = 1 - q_1 - q_3$, $q_3 = \frac{(\zeta+\gamma^2-\gamma)u^2-(\gamma-1)u^3}{(u-1)(u^2-1)}$.
KR-TRIN	$u = e^{\lambda\sigma\sqrt{\Delta t}}$, $m = 1$, $d = e^{-\lambda\sigma\sqrt{\Delta t}}$.	$q_1 = \frac{1}{2\lambda^2} + \frac{(r-\frac{1}{2}\sigma^2)\sqrt{\Delta t}}{2\lambda\sigma}$, $q_2 = 1 - \frac{1}{\lambda^2}$, $q_3 = \frac{1}{2\lambda^2} - \frac{(r-\frac{1}{2}\sigma^2)\sqrt{\Delta t}}{2\lambda\sigma}$, $\zeta = e^{2r\Delta t}\big(e^{\sigma^2\Delta t} - 1\big)$.
Tian-TRIN	$u = \varpi + \sqrt{\varpi^2 - m^2}$, $m = \gamma\zeta^2$, $\gamma = e^{r\Delta t}$, $\zeta = e^{\sigma^2\Delta t}$, $d = \varpi - \sqrt{\varpi^2 - m^2}$, $\varpi = \frac{\gamma}{2}(\zeta^4 + \zeta^3)$.	$q_1 = \frac{md-\gamma(m+d)+\gamma^2\zeta}{(u-d)(u-m)}$, $q_2 = \frac{\gamma(u+d)-ud-\gamma^2\zeta}{(u-m)(m-d)}$, $q_3 = \frac{um-\gamma(u+m)+\gamma^2\zeta}{(u-d)(m-d)}$.
Quadrinomial lattice (two factors)		
Boyle-TRIN2	$u_1 = e^{\sigma_1\sqrt{\Delta t}}$, $d_1 = e^{-\sigma_1\sqrt{\Delta t}}$, $m_1 = 1, m_2 = 1$, $u_2 = e^{\sigma_2\sqrt{\Delta t}}$, $d_2 = e^{-\sigma_2\sqrt{\Delta t}}$.	$q_1 = \frac{1}{4}\left(1+\rho+\sqrt{\Delta t}\left(\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} + \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2}\right)\right)$, $q_2 = \frac{1}{4}\left(1-\rho+\sqrt{\Delta t}\left(\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} - \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2}\right)\right)$, $q_3 = \frac{1}{4}\left(1-\rho+\sqrt{\Delta t}\left(-\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} + \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2}\right)\right)$, $q_4 = \frac{1}{4}\left(1+\rho+\sqrt{\Delta t}\left(-\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} - \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2}\right)\right)$.
Pentanomial lattice (two factors)		
KR-TRIN2	$u_1 = e^{\lambda\sigma_1\sqrt{\Delta t}}$, $d_1 = e^{-\lambda\sigma_1\sqrt{\Delta t}}$, $m_1 = 1, m_2 = 1$, $u_2 = e^{\lambda\sigma_2\sqrt{\Delta t}}$, $d_2 = e^{-\lambda\sigma_2\sqrt{\Delta t}}$.	$q_1 = \frac{1}{4}\left(\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} + \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2} + \frac{\rho}{\lambda^2}\right)\right)$, $q_2 = \frac{1}{4}\left(\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} - \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2} - \frac{\rho}{\lambda^2}\right)\right)$, $q_3 = \frac{1}{4}\left(\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(-\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} - \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2} + \frac{\rho}{\lambda^2}\right)\right)$, $q_4 = \frac{1}{4}\left(\frac{1}{\lambda^2} + \frac{\sqrt{\Delta t}}{\lambda}\left(-\frac{r-\frac{1}{2}\sigma_1^2}{\sigma_1} + \frac{r-\frac{1}{2}\sigma_2^2}{\sigma_2} - \frac{\rho}{\lambda^2}\right)\right)$, $q_5 = 1 - \frac{1}{\lambda^2}$.

Note: CRR [Cox et al., 1979]; Tian-BIN, Tian-TRIN [Tian, 1993]; Haahtela-BIN [Haahtela, 2010]; Boyle-TRIN [Boyle, 1988]; Boyle-TRIN2 [Boyle et al., 1989]; KR-TRIN, KR-TRIN2 [Kamrad and Ritchken, 1991].

many conditions and need to be estimated independently before building up the price lattice. For the purpose of this study, a direct censor on transition probabilities of each node would be more efficient, as proposed by Nelson and Ramaswamy [1990]. Our suggested pricing method in Chapter 5 is based on this concept.

Chapter 3

Optimal Pricing and Allocation of Display Inventories

This chapter discusses an optimal dynamic model for unifying programmatic guarantee (PG) and real-time bidding (RTB) in display advertising. Section 3.1 introduces the background and the overview of this study. Section 3.2 formulates the problem, discusses our assumptions and provides a solution. Section 3.3 presents the results of our experimental evaluation and Section 3.4 summarises the chapter.

3.1 Introduction

Over the last few years, the demand for automation, integration and optimisation has been the key driver for making online advertising one of the fastest advancing industries. In display advertising, a significant development is the emergence of RTB, which allows buying and selling display impressions in real-time and even a single impression at a time [Google, 2011, Yuan et al., 2013]. Yet, despite the strong growth of RTB, according to eMarketer [2013], 75% of publishers' revenue in 2012 still came from 20% guaranteed inventories, which were mainly sold through direct sales by negotiation.

Guaranteed inventories stand for the guaranteed contracts sold by top tier websites. Generally, they are [Dunaway, 2012, OpenX, 2013]:

- Highly viewable because of good position and size;
- Rich in the first-party data for behaviour targeting;
- Flexible in format, size, device, etc.;
- Audited content for brand safety.

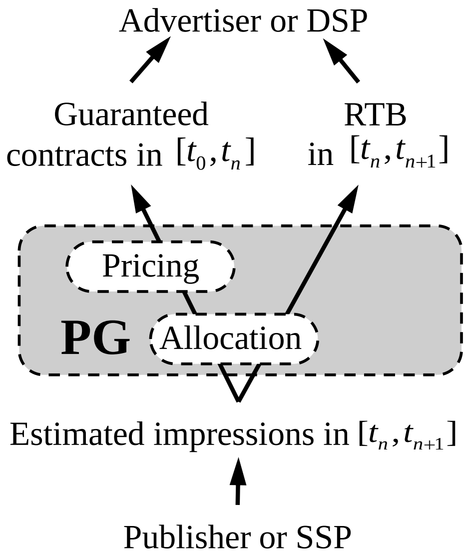

Figure 3.1: Systematic view of PG and RTB in display advertising, where $[t_0, t_n]$ is the time period that a publisher (or SSP) sells the guaranteed impressions which will be created in future period $[t_n, t_{n+1}]$.

Therefore, it is not surprising that guaranteed inventories are normally sold in bulk at high prices in advance than those sold on the spot market (i.e., RTB).

Programmatic guarantee (PG), sometimes called programmatic reserve or programmatic premium [Dunaway, 2012, OpenX, 2013], is a new concept that has gained much attention recently. Notable examples of some early services on the market are `iSOCKET.com`, `BuySellAds.com` and `ShinyAds.com`. It is essentially an allocation and pricing engine for publishers or SSPs that brings the automation into the selling of guaranteed inventories apart from RTB. Figure 3.1 illustrates how PG works for a publisher (or SSP) in display advertising. For a specific ad slot or user tag, the estimated total impressions in a future period can be evaluated and allocated algorithmically at the present time between the guaranteed market and the spot market. Impressions in the former are sold in advance via guaranteed contracts until the delivery date while in the latter are auctioned off in RTB. Unlike the traditional way of selling guaranteed contracts, there is no negotiation process between publisher and advertiser. The guaranteed price (i.e., the fixed per impression price) will be listed in ad exchanges dynamically like the posted stock price in financial exchanges. Advertisers (or DSPs) can see a guaranteed price at a time, monitor the price changes over time and purchase the needed impressions directly at the corresponding guaranteed prices a few days, weeks or months earlier before the delivery date.

Developing a revenue maximisation model for the programmatic guarantee is sophisticated and challenging. We need to solve the problem of selling unstorable impressions in advance. Similar problems have been studied in many other industries. Examples include retailers selling fashion and seasonal goods and airline companies selling flight tickets [Talluri and van Ryzin, 2005]. However, in display advertising, impressions are with uncertain salvage values because they can be auctioned off in real-time on the delivery date. The combination with RTB makes our work interesting and novel.

Several economic assumptions are made in our study. We consider that future supply and demand of impressions from an ad slot (or user tag) can be well estimated and assume that advertisers' purchase behaviour of guaranteed contracts are determined by both the guaranteed price and the time interval between the purchase time and the impression delivery date. For RTB, we consider the sealed-bid second price auction and discuss both probabilistic and empirical distributions of advertisers' bids. Under the above assumptions, an algorithmic framework is developed which gives out a functional form of the dynamic optimal price and computes the optimal amount of future impressions to sell in advance.

The development of this chapter is evaluated with two RTB datasets. Advertisers bidding behaviours in RTB are investigated and we find that the developed model adopts different strategies in pricing and allocation of impressions according to the level of competition on the spot market. If the spot market in future is less competitive, a small amount of impressions would be sold via guaranteed contracts at low prices. The maximised revenue is mainly contributed by the spot market because there is a significant growth in the expected price of auctions in the future. In a highly competitive market, the model allocates more future impressions into guaranteed contracts at high prices and the maximised revenue mainly comes from guaranteed selling. Under either situation, the revenue can be maximised successfully.

3.2 The Model

We consider there is a premium ad slot on a publisher's webpage. If there is a user comes to this webpage, the ad slot can generate a chance of ad view, usually referred

Table 3.1: Summary of key notations in Chapter 3.

Notation	Description
$t_0, \dots, t_{n+1}$	Discrete time points: $[t_0, t_n]$ is the period to sell the guaranteed impressions; $[t_n, t_{n+1}]$ is the period that the estimated impressions should be created, auctioned off (in RTB) and delivered.
$t \in [0, T]$	Continuous time where $t_0 = 0$, $t_n = T$.
$\tau \in [0, T]$	Remaining time to the impression delivery period, where $\tau = T - t$.
Q	Estimated total demanded impressions for the ad slot in $[t_n, t_{n+1}]$.
S	Estimated total supplied impressions for the ad slot in $[t_n, t_{n+1}]$.
$p(\tau)$	Guaranteed price to sell an impression when the remaining time till the delivery period is τ.
$\theta(\tau, p(\tau))$	Proportion of those who want to purchase an impression in advance at τ and at $p(\tau)$.
$f(\tau)$	Density function so that the number of those who want to purchase in advance in $[\tau, \tau + d\tau]$ is $f(\tau)d\tau$.
ω	Probability that the publisher fails to deliver a guaranteed impression in the delivery period.
κ	Size of penalty: if the publisher fails to deliver a guaranteed impression that is sold at $p(\tau)$, he needs to pay $\kappa p(\tau)$ penalty to the advertiser.
ξ	Number of advertisers who need an impression in RTB.
$\phi(\xi)$	Expected payment price of an impression in RTB for the given ξ.
$\psi(\xi)$	Expected risk of an impression in RTB for the given ξ.
λ	Level of risk aversion for advertisers.
$\pi(\xi)$	Expected winning bid of an impression in RTB for the given ξ.

to as an impression. In RTB, an impression is auctioned off simultaneously once a user comes and the winning bidder (i.e., the advertiser) has his ad displayed to the user [Google, 2011, Yuan et al., 2013]. Suppose that the publisher can estimate supply and demand of impressions from an ad slot (or user tag[1]) from historical transactions and plan to sell some of the future impressions via guaranteed contracts in advance in order to maximise the revenue. We consider an environment that is risk-averse and both publisher and advertiser make their strategies by maximizing their expected utilities [Bhalgat et al., 2012]. In other words, the advertiser is willing to pay a higher price for a fixed number of future impressions if the delivery is guaranteed. This gives the publishers an additional possibility of increasing their revenue by pre-selling some future impressions, apart from the price discrimination over time.

[1]Group of ad slots which target specific types of users.

3.2.1 Problem Formulation

The optimisation problem can be expressed as follows:

$$\max \left\{ \underbrace{\int_0^T (1-\omega\kappa)p(\tau)\theta(\tau,p(\tau))f(\tau)d\tau}_{\substack{G = \text{Expected total revenue from guaranteed selling minus} \\ \text{expected penalty of failling to delivery}}} + \underbrace{\left(S - \int_0^T \theta(\tau,p(\tau))f(\tau)d\tau\right)\phi(\xi)}_{H = \text{Expected total revenue from RTB}} \right\}, \tag{3.1}$$

$$\textbf{s.t.} \quad p(0) = \begin{cases} \phi(\xi) + \lambda\psi(\xi), & \text{if } \pi(\xi) \geq \phi(\xi) + \lambda\psi(\xi) \\ \pi(\xi), & \text{if } \pi(\xi) < \phi(\xi) + \lambda\psi(\xi), \end{cases} \tag{3.2}$$

where

$$\xi = \frac{\text{Remaining demand in } [t_n, t_{n+1}]}{\text{Remaining supply in } [t_n, t_{n+1}]} = \frac{Q - \int_0^T \theta(\tau,p(\tau))f(\tau)d\tau}{S - \int_0^T \theta(\tau,p(\tau))f(\tau)d\tau}.$$

The notations are given in Table 3.1. The publisher's expected total revenue contains: the expected revenue from guaranteed impressions sold during $[0, T]$; the expected penalty of failing to delivery guaranteed impressions in $[t_n, t_{n+1}]$; the expected revenue from RTB in $[t_n, t_{n+1}]$; and the price constraint that ensures the advertisers' willingness to buy guaranteed impressions. Eq. (3.2) shows that an advertiser's decision of buying either a guaranteed or non-guaranteed impression depends on the expected payment price and his level of risk-aversion. For simplicity and without loss of generality, each guaranteed impression is considered as a single guaranteed contract. This setting can be extended to a bulk sale in practice.

The solution to the above optimisation problem appears a bit complicated as it needs to answer how many future impressions to sell and at what prices to sell. Before discussing the solution, several assumptions, such as the distribution of bids in RTB and the advertisers' purchase behaviour in advance, have to be made.

3.2.2 Distribution of Bids in RTB

Advertisers bid for individual impressions separately in RTB [Google, 2011, Yuan et al., 2013]. Therefore, the following second-price auction is considered: for a single impression from a specific ad slot (or user tag), advertisers submit sealed bids to the publisher (or SSP), and the highest bidder wins the impression but finally pays at the bid next to him.

Either probabilistic or empirical distribution of bids in RTB can be discussed. Bidders are assumed to be symmetric in probabilistic method. Therefore, advertisers would reveal their preference and truthfully offer bids [Edelman et al., 2007, Varian, 2007, Narahari et al., 2009]. In this research, we adopt the settings used by Lahaie and Pennock [2007] and Ostrovsky and Schwarz [2011], where bids are assumed to follow a log-normal distribution, denoted by $X \sim \mathbf{LN}(\mu, \sigma^2)$. Then, the expected per impression payment price from a second-price auction is

$$\phi(\xi) = \int_0^\infty x\xi(\xi-1)g(x)\Big(1-F(x)\Big)\Big(F(x)\Big)^{\xi-2}dx, \tag{3.3}$$

where $g(x)$ and $F(x)$ are the log-normal density function and its cumulative distribution function, respectively, given by

$$g(x) = \frac{1}{x\sigma\sqrt{2\pi}}\, e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}}, \tag{3.4}$$

$$F(x) = \frac{1}{2} + \frac{1}{\sqrt{\pi}}\int_0^{\frac{\ln(x)-\mu}{\sqrt{2}\sigma}} e^{-z^2}dz, \tag{3.5}$$

so that $\xi(\xi-1)g(x)(1-F(x))(F(x))^{\xi-2}$ represents the probability that if an advertiser is the second highest bidder, then one of the $\xi-1$ other advertisers must bid at least as much as he does and all of the $\xi-2$ other advertisers have to bid no more than he does. We can check if the bids follow the log-normal distribution by the Kolmogorov-Smirnov (K-S) [Smirnov, 1948] and the Jarque-Bera (J-B) [Jarque and Bera, 1980] statistics (see Table 3.6). Once the log-normal distribution is met, $\phi(\xi)$ can be estimated numerically because the values of $g(x)$ and $F(x)$ in each integration increment can be calculated.

If the bids do not follow the log-normal distribution, empirical methods can be used to compute $\phi(\xi)$. Simply, for an ad slot (or user tag), the winning payment prices

Algorithm 3.1 Estimation of $\phi(\xi)$ by using the robust locally weighted regression (RLWR) method [Cleveland, 1979].

function RLWRSOLVE(ξ)
// (ξ_j, ϕ_j), $j = 1, \ldots, m$, are the learning data with size m.
$\widehat{\boldsymbol{\beta}}(\xi) \leftarrow \text{argmin} \sum_{j=1}^{m} \varpi_j(\xi)(\phi_j - \beta_0 - \beta_1\xi_j - \ldots - \beta_d \xi_j^d)^2$,
where $\varpi_j(\xi) \leftarrow \begin{cases} \left(1 - \left|\frac{\xi-\xi_j}{h(\xi)}\right|^3\right)^3, & \text{if } \left|\frac{\xi-\xi_j}{h(\xi)}\right| < 1, \\ 0, & \text{if } \left|\frac{\xi-\xi_j}{h(\xi)}\right| \geq 1. \end{cases}$
// $h(\xi)$ is the distance from ξ to the most distant neighbour of ξ within the span
// we choose $d = 2$.
$\widehat{\phi} \leftarrow \sum_{k=0}^{d} \widehat{\beta}_k(\xi)\xi^k$.
loop $i \leftarrow 1$ to $\hat{n}$ // repeat the update in $\hat{n}$ iterations
$\boldsymbol{\epsilon} \leftarrow \boldsymbol{\phi} - \widehat{\phi}$, $\chi(\boldsymbol{\epsilon}) \leftarrow$ median($|\ \boldsymbol{\epsilon}\ |$).
for $j \leftarrow 1$ to m **do**
$\varpi_j(\xi) \leftarrow \begin{cases} \left(1 - \left(\frac{e_j}{6\chi(\boldsymbol{\epsilon})}\right)^2\right)^2, & \text{if } |\ \epsilon_j\ | < 6\chi(\boldsymbol{\epsilon}), \\ 0, & \text{if } |\ \epsilon_j\ | \geq 6\chi(\boldsymbol{\epsilon}). \end{cases}$
end for
$\widehat{\boldsymbol{\beta}}(\xi) \leftarrow \text{argmin} \sum_{j=1}^{m} \varpi_j(\xi)(\phi_j - \beta_0 - \beta_1\xi_j - \ldots - \beta_d \xi_j^d)^2$.
$\widehat{\phi} \leftarrow \sum_{k=0}^{d} \widehat{\beta}_k(\xi)\xi^k$.
end loop
return $\phi(\xi) \leftarrow \widehat{\phi}$.
end function

are trained to develop a regression model that explains their correlation to the level of demand. Here we use the robust locally weighted regression (RLWR) method [Cleveland, 1979] (see Algorithm 3.1 and an empirical example in Section 3.3.4). Other statistical learning and forecasting methods can be developed to estimate $\phi(\xi)$, but they are not further discussed here.

3.2.3 Risk Aversion and Purchase Behaviour

Eq. (3.2) tells that at time T an advertiser's decision between guaranteed and non-guaranteed channels are indifferent. In this research, the advertisers' arrival is not modelled as a stochastic process [Gallego and van Ryzin, 1994], instead, we consider that the total demand for future impressions is deterministic but can be shift from future to present. The possibility of this shift is because advertisers are assumed to be risk-averse.

Under our risk aversion settings, $\pi(\xi)$ and $\psi(\xi)$ can be estimated by the RLWR method, and λ can be set as any non-negative number. First, the estimation of $\pi(\xi)$ is as same as Algorithm 3.1 while we consider the highest bids (per transaction) rather than the payment prices (per transaction). Second, the estimation of $\psi(\xi)$ is slightly different. We compute a series of standard deviations of daily winning payment prices and use Algorithm 3.1 to compute $\psi(\xi)$ for the given demand level. Third, advertisers' risk-averse preference are not same; therefore, λ can be regarded as the average risk-aversion level of all advertisers or of key advertisers (we consider the former in the experiments). The larger λ the more risk-averse advertisers are. More detailed discussion about the estimation of $\pi(\xi)$, $\psi(\xi)$ and λ is given in Section 3.3.4.

Similar to flight tickets booking [Anjos et al., 2004, 2005, Malighetti et al., 2009], we have the following two economic assumptions on demand:

Assumption 3.1 *Demand is negatively correlated with guaranteed price as advertisers would buy less impressions if price increases. Given* τ *and* $0 \leq p_1 \leq p_2$*, then* $\theta(\tau, p_1) \geq \theta(\tau, p_2)$*, subject to the boundary condition* $\theta(\tau, 0) = 1$.

Assumption 3.2 *Demand is negatively correlated with the time interval between purchase and delivery because more advertisers' would want to buy impressions when the delivery date is approached. Given* p *and* $0 \leq \tau_2 \leq \tau_1$*, then* $\theta(\tau_2, p) \geq \theta(\tau_1, p)$.

We adopt the functional forms of demand proposed by Anjos et al. [2004] (which were used in flight tickets booking):

$$\theta(\tau, p(\tau)) = e^{-\alpha p(\tau)(1+\beta\tau)}, \tag{3.6}$$

$$f(\tau) = \zeta e^{-\eta\tau}, \tag{3.7}$$

where α is the level of price effect, β and η are the levels of time effect, and the demand density rises to a peak ζ on the delivery date. Therefore, $f(\tau)d\tau$ is the number who would like to purchase in advance, and $\theta(\tau, p(\tau))$ is the proportion of those who want to purchase an impression in advance at time $T - \tau$ and at price $p(\tau)$.

3.2.4 Optimal Dynamic Prices

The optimisation problem in Eq. (3.1) can be solved by Algorithm 3.2. We simulate many values of $\gamma_i \in [0, 1], i = 1, \ldots, m$. For each given γ_i, we solve the optimisation problem in Eq. (3.8), find the optimal series of guaranteed prices, and calculate the optimal total revenue R_i. Then, in the global comparison, we can find the optimal γ^* that generates the maximum value of total revenue.

Algorithm 3.2 Solution to Eq. (3.1).

function $\text{PGS}\textsc{olve}(\alpha, \beta, \zeta, \eta, \omega, \kappa, \lambda, S, Q, T)$
  $\boldsymbol{t} \leftarrow [t_0, \cdots, t_n], 0 = t_0 < t_1 < \cdots < t_n = T.$
  $\boldsymbol{\tau} \leftarrow T - \boldsymbol{t}, m \leftarrow 500.$
  loop $i \leftarrow 1$ to m
    $\gamma_i \leftarrow \texttt{RandomUniformGenerate}([0, 1])$
    $\int_0^T \theta(\tau, p(\tau)) f(\tau) d\tau \leftarrow \gamma_i S$
    $\xi_i \leftarrow (Q - \gamma_i S)/(S - \gamma S)$
    $H_i \leftarrow (1 - \gamma_i) S \phi(\xi_i)$
    $G_i \leftarrow \int_0^T (1 - \omega\kappa) p(\tau) \theta(\tau, p(\tau)) f(\tau) d\tau$

$$\boldsymbol{p}_i \leftarrow \mathbf{arg\,max}\, G_i, \tag{3.8}$$

$$\textbf{s.t.} \int_0^T \theta(\tau, p(\tau)) f(\tau) d\tau = \gamma_i S, \tag{3.9}$$

$$p(0) = \begin{cases} \phi(\xi_i) + \lambda\psi(\xi_i), \text{if } \pi(\xi_i) \geq \phi(\xi_i) + \lambda\psi(\xi_i), \\ \pi(\xi_i), \text{if } \pi(\xi_i) < \phi(\xi_i) + \lambda\psi(\xi_i). \end{cases} \tag{3.10}$$

    $R_i \leftarrow \mathbf{max} G_i + H_i$
  end loop
  $\gamma^* \leftarrow \mathbf{arg\,max}_{\gamma_i \in \Omega(\gamma)} \{R_1, \ldots, R_m\}$
  $\boldsymbol{p}^* \leftarrow \mathbf{arg\,max}_{\boldsymbol{p}_i \in \Omega(\boldsymbol{p})} \{R_1, \ldots, R_m\}$
  return $\gamma^*, \mathbf{p}^*$
end function

Let us discuss how to solve the optimisation problem in Eq. (3.8). We consider the following Lagrangian:

$$\mathcal{L}(\widetilde{\lambda}, p(\tau)) = \int_0^T (1 - \omega\kappa) p(\tau) \theta(\tau, p(\tau)) f(\tau) d\tau + \widetilde{\lambda}\bigg(\gamma_i S - \int_0^T \theta(\tau, p(\tau)) f(\tau) d\tau\bigg), \tag{3.11}$$

where $\widetilde{\lambda}$ is the Lagrange multiplier. The Euler-Lagrange condition is $\partial\mathcal{L}/\partial p = 0$. For $\tau \in (0, T]$, we have

$$(1 - \omega\kappa)\theta(\tau, p(\tau)) + \Big((1 - \omega\kappa)p(\tau) - \widetilde{\lambda}\Big)\frac{\partial\theta(\tau, p(\tau))}{\partial p(\tau)} = 0. \tag{3.12}$$

Substituting Eq. (3.6) into Eq. (3.12) then gives the formula of the optimal guaranteed price:

$$p(\tau) = \frac{\widetilde{\lambda}}{1 - \omega\kappa} + \frac{1}{\alpha(1 + \beta\tau)}. \tag{3.13}$$

Consider a small time step $d\widetilde{\tau}$, then in $[0, 0 + d\widetilde{\tau}]$, there are $\theta(0, p(0)) f(0) d\widetilde{\tau}$ demand fulfilled. Therefore, we have

$$\int_{d\widetilde{\tau}}^{T} \theta(\tau, p(\tau)) f(\tau) d\tau = \gamma_i S - \theta(0, p(0)) f(0) d\widetilde{\tau} \tag{3.14}$$

By substituting Eqs. (3.2) and Eqs. (3.6)-(3.7) into Eq. (3.14), we have

$$\frac{-\zeta(1-\omega\kappa)e^{-\left(\frac{\alpha\widetilde{\lambda}}{1-\omega\kappa}+1\right)}}{\alpha\widetilde{\lambda}\beta+(1-\omega\kappa)\zeta}\left(e^{-\left(\frac{\alpha\widetilde{\lambda}\beta}{1-\omega\kappa}+\eta\right)T} - e^{-\left(\frac{\alpha\widetilde{\lambda}\beta}{1-\omega\kappa}+\eta\right)d\widetilde{\tau}}\right) = \gamma_i S - e^{-\alpha p(0)} d\widetilde{\tau}. \tag{3.15}$$

Eq. (3.15) shows that the value of $\widetilde{\lambda}$ is dependent on $\gamma_j S$ and other parameters. However, the explicit solution of $\widetilde{\lambda}$ cannot be deduced. The value of $\widetilde{\lambda}$ can be estimated by using numerical methods, e.g. the Newton-Raphson method. Eq. (3.13) can then be rewritten as follows

$$p(\tau) = \frac{\widetilde{\lambda}(\alpha, \beta, \zeta, \eta, \omega, \kappa, \gamma_i S)}{1-\omega\kappa} + \frac{1}{\alpha(1+\beta\tau)}. \tag{3.16}$$

The notation $\widetilde{\lambda}(\alpha, \beta, \zeta, \eta, \omega, \kappa, \gamma_i S)$ represents the dependency relationship among $\widetilde{\lambda}$ and other parameters. Figure 3.2 gives a numerical investigation on the relationships between $p(\tau)$ and model parameters. Recall that in Eqs. (3.6)-(3.7) a large value of α means advertisers are price sensitive; therefore, $p(\tau)$ decreases if α increases. Similar negative correlations are with β and η. These two parameters describe the time effect on advertisers' willingness to purchase. The model thus encourages advertisers to purchase in advance by selling guaranteed contracts at low prices. Conversely, the optimal price is positively correlated with ζ because the parameter shows the maximum number of advertisers that would be willing to buy guaranteed impressions at a time point. More advertisers means more competition; therefore, more advertisers would purchase in advance in order to secure the targeted impressions. In such a situation, the model gives out high guaranteed prices and allocates more impressions to guaranteed contracts. While the expected penalty $\omega\kappa$ has less effect on price, the larger $\omega\kappa$ the higher $p(\tau)$. It is worth noting that ω and κ are considered as given parameters because: (i) κ can be set by negotiation between publisher and advertiser; (ii) ω can be estimated[2] and

[2] ω can be approximated by the percentage of guaranteed impressions that the publisher failed to deliver.

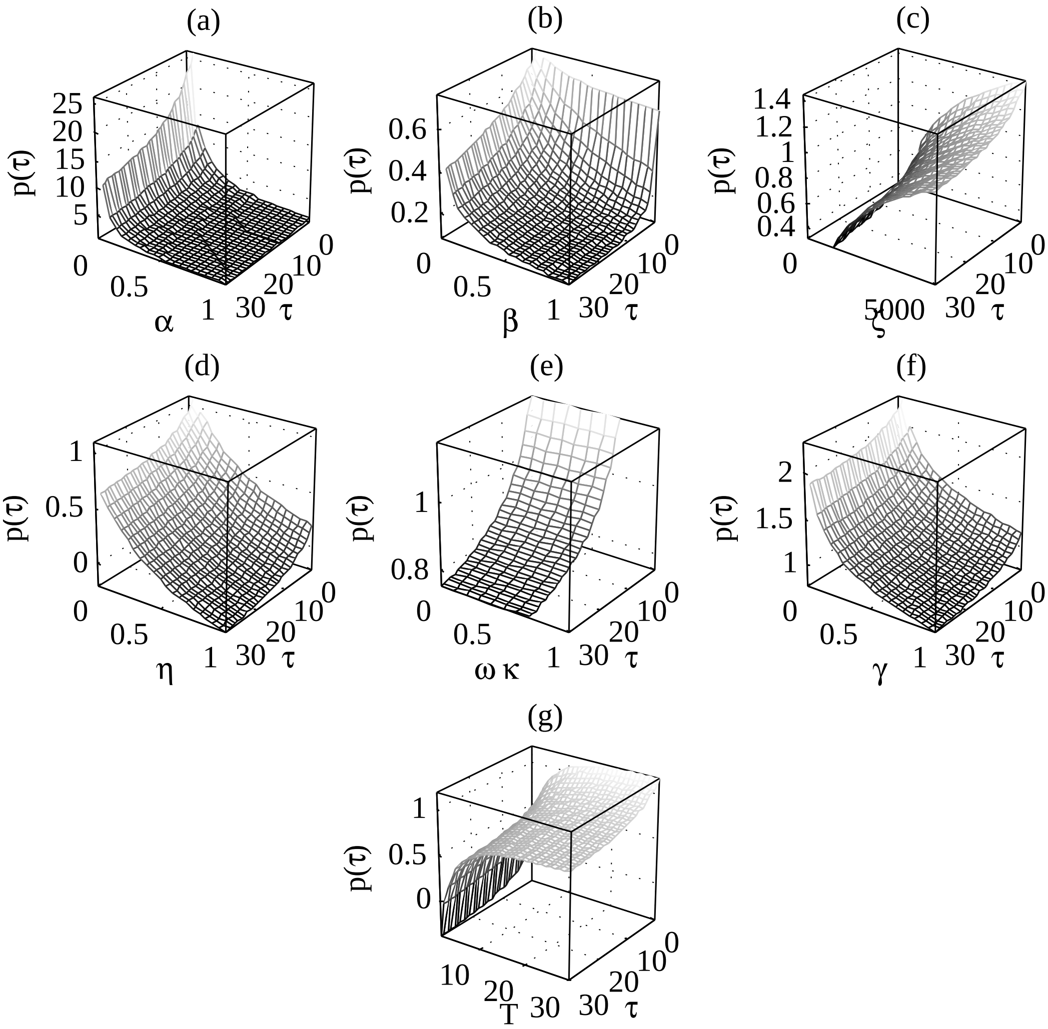

Figure 3.2: The impact of model parameters on the guaranteed selling prices: α, β, ζ, η are defined in Eqs. (3.6) & (3.7); $\omega\kappa$ is the expected size of penalty; γ is the percentage of estimated future impressions to sell in advanced; T is the length of guaranteed selling period; τ is the remaining time to the delivery date; $p(\tau)$ is the guaranteed selling price at τ.

updated once the PG system runs for a certain period of time. Here we set $\omega = 0.05$, $\kappa = 1$. With less and less supplied impressions to sell on the market, the price $p(\tau)$ increases. The total length of time period to sell guaranteed contracts positively affects the guaranteed price, the longer T, the higher the $p(\boldsymbol{\tau})$.

3.3 Experiments

We describe our datasets in Section 3.3.1, investigate the RTB campaigns in Sections 3.3.2-3.3.3, discuss the estimation of model parameters in Sections 3.3.4-3.3.5, and evaluate the performance of revenue maximisation in Section 3.3.6.

Table 3.2: Summary of RTB datasets.

Dataset	SSP	DSP
Time period (date)	08/01/2013-14/02/2013	19/10/2013-27/10/2013
No. of ad slots	31	53571
No. of user tags	NA	69
No. of advertisers	374	4
No. of impressions	6646643	3158171
No. of bids	33043127	11457419
Bid quote	GBP/CPM	CNY/CPM

Table 3.3: Experimental design of the SSP dataset.

	Time period (date)
Training set	08/01/2013-13/02/2013
Development set	08/01/2013-14/02/2013
Test set	14/02/2013

3.3.1 Datasets

We use two different RTB datasets: one from a medium-sized SSP in the UK and the other from a DSP in China. Table 3.2 shows a brief summary of these two datasets. The SSP dataset is used throughout the whole experiments while the DSP dataset is used for further exploring advertisers' strategies in RTB. In these two datasets, all the bids are expressed as CPM.

Table 3.3 illustrates our experimental design. The SSP dataset is divided into one training set, one development set and one test set. In the training set, we investigate RTB campaigns and estimate model parameters. In the development set, we use the discussed model to allocate and price the impressions that are created on 14/02/2013. Guaranteed contracts are sold over the period from 08/01/2013 to 13/02/2013 and the rest impressions are auctioned off on the delivery date 14/02/2013. In the development set, we simulate the transactions of guaranteed contracts and calculate the remaining campaigns of RTB on 14/02/2013. The test set contains the actual bids and winning payment prices of 14/02/2013, which is used to evaluate the revenue maximisation performance. Note that time periods of training and development sets can be different. For example, the development period can be a few days/weeks later than the training period. However, this requires a number of forecasting methods to estimate all the model parameters (features). As our primary intention here is not to discuss better forecasting methods, we choose a learning period that is close to the impression delivery date so

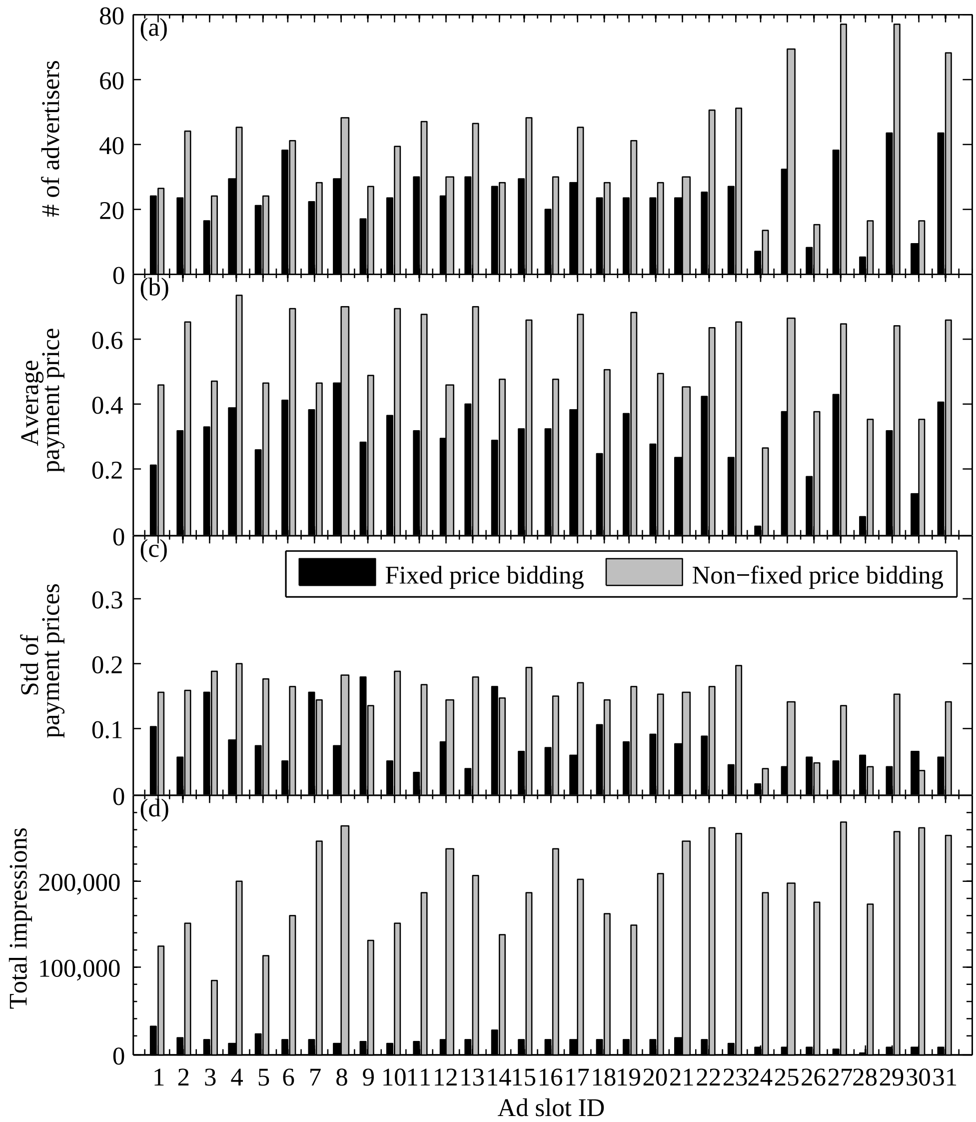

Figure 3.3: Overview of statistics for the winning advertisers from the SSP dataset in the training period.

that the learned parameters are more accurate for the evaluation purpose.

3.3.2 Bidding Behaviours

We first examine if selling guaranteed impressions in advance can be a viable way to segment advertisers according to their bids, and then discuss how much of revenue growth can be expected.

Let us first look at advertisers behaviours in RTB. From the SSP dataset, we find that advertisers mainly join auctions in the morning from 6am to 10am. It is the time period that supplied impressions arrive peak. We also find that the winning advertisers' final payments are much less than their bids. Figure 3.3 provides some descriptive

Table 3.4: Summary of winning advertisers' statistics from the SSP dataset in the training period: the numbers in the brackets represent how many advertisers who use the combined bidding strategies.

Bidding strategy	No. of advertisers	No. of impressions	Average change rate of payment prices	Ratio of payment price to winning bid
Fixed price	188 (51)	454681	188.85%	43.93%
Non-fixed price	200 (51)	6068908	517.54%	58.94%

Table 3.5: Summary of advertisers' winning campaigns from the DSP dataset. All advertisers use the fixed price bidding strategy. Each user tag contains many ad slots and an ad slot is sampled from the dataset only if the advertiser wins more than 1000 impressions from it.

Advertiser ID	No. of user tags	No. of ad slots	No. of impressions won	Average change rate of payment prices	Ratio of payment price to winning bid
1	69	635	196831	58.57%	36.07%
2	69	428	144272	58.94%	34.68%
3	69	1267	123361	79.24%	30.89%
4	65	15	3139	104.19%	22.32%

statistics about this finding across all 31 ad slots. Winning advertisers are divided into two groups. The first group contains those who always offer a fixed bid; the second group contains those who frequently change their bids. Figure 3.3 shows that more winning advertisers adopt the non-fixed price bidding in RTB. They intend to offer higher bids on each impression, endure more variance in payment prices due to the second price auction, and obtain more impressions. The second price auction in RTB provides an opportunity for making more revenue by selling impressions in advance: (i) a risk-averse advertiser is willing to buy in advance to lock in the price; (ii) the publisher would be able to increase the price for the guaranteed contracts by charging advertisers their private valuations rather than the second price bids. The question is how big the difference between the top bids and actually payments (the second price). Table 3.4 shows that the publisher can expect 100% increase in revenue because the current average ratio of actual payment price (the second price) to winning bid (the first price) is about 50%.

We further examine the DSP dataset, and find all four advertisers use the fixed price strategy in their bidding. This might be because the DSP itself adopts the fixed price strategy for these 4 advertisers. While the DSP dataset itself is biased, we can still take a look at the average volatility of the advertisers's payment prices and the average ratio of payment price to winning bid. The DSP dataset shows that advertisers actually bid for user tags instead of specific ad slots. Each user tag contains a set of ad slots that have similar features so that advertisers are able to target a certain group of users. Therefore, RTB is an auction mechanism with user targeting. Consider in a user tag the advertiser's bids are not well distributed among ad slots, we only investigate the ad slots where the advertiser wins more than 1000 impressions. Table 3.5 confirms our earlier statement from a buy side perspective. Even using the fixed price bidding strategy, advertisers' payment prices are volatile (more than 50% from each impression). In fact, these advertisers can afford more to reduce the risk because the current payment prices are much lower than their private valuations (around 30% across 4 advertisers).

3.3.3 Supply and Demand

Figure 3.4 presents some descriptive statistics about supply and demand of all 31 ad slots from the SSP dataset in the training set. The ad slots have the same daily supply levels as well as their upper and lower bounds. However, the levels of daily demand are significantly different: `AdSlot25`, `AdSlot27`, `AdSlot29` and `AdSlot31` are in higher demand than others, about 9 bidders per impression while the rest ad slots have the average value around 5. As shown in Figure 3.5, we take the average distance [Han et al., 2011] in ξ as the metric to cluster ad slots and obtain two groups. Note that ξ significantly deviates from its mean value in a day's period because many more advertisers join RTB at peak time from 6am to 10am, as shown in Figure 3.6. In these hours, ξ is 118.96% higher than other hours. We can develop regression or time-series models to estimate Q and S on delivery date; however, this is not a significant part of our study so we consider them as given parameters.

3.3.4 Bids and Payment Prices

Once ξ is given, either probabilistic or empirical models can be used to estimate the corresponding payment price $\phi(\xi)$ in RTB. In probabilistic models, bidders are assumed to be symmetric, whose bids follow a log-normal distribution. However, the distribution tests shown in Table 3.6 reveal the fact that actual bids in RTB are not log-normally distributed. This confirms the statement that advertisers in the real-world environment are not symmetric. They may frequently change their bids for unclear reasons. Therefore, we use the empirical method to estimate $\phi(\xi)$ as well as $\psi(\xi)$ and $\pi(\xi)$.

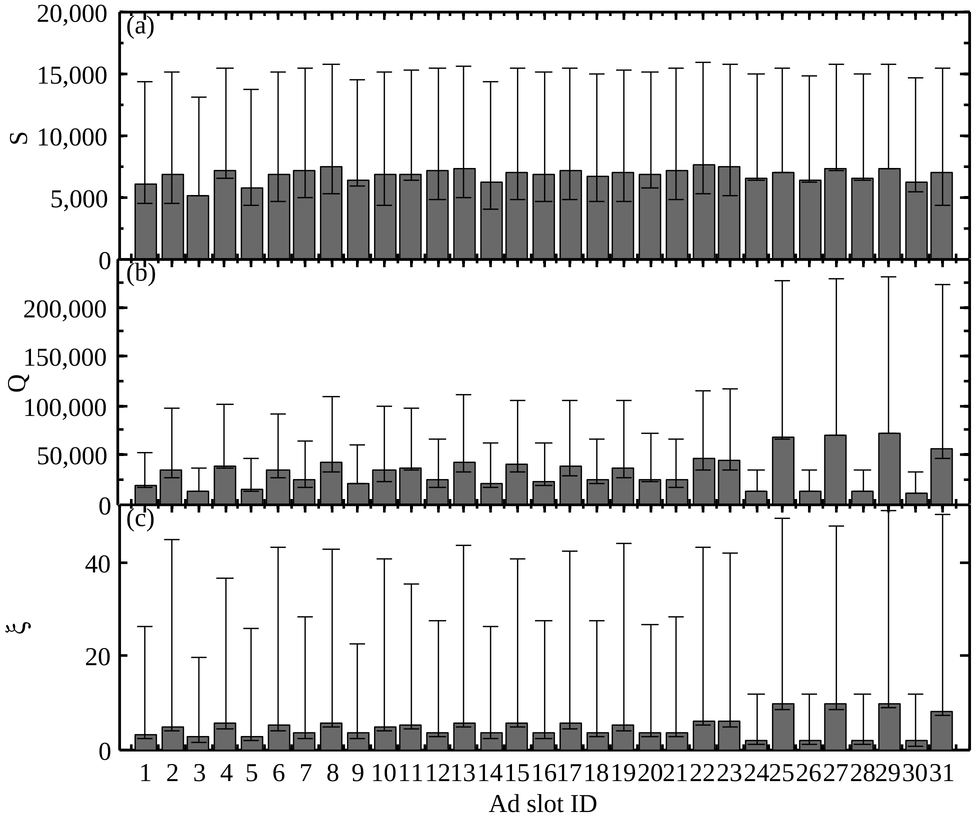

Figure 3.4: Overview of daily supply and demand of ad slots in the SSP dataset in the training period: S is the number of total supplied impressions; Q is the number of total demand impressions; ξ is the per impression demand (i.e., the number of advertisers who bid for an impression).

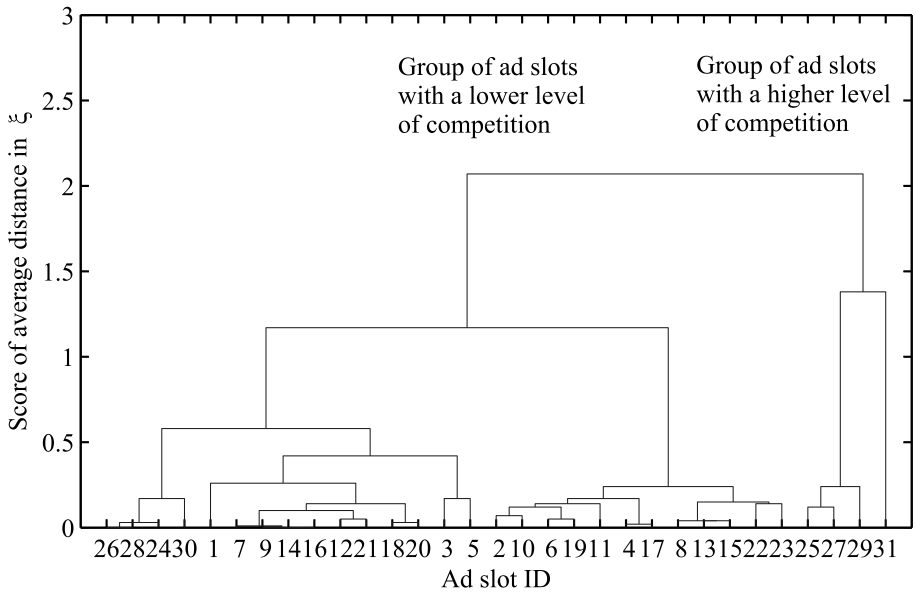

Figure 3.5: Hierarchical cluster tree of ad slots in the SSP dataset where the cluster metric is average distance in the per impression demand ξ.

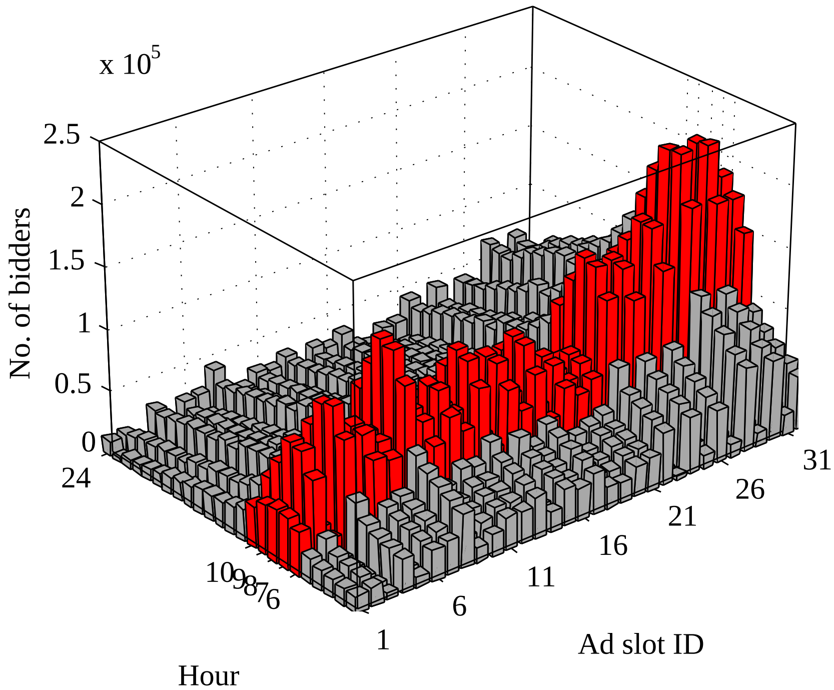

Figure 3.6: Overview of advertisers' hourly arrival per day, where the red shaded bar represents the peak hour.

Table 3.6: Summary of bids distribution tests, where the numbers in the Kolmogorov-Smirnov (K-S) and Jarque-Bera (J-B) tests represent the percentage of tested auctions that have lognormal bids.

Group of ad slots	No. of auctions where $\xi \geq 30$	K-S test	J-B test
Low competition	286	0.00%	0.00%
High competition	15702	0.00%	0.00%

Table 3.7: Comparison of estimations between the empirical distribution and the actual bids: $\phi(\xi)$ is the expected payment price; $\psi(\xi)$ is the standard deviation of payment prices; $\pi(\xi)$ is the expected winning bid; and the per impression demand $\xi = Q/S$.

Group of ad slots	Difference in $\phi(\xi)$	Difference in $\psi(\xi)$	Difference in $\pi(\xi)$
Low competition	14.35%	814.45%	24.43%
High competition	6.23%	11.25%	1.22%

Figure 3.7 illustrates an example of our empirical distribution method for `AdSlot25`. In the learning set, each winning price can be plot against the demand level ξ. We then use Algorithm 3.1 to compute $\phi(\xi)$. As described earlier, $\psi(\xi)$ and $\pi(\xi)$ are obtained numerically in the similar manner. In our experiments, 10% span of smoothing is allowed. As shown in Figure 3.7, $\phi(\xi)$ and $\pi(\xi)$ are increasing with ξ while $\psi(\xi)$ shows a quadratic pattern on ξ. Once ξ is given, we can calculate the

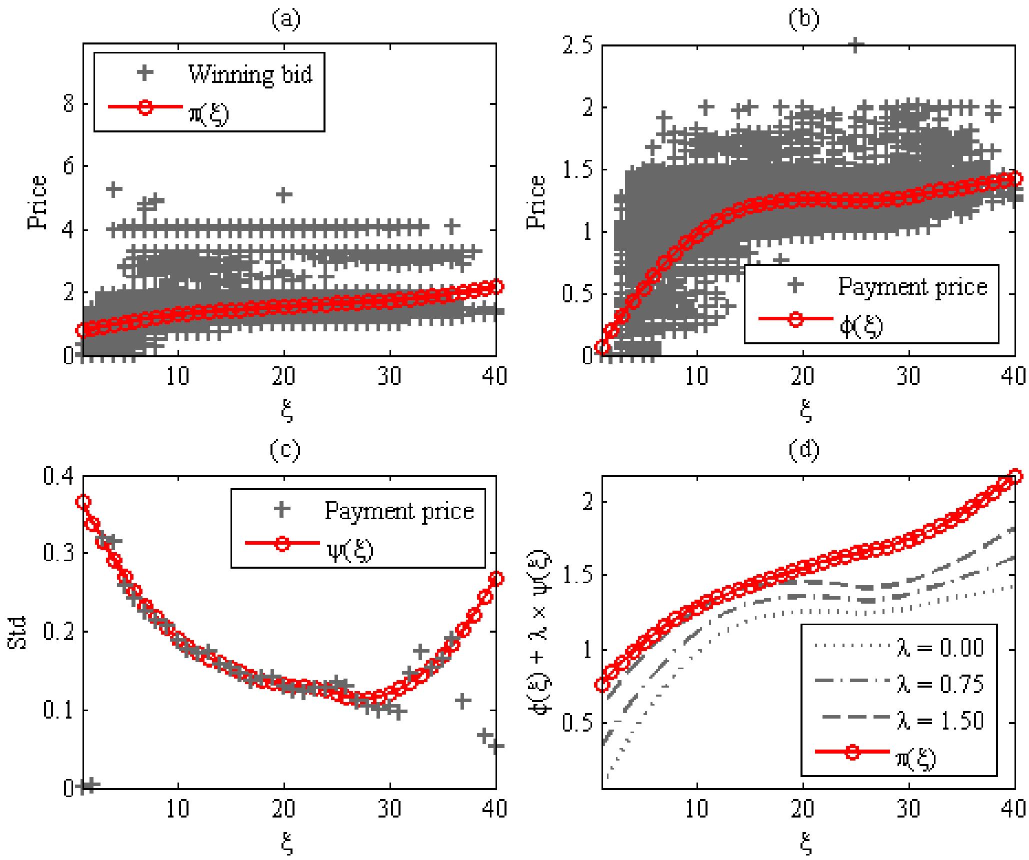

Figure 3.7: Empirical example of estimating $\pi(\xi)$, $\phi(\xi)$ and $\psi(\xi)$ for AdSlot25 from historical bids, where ξ is the per impression demand, $\pi(\xi)$ is the expected winning bid, $\phi(\xi)$ is the expected payment price and $\psi(\xi)$ is the standard deviation of payment prices.

value of the terminal condition $p(0)$ by Eq. (3.2). Figure 3.7 also confirms our earlier statement on λ. Advertisers are not risk-averse if $\lambda = 0$; and they are risk-sensitive for a large λ. In our experiments, we set $\lambda = 1$.

Table 3.7 examines the forecast performance of empirical method and compares the estimated values of $\phi(\xi)$, $\psi(\xi)$, $\pi(\xi)$ to the results of actual bids in the test set. The estimations of $\phi(\xi)$ and $\pi(\xi)$ are much better accurate than that of $\psi(\xi)$. We find that the weak estimations of $\psi(\xi)$ mainly come from AdSlot24, AdSlot26, AdSlot28 and AdSlot30. Their average per impression demand (in both learning and test sets) are around 1.3. As also shown in Figure 3.7, the lower ξ the larger $\psi(\xi)$. Therefore, for the ad slots with a very low level of competition, we set $p(0) = \pi(\xi)$.

3.3.5 Demand for Guaranteed Impressions

The advertisers' purchase behaviour of guaranteed impressions is modeled by parameters α, β, ζ, η as well as be restricted by the expected risk-aversion cost $\phi(\xi) + \lambda\psi(\xi)$.

Here we discuss how to learn the values of α and ζ.

If we only consider the price effect, we can create the function $c(p) = e^{-\alpha p}$ from Eq. (3.6) to represent the probability that an advertiser would like to buy an impression at price p when $\tau = 0$. In RTB, this probability can be learned from the data by investigating the inverse cumulative distribution function (CDF) of all bids, denoted by $z(x) = 1 - F(x)$. For a same domain space of p and x, we can have two series of probabilities $c(p)$ and $z(x)$. Therefore, α can be calibrated as the value that gives the smallest root mean square error (RMSE) between $c(p)$ and $z(x)$. Figure 3.8 illustrates an empirical example of this calibration graphically for `AdSlot25` where the estimated $\alpha = 1.72$.

The values of ζ can also be calibrated from data. Consider a small time step $d\widetilde{\tau}$, then we have the following inequality

$$e^{-\alpha p(1+\beta\times 0)}\zeta e^{-\eta\times 0}d\widetilde{\tau} \leq Q \times (1 - F(\{x \geq p\})). \tag{3.17}$$

If $d\widetilde{\tau} = 1$, then we can have $\zeta = Q \times (1 - F(\{x \geq p\}))/e^{-\alpha p}$.

It is difficult to learn the values of parameters β and η given our current datasets. The two parameters represent the time effect on advertiser's buy behaviour of guaranteed impressions. Here we simply adopt the initial parameter settings used in the flight tickets booking system [Anjos et al., 2004, Malighetti et al., 2009] and set $\beta = \eta = 0.2$. These two parameters can be then updated if the PG system runs for a certain period of time. By having the values of all the model parameters, we can construct the demand surface for a certain range of price series. Figure 3.9 presents a demand surface that satisfies Assumptions 3.1-3.2. It is convex in the guaranteed price and in the time interval between the purchase time and the delivery date.

3.3.6 Revenue Analysis

Two empirical examples are first presented to illustrate how the developed model works with different levels of competition. The overall results are then provided.

Figure 3.10 shows an example of a less competitive market. The learned average per impression demand on `AdSlot14` is about 3.39 (in the test set the actual $\xi = 6.21$). In such a market, advertisers would be less willing to purchase future impressions in advance because they think they can obtain the targeted impressions at lower payment

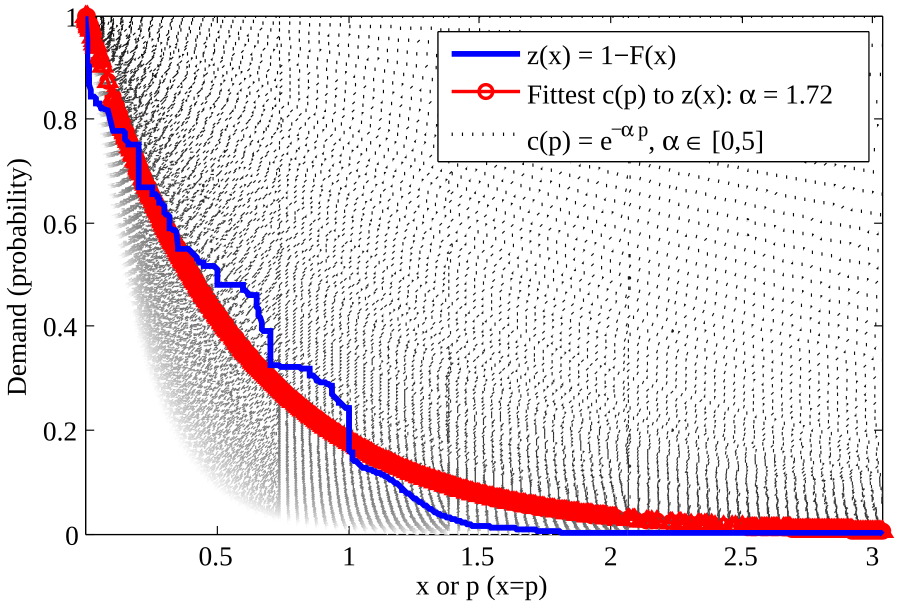

Figure 3.8: Empirical example of estimating the value of α for `AdSlot25`, where α is calculated based on the smallest RMSE between the inverse function of empirical CDF of bids $z(x) = 1 - F(x)$ and the function $c(p) = e^{-\alpha p}$.

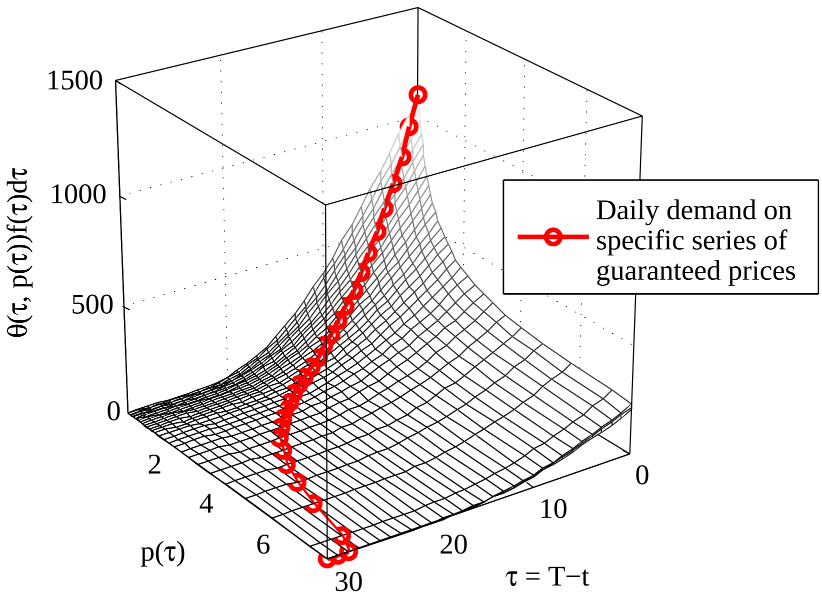

Figure 3.9: Numerical example of demand surface for guaranteed impressions of an ad slot, where $\theta(\tau, p(\tau))f(\tau)d\tau$ represents the number of advertisers who will buy guaranteed impressions at $p(\tau)$ and in $[\tau, \tau + d\tau]$; other parameters are $\alpha = 1.85; \beta = 0.01; \zeta = 2000; \eta = 0.01; T = 30$.

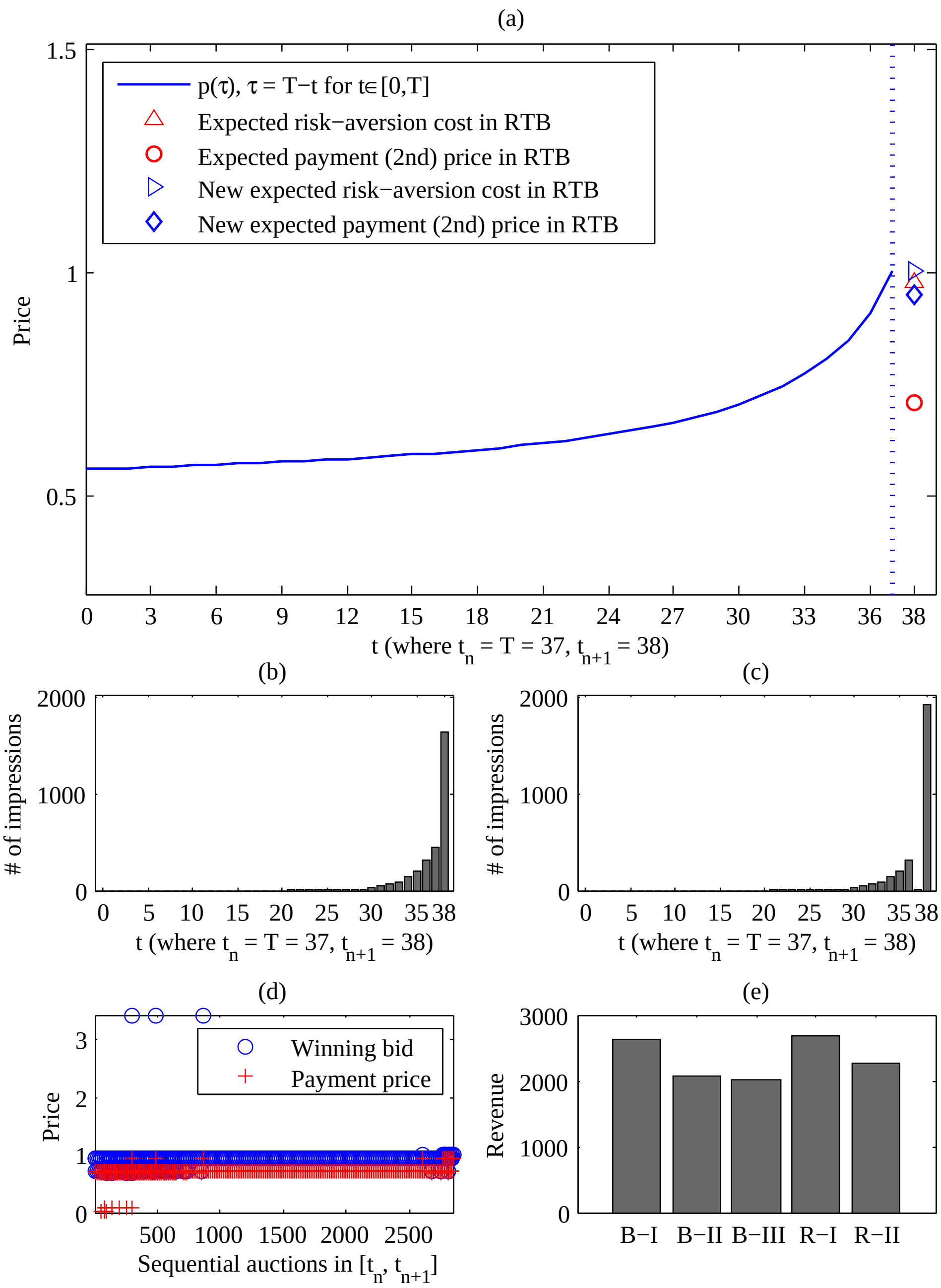

Figure 3.10: Empirical example of `AdSlot14`: (a) the optimal dynamic guaranteed prices; (b) the estimated daily demand; (c) the daily demand calculated based on the actual bids in RTB on the delivery date; (d) the winning bids and payment price in RTB on the delivery date; (e) the comparison of revenues [see Table 3.8 for summary of notations `B-I`, `B-II`, `B-III`, `R-I`, `R-II`]. The parameters are: $Q = 17691; S = 2847; \alpha = 2.0506; \beta = 0.2; \zeta = 442; \eta = 0.2; \omega = 0.05; \kappa = 1; \gamma = 0.4240; \lambda = 2$.

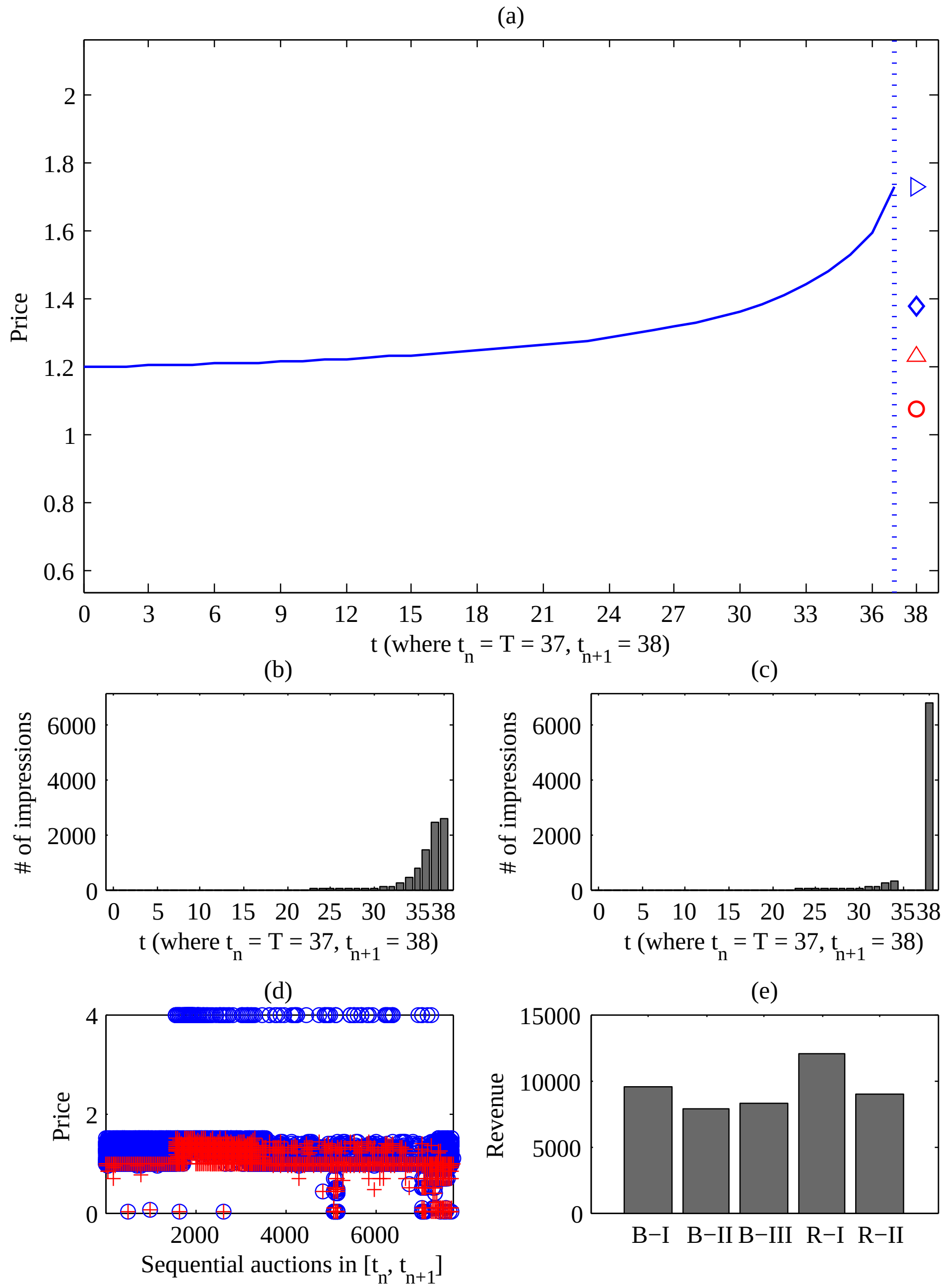

Figure 3.11: Empirical example of `AdSlot27`: (a) the optimal dynamic guaranteed prices; (b) the estimated daily demand; (c) the daily demand calculated based on the actual bids in RTB on the delivery date; (d) the winning bids and payment price in RTB on the delivery date; (e) the comparison of revenues [see Table 3.8 for summary of notations `B-I`, `B-II`, `B-III`, `R-I`, `R-II`]. The parameters are: $Q = 89126; S = 7678; \alpha = 1.7932; \beta = 0.2; \zeta = 2466; \eta = 0.2; \omega = 0.05; \kappa = 1; \gamma = 0.66; \lambda = 2$.

Table 3.8: Summary of plot notations in Figures 3.10 & 3.11.

Calculated revenue:	
`R-I`	Optimal total revenue calculated based on the estimated demand.
`R-II`	Optimal total revenue calculated based on the actual bids in the test set.
Baseline:	
`B-I`	RTB revenue calculated based on actual winning bids in the test set.
`B-II`	RTB revenue calculated based on actual payment prices in the test set.
`B-III`	Estimated RTB revenue based on the learned empirical distribution.

Table 3.9: Summary of revenue evaluation of all 31 ad slots in the SSP dataset.

Group of ad slots	Performance of revenue maximisation		
	Estimated revenue increase	Actual revenue increase	Difference of RTB revenue between estimation & actual payment
Low competition	31.06%	8.69%	13.87%
High competition	31.73%	21.51%	6.23%

Group of ad slots	Performance of price discrimination	
	Ratio of actual 2nd price reve to actual 1st price reve	Ratio of actual optimal reve to actual 1st price reve
Low competition	67.05%	81.78%
High competition	78.04%	94.70%

prices. The model finally allocates 42.40% of future impressions to the guaranteed contracts. In the meantime, the calculated guaranteed prices are not expensive. The prices start with a value lower than the expected payment price from RTB and steadily increases into the level that is close to the maximum value of advertisers' bids. In Figure 3.10, we find that our forecasting values are close to the actual campaigns because the estimated RTB revenue `B-III` is almost the same as the actual RTB revenue `B-II`. Therefore, the estimated advertisers' demand for guaranteed impressions is similar to the actual daily demand (see Figure 3.10(b)&(c)). We also test the guaranteed selling with the actual bids in the test set and find that the calculated revenue `R-II` is still higher than actual second-price RTB revenue `B-II`. This shows that the developed model successfully segments advertisers.

Figure 3.11 describes an example of a competitive market, where the learned average per impression demand for `AdSlot27` is 9.63 (in the test set the actual $\xi = 11.61$). More advertisers would be willing to purchase guaranteed impressions in advance because of the increased level of competition and risk. The model finally allocates 66% future impressions to guaranteed contracts and suggests higher prices at the beginning of guaranteed selling. The estimated total revenue is maximised (i.e., `R-I` > `B-III`); the optimal revenue calculated by the actual bids is more than the actual second-price RTB revenue (i.e., `R-II` > `B-II`).

The overall results are presented in Table 3.9. The revenues calculated based on the estimated demand are always maximised. If we use the actual bids to calculate the demand at the given guaranteed prices, we still have increased revenues (compared to the actual second-price auction market). The results successfully validate the developed model and we find the model performs better in a more competitive environment. This is because the publisher's revenue is actually maximised by the price discrimination, and in a competitive market there are more risk-averse advertisers to segment. With more and more risk-averse advertisers buying the guaranteed impressions, the increased total revenue will approximate advertisers' private evaluations.

3.4 Summary

In this chapter, we investigated a dynamic model for a publisher (or SSP) who engages in RTB to provide the guaranteed delivery of display impressions. We not only designed the mechanism tailored to RTB but also explored its feasibility and performance by analysing the real datasets. Our experimental evaluation successfully validated the developed model as the publisher can receive increased revenues. This work opens several directions for future research. First, we can further consider stochastic supply and demand in revenue maximisation. Second, a parametric updating framework for multi-period pricing and allocation would be of interest.

Chapter 4

Multi-Keyword Multi-Click Ad Options for Sponsored Search

This chapter proposes an ad option tailored for the unique environment of sponsored search, where multiple candidate keywords and a certain number of required clicks are considered. Section 4.1 introduces the background and the problem. Section 4.2 describes the option structure and usage. We then discuss the option pricing methods in Section 4.3 and analyse the revenue effects in Section 4.4. Section 4.5 presents our experimental evaluation and Section 4.6 summarises the chapter. Several important mathematical results are provided in Section 4.7.

4.1 Introduction

Sponsored search is an important form of online advertising. A search engine sells ad slots in the search engine results pages (SERPs) generated in response to a user query. Along with the click on the search button, the query term is what triggers the results on the SERP to appear. The SERP has two types of result listings in response to the submitted query: organic results and paid results. *Organic search results* are the Web page listings that most closely match the user's search query based on relevance. *Paid results* are basically online ads – the companies who have paid to have their Web pages displayed for certain keywords, so such listings show up when an user submits a search query containing those keywords. The price of an ad slot is usually determined by a keyword auction [Jansen, 2011, Börgers et al., 2013, Qin et al., 2014] such as the widely used *generalized second price (GSP) auction* [Edelman et al., 2007, Varian, 2007]. In the GSP auction, advertisers bid on keywords present in the query, and the highest bidder pays the price associated with the bid next to him.

Despite the success of keyword auctions, there are two major drawbacks. First, the uncertain and volatile bids make it difficult for advertisers to predict their campaign costs and thus complicate their business planning [Wang and Chen, 2012]. Second, the "pay-as-you-go" nature of the auction mechanisms does not encourage a stable relationship between advertiser and search engine [Jank and Yahav, 2010] – an advertiser can switch from one search engine to another in the next bidding at near-zero cost.

To alleviate these problems, we propose a *multi-keyword multi-click ad option* in this chapter. It is essentially a contract between an advertiser and a search engine. It consists of a non-refundable upfront fee, known as the *option price*, paid by the advertiser, in return for the right, but not the obligation, to subsequently purchase a fixed number of clicks for particular keywords for pre-specified fixed cost-per-clicks (CPCs) during a specified period of time. From the advertiser's perspective, fixing the CPCs significantly reduces the uncertainty in cost of advertising campaigns. Moreover, for a keyword, if the spot CPC set by keyword auction falls below the fixed CPC, the advertiser is not obligated to exercise the option, but can, instead, participate in keyword auctions. Therefore, the option can be considered as an "insurance" that establishes an upper limit on the cost of advertising campaigns. From the search engine's perspective, the proposed option is not only an additional guaranteed service provided for advertisers. We show that the search engine can, in fact, increase the expected revenue in the process of selling an option. Also, the option covers a specific period of time should encourage a more stable relationship between advertiser and search engine. An important question for us is to determine the option price and the fixed CPCs associated with candidate keywords in the advertiser's request list. Clearly if the option is priced too low, then significant loss in revenue may ensure. Moreover, this may create an arbitrage opportunity where the buyer of the option sells the clicks their targeted keywords to gain extra profits. Conversely, if the option is priced too high, then the advertiser will not purchase it. In this chapter we consider a risk-neutral environment and price the option under the no-arbitrage objective [Wilmott, 2006, Björk, 2009]. We use the Monte Carlo method to price the option with many keywords and show the closed-form pricing formulas for the cases of single and two keywords. Further, the effects of ad options on the search engine's revenue is analysed.

This chapter has three major contributions. First, we propose a new way to pre-sell ad slots in sponsored search which provides flexible guaranteed deliveries to advertisers. It naturally complements the current keyword auction mechanism and offers both advertiser and search engine an effective risk mitigation tool to deal with the bid price fluctuation. Although the proposed ad option belongs to a family of exotic options, its payoff function differs from existing exotic options that we know from finance and other industries (see Table 2.1 for detailed comparisons): it can be exercised not once but multiple times during the contract period; it is not for a single keyword but multiple keywords and each keyword has its own fixed CPC; it allows its buyer to choose which keyword to advertise at the corresponding fixed CPC later during the contract period. Second, we discuss a generalized pricing method for the proposed ad option (see Algorithm 4.1) to deal with the high dimensionality. Third, we demonstrate that, compared to keyword auctions, a search engine can have an increased expected revenue by selling an ad option.

4.2 Flexible Guaranteed Deliveries via Multi-Keyword Multi-Click Ad Options

We use the following example to illustrate our idea. Suppose that a computer science department creates a new master degree programme Web Science and Big Data Analytics and is interested in search advertising based around relevant search terms such as ‘MSc Web Science’, ‘MSc Big Data Analytics’ and ‘Data Mining’ etc. The campaign is to start immediately and last for three months and the goal is to generate at least 1000 clicks on the ad which directs users to the programme’s homepage. The department (i.e., advertiser) does not know how the clicks will be distributed among the keywords, nor how much the campaign will cost if based on keyword auctions. However, with the ad option, the advertiser can submit a request to the search engine to lock-in the advertising cost. The request consists of the candidate keywords, the overall number of clicks needed, and the duration of the contract. The search engine responds with a price table for the option, as shown in Figure 4.1. It contains the option price and the fixed CPC for each keyword. The CPCs are fixed yet different across the candidate keywords. The contract is entered into when the advertiser pays the option price.

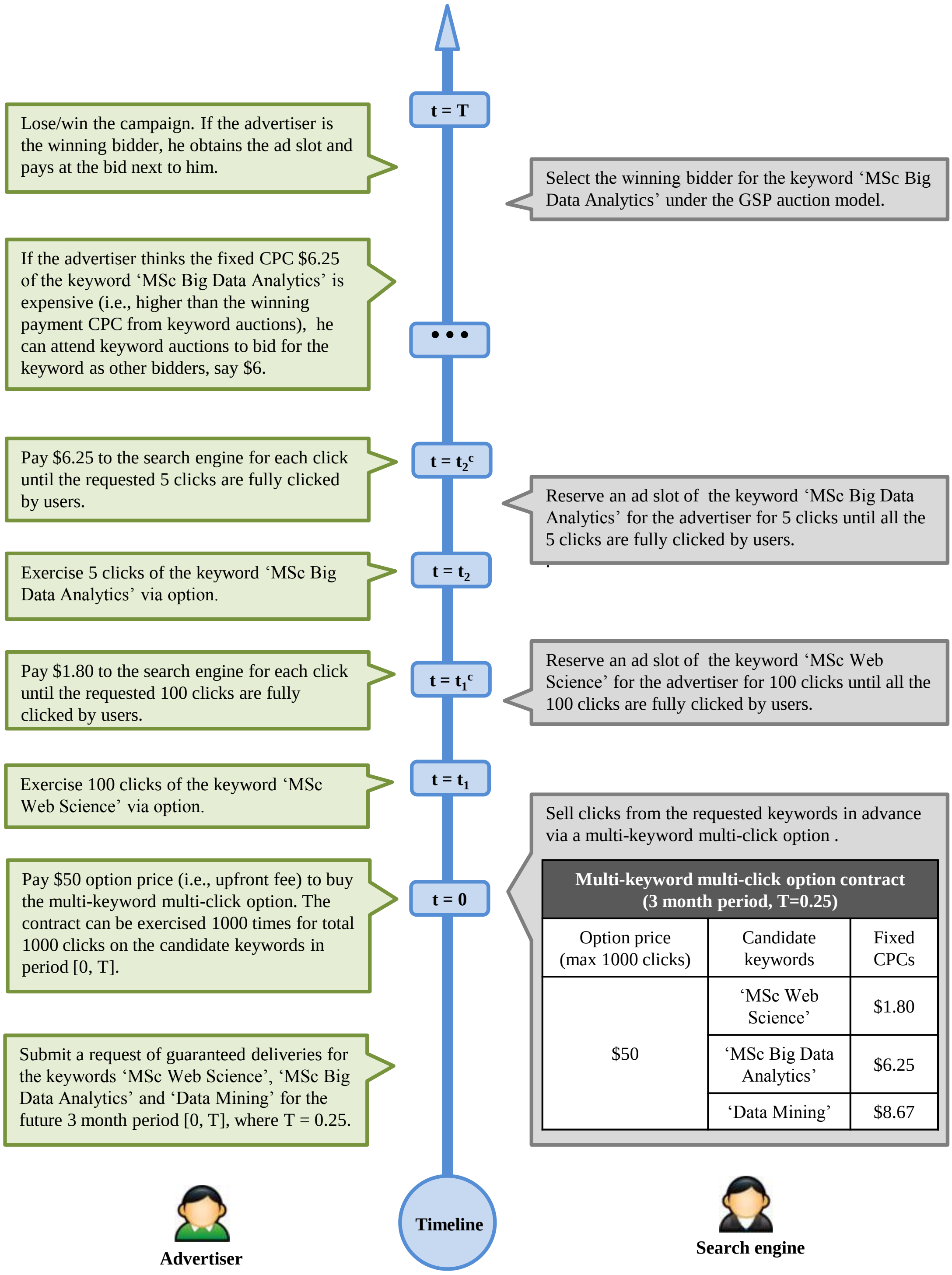

Figure 4.1: Schematic view of buying, selling and exercising a multi-keyword multi-click ad option for sponsored search.

During the contract period $[0, T]$, where T represents the contract expiration date (in terms of year format and is three months in this example), the advertiser has the right, at any time, to exercise portions of the contract, for example, to buy a requested number of clicks for a specific keyword. This right expires after time T or when the total number of clicks have been purchased, whichever is sooner. For example, at time $t_1 \leq T$ the advertiser may exercise the right for 100 clicks on the keyword 'MSc Web Science'. After receiving the exercise request, the search engine immediately reserves an ad slot for the keyword for the advertiser until the ad is clicked by a 100 times. In our current design, the search engine decides which rank position the ad should be displayed as long as the required number of clicks is fulfilled - we assume there are adequate search impressions within the period. It is also possible to generalise the study in this research and define a rank specific option where all the parameters (CPCs, option prices etc.) become rank specific.

The advertiser can switch among the candidate keywords and also monitor the keyword auction market. If, for example, the CPC for the keyword 'MSc Big Data Analytics' drops below the fixed CPC, then the advertiser may choose to participate in the auction rather than exercise the option for the keyword. If later in the campaign, the spot price for the keyword 'MSc Big Data Analytics' exceeds the fixed CPC, the advertiser can then exercise the option.

The above example illustrates the flexibility of the proposed ad option. Specifically, (i) the advertiser does not have to use the option and can participate in keyword auctions as well, (ii) the advertiser can exercise the option at any time during the contract period, (iii) the advertiser can exercise the option up to the maximum number of clicks, (iv) the advertiser can request any number of clicks in each exercise provided the accumulated number of exercised clicks does not exceed the maximum number, and (v) the advertiser can switch among keywords at each exercise with no additional cost. Of course, this flexibility complicates the pricing of the option, which is discussed next.

4.3 Option Pricing Methods

The proposed multi-keyword multi-click ad option enables an advertiser to fix his advertising cost beforehand, yet leave the decision of selecting suitable keywords and the exact timing to place the ad to later. Since the advertiser enjoys great flexibility in sponsored search, there is an intrinsic value associated with an ad option and it buyer needs to pay an upfront option price first. In the following discussion, we focus on calculating a fair upfront option price for the given option specifications, such as the candidate keywords, the current winning payment prices and volatility of these keywords, the length contract period, the risk-less bank interest rate, the preferred fixed CPCs. Recall that Table 2.1 presents the payoff functions of the proposed ad option. We discuss the option pricing for the first payoff function here and in Section 4.3.4 we briefly explain how the introduced option pricing methods can be applied to the second payoff function.

4.3.1 Underlying Stochastic Model

The winning payment CPC of the candidate keyword K_i (for a specific slot/position) at time t is denoted by $C_i(t)$, and whose movement can be described by a multivariate Geometric Brownian Motion (GBM) [Samuelson, 1965a]:

$$dC_i(t) = \mu_i C_i(t)dt + \sigma_i C_i(t)dW_i(t), \quad i = 1, \ldots, n, \tag{4.1}$$

where μ_i and σ_i are constant drift and volatility of the CPC respectively, and $W_i(t)$ is a standard Brownian motion satisfying the conditions:

$$\begin{aligned}\mathbb{E}(dW_i(t)) &= 0,\\ \mathrm{var}(dW_i(t)) &= \mathbb{E}(dW_i(t)dW_i(t)) = dt,\\ \mathrm{cov}(dW_i(t), dW_j(t)) &= \mathbb{E}(dW_i(t)dW_j(t)) = \rho_{ij}dt,\end{aligned}$$

where ρ_{ij} is the correlation coefficient between keywords K_i and K_j, such that $\rho_{ii} = 1$ and $\rho_{ij} = \rho_{ji}$. The correlation matrix is denoted by $\mathbf{\Sigma}$, so that the covariance matrix is simply $\boldsymbol{M}\mathbf{\Sigma}\boldsymbol{M}$, where $\boldsymbol{M}$ is the matrix with the σ_i along the diagonal and zeros everywhere else. For the reader's convenience, detailed descriptions of notations are provided in Table 4.1.

Since the GBM assumption lays down the foundation of pricing the proposed ad option, we also provide several discussions and investigations on the GBM. In Section 4.3.4, we explain why the GBM assumption is suitable for pricing an ad option

Table 4.1: Summary of notations.

Notation	Description
r	Constant continuous (risk-less) interest rate.
T	Option expiration date.
t	Continuous time point in $[0, T]$.
m	Number of total clicks specified by an ad option.
n	Number of total number of keywords specified by an ad option.
$\boldsymbol{K}$	Keywords specified by an ad option, $\boldsymbol{K} = (K_1, \dots, K_n)$.
$\boldsymbol{F}$	Pre-specified fixed CPCs for keywords $\boldsymbol{K}$.
$\boldsymbol{C}(t)$	Winning payment CPCs for keywords $\boldsymbol{K}$ from auctions at time t.
$V(t, \boldsymbol{C}(t); T, \boldsymbol{F}, m)$	Value of an n-keyword m-click ad option at time t.
μ_i	Constant drift of CPC for keyword $K_i, i = 1, \dots, n$.
σ_i	Constant volatility of CPC for keyword $K_i, i = 1, \dots, n$.
$\boldsymbol{\Sigma}$	Price correlation matrix, in which ρ_{ij} is the correlation coefficient between ith and jth keywords, such that $\rho_{ii} = 1$ and $\rho_{ij} = \rho_{ji}$.
$\boldsymbol{M\Sigma M}$	Price covariance matrix, where $\boldsymbol{M}$ is the matrix with σ_i along the diagonal and zeros everywhere else.
$\Phi(\boldsymbol{C}(t))$	Payoff function of an ad option at time t.
π	Option price (i.e., upfront fee) of an ad option.
$\mathbf{MVN}(\boldsymbol{\mu}, \boldsymbol{M\Sigma M})$	Multivariate normal distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{M\Sigma M}$.

in sponsored search as well as indicate its limitations. In Section 4.5.2, we discuss the estimation of the GBM parameters. In Section 4.5.3, we investigate the goodness-of-fit tests with the real datasets and track the "errors" is the GBM underlying model is not valid empirically.

4.3.2 Terminal Value Pricing

To simplify the discussion and without loss of generality, the value of an n-keyword m-click ad option can be decomposed as the sum of m independent n-keyword 1-click ad options. If an advertiser buys an ad option at time 0, the option price π can be expressed as follows

$$\pi = V(0, \boldsymbol{C}(0); T, \boldsymbol{F}, m) = mV(0, \boldsymbol{C}(0); T, \boldsymbol{F}, 1), \tag{4.2}$$

where $V(0, \boldsymbol{C}(0); T, \boldsymbol{F}, m)$ represents the option value at time 0.

Our focus now centres on the n-keyword 1-click ad option. By adopting the basic economic setting [Narahari et al., 2009], we assume that an advertiser is risk-neutral. Simply, he has no preference across the candidate keywords and exercises the option for the keyword which has the maximum difference between its winning payment price and the pre-specified fixed price. This difference shows the value of the option because the advertiser is offered the right to move from the auction market to the guaranteed market. Let's consider if the advertiser exercises the option at the contract expiration date T, the option payoff can be defined as follows

$$\Phi(\boldsymbol{C}(T)) = \max\{C_1(T) - F_1, \ldots, C_n(T) - F_n, 0\}. \tag{4.3}$$

Note that the option payoff in sponsored search does not mean the direct reward but it measures the difference of advertising cost between the auction market and the guaranteed market. By having Eq. (4.3), we can see if the advertiser would like to early exercise the option by using the backward deduction method. The option value at time time $t < T$ is then

$$V(t, \boldsymbol{C}(t); T, \boldsymbol{F}, 1) = \begin{cases} \Phi(\boldsymbol{C}(t)), & \text{if early exercise,} \\ \mathbb{E}_t^{\mathbb{Q}}\big[e^{-r(T-t)}\Phi(\boldsymbol{C}(T))\big], & \text{if not early exercise,} \end{cases}$$

where r is the constant risk-less bank interest rate and $\mathbb{E}_t^{\mathbb{Q}}[\cdot]$ is the conditional expectation with respect to time t under the probability measure $\mathbb{Q}$. As we use the risk-less bank interest rate as the discounted factor, the probability measure $\mathbb{Q}$ is also called the *risk-neutral probability measure* [Björk, 2009]. Here we do not further discuss why using the risk-less bank interest rate while we provide a brief explanation in Section 4.7.2 together with introducing an alternative way of option pricing.

Let's back to the decision making problem. If the ad option is early exercised at time t, the option value is equal to its payoff $\Phi(\mathbf{C}(t))$. However, if the ad option is not exercised, the option value at time t is equal to the discounted value of the expected payoff from the expiration date T. The comparison between $\Phi(\boldsymbol{C}(t))$ and $\mathbb{E}_t^{\mathbb{Q}}\big[e^{-r(T-t)}\Phi(\boldsymbol{C}(T))\big]$ can tell us the optimal decision for the advertiser. Since the payoff function defined is convex, we then obtain the following inequality (see Section 4.7.1):

$$\Phi(\boldsymbol{C}(t)) \leq \mathbb{E}_t^{\mathbb{Q}}\big[e^{-r(T-t)}\Phi(\boldsymbol{C}(T))\big]. \tag{4.4}$$

Eq. (4.4) illustrates, to gain the maximum option value, the advertiser will not exercise the option until its expiration date. Hence, the option price should be computed at the discounted value of the expected payoff from the expiration date T. Together with

Eq. (4.2), we can obtain the option pricing formula for the n-keyword m-click ad option as follows

$$\pi = me^{-rT}\mathbb{E}_0^{\mathbb{Q}}\big[\Phi(\boldsymbol{C}(T))\big]. \tag{4.5}$$

It is worth noting that we rule out the arbitrage [Varian, 1987] between the auction market and the guaranteed market in option pricing. The concept of arbitrage can be understood as the "free lunch". As a market designer, we need to make sure that everyone obtains something by paying something so that it is fair for both buy and sell sides. Since we assume that an advertiser is risk-neutral, then it makes sense that the risk-less bank interest rate can be employed as the benchmark rate to rule out arbitrage. Eq. (4.5) can also be obtained by constructing an advertising strategy for the advertiser and the detailed discussion is provided in Section 4.7.2.

4.3.3 Solutions

Eq. (4.5) can be expanded in an integration form as follows

$$\begin{aligned}\pi = & \ me^{-rT}\big(2\pi T\big)^{-\frac{n}{2}}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}\left(\prod_{i=1}^{n}\sigma_i\right)^{-1} \\ & \times \int_0^{\infty}\cdots\int_0^{\infty}\frac{\Phi(\widetilde{\boldsymbol{C}})}{\prod_{i=1}^{n}\widetilde{C}_i}\exp\left\{-\frac{1}{2}\boldsymbol{\zeta}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\zeta}\right\}d\widetilde{\boldsymbol{C}},\end{aligned} \tag{4.6}$$

where $\boldsymbol{\zeta} = [\zeta_1,\ldots,\zeta_n]^{\mathbf{T}}$, $\zeta_i = \frac{1}{\sigma_i\sqrt{T}}\big(\ln\{\widetilde{C}_i/C_i(0)\} - (r - \frac{\sigma_i^2}{2})T\big)$, and other notations are described in Table 4.1.

Closed form solutions to Eq. (4.6) can be derived if $n \le 2$. If $n = 1$, Eq. (4.6) is equivalent to the Black-Scholes-Merton (BSM) pricing formula for an European call option [Black and Scholes, 1973, Merton, 1973]. If $n = 2$, Eq. (4.6) contains a bivariate normal distribution and the option price can be obtained by employing the pricing formula for a dual-strike European call option [Zhang, 1998]. The discussed two formulas are provided in Section 4.7.3.

If $n \ge 3$, taking integrals in Eq. (4.6) is computationally difficult. In such a case, we resort to numerical techniques to approximate the option price. Algorithm 4.1 illustrates our Monte Carlo method. Let's consider $\widetilde{n}$ number of simulations, and for

Algorithm 4.1 Pricing a multi-keyword multi-click ad option via Monte Carlo simulation. Detailed notations are provided in Table 4.1.

```
function OptionPricingMC(K, C(0), Σ, M, m, r, T)
    ñ ← 1000; # Number of simulated paths;
    for k ← 1 to ñ do
        [z_{1,k}, ..., z_{n,k}] ← GeneratingMultiNoise(MVN[0, MΣM])
        for i ← 1 to n do
            C_{i,k} ← C_i(0) exp{(r − ½σ_i²)T + σ_i z_{i,k} √T}.
        end for
        G_k ← Φ([C_{1,k}, ..., C_{n,k}]).
    end for
    π ← m e^{−rT} E_0[Φ(C(T))] ≈ m e^{−rT} ( (1/ñ) Σ_{k=1}^{ñ} G_k ).
    return π
end function
```

each simulation, we generate a vector of multinormal noise and then calculate the CPCs at time T. Eq. (4.4) shows that there is no need to generate the whole paths in each simulation as we only consider the CPCs on the expatriation date in the calculation of option payoff. Hence, by having $\widetilde{n}$ payoffs at time T, the option price π can be then approximated numerically, and Algorithm 4.1 is lightweight and computationally fast.

4.3.4 Discussion

Like other methods based on the GBM assumption, the candidate keywords' prices may not follow it empirically because some time series features, such as jumps and volatility clustering, cannot be captured effectively [Marathe and Ryan, 2005]. However, the GBM assumption is still a good choice for pricing ad options in sponsored search. First, in our data analysis (see Section 4.5.3.1), we find that several keywords' winning payment CPCs satisfy the GBM assumption. Second, for the cases that the GBM assumption is not valid empirically (see Section 4.5.3.2), we find that the pricing model is reasonably robust as the identified arbitrage values in many experimental groups are small. Third, our dataset might be biased. However, other previous research in keyword auctions support the GBM assumption: Lahaie and Pennock [2007] tested the log-normality of bids on Yahoo! search advertising data and gave the estimated distribution parameters; Ostrovsky and Schwarz [2011] performed a field experiments based on the log-normal bids on Yahoo! search advertising platform; Pin and Key [2011] observed random bids from Microsoft Bing and simulated similar bids based on the log-normal distribution. Since in these research the advertisers' bids are tested across auctions, the winning payment prices over time should also satisfy the log-normal distribution (because they are the second-highest bids from auctions). Recall that in the GBM assumption, the difference between two logarithms of winning payment prices follows a time dependent normal distribution. If we consider the average daily winning payment price as the underlying variable, these previous work can provide the distribution hypothesis tests to support the GBM assumption in sponsored search. However, for display advertising, the GBM assumption is mostly not valid. This has been indicated by Chen et al. [2014c] and Yuan et al. [2014].

In addition, the option payoff defined in Eq. (4.3) can be used for both keyword exact match and keyword broad match settings. It depends on the type of the winning payment prices used. Also as described in Table 2.1, if we only have the exactly matched $\boldsymbol{C}(T)$, we can still construct a broad match structure for the option, similar to Eq. (4.3), the option payoff function on time T is

$$\Phi(\boldsymbol{C}(T)) = \max\left\{\sum_{i=1}^{k_1}\omega_{1i}C_{1i}(T) - F_1, \cdots, \sum_{i=1}^{k_n}\omega_{ni}C_{ni}(T) - F_n, 0\right\}. \tag{4.7}$$

where ω_{ji} is the probability that the ith broad matched keyword (i.e., the sub-phrase occurs in search queries) for the keyword K_j. Eq. (4.1) can be still used to model the underlying CPCs' movement but the selected keywords need to be uniquely distinctive from each other. For simplicity, we denote them by $\widetilde{\boldsymbol{C}}(T)$. The correlation matrix is the correlation between these distinctive underlying keywords, denoted by $\widetilde{\boldsymbol{\Sigma}}$. By having the distinctive underlying keywords in Eq. (4.1), the option price π_0 can be calculated by Algorithm 4.1.

4.4 Revenue Analysis for Search Engine

The proposed ad option can be loosely considered as a kind of insurance for an advertiser. It does not come without a cost because the advertiser needs to pay the upfront option price; therefore, the ad option is also beneficial to the search engine's revenue. In the following discussion, we analyse the effect of an ad option on the search engine's revenue. We provide a functional analysis for the 1-keyword 1-click ad option in this section and leave the empirical investigation of the n-keyword cases in Section 4.5.

Let $D(F)$ be the difference between the expected revenue from ad option and the expected revenue from only keyword auctions, we then have

$$D(F) = \underbrace{\Big(C(0)\mathcal{N}[\zeta_1] - e^{-rT}F\mathcal{N}[\zeta_2] + e^{-rT}F\Big)\mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] \geq F)}_{=\text{ Discounted value of expected revenue from option if } \mathbb{E}_0^{\mathbb{Q}}[C(T)]\geq F}$$

$$+ \underbrace{\Big(C(0)\mathcal{N}[\zeta_1] - e^{-rT}F\mathcal{N}[\zeta_2] + e^{-rT}\mathbb{E}_0^{\mathbb{Q}}[C(T)]\Big)\mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] < F)}_{=\text{ Discounted value of expected revenue from option if } \mathbb{E}_0^{\mathbb{Q}}[C(T)]<F}$$

$$- \underbrace{e^{-rT}\mathbb{E}_0^{\mathbb{Q}}[C(T)]}_{=\text{ Discounted value of expected revenue from auction}} .$$

$$= C(0)\mathcal{N}[\zeta_1] - e^{-rT}F\mathcal{N}[\zeta_2] - e^{-rT}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] - F) \times \mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] \geq F), \tag{4.8}$$

where $\mathcal{N}[\cdot]$ represents the cumulative probability of a standard normal distribution.

Let us take a look at the boundary values first. If $F = 0$, the option price π achieves its maximum value $e^{-rT}\mathbb{E}_0^{\mathbb{Q}}[C(T)]$; therefore, $D(F) \to 0$. If $\pi = 0$, the fixed CPC F is as large as possible, and $\mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] \geq F) \to 0$ and $D(F) \to 0$. Since

$$\ln\{C(T)/C(0)\} \sim \mathbf{N}\big((r - \sigma^2/2)T, \sigma^2 T\big),$$

we can have

$$\begin{aligned}
\mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] \geq F) &= \mathbb{P}\Big(C(0)\exp\{(r - \frac{1}{2}\sigma^2)T\} \geq F\Big) \\
&= \mathbb{P}\Big(\ln\{F/C(0)\} - (r - \frac{1}{2}\sigma^2)T \leq 0\Big) \\
&= \mathbb{P}\Big(\ln\{C(T)/C(0)\} - (r - \frac{1}{2}\sigma^2)T \\
&\qquad \leq \ln\{C(0)/F\} + (r - \frac{1}{2}\sigma^2)T + \sigma W(T)\Big) \\
&\approx \mathcal{N}\left[\frac{1}{\sigma\sqrt{T}}\Big(\ln\{C(0)/F\} + (r - \frac{1}{2}\sigma^2)T\Big)\right] = \mathcal{N}[\zeta_2].
\end{aligned} \tag{4.9}$$

Substituting Eq. (4.9) into Eq. (4.8) gives

$$\begin{aligned}
D(F) &= C(0)\mathcal{N}[\zeta_1] - e^{-rT}\mathbb{E}_t^{\mathbb{Q}}(C(T))\mathcal{N}[\zeta_2] \\
&\geq C(0)\mathcal{N}[\zeta_1] - e^{-rT}\mathbb{E}_t^{\mathbb{Q}}(C(T))\mathcal{N}[\zeta_1] \quad \text{(because } \mathcal{N}[\zeta_1] \geq \mathcal{N}[\zeta_2]) \\
&\geq C(0)\mathcal{N}[\zeta_1] - e^{-rT}C(0)e^{(r-\frac{1}{2}\sigma^2)T}\mathcal{N}[\zeta_1] \\
&= C(0)\mathcal{N}[\zeta_1](1 - e^{-\frac{1}{2}\sigma^2 T}) > 0, \qquad (4.10)
\end{aligned}$$

suggesting that the search engine can have an increase expected revenue if he sells the click via an option rather than through an auction. We then take the derivative of $D(F)$ with respect to F and assign its value to zero:

$$\begin{aligned}
\frac{\partial D(F)}{\partial F} &= C(0)\frac{\partial \mathcal{N}[\zeta_1]}{\partial \zeta_1}\frac{\partial \zeta_1}{\partial F} - e^{-rT}\mathcal{N}[\zeta_2] - e^{-rT}F\frac{\partial \mathcal{N}[\zeta_2]}{\partial \zeta_2}\frac{\partial \zeta_2}{\partial F} \\
&- e^{-rT}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] - F)\frac{\partial \mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] \geq F)}{\partial F} + e^{-rT}\mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] \geq F) = 0.
\end{aligned} \qquad (4.11)$$

Since $\partial \mathcal{N}(x)/\partial x = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$, the following equation holds

$$\frac{\partial \mathcal{N}[\zeta_2]}{\partial \zeta_2} \Big/ \frac{\partial \mathcal{N}[\zeta_1]}{\partial \zeta_1} = \exp\left\{\frac{1}{2}(\zeta_1^2 - \zeta_2^2)\right\} = \frac{C(0)e^{rT}}{F}. \qquad (4.12)$$

Taking the derivative of ζ_1 and ζ_2 with respect to F gives

$$\frac{\partial \zeta_1}{\partial F} = \frac{\partial \frac{1}{\sigma\sqrt{T}}\left(\ln\{C(0)/F\} + (r + \frac{1}{2}\sigma^2)T\right)}{\partial F} = -\frac{1}{F\sigma\sqrt{T}}, \qquad (4.13)$$

$$\frac{\partial \zeta_2}{\partial F} = \frac{\partial \zeta_1}{\partial F} - \frac{\partial \sigma\sqrt{T}}{\partial F} = -\frac{1}{F\sigma\sqrt{T}}. \qquad (4.14)$$

and we find that $D(F)$ achieves its maximum or minimum value at $F = \mathbb{E}_0^{\mathbb{Q}}[C(T)]$. Further taking the second derivative of $D(F)$ with respect to $F = \mathbb{E}_0^{\mathbb{Q}}[C(T)]$ gives

$$\frac{\partial^2 D(F)}{\partial F^2} = \frac{\partial \mathbb{P}(\mathbb{E}_0^{\mathbb{Q}}[C(T)] \geq F)}{\partial F} = \frac{\partial \mathcal{N}[\zeta_2]}{\partial \zeta_2}\frac{\partial \zeta_2}{\partial F} = -\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\zeta_2{}^2}\frac{1}{F\sigma\sqrt{T}} < 0.$$

Hence, if the fixed CPC is set as same as the estimated spot CPC on the contract expi-

Table 4.2: Overview of experimental settings of data.

Market	Group	Training set (31 days)	Development & test set (31 days)
US	1	25/01/2012-24/02/2012	24/02/2012-25/03/2012
	2	30/03/2012-29/04/2012	29/04/2012-31/05/2012
	3	10/06/2012-12/07/2012	12/07/2012-17/08/2012
	4	10/11/2012-11/12/2012	11/12/2012-10/01/2013
UK	1	25/01/2012-24/02/2012	24/02/2012-25/03/2012
	2	30/03/2012-29/04/2012	29/04/2012-31/05/2012
	3	12/06/2012-13/07/2012	13/07/2012-19/08/2012
	4	18/10/2012-22/11/2012	22/11/2012-24/12/2012

ration date (i.e., $F = \mathbb{E}_0^{\mathbb{Q}}[C(T)]$), the search engine can have a maximised profit.

4.5 Experiments

This section presents our evaluation results. We first describe the dataset, then conduct assumption and fairness tests, and finally investigate the effects of ad options on the search engine's revenue.

4.5.1 Data and Experimental Design

The data[1] used in the experiments are collected from Google AdWords by using its Traffic Estimation service Google. When an advertiser submits his ad keywords, budget, and other settings such as keyword match type and targeted ad location, the Traffic Estimation will return a list of data values, including estimated CPC, clicks, global impressions, local impressions and position etc. Such values are recorded for the period from 26/11/2011 to 14/01/2013, for a total of 557 keywords across US and UK markets. Note that in the data 21 keywords have missing values and 115 keywords's CPCs are all 0.

For each market, the data is split into four experimental groups and each group has one training, one development, and one test set, as illustrated in Table 4.2. The training set is used to: (i) select the keywords with non-zero CPCs; (ii) test the statistical properties of the underlying dynamic and estimate the model parameters. We then price ad options and simulate the corresponding buying and selling transactions in the development set. Finally, the test set is used as the baseline to examine the priced ad options.

[1]The data is available at:
`http://www.computational-advertising.org` [Yuan and Wang, 2012].

4.5.2 Parameter Estimation and Option Pricing

In the experiments, we use the method suggested by Wilmott [2006] to estimate the GBM parameters. For the keyword K_i, the volatility σ_i is the sample standard deviation of change rates of log CPCs and the correlation is calculated as follows

$$\rho_{ij} = \frac{\sum_{k=1}^{\widetilde{m}} \big(y_i(k) - \bar{y}_i\big)\big(y_j(k) - \bar{y}_j\big)}{\sqrt{\sum_{k=1}^{\widetilde{m}} \big(y_i(k) - \bar{y}_i\big)^2 \sum_{k=1}^{\widetilde{m}} \big(y_j(k) - \bar{y}_j\big)^2}}, \tag{4.15}$$

where $\widetilde{m}$ is the size of training data and $y_i(t_k)$ is the kth change rate of log CPCs.

Figure 4.2 illustrates an empirical example, where the candidate keywords are

$$\boldsymbol{K} = \left\{ \begin{array}{c} K_1 \\ K_2 \\ K_3 \end{array} \right\} = \left\{ \begin{array}{c} \text{‘canon cameras’} \\ \text{‘nikon camera’} \\ \text{‘yahoo web hosting’} \end{array} \right\},$$

and the model parameters are estimated as follows

$$\boldsymbol{\sigma} = \begin{pmatrix} 0.2263 \\ 0.4521 \\ 0.2136 \end{pmatrix}, \ \boldsymbol{\Sigma} = \begin{pmatrix} 1.0000 & 0.2341 & 0.0242 \\ 0.2341 & 1.0000 & -0.0540 \\ 0.0242 & -0.0540 & 1.0000 \end{pmatrix}.$$

Note that a high contextual relevance of keywords normally means that they have a high substitutional degree to each other, such as ‘canon cameras’ and ‘nikon camera’, whose CPCs move in the same direction with correlation 0.2341. The other keyword ‘yahoo web hosting’ is contextually less relevant to the formers and also has very low price correlations to them. The example also shows that the contextual relevance of keywords has an impact on their CPCs movement.

Based on the estimated parameters, we draw a sample of simulated paths of a 3-dimensional GBM in Figure 4.2(a) for 31 days (where the x-axis is expressed in terms of year value). Recall that the option payoff at any time t in the contract lifetime is $\max\{C_1(t) - F_1, \ldots, C_n(t) - F_n, 0\}$. In Figure 4.2(b), we plot the price difference between the spot CPC and the fixed CPC of each candidate keyword (i.e., $C_i(t) - F_i$, $i = 1, \ldots, n$) and also indicate the corresponding option daily payoffs (shown by the cyan curve). It suggests that switching between keywords would help the advertiser

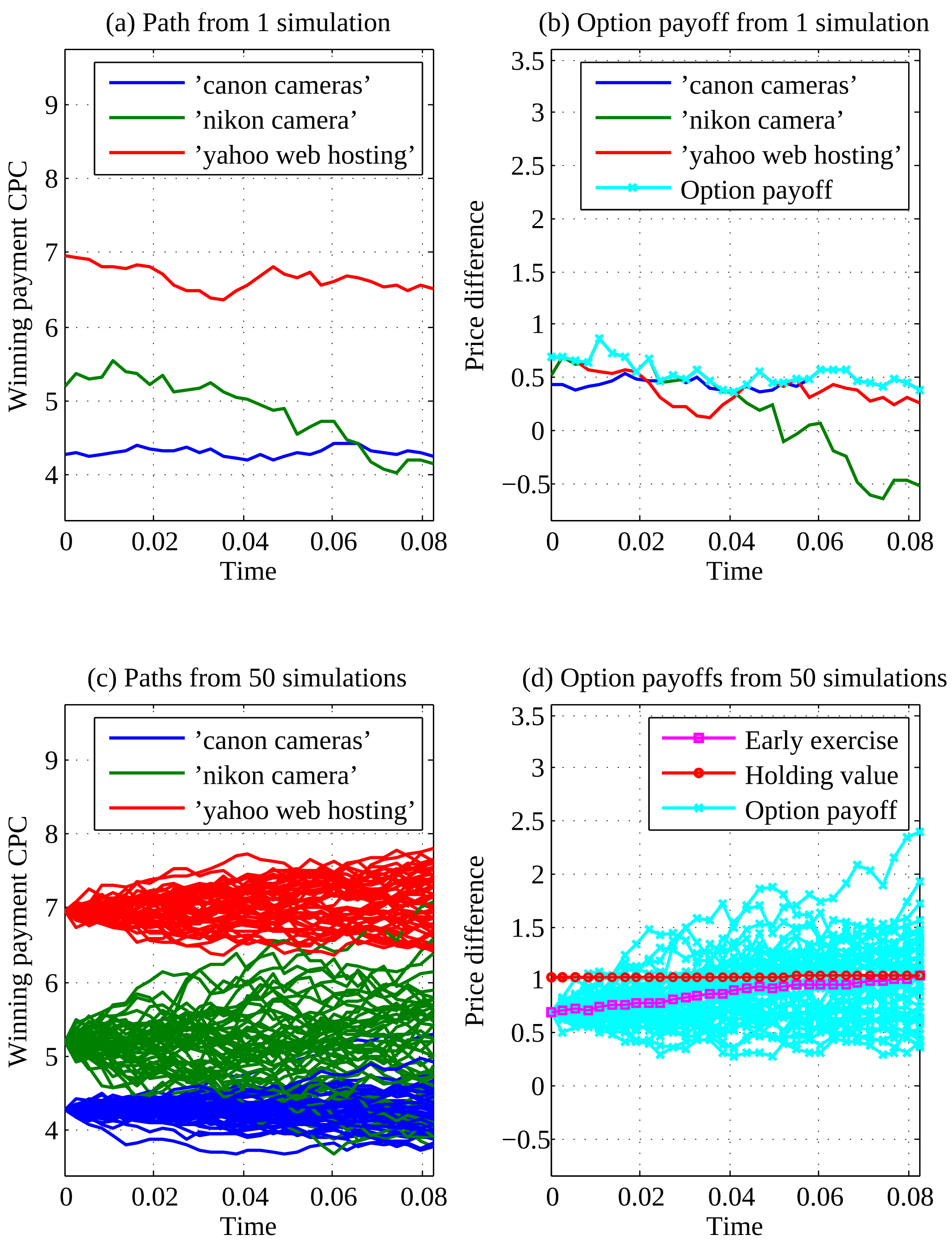

Figure 4.2: Empirical example of generating paths under a GBM for a 3-keyword 1-click option and calculating the corresponding payoffs: K_1 = 'canon cameras', K_2 = 'nikon camera', K_3 = 'yahoo web hosting', $F_1 = 3.8505$, $F_2 = 4.6704$ and $F_3 = 6.2520$.

to maximise the benefits of the ad option. Repeating the above simulations 50 times generates 50 simulated vales of each keyword for each day, as shown in Figure 4.2(c). We then calculate 50 option payoffs and their daily mean values to obtain the final option price, as shown in Figure 4.2(d).

To examine the fairness (i.e., no-arbitrage) of the calculated option price, we can construct a risk-less value difference process by delta hedging $\partial V/\partial C_j$ and check if any arbitrage exists (see Section 4.7.2 or [Wilmott, 2006]). The hedging delta of the 1-keyword 1-click ad option can be obtained as follows

$$\frac{\partial V}{\partial C} = \mathcal{N}(\zeta_1) + C(0)\mathcal{N}(\zeta_1)'\frac{\partial \zeta_1}{\partial C} - Fe^{-rT}\mathcal{N}(\zeta_2)'\frac{\partial \zeta_2}{\partial C}, \tag{4.16}$$

and

$$\zeta_1 = \frac{1}{\sigma\sqrt{T}}\big(\ln\{C(0)/F\} + (r + \frac{\sigma^2}{2})T\big),$$
$$\zeta_2 = \zeta_1 - \sigma\sqrt{T}.$$

Therefore, $\partial\zeta_1/\partial C = \partial\zeta_2/\partial C$. Since

$$\mathcal{N}(\zeta_1)' = \frac{1}{\sqrt{2\pi}}e^{-\frac{\zeta_1^2}{2}}, \tag{4.17}$$

$$\mathcal{N}(\zeta_2)' = \frac{1}{\sqrt{2\pi}}e^{-\frac{(\zeta_1-\sigma\sqrt{T})^2}{2}} = \frac{C(0)}{F}e^{rT}\mathcal{N}(\zeta_1)', \tag{4.18}$$

then $\partial V/\partial C = \mathcal{N}(\zeta_1)$.

For the n-keyword 1-click option, the hedging delta of each keyword can be computed by the Monte Carlo method, i.e., $\partial V/\partial C_i = \mathbb{E}^{\mathbb{Q}}[\partial V(T, C(T))/\partial C_i(T)]$. According to Section 4.7.2, we can define the 31-day growth rate of the value difference process as $\widetilde{\gamma} = \Big(\Pi(t_{31}) - \Pi(t_0)\Big)/\Pi(t_0)$, and compare $\widetilde{\gamma}$ to the risk-less bank interest rate $r = 5\%$ (equivalent to $\widetilde{r} = 4.12\%$ per 31 days return[2]). The arbitrage detection criteria is

$$|\widetilde{\gamma} - \widetilde{r}| \leq \varepsilon \text{ ? } \textbf{arb doesn't exist : arb exists}, \tag{4.19}$$

where the notation ε is the model variation threshold (and we set $\varepsilon = 5\%$ in experi-

[2]The relationship between the continuous compounding r and the return per 31 days $\widetilde{r}$ is: $1 + \widetilde{r} = e^{r\times 30/365}$ [Wilmott, 2006].

Table 4.3: Test of arbitrage for ad options under the GBM assumption: n is the number of keywords, N is the number of options priced in a group, $\mathbb{P}(\alpha)$ is percentage of options in a group found arbitrage, and the $\mathbb{E}[\alpha]$ is the average arbitrage value of the options found arbitrage, where the arbitrage α is defined by Eq. (4.20) and the risk-less bank interest rate $r = 5\%$.

n	Group	US market			UK market		
		N	$\mathbb{P}(\alpha)$	$\mathbb{E}[\alpha]$	N	$\mathbb{P}(\alpha)$	$\mathbb{E}[\alpha]$
1	1	94	0.00%	0.00%	76	0.00%	0.00%
	2	64	0.00%	0.00%	45	0.00%	0.00%
	3	94	1.06%	0.75%	87	0.00%	0.00%
	4	112	0.89%	-0.37%	53	0.00%	0.00%
2	1	47	4.26%	1.63%	38	0.00%	0.00%
	2	32	9.38%	0.42%	22	4.55%	13.41%
	3	47	4.26%	0.84%	43	4.65%	0.82%
	4	56	5.36%	3.44%	26	23.08%	-6.22%
3	1	31	0.00%	0.00%	25	4.00%	0.00%
	2	21	4.76%	-1.38%	15	0.00%	0.00%
	3	31	0.00%	0.00%	29	3.45%	-1.12%
	4	37	10.81%	3.87%	17	35.29%	-2.54%

ments). Then the identified arbitrage α is defined as the excess return, that is

$$\alpha = \begin{cases} \widetilde{\gamma} - (\widetilde{r} + \varepsilon), & \text{if } \widetilde{\gamma} \geq \widetilde{r} + \varepsilon, \\ \widetilde{\gamma} - (\widetilde{r} - \varepsilon), & \text{if } \widetilde{\gamma} \leq \widetilde{r} - \varepsilon. \end{cases} \tag{4.20}$$

Hence, a positive α means that the advertiser buys an option can obtain arbitrage while a negative α indicates the case of making arbitrage by selling an option.

Table 4.3 presents the overall results of our arbitrage test based on the GBM. We generate paths for candidate keywords with 100 simulations and examine the options price using delta hedging. There are 99.76% (1-keyword), 93.06% (2-keyword) and 92.71% (3-keyword) options fairly priced. Only a small number of options exhibits arbitrage and most of the mean arbitrage values lie within 5%, such as shown in Figure 4.3. The existence of small arbitrage may be due to two reasons. First, the stability of process simulations in both option pricing and arbitrage test. Second, the candidate keywords are randomly selected for the 2-keyword and 3-keyword options. The significant differences on the absolute prices these keywords can generates a large variation of calculated option payoffs, which then trigger arbitrage.

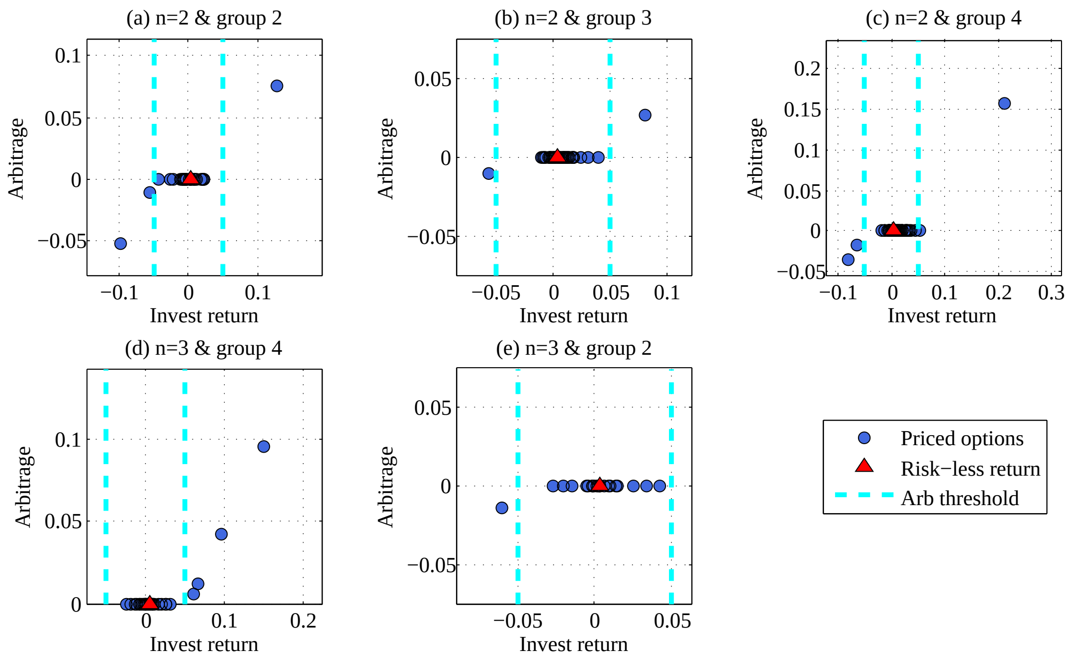

Figure 4.3: Empirical example of the arbitrage analysis under the GBM dynamic for the US market.

4.5.3 Model Validation and Robustness Test

We now examine the GBM assumption and investigate if arbitrage exists when the keywords in an option do not follow the GBM movement.

4.5.3.1 Checking the GBM Assumption

To validate the GBM assumption, two validation conditions are tested [Marathe and Ryan, 2005]: (i) the normality of change rates of log CPCs; and (ii) the independence from previous data. Normality can be either checked graphically by histogram/Q-Q plot or verified statistically by the Shapiro-Wilk test [Shapiro and Wilk, 1965]. To examine independence, we employ the autocorrelation function (ACF) [Tsay, 2005] and the Ljung-Box statistic [Ljung and Box, 1978]. Figure 4.4 provides an empirical example of the keyword ‘canon 5d’. Figure 4.4 (a)-(b) exhibit the movement of CPCs and log change rates while Figure 4.4 (c)-(d) show that the stated two conditions are satisfied in this case.

We check the discussed two conditions with the training data. As shown in Figure 4.5, there are 14.25% and 17.20% of keywords in US and UK markets that satisfy the GBM assumption, respectively. Thus 15.73% of keywords can be effectively priced into an option based on the GBM. It is worth mentioning that not all keywords follow

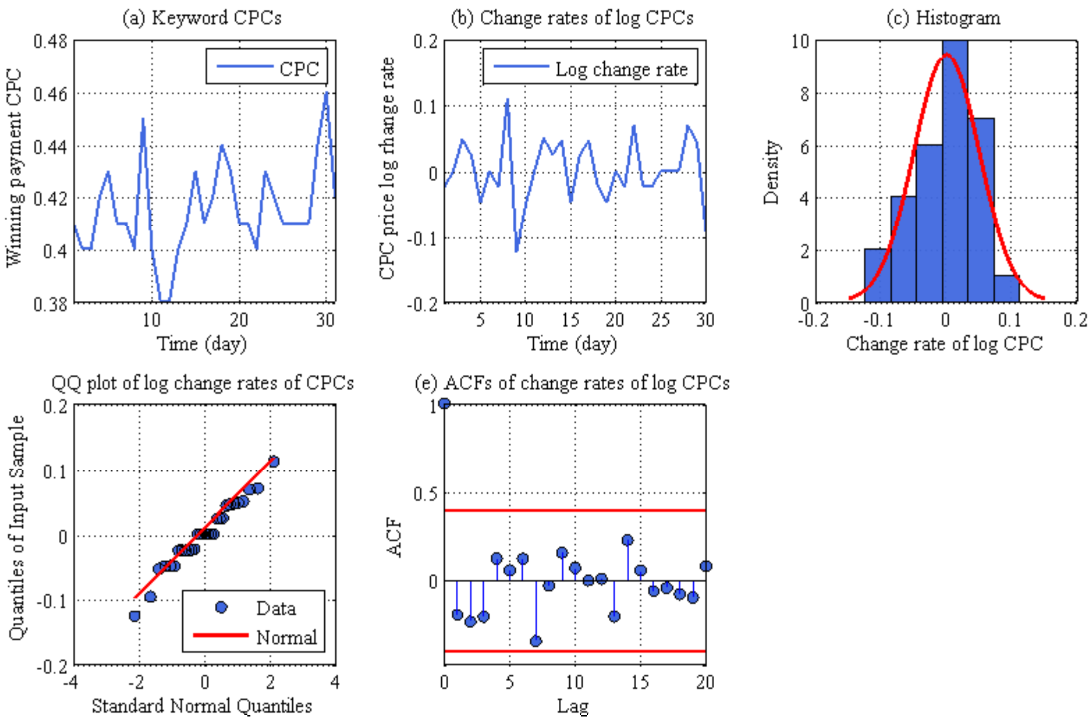

Figure 4.4: Empirical example of the GBM assumptions checking for the keyword ‘canon 5d’, where the Shapiro-Wilk test is with p-value 0.2144 and the Ljung-Box test is with p-value 0.6971.

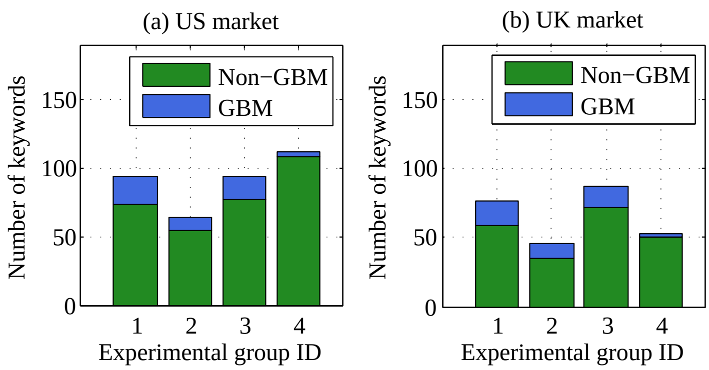

Figure 4.5: Overview of the GBM assumption checking for all candidate keywords of experimental groups in both US and UK markets.

the GBM. Next, we examine the robustness of pricing model and investigate the arbitrage based on non-GMB models.

4.5.3.2 Examining Arbitrage for Non-GBM Dynamics

Several popular stochastic processes (together with the real data) are tested to check the arbitrage in option pricing. Table 4.4 shows the candidate models and each model can

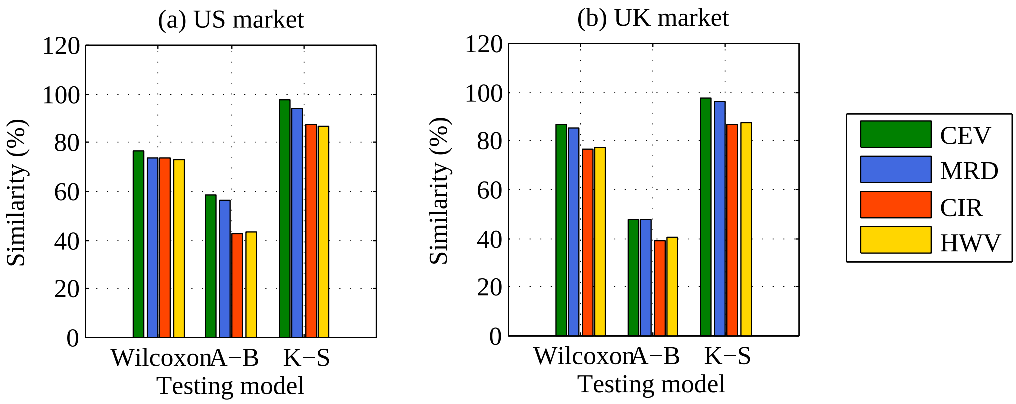

Figure 4.6: Overview of model similarity testing: Wilcoxon test, Ansari-Bradley (A-B) test and Two-sample Kolmogorov-Smirnov (K-S) test.

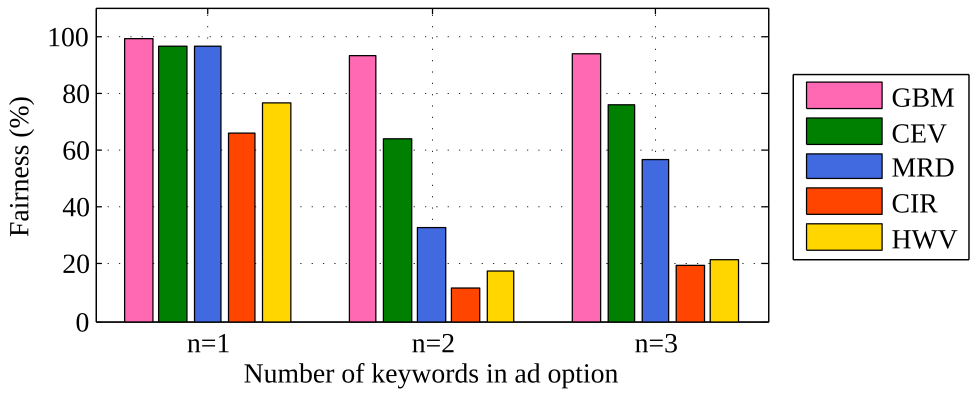

Figure 4.7: Overview of pricing model robust tests.

Table 4.4: Other underlying models or dynamics examined, where the parameter k_i is set 0.5 and the rest of parameters are estimated from the training data.

Dynamic	Stochastic differential equation (SDE)
CEV	$dC_i(t) = \mu_i C_i(t)dt + \sigma_i(C_i(t))^{1/2}dW_i(t)$
MRD	$dC_i(t) = k_i(\mu_i - C_i(t))dt + \sigma_i(C_i(t))^{1/2}dW_i(t)$
CIR	$dC_i(t) = k_i(\mu_i - C_i(t))dt + (\sigma_i)^{1/2}C_i(t)dW_i(t)$
HWV	$dC_i(t) = k_i(\mu_i - C_i(t))dt + \sigma_i dW_i(t)$

Note: the Constant elasticity of variance (CEV) model [Cox and Ross, 1976]; the Mean-reverting drift (MRD) model [Wilmott, 2006]; the Cox-Ingersoll-Ross (CIR) model [Cox et al., 1985]; the Hull-White/Vasicek (HWV) model [Hull and White, 1990].

capture certain features of time series data, such as mean-reversion, constant volatility and square root volatility [Wilmott, 2006]. The arbitrage tests here are slightly different from that of GBM. We estimate the model parameters from the actual data in the test sets instead of the learning sets and treat the actual data as one single path of each model. Hence, the simulated data has the same drift, volatility and correlations as the test data. We are now able to examine the arbitrage multiple times when the real-world environment does not follow GBM. Also, for the candidate models, hypothesis tests are used to check if the simulated path and actual data come from a same distribution. These tests include the Wilcoxon test [Wilcoxon, 1945], Ansari-Bradley test [Mood et al., 1974] and Two-sample Kolmogorov-Smirnov test [Justel et al., 1997]. Figure 4.6 summarizes the results of models' goodness-of-fit tests, where the y-axis represents the mean percentage of simulated paths not rejected by the hypothesis tests. Even though the three tests give different absolute percentages, the dynamics' performance is similar and consistent: the CEV model has the best simulations for the actual data, followed by the MRD model; the CIR and HWV models are very close.

Table 4.5 presents the arbitrage testing results for non-GBM dynamics, where most of experimental groups exhibit arbitrage. The CEV model gives the best no-arbitrage performance, showing that 78.65% of CEV-based keywords can be fairly priced by using the GBM-based option pricing model. About 53.05% of CIR and about 43% of MRD or HWV based options have no arbitrage. For single-keyword options, the fairness percentage is more than 85% across all groups. However, this figure drops to around 38% for multi-keyword options (36.27% for 2-keyword options and 42% for 3-keyword options). For the identified arbitrage, many groups (especially single-keyword options) show small arbitrage values around 10% while arbitrage explodes in some groups.

In summary, Tables 4.3 and 4.5 illustrate that our option pricing methods are effective and reasonably robust for the real sponsored search data. As shown in Figure 4.7, when the keywords's price follow a GBM (15.73%), the pricing model ensures that 95.17% of ad options are fairly priced under the 5% arbitrage precision. For the non-GBM keywords, the CEV model is the best performance model, giving 78.65% of fairness; the CIR model is worst performance model and is with only 31.97% of fairness. Overall, the best expected fairness for all keywords is 81.25% while the worst is 41.91%. We find that the increase of the number of candidate keywords in an ad option increases the likelihood of arbitrage. This is confirmed by the fact that expected fairness drops from 86.83% (99.76% GBM and 83.60% non-GBM for 1-keyword options) to 43.69% (2-keyword options) and 53.39% (3-keyword options), respectively.

Table 4.5: Overview of delta hedging arbitrage testing for non-GBM dynamics. Notations are described in Table 4.3.

Market	n	Group	N	real data + CEV simu		real data + MRD simu		real data + CIR simu		real data + HWV simu	
				$\mathbb{P}$(arb)	Mean arb	$\mathbb{P}$(arb)	Mean arb	$\mathbb{P}$(arb)	Mean arb	$\mathbb{P}$(arb)	Mean arb
US	1	1	74	2.70%	-1.97%	8.11%	-0.38%	75.68%	1.94%	56.76%	-0.01%
		2	55	0.00%	0.00%	1.82%	1.80%	16.36%	1.86%	9.09%	-0.93%
		3	77	9.09%	-10.26%	5.19%	-7.93%	42.86%	0.85%	23.38%	-1.51%
		4	108	3.70%	1.38%	2.78%	4.27%	7.41%	3.99%	8.33%	2.84%
	2	1	37	24.32%	3.63%	81.08%	-4.09%	97.30%	-16.24%	97.30%	-13.24%
		2	27	37.04%	6.01%	70.37%	5.36%	85.19%	11.01%	85.19%	10.51%
		3	38	31.58%	5.97%	31.58%	-0.41%	73.68%	-6.91%	57.89%	-5.96%
		4	54	29.63%	5.95%	81.48%	6.61%	94.44%	16.78%	94.44%	16.25%
	3	1	24	45.83%	-1.36%	79.17%	-6.04%	100.00%	-19.99%	100.00%	-17.44%
		2	18	11.11%	-2.00%	22.22%	-5.11%	55.56%	-4.39%	72.22%	-2.22%
		3	25	24.00%	7.01%	32.00%	-3.71%	84.00%	-11.14%	76.00%	-10.26%
		4	36	16.67%	3.91%	30.56%	2.66%	83.33%	2.73%	88.89%	3.38%
UK	1	1	58	0.00%	0.00%	0.00%	0.00%	74.14%	1.95%	55.17%	-1.60%
		2	35	0.00%	0.00%	2.86%	1.51%	22.86%	1.65%	14.29%	2.29%
		3	72	5.56%	-5.78%	1.39%	-10.38%	29.17%	1.03%	18.06%	-0.32%
		4	50	4.00%	5.55%	6.00%	4.47%	10.00%	4.56%	8.00%	3.45%
	2	1	29	37.93%	-0.76%	62.07%	-5.52%	89.66%	-14.55%	72.41%	-11.39%
		2	17	47.06%	2.18%	82.35%	5.00%	100.00%	9.87%	100.00%	8.62%
		3	36	19.44%	2.71%	33.33%	-1.78%	75.00%	-5.24%	61.11%	-3.58%
		4	25	64.00%	6.71%	96.00%	9.01%	92.00%	21.17%	92.00%	20.04%
	3	1	19	26.32%	-1.56%	84.21%	-5.21%	100.00%	-16.33%	78.95%	-16.34%
		2	11	18.18%	0.40%	18.18%	-1.09%	63.64%	-1.28%	63.64%	-1.05%
		3	24	16.67%	3.45%	25.00%	-1.61%	79.17%	-9.14%	66.67%	-9.23%
		4	16	37.50%	7.83%	43.75%	7.37%	81.25%	0.68%	81.25%	8.64%

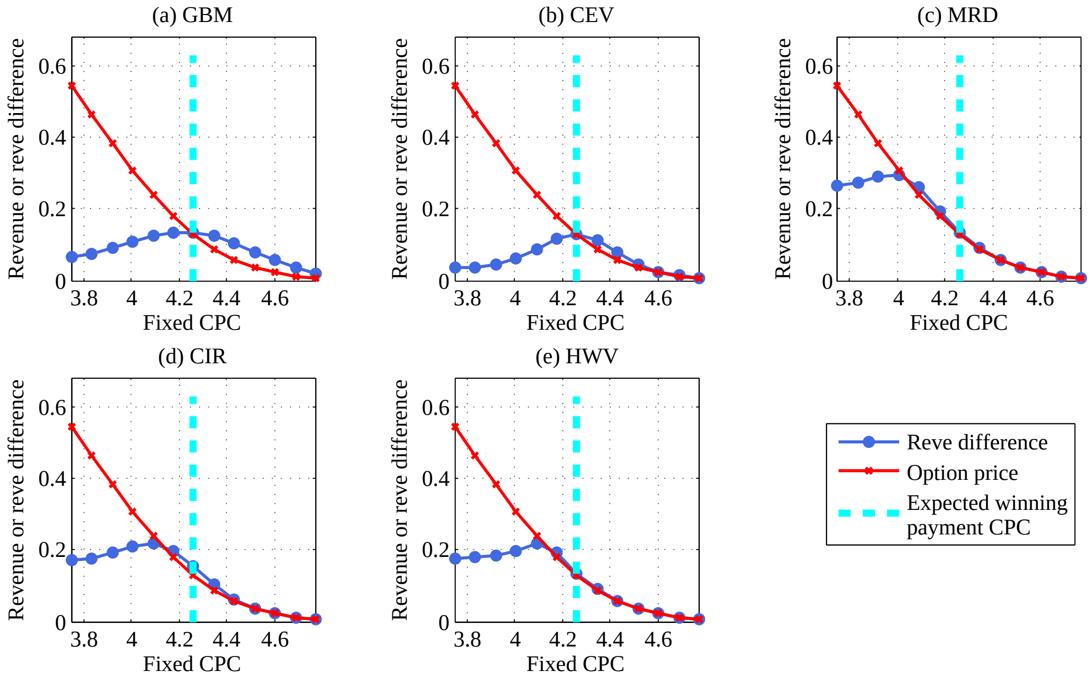

Figure 4.8: Empirical example of analysing the search engine's revenue for the keyword 'canon cameras'.

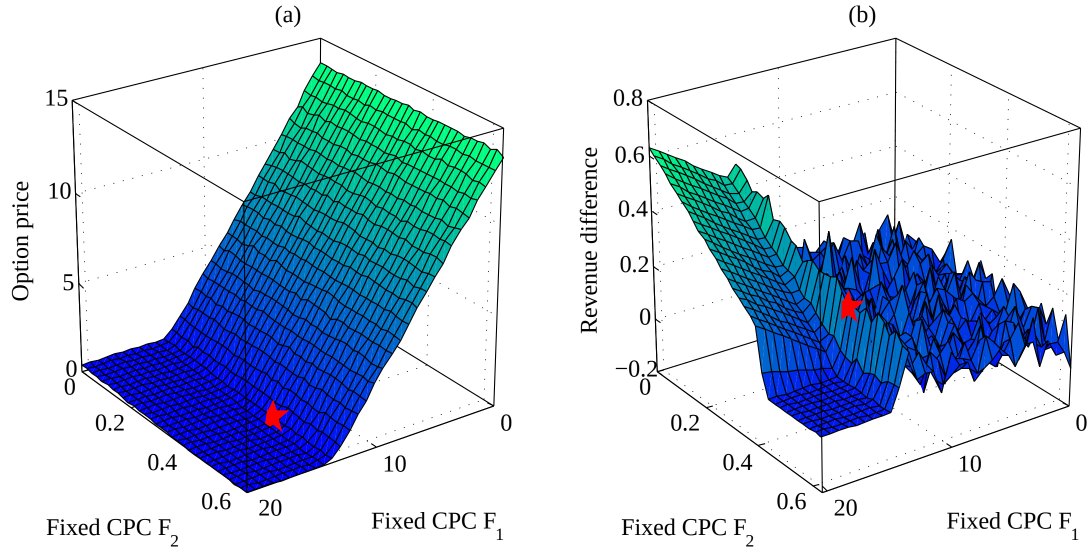

Figure 4.9: Empirical example of analysing the search engine's revenue for the keywords 'non profit debt consolidation' and 'canon 5d", where $\rho = 0.0259$.

4.5.4 Effects on Search Engine's Revenue

Let us start with the case of 1-keyword options. The example of keyword 'canon cameras' in Figure 4.8(a) illustrates (other keywords exhibit the similar pattern) the conclusions from our theoretical analysis in Section 4.4 that (i) the revenue difference between option and auction is always positive and (ii) that when the fixed CPC $F = \mathbb{E}_t^{\mathbb{Q}}[C(T)]$, the revenue difference $D(F)$ achieves its maximum and the two boundary values are approximately zero.

The non-GBM cases are further examined in Figure 4.8(b)-(e), which show that when the fixed CPC is close to zero, the revenue difference $D(F) \to 0$. This is because when the fixed CPC approximates zero, it is almost certain that the option will be used in the contract period. As such, the only income for the keyword is from the option price, which in this case is close to the CPC in the auction market (discounted back to t=0). On the other hand, if the fixed CPC is very high, it is almost certain that the option won't be used. In this case, the option price $\pi \to 0$ and the probability of exercising the option $\mathbb{P}(\mathbb{E}_t^{\mathbb{Q}}[C(T)] \geq F) \to 0$. Hence, $D(F)$ is zero. However, under the non-GBM dynamics, the point $F = \mathbb{E}_t^{\mathbb{Q}}[C(T)]$ is not the optimal value that gives the maximum $D(F)$, which indicates that arbitrage may occur.

Next, Figure 4.9 illustrates an empirical example a 2-keyword ad option. The candidate keywords are 'non profit debt consolidation' and 'canon 5d". Figure 4.9(a) tells that the higher the fixed CPCs the lower is the option price (even though the option price is less sensitive to the keyword 'canon 5d") and it achieves the maximum when all the fixed CPCs are zeros. This monotone results are as same as the 1-keyword options. Figure 4.9(b) then shows the revenue difference curve of the search engine, where the red star represents the value where $F_1 = \mathbb{E}_t^{\mathbb{Q}}[C_1(T)]$ and $F_2 = \mathbb{E}_t^{\mathbb{Q}}[C_2(T)]$. The expected revenue differences are all non-negative, showing that this 2-keyword ad option is beneficial to the search engine's revenue. However, the red star point is not the maximum difference revenue. This is different to we see in 1-keyword ad options.

For higher dimensional ad options (i.e., $n \geq 3$), we cannot graphically examine the revenue difference. However, based on the earlier discussions, we can summarize two properties. First, there are boundary values of the revenue differences. If every $F_i \to 0$, $D(\mathbf{F}) \to 0$; and if every $F_i \to \infty$, $D(\mathbf{F}) \to 0$. Second, there exists a maximum revenue difference value even though this may not at the point where $F_i = \mathbb{E}_t^{\mathbb{Q}}[C_i(T)]$.

Hence, compared to only keyword auctions, proper setting the fixed CPCs can increase the search engine's expected revenue.

4.6 Summary

In this chapter, we proposed a novel framework to provide flexible guaranteed deliveries for sponsored search, from which both buy and sell sides can benefit. On the buy side, advertisers are able to secure a certain number of clicks from their targeted keywords in the future and can decide how to advertise later. They can be released from auction campaigns and can manage price risk under the given budgets. On the sell side, search engine can sell the future clicks in advance and can receive a more stable and increased expected revenue over time. In addition, advertisers would be more loyal to a search engine due to the contractual relationships, which has the potential to boost the search engine's revenue on the long run.

We also believe that the proposed ad options will soon be welcomed by the sponsored search market. Several similar but different developments appeared in the display digital markets are able to support our point of view. They are:

09/2013: AOL's Programmatic Upfront[3].

03/2013: OpenX Programmatic Guarantee [OpenX, 2013].

10/2012: Adslot Media's Programmatic Direct Media Buying[4].

10/2012: Shiny Ads Direct's End-to-end Programmatic Direct Advertising Platform[5].

10/2012: iSOCKET's Programmatic Direct[6].

Our work differs to the above developments in many aspects. First, we focus on sponsored search while they are for display advertising. Second, the proposed ad options provide flexible guaranteed deliveries (e.g., multi-keyword targeting, multi-click exercise, early exercise, no obligation of exercise) while other recent developments do not provide such new features.

Our work leaves several directions for future research. First, to address the limitations of GBM, other stochastic processes tailored to some specific keywords are worth

[3] http://www.aolnetworks.com/programmaticupfront.
[4] https://media.adslot.com/marketplace.
[5] http://www.shinyads.com/solutions.
[6] https://www.isocket.com/programmatic-direct.

studying, such as the jump-diffusion model [Kou, 2002]. The most challenging part of this future research is that the underlying model is multi-dimensional and needs to be computational fast. Second, it would be interesting to discuss an optimal pricing and allocation model of ad options so that a search engine can algorithmically manipulate the limited future clicks in front of uncertain demand. Third, the game-theoretical pricing of ad options can be another direction.

4.7 Chapter Appendix

4.7.1 Proof of the No-Early Exercise Property for the Ad Option

Eq. (4.3) can be rewritten as $\Phi(\boldsymbol{x}) = \max\{\boldsymbol{x} - \boldsymbol{f}, 0\}$, where $\boldsymbol{x} = [x_1, \ldots, x_n]^{\mathbf{T}}$ and $\boldsymbol{f} = [f_1, \ldots, f_n]^{\mathbf{T}}$. It is not difficult to find that $\Phi(\boldsymbol{x})$ is multivariate convex. Let $0 \leq \lambda \leq 1$ and let $\boldsymbol{y} = [y_1, \ldots, y_n]^{\mathbf{T}}$, if the elements of vector $\boldsymbol{a} = \boldsymbol{y} - \boldsymbol{x}$ are all non-negative, we have

$$\Phi\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) \leq \lambda\Phi(\boldsymbol{x}) + (1-\lambda)\Phi(\boldsymbol{y}).$$

If taking $\boldsymbol{y} = [0, \ldots, 0]^{\mathbf{T}}$, and using the fact that $\Phi(\mathbf{0}) = 0$, we obtain

$$\Phi(\lambda\boldsymbol{x}) \leq \lambda\Phi\big(\boldsymbol{x}\big), \quad \text{for all } x_i \geq 0,\ 0 \leq \lambda \leq 1.$$

For $0 \leq s \leq t \leq T$, we have $0 \leq e^{-r(t-s)} \leq 1$, and then

$$\begin{aligned}
\mathbb{E}_s^{\mathbb{Q}}\big[e^{-r(t-s)}\Phi\big(\boldsymbol{X}(t)\big)\big] &\geq \mathbb{E}_s^{\mathbb{Q}}\big[\Phi\big(e^{-r(t-s)}\boldsymbol{X}(t)\big)\big] \\
&\geq \Phi\big(\mathbb{E}_s^{\mathbb{Q}}\big[e^{-r(t-s)}\boldsymbol{X}(t)\big]\big) \ \text{(by the Jenen's Inequality)} \\
&= \Phi\big(e^{rs}\mathbb{E}_s^{\mathbb{Q}}\big[e^{-rt}\boldsymbol{X}(t)\big]\big),
\end{aligned}$$

where $\mathbb{E}_s^{\mathbb{Q}}[\cdot]$ is the conditional expectation with respect to time s under the risk-neutral probability $\mathbb{Q}$. As $e^{-rt}\boldsymbol{X}(t)$ is a martingale under $\mathbb{Q}$ [Björk, 2009], we have

$$\Phi\big(e^{rs}\mathbb{E}_s^{\mathbb{Q}}\big[e^{-rt}\boldsymbol{X}(t)\big]\big) = \Phi\big(e^{rs}e^{-rs}\boldsymbol{X}(s)\big) = \Phi\big(\boldsymbol{X}(s)\big).$$

Therefore, we obtain $\mathbb{E}_s^{\mathbb{Q}}\big[e^{-r(t-s)}\Phi\big(\boldsymbol{X}(t)\big)\big] \geq \Phi\big(\boldsymbol{X}(s)\big)$, showing that $e^{-rt}\Phi\big(\boldsymbol{X}(t)\big)$ is a sub-martingale under $\mathbb{Q}$. This tells that we can price the proposed ad option as same as its European structure, focusing on the payoff on the contract expiration date. For the detailed definitions of martingale and sub-martingale, see Björk [2009].

4.7.2 Derivation of the Ad Option Pricing Formula

As the proposed ad option complements the existing keyword auctions, there may exist a situation that some advertisers make guaranteed profits from the difference of costs between the option and auctions without taking any risk. This situation is called the arbitrage opportunity [Varian, 1987]. Therefore, we must fairly evaluate the option so that arbitrage is eliminated.

We consider the advertiser buys a n-keyword m-click ad option. At time t, the difference between the option value and the market value of candidate keywords can be expressed as

$$\Pi(t) = V(t, \boldsymbol{C}(t); \boldsymbol{F}, T, m) - \sum_{i=1}^{n} \psi_i(t) C_i(t), \tag{4.21}$$

where $\psi_i(t)$ represents the number of clicks needed for the keyword K_i such that $\sum_i \psi_i(t) = m$. Here we call $\Pi(t)$ as the *value difference process*. As in Eq. (4.3) we consider the value of a n-keyword m-click option as the sum of m independent n-keyword 1-click options, for the mathematical convenience, we can rewrite Eq. (4.21) as follows

$$\Pi(t) = m\left(V(t, \boldsymbol{C}(t); \boldsymbol{F}, T, 1) - \sum_{i=1}^{n} \Delta_i C_i(t) \right), \tag{4.22}$$

where Δ_i represents the probability that the click goes for the keyword K_i and $\sum_{i=1}^{n} \Delta_i = 1$. The changes of Π over a very short period of time dt is

$$d\Pi(t) = m\left(\frac{\partial V}{\partial t} dt + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_i \sigma_j \rho_{ij} C_i C_j \frac{\partial^2 V}{\partial C_i \partial C_j} dt + \sum_{i=1}^{n} \frac{\partial V}{\partial C_i} dC_i - \sum_{i=1}^{n} \Delta_i dC_i \right). \tag{4.23}$$

We can remove the uncertain components in $d\Pi(t)$ if choosing $\Delta_i = \partial V / \partial C_i$. This is called *delta hedging* in financial option pricing [Wilmott, 2006]. Therefore, $\Pi(t)$ is now a risk-less process over time

$$d\Pi(t) = m\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i\sigma_j\rho_{ij}C_iC_j\frac{\partial^2 V}{\partial C_i \partial C_j}\right)dt. \tag{4.24}$$

Consider if the advertiser has no initial fund. He borrows the money from a bank at the risk-less bank interest rate r. Then the interest of this debt is

$$d\Pi(t) = r\Pi(t)dt = rm\left(V - \sum_{i=1}^{n}\frac{\partial V}{\partial C_i}C_i\right)dt. \tag{4.25}$$

Eqs. (4.24) and (4.25) should be equal otherwise arbitrage exists. If the risk-less growth rate of the value difference process is larger than the risk-less bank interest rate, the advertiser can obtain arbitrage by: (i) borrowing the money from bank at interest rate r to buy an ad option first; (ii) selling the ad option later to repay the bank interest. In the case when the risk-less growth rate of the value difference process is smaller than the risk-less bank interest rate, the advertiser can obtain the risk-less surplus by: (i) selling short an ad option first and deposit the revenue into bank; (ii) using the deposit money to buy the clicks of underlying keywords later. In either case, the advertiser can finally receive a risk-less surplus; therefore, arbitrage exists.

Solving Eqs. (4.24)-(4.25) gives a parabolic partial differential equation (PDE) for the no-arbitrage equilibrium as follows

$$\frac{\partial V}{\partial t} + r\sum_{i=1}^{n}\frac{\partial V}{\partial C_i}C_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 V}{\partial C_i \partial C_j}\sigma_i\sigma_j\rho_{ij}C_iC_j - rV = 0.$$

The PDE satisfies the boundary condition in Eq. (4.3). We can employing the multidimensional Feynman-Kač stochastic representation [Björk, 2009] to obtain the solution

$$V(t, \boldsymbol{C}(t); \boldsymbol{F}, T, 1) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}[\Phi(\boldsymbol{C}(T))],$$

where $\mathbb{E}_t^{\mathbb{Q}}[\cdot]$ is the conditional expectation with respect to time t under the risk-neutral probability $\mathbb{Q}$. Under this, the process $C_i(t)$ is rewritten as

$$dC_i(t) = rC_i(t)dt + \sigma_iC_i(t)dW_i^{\mathbb{Q}}(t),$$

where $W_i^{\mathbb{Q}}(t)$ is the standard Brownian motion under $\mathbb{Q}$. Therefore, the option price π_0 can be calculated by

$$\pi_0 = V(0, \boldsymbol{C}(0); \boldsymbol{F}, T, m) = mV(0, \boldsymbol{C}(0); \boldsymbol{F}, T, 1) = me^{-rT}\mathbb{E}_0^{\mathbb{Q}}[\Phi(\boldsymbol{C}(T))].$$

4.7.3 Option Pricing Formulas for Special Cases

If there is only one candidate keyword (i.e. $n = 1$), Eq. (4.6) is equivalent to the Black-Scholes-Merton (BSM) pricing formula for an European call option [Black and Scholes, 1973, Merton, 1973], so we have

$$\pi_0 = mC(0)\mathcal{N}[\zeta_1] - mFe^{-rT}\mathcal{N}[\zeta_2], \tag{4.26}$$

where $\zeta_1 = \frac{1}{\sigma\sqrt{T}}\Big(\ln\{C(0)/F\} + (r + \frac{\sigma^2}{2})T\Big)$ and $\zeta_2 = \zeta_1 - \sigma\sqrt{T}$.

If there are two candidate keywords (i.e. $n = 2$), Eq. (4.6) contains a bivariate normal distribution. We can use the formula from a dual-strike European call option [Zhang, 1998] to calculate the option price, given by

$$\begin{aligned}\pi_0 = {} & mC_1(0)\int_{-\infty}^{\zeta_1+\sigma_1\sqrt{T}} f(u)\mathcal{N}\left[\frac{q_1(u+\sigma_1\sqrt{T}) - \rho\sigma_1\sqrt{T} + \rho u}{\sqrt{1-\rho^2}}\right]du \\ & + mC_2(0)\int_{-\infty}^{\zeta_2+\sigma_2\sqrt{T}} f(v)\mathcal{N}\left[\frac{q_2(u+\sigma_2\sqrt{T}) - \rho\sigma_1\sqrt{T} + \rho v}{\sqrt{1-\rho^2}}\right]dv \\ & - me^{-rT}\left(F_1\int_{-\infty}^{\zeta_1} f(u)\mathcal{N}\left[\frac{q_1(u) + \rho u}{\sqrt{1-\rho^2}}\right]du + F_2\int_{-\infty}^{\zeta_2} f(v)\mathcal{N}\left[\frac{q_2(v) + \rho v}{\sqrt{1-\rho^2}}\right]dv\right),\end{aligned} \tag{4.27}$$

where

$$q_1(u) = \frac{1}{\sigma_2\sqrt{T}}\left(\ln\left\{\frac{F_2 - F_1 + C_1(0)e^{(r-\frac{1}{2}\sigma_1^2)T - u\sigma_1\sqrt{T}}}{C_2(0)}\right\} - (r - \frac{1}{2}\sigma_2^2)T\right),$$

$$q_2(u) = \frac{1}{\sigma_1\sqrt{T}}\left(\ln\left\{\frac{F_1 - F_2 + C_2(0)e^{(r-\frac{1}{2}\sigma_2^2)T - v\sigma_2\sqrt{T}}}{C_1(0)}\right\} - (r - \frac{1}{2}\sigma_1^2)T\right),$$

$$\zeta_1 = \frac{1}{\sigma_1\sqrt{T}}\left(\ln\{C_1(0)/F_1\} + (r - \frac{1}{2}\sigma_1^2)T \right),$$

$$\zeta_2 = \frac{1}{\sigma_2\sqrt{T}}\left(\ln\{C_2(0)/F_2\} + (r - \frac{1}{2}\sigma_2^2)T \right).$$

Eq. (4.27) appears somewhat complicated and we can further approximate the option price by using some polynomial functions. More detailed discussions are provided by Zhang [1998].

Chapter 5

Lattice Methods for Pricing Display Ad Options

This chapter studies an ad option for display advertising. Section 5.1 introduces the background and indicates the problem. Section 5.2 investigates several one-factor lattice methods reviewed in Table 2.2, which are all based on the GBM underlying model. Section 5.3 discusses our proposed lattice method for pricing a display ad option with the SV underlying model. We present several experimental results in Section 5.4 and summarise the chapter in Section 5.5. Some important mathematical results are provided in Section 5.6.

5.1 Introduction

Options, as a concept, have been introduced recently into online advertising to solve the non-guaranteed delivery problem as well as provide advertisers with greater flexibility. Moon and Kwon [2010] focused on an option for advertisers to make a choice between CPM and CPC, whereas Wang and Chen [2012] and Chen et al. [2014a] proposed ad options between buying and non-buying the future impressions. In practice, the latter has been implemented as a "First Look" tactic that is widely offered by publishers who offer prioritised access to selected advertisers within an open RTB market environment [Yuan et al., 2013]. Instead of the winning impression going to the highest bid in RTB, "First Look" affords first the right of refusal for an impression within an exchange based on a pre-negotiated floor or fixed price. If a buyer requests it, he is guaranteed to win the impression. Formally, an ad option is a contract in which an advertiser can have a right but not obligation to purchase future impressions or clicks from a specific ad slot or keyword at a pre-specified price. The pre-negotiated price is usually called the strike price in finance. In display advertising, the price can be charged as either a CPM or CPC depending on the underlying ad format. The corresponding winning payment price of impressions or clicks from real-time auctions is called the underlying price. The publisher or search engine grants this right in exchange for a certain amount of upfront fee, called the option price. Options are more flexible than guaranteed contracts as on the delivery date, if the advertiser thinks that the spot market is more beneficial, he can join online auctions as a bidder and his cost of not using an ad option is only the option price.

When evaluating ad options, the previous research [Wang and Chen, 2012, Chen et al., 2014a] is mostly restricted in their usage to those situations where the underlying price follows a GBM [Samuelson, 1965b]. According to Yuan et al. [2013], Yuan et al. [2014] and Chen et al. [2014b], there are only a very small number of ad slots or keywords whose CPM or CPC satisfies the GBM assumption. Therefore, the previous studies fail to provide an effective unified framework that covers general situations.

In this chapter, we address the issue and provide a more general pricing framework. Our option pricing is based on lattice methods and uses a stochastic volatility (SV) model to describe the underlying price movement for the cases where the GBM assumption is not valid. Based on the SV model, a censored binomial lattice is constructed for option pricing. We also examine several binomial and trinomial lattice methods to price a display ad option with the GBM underlyings and deduce a close-form solution to examine the convergence performance of these lattice methods. Our developments are validated by experiments using real advertising data. We examine the fitness of the underlying model, and illustrate that the options provide a more flexible way of selling and buying ads. In particular, we show that an advertiser can have better deliveries in a bull market, where the underlying price increases. On the other hand, a publisher or search engine is able to reduce the revenue volatility over time. In a bear market, where the underlying price decreases, there is a growth in total revenue.

5.2 Lattices for the GBM Underlying Model

This section introduces the basic settings of the lattice based option pricing framework in the context of display advertising. We examine the previous lattice methods based

Table 5.1: Summary of key notations in Chapter 5.

Notation	Description
T	Option expiration date (in terms of year).
n	Total number of time steps and the length of each time step is $\Delta t = T/n$.
$\widehat{r}, r$	Constant risk-less interest rate: $\widehat{r}$ is the discrete-time interest rate in Δt and $\widetilde{r} = 1 + \widehat{r}$; r is a continuous-time interest rate such that $e^{r\Delta t} = \widetilde{r}$.
u, m, d	State transition scale in upward, unchanged and downward movement.
$q_1, \ldots, q_n$	Risk-neutral transition probability, nodes are labelled from top to bottom.
$Q^{\{i\}}(t_k)$	Risk-neutral probability on each node, $i = 1, \ldots, k+1$.
$M_i, M(t)$	M_i is CPM at time step i, $i = 0, \ldots, n$; $M(t)$ is CPM at time t.
$C_i, C(t)$	C_i is CPC at time step i; $C(t)$ is CPC at time t.
H	Constant CTR.
R_i	Revenue in time step i, $i = 0, \ldots, n$ (see Section 5.6.1).
Φ_n	Option payoff on the expiration date (i.e., the time step n).
F^M, F^C	Strike price in terms of CPM, CPC.
π_0	Option price at time 0 (i.e., the time step 0).
$\mu, \sigma, \sigma(t)$	Constant drift, constant volatility and stochastic volatility for the underlying price.
κ, θ, δ	Constant speed, long-term mean, and volatility for the stochastic volatility model.

on the GBM assumption (see Table 2.2) and provide a comparative analysis of their convergence performances to a closed-form pricing formula (see Section 5.6.4). For the reader's convenience, the key notations used throughout the chapter are described in Table 5.1. It is worth mentioning that we here discuss the case where an ad option allows its buyer to pay a fixed CPC for display impressions. Therefore, the strike price of the option is the fixed CPC and the underlying price is the uncertain winning payment CPM from RTB, where each single impression being auctioned off is paid at the second highest bid [Google, 2011, Yuan et al., 2013]. Other ad option cases can be discussed in the same manner. For example, the case where an ad option allows its buyer to pay a fixed CPM for display impressions, or the case where an ad option allows its buyer to pay a fixed CPM or CPC for clicks.

5.2.1 Binomial Lattice

Suppose that an advertiser buys a display ad option in time 0 which allows him to purchase several impressions from a publisher's ad slot in time 1 at a fixed CPC, denoted by F^C. As impressions are normally auctioned off at a CPM value, the underlying price is the winning payment CPM from RTB, denoted by M_i, $i = 0, 1$. In time 1, the underlying price may rise or fall, denoted by $M_1^{\{u\}}$ or $M_1^{\{d\}}$. Let's consider the upward case. If $M_1^{\{u\}}/(1000H) \geq F^C$, the advertiser will exercise the option; if $M_1^{\{u\}}/(1000H) < F^C$, he will not exercise the option but join RTB instead. Note that H represents a constant CTR; therefore, the underlying and strike prices can be compared on the same measurement basis. Mathematically, we use the option payoff function $\Phi_1^{\{u\}}$ to describe the above decision making, $\Phi_1^{\{u\}} := \max\{M_1^{\{u\}}/(1000H) - F^C, 0\}$. Similarly, if the winning payment CPM moves downward, the option payoff $\Phi_1^{\{d\}} := \max\{M_1^{\{d\}}/(1000H) - F^C, 0\}$.

We adopt a general economic setting and assume that the advertiser is risk-neutral so he exercises the ad option only if the option payoff is maximised [Wilmott, 2006, Narahari, 2014]. In finance, the so-called risk-neutral probability measure [Björk, 2009] is defined by the statement that the expected risky return of an asset is equal to a risk-less bank interest return. In the online advertising environment, the risk-neutral probability measure $\mathbb{Q} = (q, 1-q)$ satisfies the following equation

$$\widetilde{r}M_0 \equiv quM_0 + (1-q)dM_0, \tag{5.1}$$

where $\widetilde{r} = (1+\widehat{r})$ is the risk-less return over the period from time 0 to time 1, $u = M_1^{\{u\}}/M_0$ and $d = M_1^{\{d\}}/M_0$ are the movement scales of CPM. Therefore, we can obtain the risk-neutral transition probability $q = (\widetilde{r} - d)/(u - d)$. Note that here q equals to q_1 in Table 2.2, which describes the probability that CPM moves upward in time 1. Since the option value can be considered as a bivariate function of time and underlying price, the option value at time 0 can be obtained by discounting the expected option value at time 1 under $\mathbb{Q} = (q, 1-q)$ [Björk, 2009, see Martingale]. The option value at time 1 is actually the option payoff; therefore, the option price at time 0 can be obtained by discounting the expected payoff, that is

$$\pi_0 = \widetilde{r}^{-1}\mathbb{E}_{\mathbb{Q}}[\Phi_1] = \widetilde{r}^{-1}\Big(q\Phi_1^{\{u\}} + (1-q)\Phi_1^{\{d\}}\Big). \tag{5.2}$$

This option price π_0 is fair because it rules out arbitrage [Varian, 1987, Björk, 2009]. Arbitrage means that an advertiser can obtain a profit larger or smaller than the risk-less bank interest rate with certainty. Consider if the option price is overestimated, i.e., $\pi_0 > \widetilde{r}^{-1}(q\Phi_1^{\{u\}} + (1-q)\Phi_1^{\{d\}})$, the advertiser can sell short an ad option at time 0 and save the money into bank to get the risk-less profit $\widetilde{r}\pi_0 - (q\Phi_1^{\{u\}} + (1-q)\Phi_1^{\{d\}})$.

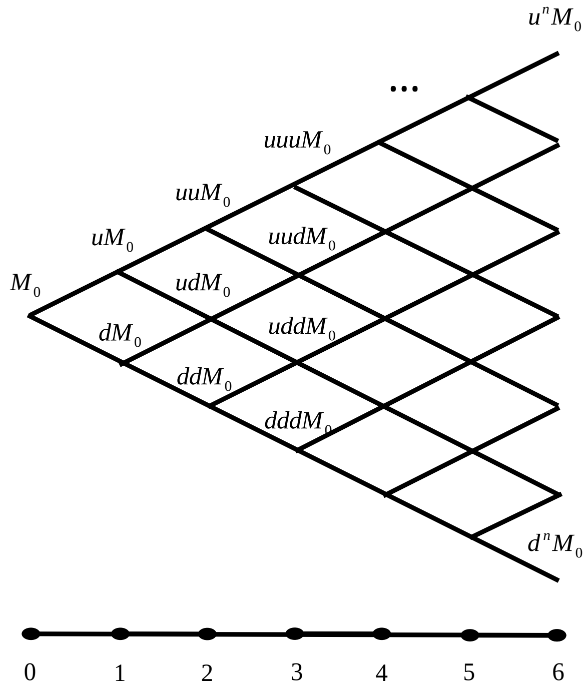

Figure 5.1: Binomial lattice for CPM. Detailed description of notations is given in Table 5.1.

Converse strategies can be used to obtain arbitrage if the option price is underestimated. Up to this point, we have discussed the option pricing framework that is the one-step binomial method, initially proposed by Sharpe [1978]. Eq. (5.2) can also be derived from the perspective of a publisher who wants to hedge the revenue risk incurred from CPM changes (see Section 5.6.1 for more details).

For a multi-step binomial lattice, as shown in Figure 5.1, the possible values of CPM and the corresponding risk-neutral transition probabilities can be estimated directly by investigating various combinations of each one-step model, so the option price π_0 can be obtained as follows

$$
\begin{aligned}
\pi_0 &= \widetilde{r}^{-n}\left(\sum_{j=0}^{n} f_n^{\{j\}}\Phi_n^{\{j\}}\right) \quad \text{(where } f_n^{\{j\}} \text{ is the risk-netural probability for } \Phi_n^{\{j\}}\text{)} \\
&= \widetilde{r}^{-n}\left(\sum_{j=0}^{n}\begin{pmatrix} n \\ j \end{pmatrix} q^j(1-q)^{n-j}\max\left\{\frac{u^j d^{n-j}M_0}{1000H} - F^C, 0\right\}\right).
\end{aligned}
$$

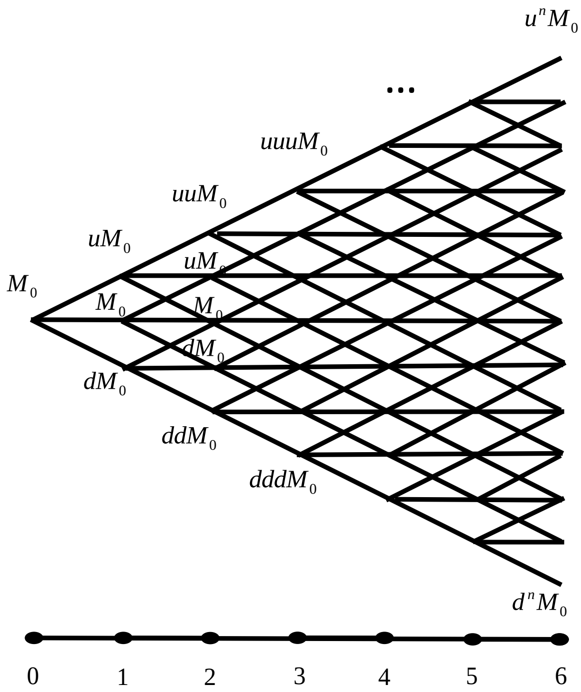

Figure 5.2: Trinomial lattice for CPM. Detailed description of notations is given in Table 5.1.

If for any $j \geq j^*$, $u^j d^{n-j} M_0/(1000H) \geq F^C$, then

$$\begin{aligned}
\pi_0 &= \widetilde{r}^{-n}\left(\sum_{j=0}^{n}\binom{n}{j} q^j(1-q)^{n-j}\frac{u^j d^{n-j} M_0}{1000H} - F^C\sum_{j=0}^{n}\binom{n}{j} q^j(1-q)^{n-j}\right) \\
&= \frac{M_0}{1000H}\sum_{j=j^*}^{n}\binom{n}{j}\left(\frac{qu}{\widetilde{r}}\right)^j\left(\frac{(1-q)d}{r}\right)^{n-j} - F^C\widetilde{r}^{-n}\sum_{j=j^*}^{n}\binom{n}{j} q^j(1-q)^{n-j} \\
&= \frac{M_0}{1000H}\sum_{j=j^*}^{n}\binom{n}{j}\widetilde{q}^j(1-\widetilde{q})^{n-j} - F^C\widetilde{r}^{-n}\sum_{j=j^*}^{n}\binom{n}{j} q^j(1-q)^{n-j} \\
&= \frac{M_0}{1000H}\psi(j^*, n, \widetilde{q}) - F^C\widetilde{r}^{-n}\psi(j^*, n, q), \qquad (5.3)
\end{aligned}$$

where $\widetilde{q} = q \times (u/\widetilde{r})$. If each time step $\Delta t = T/n$ is sufficiently small, a continuous-time closed-form formula for π_0 can be obtained (see 5.6.2 for more details), which is very similar to the BSM option pricing formula [Black and Scholes, 1973, Merton, 1973].

5.2.2 Trinomial Lattice

Figure 5.2 shows a trinomial lattice. There are 6 parameters: u, m, d are state movement scales; q_1, q_2, q_3 are the corresponding risk-neutral transition probabilities. These parameters uniquely determine the movement of CPM, which then determines a unique value of an ad option written on CPM. They must be restricted such that the constructed trinomial lattice converges to the log-normal distribution of CPM in continuous time (i.e., the GBM assumption). The moment matching technique [Cox et al., 1979] can be used to define the basic restrictions as follows:

$$q_1 + q_2 + q_3 = 1, \tag{5.4}$$

$$q_1 u + q_2 m + q_3 d = \gamma = e^{r\Delta t}, \tag{5.5}$$

$$q_1 u^2 + q_2 m^2 + q_3 d^2 = \gamma^2 \zeta = e^{2r\Delta t} e^{\sigma^2 \Delta t} \tag{5.6}$$

where $0 \leq q_1, q_2, q_3 \leq 1$. Since there are 6 parameters, 3 additional equations are necessary to define a unique solution. In this research, we examine the additional conditions discussed by previous research [Boyle, 1988, Kamrad and Ritchken, 1991, Tian, 1993] and use the same settings to price a display ad option. For the sake of completeness, a simple algorithm is provided in Section 5.6.3, which describes how to construct a trinomial lattice for the underlying price and then how to calculate the option value backward iteratively.

5.2.3 Discussion

The main results of related binomial and trinomial lattices are presented in Table 2.2. In Figure 5.3, we compare the convergence performance of these lattice methods for pricing a display ad option with the GBM underlying. The BSM-like closed-form solution (see Eq. (5.34) in Section 5.6.2) is used as the gold standard to examine how quickly that the option price calculated based on a lattice can approximate its closed-form value. Figure 5.3(a) illustrates the situation where the option value at time 0 is in the money (i.e., $M_0/(1000H) \geq F^C$) and Figure 5.3(b) shows the out of the money case (i.e., $M_0/(1000H) < F^C$). Several findings are worth mentioning here. First, the convergence rate of the trinomial lattice is faster than that of the binomial lattice; however, more nodes need to be computed for the former, i.e., $(n+1)^2$ nodes for the trinomial lattice while there are only $(n+1)(n+2)/2$ nodes for binomial lattice. Second, we find that the Tian-TRIN [Tian, 1993] model has a better convergence performance than the others.

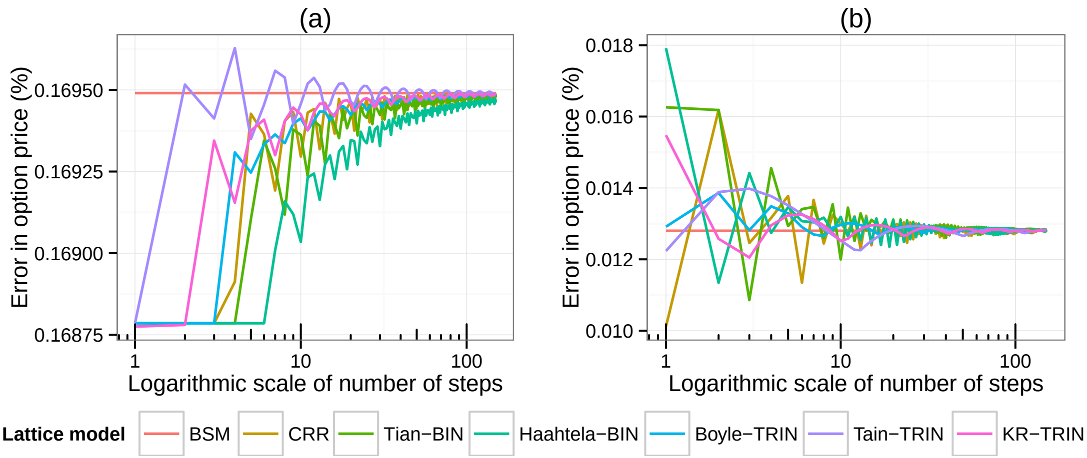

Figure 5.3: Comparison of the convergence performance of the binomial and trinomial lattices for pricing a display ad option with the GBM underlying: (a) the option value at time 0 is in the money where $M_0 = 2$, $F^C = 0.005$, $CTR = 0.3$, $r = 0.05$, $T = 31/365$ and $\sigma = 0.5$; and (b) the option value at time 0 is out of the money where $M_0 = 2$, $F^C = 0.075$, $CTR = 0.3$, $r = 0.05$, $T = 31/365$ and $\sigma = 0.5$. Detailed descriptions of notations are provided in Table 5.1.

5.3 Censored Binomial Lattice for the SV Underlying

When the GBM assumption is not valid empirically, the SV model can be used to describe the underlying price movement. Let us extend the case whereby an ad option allows its buyer to pay a fixed CPC for display impressions. The SV model for the uncertain winning payment CPM (i.e., $M(t)$) can be expressed as follows:

$$dM(t) = \mu M(t)dt + \sigma(t)M(t)dW(t), \tag{5.7}$$

$$d\sigma(t) = \kappa(\theta - \sigma(t))dt + \delta\sqrt{\sigma(t)}dZ(t), \tag{5.8}$$

where μ is the constant drift of CPM, $\sigma(t)$ is the volatility of CPM, $W(t)$ and $Z(t)$ are the standard Brownian motion under the real world probability measure $\mathbb{P}$, and κ, θ, δ are the volatility parameters. The drift factor $\kappa(\theta - \sigma(t))$ ensures the mean reversion of $\sigma(t)$ towards its long-term value θ. The volatility factor $\delta\sqrt{\sigma(t)}$ avoids the possibility of negative $\sigma(t)$ for all positive values of κ and θ.

Let $X(t) = \ln(M(t))$, the following risk-neutral form of Eq. (5.7) can be obtained (see Section 5.6.5.1 for more details):

$$dX(t) = \left(r - \frac{\sigma^2(t)}{2}\right)dt + \sigma(t)dW^{\mathbb{Q}}(t), \tag{5.9}$$

Algorithm 5.1 Censored binomial lattice method for pricing a display ad option with the SV underlying. Detailed description of notations is provided in Table 5.1.

```
function OptionPricingCensoredBinLattice(M_0, σ_0, κ, θ, δ, H, T, n, r, F^C)
    Δt ← T/n; r̃ ← e^{rΔt};
    for k ← 0 to n − 1 do
        i ∈ nodes in time step k
        if i = 1 then
            Step ①;
        else
            Step ②;
        end if
    end for
    π_0 ← Eq. (5.21) (see Step ③);
end function
```

where r is the constant continuous-time risk-less interest rate and $W^{\mathbb{Q}}$ is a standard Brownian motion under the risk-neutral probability measure $\mathbb{Q}$. The process $X(t)$ can be weakly approximated by a series of binomial processes, say $\widetilde{X}(t_i), i = 1, \ldots, n$. The approximation conditions are discussed by Nelson and Ramaswamy [1990] (see 5.6.4 for more details).

In Algorithm 5.1, we present our method of calculating the option price for a display ad option whose underlying is the SV model. Simply, a binomial lattice for $\widetilde{X}(t_i)$ is first constructed to approximates $X(t)$ weakly. The lattice is constructed from time step 0 to time step n, and at each time step, nodes are calculated from top to bottom. In the following discussion, we explain the details of Steps ①-③. Figure 5.4 illustrates the calculation from time step k to time step $k+1$.

Step ① We start the estimation from the first node $\widetilde{X}^{\{1\}}(t_k)$ in Figure 5.4, whose two successors can be expressed as follows

$$\widetilde{X}^{\{1,u\}}(t_k + \Delta t) = (J^{\{1\}}(t_k) + 1)\sigma(t_k + \Delta t)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(t_k + \Delta t)}{2}\right)\Delta t, \quad (5.10)$$

$$\widetilde{X}^{\{1,d\}}(t_k + \Delta t) = (J^{\{1\}}(t_k) - 1)\sigma(t_k + \Delta t)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(t_k + \Delta t)}{2}\right)\Delta t, \quad (5.11)$$

where $J^{\{1\}}(t_k)\sigma(t_k + \Delta t)\sqrt{\Delta t}$ is the point on the grid closest to $\widetilde{X}^{\{1\}}(t_k)$, given by

$$J^{\{1\}}(t_k) = \inf_{J^* \in \mathbb{N}} \left| J^* \times \sigma(t_k + \Delta t)\sqrt{\Delta t} - \widetilde{X}^{\{1\}}(t_k) \right|, \quad (5.12)$$

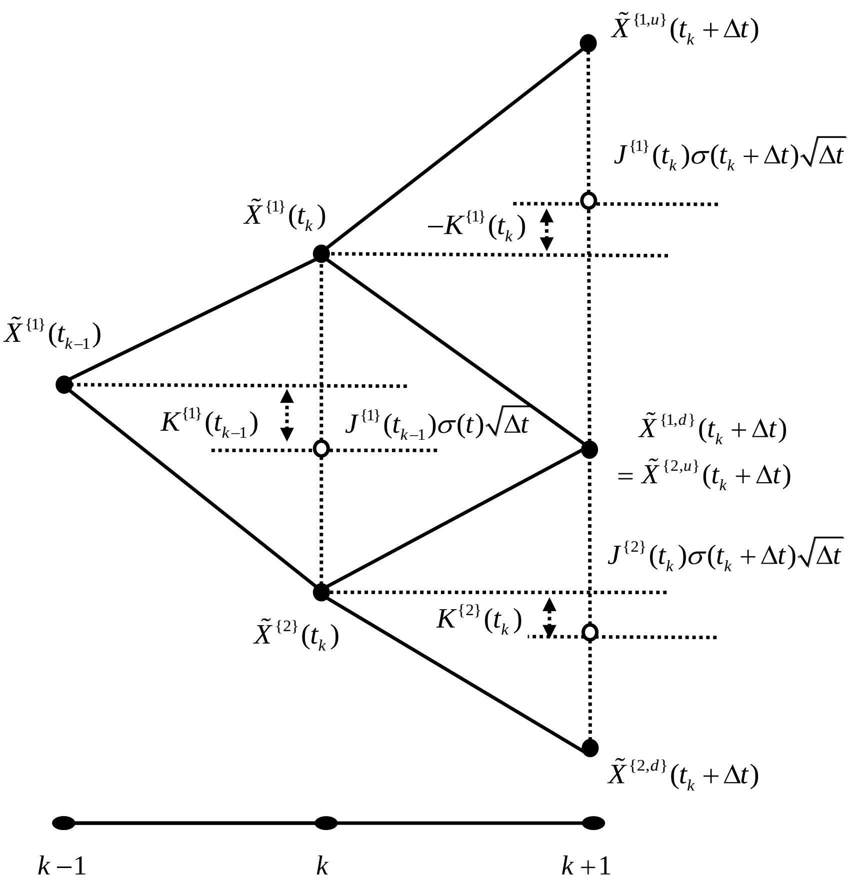

Figure 5.4: Censored binomial lattice for the SV underlying. Detailed description of notations is provided in Table 5.1.

and $\sigma(t_k + \Delta t)$ can be estimated by (see Section 5.6.5.2)

$$\sigma(t_k + \Delta t) = \sigma(t_0)e^{-\kappa(t_k+\Delta t)} + \theta(1 - e^{-\kappa(t_k+\Delta t)}). \tag{5.13}$$

Eqs. (5.10)-(5.11) verify the conditions that a binomial lattice can be used to approximate a general diffusion process (see Eqs. (5.37)-(5.39) in Section 5.6.4). Eqs. (5.10)-(5.11) can be rewritten in terms of their conditional increments as follows:

$$\widetilde{X}^{\{1,u\}}(t_k + \Delta t) - \widetilde{X}^{\{1\}}(t_k) = \sigma(t_k + \Delta t)\sqrt{\Delta t} - K^{\{1\}}(t_k) + \left(r - \frac{\sigma^2(t_k + \Delta t)}{2}\right)\Delta t, \tag{5.14}$$

$$\widetilde{X}^{\{1,d\}}(t_k + \Delta t) - \widetilde{X}^{\{1\}}(t_k) = -\sigma(t_k + \Delta t)\sqrt{\Delta t} - K^{\{1\}}(t_k) + \left(r - \frac{\sigma^2(t_k + \Delta t)}{2}\right)\Delta t, \tag{5.15}$$

where $K^{\{1\}}(t_k)$ is the grid adjusting parameter for the successors of the first node at time t_k. As shown in Figure 5.4, the value of $K^{\{i\}}(t_k)$, $i = 1, 2, \ldots, 2k-1$, can be either positive or negative, To satisfy the approximation condition set in Eq. (5.36) (see Section 5.6.4), the following equation holds:

$$\mathbb{E}\Big[\widetilde{X}^{\{1\}}(t_k+\Delta t) - \widetilde{X}^{\{1\}}(t_k) \mid \mathcal{F}(t_k)\Big] = \left(r - \frac{\sigma^2(t_k+\Delta t)}{2}\right)\Delta t \tag{5.16}$$

Then, the following system of equations is obtained

$$\begin{aligned}\Big(\sigma(t_k+\Delta t)\sqrt{\Delta t} - K^{\{1\}}(t_k)\Big)\frac{q_1^{\{1\}}(t_k)}{Q^{\{1\}}(t_k)} + \Big(-\sigma(t_k+\Delta t)\sqrt{\Delta t} - K^{\{1\}}(t_k)\Big)\frac{q_2^{\{1\}}(t)}{Q^{\{1\}}(t)} &= 0,\\ q_1^{\{1\}}(t_k) + q_2^{\{1\}}(t_k) &= Q^{\{1\}}(t),\end{aligned}$$

where $q_1^{\{1\}}(t_k)$ and $q_2^{\{1\}}(t_k)$ are the risk-neutral probabilities that the successor of the first node at time t_k rises or falls in time $t_k+\Delta t$, and $Q^{\{1\}}(t_k)$ is the risk-neutral probability for the first node at time t_k. Solving the above equations gives

$$\begin{aligned}q_1^{\{1\}}(t_k) &= \begin{cases}\frac{Q^{\{1\}}(t_k)}{2}\left(1+\frac{K^{\{1\}}(t_k)}{\sigma(t_k+\Delta t)\sqrt{\Delta t}}\right), & \text{if } 0 \le \frac{Q^{\{1\}}(t_k)}{2}\left(1+\frac{K^{\{1\}}(t_k)}{\sigma(t_k+\Delta t)\sqrt{\Delta t}}\right) \le Q^{\{1\}}(t_k),\\ 0, & \text{if } \frac{Q^{\{1\}}(t_k)}{2}\left(1+\frac{K^{\{1\}}(t_k)}{\sigma(t_k+\Delta t)\sqrt{\Delta t}}\right) < 0,\\ Q^{\{1\}}(t_k), & \text{if } \frac{Q^{\{1\}}(t_k)}{2}\left(1+\frac{K^{\{1\}}(t_k)}{\sigma(t_k+\Delta t)\sqrt{\Delta t}}\right) \ge Q^{\{1\}}(t_k),\end{cases}\\ &= \max\left\{0, \min\left\{Q^{\{1\}}(t_k), \frac{Q^{\{1\}}(t_k)}{2}\left(1+\frac{K^{\{1\}}(t_k)}{\sigma(t_k+\Delta t)\sqrt{\Delta t}}\right)\right\}\right\}, \qquad (5.17)\\ q_2^{\{1\}}(t_k) &= Q^{\{1\}}(t_k) - q_1^{\{1\}}(t_k). \qquad (5.18)\end{aligned}$$

Eqs. (5.17) and (5.18) show that transition probabilities $q_1^{\{1\}}(t_k)$ and $q_2^{\{1\}}(t_k)$ are censored in the approximation.

Step ② We then move to other nodes and construct their successors in the same manner. However, as some nodes in the next step are recombining, the following equations hold for $1 \le i \le k$:

$$\begin{aligned}\widetilde{X}^{\{i,d\}}(t_k+\Delta t) &= (J^{\{i\}}(t_k)-1)\sigma(t_k+\Delta t)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(t_k+\Delta t)}{2}\right)\Delta t,\\ = \widetilde{X}^{\{i+1,u\}}(t_k+\Delta t) &= (J^{\{i+1\}}(t_k)+1)\sigma(t_k+\Delta t)\sqrt{\Delta t} + \left(r - \frac{\sigma^2(t_k+\Delta t)}{2}\right)\Delta t,\end{aligned}$$

therefore, $J^{\{i+1\}}(t_k) = J^{\{i\}}(t_k) - 2$ and $K^{\{i+1\}}(t_k) = J^{\{i+1\}}(t_k)\sigma(t_k + \Delta t)\sqrt{\Delta t} - \widetilde{X}^{\{i+1\}}(t_k)$. The transition probabilities for the node $\widetilde{X}^{\{i+1\}}(t_k)$ can be given by

$$q_1^{\{i\}}(t_k) = \max\left\{0, \min\left\{Q^{\{i\}}(t_k), \frac{Q^{\{i\}}(t_k)}{2}\left(1 + \frac{K^{\{i\}}(t_k)}{\sigma(t_k + \Delta t)\sqrt{\Delta t}}\right)\right\}\right\}, \tag{5.19}$$

$$q_2^{\{i\}}(t_k) = Q^{\{i\}}(t_k) - q_1^{\{i\}}(t_k). \tag{5.20}$$

Step ③ We follow the calculation steps ①-② for each time step until the contract expiration date, and finally obtain $Q^{\{i\}}(t_n)$ and $\widetilde{X}^{\{i\}}(t_n)$, for all nodes ($i = 1, \ldots, n+1$) at time step n. Then the option price can be obtained as follows:

$$\pi_0 = \widetilde{r}^{-n} \sum_{i=1}^{n+1} Q^{\{i\}}(t_n) \max\left\{\frac{1}{1000H} e^{\widetilde{X}^{\{i\}}(t_n)} - F^C, 0\right\}. \tag{5.21}$$

Similar to Algorithm 5.2, the option value can also be calculated recursively over the lattice.

In the above discussion, we follow Florescu and Viens [2008] to construct the binomial lattice and use variables $K^{\{i\}}(t_k)$ and $J^{\{i\}}(t_k)$ to tune the grid so that the constructed framework is recombining. In the meantime, it satisfies the approximation conditions proposed by Nelson and Ramaswamy [1990]. We here use a modified mean-reverting process (i.e., the Cox-Ingersoll-Rubinstein (CIR) model [Cox et al., 1985]) for the volatility underlying so that the volatility will always be non-negative. We also simplify the calculation of node parameters in Florescu and Viens [2008]. Since the transition probabilities are censored directly at each node, $K^{\{i\}}(t_k)$, $J^{\{i\}}(t_k)$ and $Q^{\{i\}}(t_k)$ can be calculated sequentially from top to bottom alongside the lattice construction for the underlying price. Once the upper node is calculated, it can be used to update the value of its lower node. Hence, the risk-neutral probability distribution $Q^{\{i\}}(t_k)$ for each node can be quickly computed as follows:

$$Q^{\{i\}}(t_k + \Delta t) = \begin{cases} q_1^{\{1\}}(t_k), & \text{if } i = 1, \\ q_2^{\{i-1\}}(t_k) + q_1^{\{i\}}(t_k), & \text{if } 1 < i < k+1, \\ q_2^{\{k+1\}}(t_k), & \text{if } i = k+1, \end{cases}$$

and $Q(t_0) = 1$.

Figure 5.5 presents an empirical example of constructing a censored binomial lattice for pricing a display ad option written on an ad slot from a SSP in the UK. The given values of the model parameters are estimated from the training data. Figure 5.5(a) shows a censored binomial lattice for the underlying CPM and Figure 5.5(b) illustrates how the option value is calculated backward iteratively from the expatriation date to time 0. For the sake of comparison, Figure 5.6 illustrates the binomial lattices constructed by the CRR model with the same parameter settings. Obviously, the changing volatility can be found in Figure 5.5(a) while 5.6(a) exhibits a constant volatility over time. We find that the option price given by the SV model is slightly smaller than that of the CRR model. This is because the long-term mean value of volatility is 0.2959, smaller than its initial value 0.8723. Therefore, the drift drags the volatility downside to its long-term level and the option value based on the SV model contains less risk than the CRR model.

5.4 Experiments

Our experimental results are presented in this section. We examine the GBM assumption with the real advertising data, compare the goodness-of-fit of the underlying models, analyse if an advertiser can have better deliveries under a fixed daily budget, and discuss the effects on the publisher's (or search engine's) revenue.

5.4.1 Datasets and Experimental Design

The following two datasets are used in the experiments (see Table 5.2): a RTB dataset from a SSP in the UK; and a sponsored search dataset from Google AdWords. The RTB dataset contains all advertisers' bids and the corresponding winning payment CPMs (per transaction). The Google dataset is obtained by using Google's Traffic Estimation service [Yuan and Wang, 2012], in which we remove 21 keywords that have over 30% missing values and also 115 keywords whose CPCs are all zero.

Tables 5.3-5.4 illustrate our experimental settings. Each dataset is divided into several experimental groups, each of which consists of one training, one development and one test set. The model parameters are estimated in the training set. Display ad options are priced in the development set. The actual bids in the test set are used to examine the priced options. The default value of CTR in the experiments is set to 0.03.

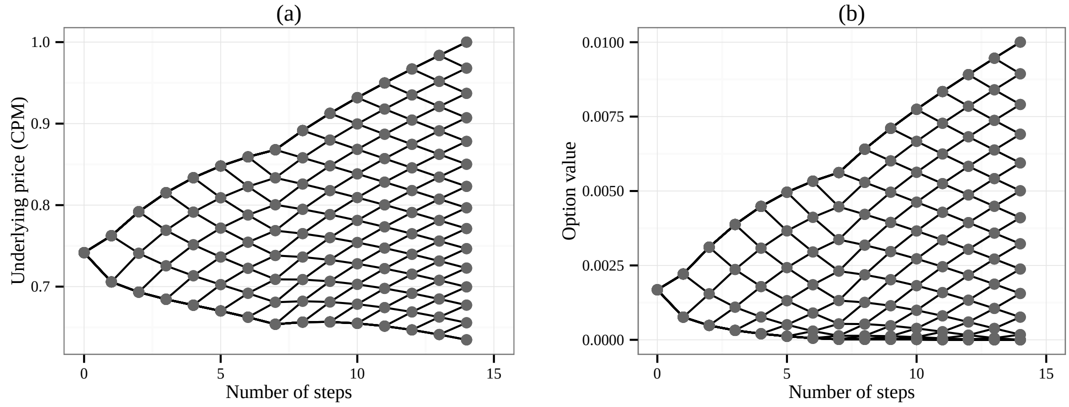

Figure 5.5: Empirical example of binomial lattices for an ad slot from the SSP dataset: (a) the censored binomial lattice for CPM based on the SV model, where $r = 0.05, T = 0.0384, n = 14, CPM = 0.7417, \sigma_0 = 0.8723, \kappa = 96.4953, \theta = 0.2959, \delta = 14.9874$; (b) the censored binomial lattice for the option value. The model parameters are estimated based on the training data.

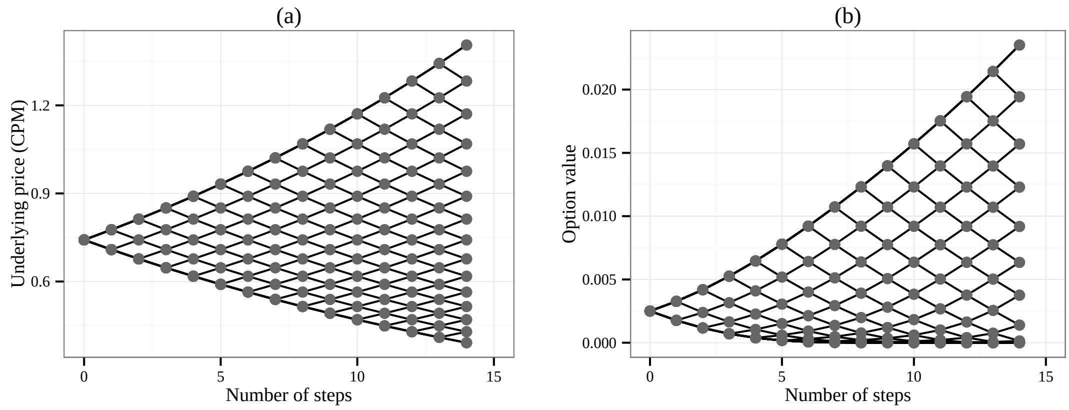

Figure 5.6: Example of binomial lattices for the same ad slot in Figure 5.5: (a) the CRR binomial lattice for CPM based on the GBM model, where $r = 0.05, T = 0.0384, n = 14, CPM = 0.7417, \sigma_0 = 0.8723$. Here we use the same parameters' values in Figure 5.5; (b) the CRR binomial lattice for the option value.

5.4.2 Fitness of GBM and SV Models

The following two conditions hold if the GBM assumption is valid empirically [Marathe and Ryan, 2005]: (i) the normality of the logarithm ratios of the winning payment price[1]; and (ii) the independence of the logarithm ratios from the previous data. Normality can be graphically checked by a histogram and Q-Q plot, and be statistically verified by the Shapiro-Wilk test [Shapiro and Wilk, 1965]. To examine the independence, we use the autocorrelation function (ACF) [Tsay, 2005] and the

[1]The logarithm ratio of winning payment price L_i is defined by $L_i = \ln(M_{i+1}/M_i)$ or $L_i = \ln(C_{i+1}/C_i)$.

Table 5.2: Summary of datasets for experiments.

Dataset	SSP	Google AdWords
Period	08/01/2013 - 14/02/2013	26/11/2011 - 14/01/2013
No. of ad slots or keywords	31	557
No. of advertisers	374	$\times$
No. of impressions	6646643	$\times$
No. of bids	33043127	$\times$
Winning payment price	$\surd$	$\surd$
Bid quote	GBP/CPM	GBP/CPC

Table 5.3: Experimental settings of the SSP dataset.

Training set (31 days)	Development & test set (7 days)
08/01/2013-07/02/2013	08/02/2013-14/02/2013

Table 5.4: Experimental settings of the Google AdWords dataset.

Market	Group	Training set (31 days)	Development & test set (31 days)
US	1	25/01/2012-24/02/2012	24/02/2012-25/03/2012
	2	30/03/2012-29/04/2012	29/04/2012-31/05/2012
	3	10/06/2012-12/07/2012	12/07/2012-17/08/2012
	4	10/11/2012-11/12/2012	11/12/2012-10/01/2013
UK	1	25/01/2012-24/02/2012	24/02/2012-25/03/2012
	2	30/03/2012-29/04/2012	29/04/2012-31/05/2012
	3	12/06/2012-13/07/2012	13/07/2012-19/08/2012
	4	18/10/2012-22/11/2012	22/11/2012-24/12/2012

Ljung-Box statistic [Ljung and Box, 1978]. If the winning payment price satisfies the GBM assumption, we evaluate the ad option by using the Tian-TRIN model (or the BSM-like closed-form formula). If the GBM assumption is not valid empirically, we develop a SV model and price the ad option by using the censored binomial lattice method.

Figure 5.7 presents an empirical example of testing the GBM assumption for an ad slot from the SSP dataset, where the underlying winning CPM cannot be described accurately as a GBM. In fact, none of the 31 ad slots in the SSP dataset satisfy the GBM model. Therefore, we use the SV model for the ad slots in the SSP dataset. Figure 5.8 presents an example of a keyword from the Google dataset. The keyword's winning CPC satisfies the GBM assumption. The log-normality of CPC is validated in Figure 5.8(a)-(c) and the independence is confirmed by Figure 5.8(d). The overview results of the Google dataset is shown in Figure 5.9. There are 14.25% and 17.20% of

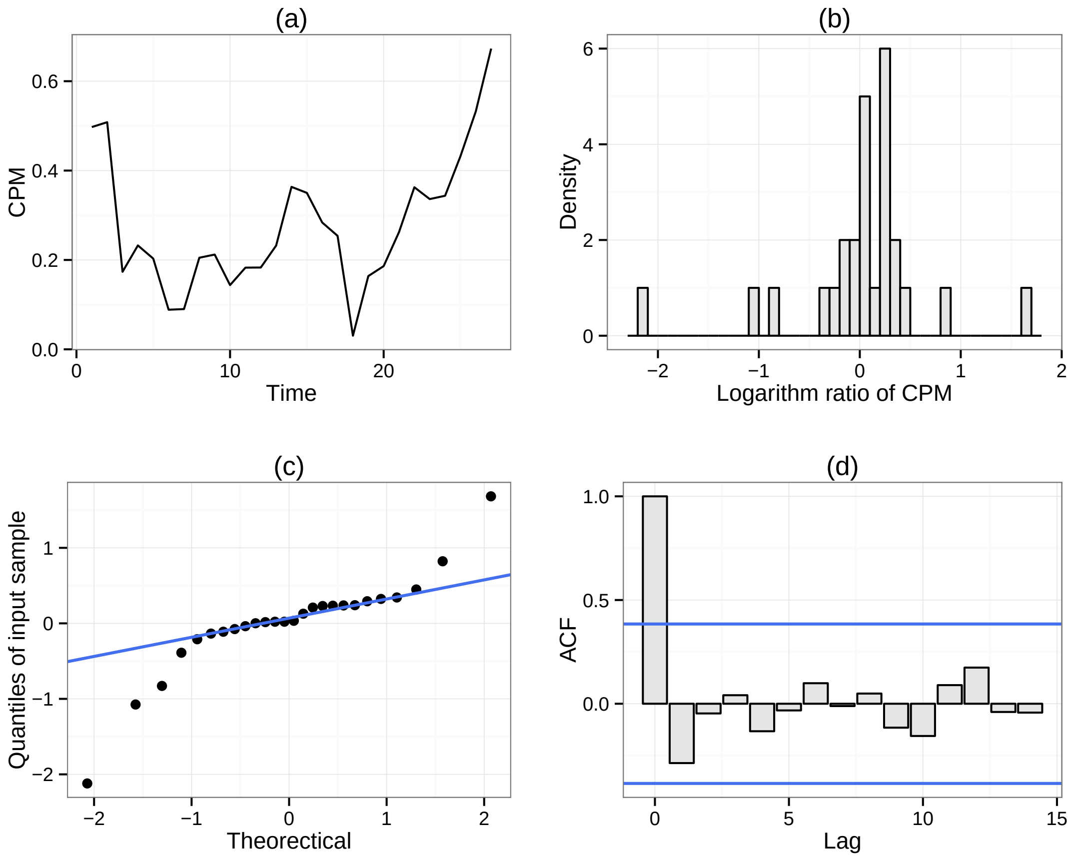

Figure 5.7: Empirical example of testing the GBM conditions on an ad slot from the SSP dataset: (a) the plot of the average daily winning payment CPMs from auctions; (b) the histogram of the logarithm ratios of the CPM, i.e., $\ln(M_{i+1}/M_i)$, $i = 1, \ldots, n-1$; (c) the QQ plot of the logarithm ratios; (d) the plot of the ACFs of the logarithm ratios. The Shapiro-Wilk test is with p-value 0.0009 and the Ljung-Box test is with p-value 0.1225.

the keywords in the US and UK markets respectively that can be accurately described by the GBM model. We will price the remaining keywords using the SV model.

Figure 5.10 gives an empirical example of the model fitness test for the situation where the GBM assumption is not valid. We give three different instances of simulated paths from the GBM and SV models for the same keyword. Figure 5.10(a),(c),(d) compares the simulations from these two models with the actual winning payment CPCs in real-time auctions. The smooth movement pattern of these three instances is examined in Figure 5.10(b),(d),(f). We find that the SV model has a better fitness to the data. In addition, we use the Euclidean distance (also called the L-2 distance) to examine the similarity of a simulated path and the test data. The overall results of the ad slots and keywords in our datasets are presented in Tables 5.5-5.6, showing that the SV model has a general better fitness to the real data.

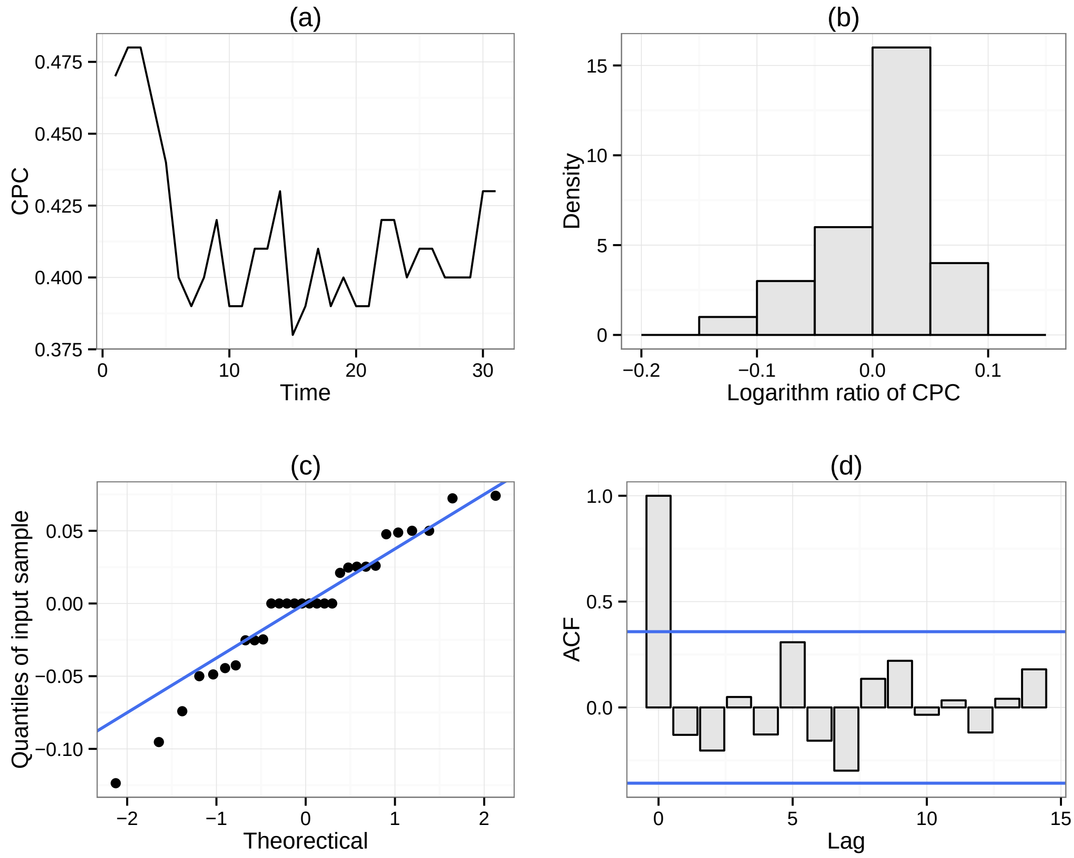

Figure 5.8: Empirical example of testing the GBM conditions on the keyword "canon 5d" from the Google AdWords dataset: (a) the plot of average daily winning payment CPCs; (b) the histogram of logarithm ratios of CPC, i.e., $\ln(C_{i+1}/C_i)$, $i = 1, \dots, n-1$; (c) the QQ plot of the logarithm ratios; (d) the plot of the ACFs of the logarithm ratios. The Shapiro-Wilk test is with p-value 0.2144 and the Ljung-Box test is with p-value 0.6971.

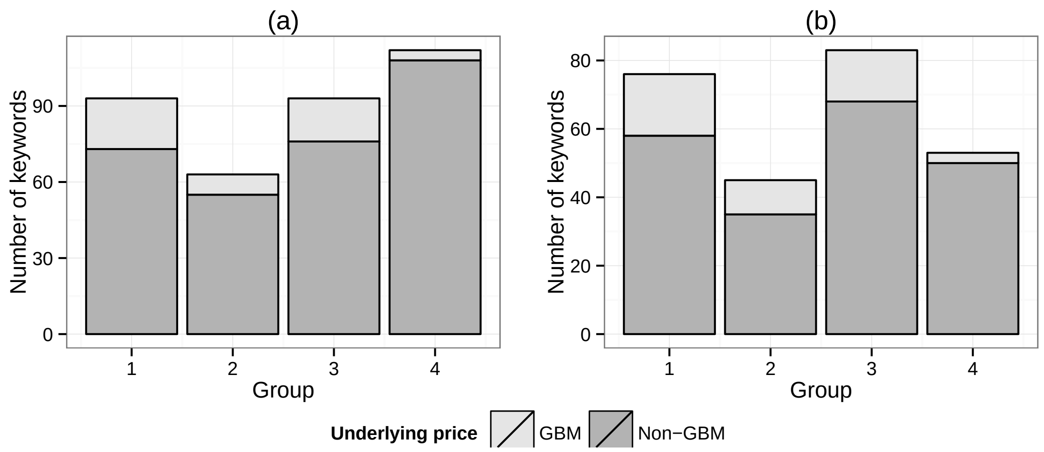

Figure 5.9: Summary of the GBM conditions test for all keywords in the Google AdWords dataset in (a) the US market; and (b) the UK market.

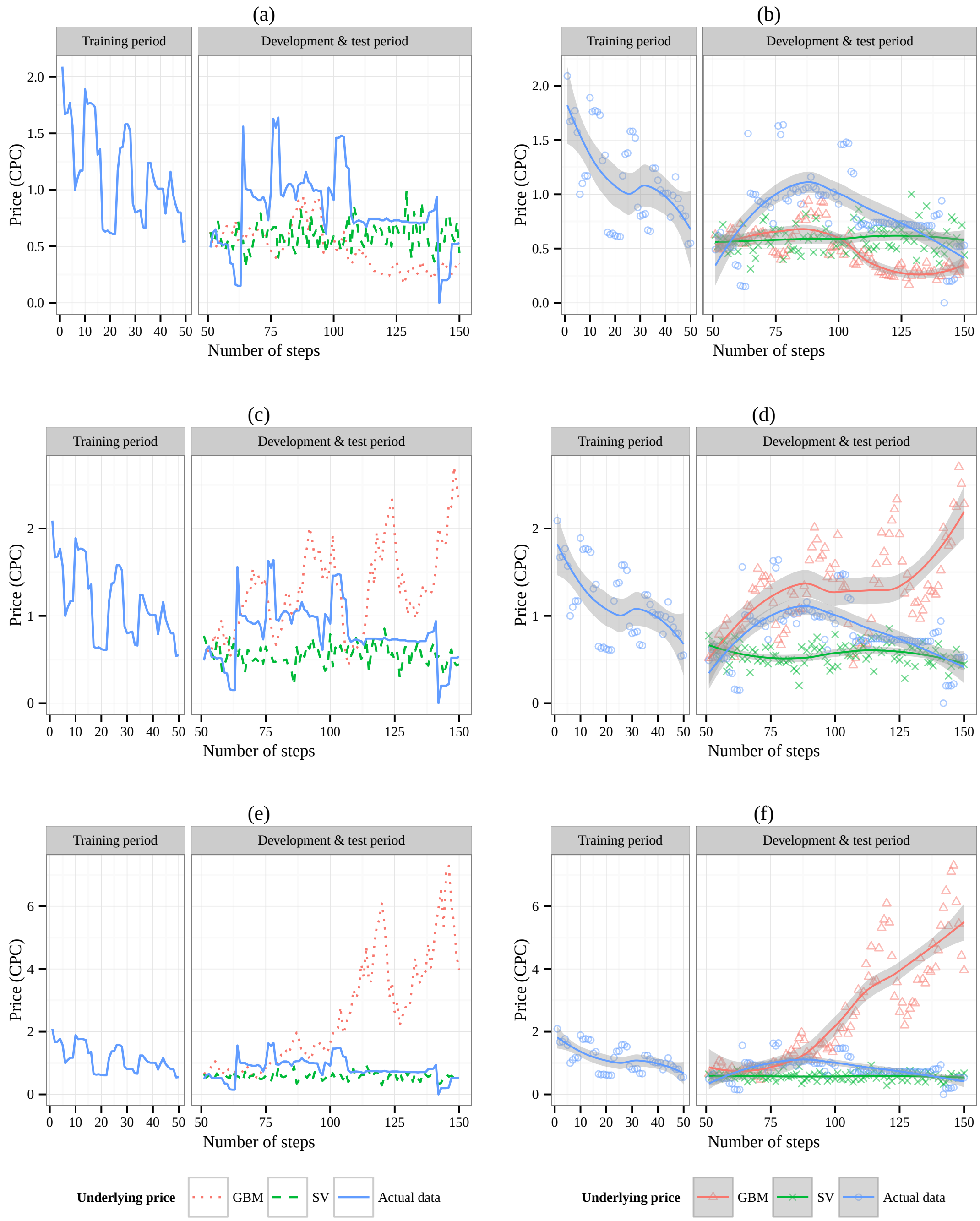

Figure 5.10: Empirical example of comparing the fitness of GBM and SV models to the keyword “kinect for xbox 360” from the Google AdWords dataset. The training period is from time step 1 to 50, the development and test periods are from time step 51 to 150. Plot (a), (c), (e) illustrates three instances of simulated paths from the estimated GBM and SV, respectively. Plot (b), (d), (f) provides the corresponding smooth pattern and confidence interval of plot (a), (c), (e).

Table 5.5: Comparing the model fitness for all 31 ad slots in the SSP dataset. L-2 distance is the Euclidean distance, and the number represents the percentage of ad slots which shows that the SV model has a better fitness (i.e., a smaller L-2 distance).

Training set (31 days)	Development & test set (7 days)	L2 distance of simu paths	L2 distance of smoothed simu paths
08/01/2013-07/02/2013	08/02/2013-14/02/2013	54.8387%	67.7419%

Table 5.6: Comparing the model fitness for the non-GBM keywords in the Google AdWords dataset. L-2 distance is the Euclidean distance, and the number represents the percentage of non-GBM keywords which shows that the SV model has a better fitness (i.e., a smaller L-2 distance).

Market	Group	Training set	Development & test set (31 days)	L2 distance of simu paths	L2 distance of smoothed simu paths
US	1	25/01/2012-24/02/2012	24/02/2012-25/03/2012	82.8571%	80.0000%
	2	30/03/2012-29/04/2012	29/04/2012-31/05/2012	94.8718%	96.1538%
	3	10/06/2012-12/07/2012	12/07/2012-17/08/2012	64.2857%	64.2857%
	4	10/11/2012-11/12/2012	11/12/2012-10/01/2013	98.1481%	100.0000%
UK	1	25/01/2012-24/02/2012	24/02/2012-25/03/2012	96.3636%	90.9091%
	2	30/03/2012-29/04/2012	29/04/2012-31/05/2012	98.2456%	94.7368%
	3	12/06/2012-13/07/2012	13/07/2012-19/08/2012	58.0645%	67.7419%
	4	18/10/2012-22/11/2012	22/11/2012-24/12/2012	72.2222%	80.5556%

5.4.3 Delivery Performance for Advertiser

Tables 5.7-5.8 provide an empirical example that compares an advertiser's delivery performance between RTB and ad options. Tables 5.7 shows the advertiser's delivery performance in RTB with a fixed daily budget. If the supplied impressions are at same levels and if the average winning payment CPMs increase, the advertiser will receive fewer impressions. In Table 5.8, the advertiser buys several display ad options in advance. Consider if he purchases an ad option with expiration date 08/02/2013, he has the right to secure impressions that will be created on 08/02/2013 at a fixed CPC. Here we consider the advertiser uses the daily budget from the corresponding delivery date to pay the upfront option price. Therefore, as shown in Table 5.8, the advertiser's advertising strategy is to purchase as many ad options in advance as possible, and the remaining daily budgets will be used on the corresponding delivery dates. Actual bids from RTB are used to simulate the real-time feeds of the spot market, so if the market value of a click is higher than the fixed payment, the advertiser will use ad options to secure the needed clicks and then pay the fixed CPCs accordingly. Otherwise, the advertiser will obtain the equivalent clicks from RTB. Our example shows a "bull market" where the average spot CPM in the test set is far higher than the initial CPM. Therefore, the bought ad options would be actively used by the advertiser to purchase the clicks. Compared to Table 5.7, the advertiser can receive more clicks (increased by 20.92%) in a bull market via ad options.

The overall results are presented in Tables 5.9-5.10. For the SSP dataset, we consider the ad options that allow advertisers to pay a fixed CPC to purchase impressions of targeted ad slots. For the Google dataset, we consider the ad options that allow advertisers to pay a fixed CPM to purchase clicks of their targeted keywords. To summarise, we find that an advertiser's daily budget can be used more effectively in a bull market and that his delivery increases as well. The advertiser's average cost spent on each impression or click is reduced. In a bear market (i.e., the underlying price decreases), the advertiser will use the ad options less (and sometimes not at all) and the maximum cost is just the option price. It is worth noting that here we consider the ad options are in the money at time 0 (i.e., the strike price is less than the current underlying price). In Table 5.7, there are 4 ad slots that exhibit somewhat bear markets. However, these 4 ad slots do not receive enough bids in the test set and the actual winning payment CPMs

Table 5.7: Empirical example of an advertiser's delivery of an ad slot from the SSP dataset in RTB (Note: CTR is 0.03 and the non-integer numbers are displayed at 4 digits after the decimal point while in computing we consider 25-digit scale).

Day	0	1	2	3	4	5	6	7	Total
Date	07/02/2013	08/02/2013	09/02/2013	10/02/2013	11/02/2013	12/02/2013	13/02/2013	14/02/2013	
Average payment CPM	0.7427	0.9585	0.9770	0.9666	0.8754	0.8513	0.8294	0.9903	
No. of total impressions generated		8298	8277	8190	7971	8097	8201	3812	52846
Budget		5.0000	5.0000	5.0000	5.0000	5.0000	5.0000	5.0000	35.0000
No. of impressions received		5210	5113	5166	5711	5867	6028	3812	36907
No. of clicks received		156	153	154	171	176	180	114	1104
Used budget		5.0000	5.0000	5.0000	4.9830	5.0000	4.9926	3.7748	33.7504

Table 5.8: Empirical example of an advertiser's delivery of buying ad options for an advertisement slot in the SSP dataset (Note: CTR is 0.03 and the non-integer numbers are displayed at 4 digits after the decimal point while in computing we consider 25-digit scale).

Day	Date	Average payment CPM	No. of total impressions generated	Budget	Remaining budget	No. of options	Expiration date	Option price	Strike price CPC	No. of options exercised	No. of impressions received	No. of clicks received	Used budget
0	07/02/2013	0.7427				201	08/02/2013	0.0025	0.0223				0.4982
		0.7427				201	09/02/2013	0.0025	0.0223				0.4988
		0.7427				201	10/02/2013	0.0025	0.0223				0.5117
		0.7427				201	11/02/2013	0.0026	0.0223				0.5192
		0.7427				200	12/02/2013	0.0026	0.0223				0.5271
		0.7427				200	13/02/2013	0.0027	0.0223				0.5379
		0.7427				199	14/02/2013	0.0027	0.0223				0.5427
1	08/02/2013	0.9585	8298	5.0000	4.5018	201	08/02/2013		0.0223	201	6816	204	4.5013
2	09/02/2013	0.9770	8277	5.0000	4.5012	201	09/02/2013		0.0223	201	6770	203	4.5011
3	10/02/2013	0.9666	8190	5.0000	4.4883	201	10/02/2013		0.0223	201	6742	202	4.4878
4	11/02/2013	0.8754	7971	5.0000	4.4808	201	11/02/2013		0.0223	199	6836	205	4.4298
5	12/02/2013	0.8513	8097	5.0000	4.4729	200	12/02/2013		0.0223	200	6776	203	4.4727
6	13/02/2013	0.8294	8201	5.0000	4.4621	200	13/02/2013		0.0223	197	6792	204	4.4616
7	14/02/2013	0.9903	3812	5.0000	4.4573	199	14/02/2013		0.0223	114	3812	114	2.5463
Total			52846	35.0000							44544	1335	33.0362

Table 5.9: Overview of the improvement in delivery performance by using ad options for all ad slots in the SSP dataset.

	Bull market	Bear market
Change on used budget (%)	-8.7878%	–
Change on delivery of impressions (%)	6.1781%	–

Table 5.10: Overview of the improvement in delivery performance by using ad options for keywords in the Google AdWords dataset.

Market	Group	Change in used budget (%)		Change in delivery of impressions (%)	
		Bull market	Bear market	Bull market	Bear market
US	1	0.3447%	2.3438%	9.3050%	-0.1122%
	2	1.7748%	3.9687%	2.3153%	-2.6285%
	3	0.5372%	4.8567%	44.3735%	-0.0940%
	4	5.6288%	29.3626%	1.6433%	-1.0993%
UK	1	21.4285%	6.8940%	3.0717%	-0.2523%
	2	5.4426%	0.0000%	0.4419%	0.0000%
	3	10.9285%	3.8474%	28.7706%	-2.1066%
	4	6.7155%	0.1552%	16.6955%	-2.1550%

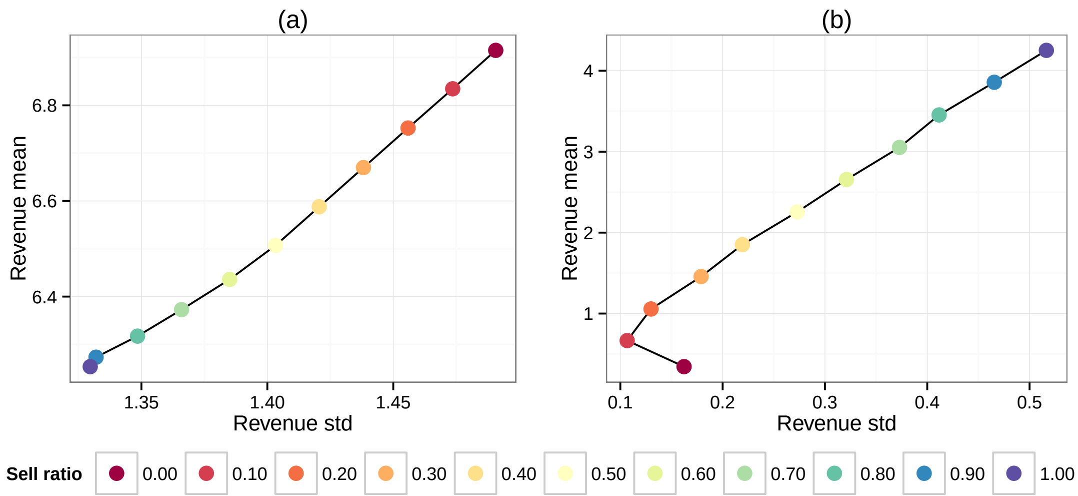

Figure 5.11: Empirical examples of the publisher's revenue: (a) from an ad slot in the bull market; and (b) from an ad slot in the bear market. The sell ratio represents the percentage of future daily impressions that are sold in advance via display ad options. Note that here the ad slot in the bear market does not receive enough bids in the test set, so we randomly simulate some underlying prices for the bear market.

are just around its floor reserve level (i.e., the CPM is £0.01 so the per impression price is £0.00001). Since these prices will seriously bias the results, we do not take them into account in the situation of a bear market.

5.4.4 Revenue Analysis for Publisher and Search Engine

The revenue for a publisher (or search engine) is examined in this section. We consider the revenue effects when a certain amount of future impressions or clicks can be sold in advance. Figure 5.11 presents two empirical examples of ad slots from the SSP dataset: one exhibits the bull market while the other shows the bear market. The sell ratio in the figure represents the percentage of future impressions that are sold in advance via display ad options; therefore, when the sell ratio equals zero, the publisher auctions off all of the future impressions in RTB. Figure 5.11(a) suggests that the publisher should sell less future impressions in advance if the future market is bull. This is because the ad options will be exercised by advertisers in the future and the obtained revenues from the fixed payment are less than these impressions' market values. Of course, the publisher can choose a certain percentage of future impressions to sell according to his level of risk tolerance or to meet other business objectives. For example, the publisher may be willing to sacrifice some revenues in order to increase the advertisers' engagement in the long run. Conversely, in a bear market, as shown in Figure 5.11(b), the publisher is advised to sell more future impressions in advance because there is more upfront income if more display ad options are sold, and in the future advertisers will not exercise the sold options. Therefore, the increased revenue comes from the option price.

Based on the above analysis, the revenue effects across all ad slots and keywords in our datasets are examined. In the experiments, the display ad options in a bull market are priced in the money while in a bear market they are priced out of the money. The sell ratio is set at 0.20 in a bull market while it is set at 0.80 in a bear market. The overall results are presented in Tables 5.11-5.12, which further confirm our analysis in the empirical examples. The average revenue is reduced in the bull market as well as the standard deviation (i.e., one kind of revenue risk). However, as described, the publisher (or search engine) may be willing to sacrifice some revenue to establish a long-term relationship with advertisers. In a bear market, the average revenue increases significantly. This is because fewer display ad options are exercised. Many premium advertisers join RTB so that the market equilibrium is almost as same as that in an environment with only auctions. Finally, the publisher (or search engine) earns the upfront payment without providing guaranteed deliveries.

Table 5.11: Overview of the improvement in revenue by selling display ad options for ad slots in the SSP dataset.

	Bull market	Bear market
Change on mean (%)	-7.1283%	726.3085%
Change on standard deviation (%)	-2.7041%	196.0547%

Table 5.12: Overview of the improvement in revenue by selling display ad options for ad slots in the Google AdWords dataset.

Market	Group	Change in mean (%)		Change in standard deviation (%)	
		Bull market	Bear market	Bull market	Bear market
US	1	-20.5880%	22.3898%	-0.6507%	9.3291%
	2	-23.2971%	17.1898%	-17.6508%	9.4175%
	3	-32.8388%	69.9113%	-21.9468%	-2.1065%
	4	-24.4710%	8.9650%	-10.6024%	95.4868%
UK	1	-8.5463%	15.4155%	4.5617%	10.4116%
	2	-20.0632%	4.3816%	-16.0239%	6.8847%
	3	-16.9050%	30.7737%	-11.4811%	-19.4625%
	4	-21.8142%	7.6342%	-19.4368%	0.3877%

5.5 Summary

In this chapter, we described a new ad option tailored to the display advertising environment. We examined several lattice methods for an ad option with the GBM underlying, and proposed a new lattice method to price it if the underlying price does not follow the GBM model. Our lattice method is based on the SV model which can capture the changing volatility and mean-reverting fact of price movement. Our developments were examined and validated by experiments using real advertising data. For future research, we are interested in developing a lattice method that can be used to price an ad option with the multivariate non-GBM underlying.

5.6 Chapter Appendix

5.6.1 Proof of Equivalence of the Option Price under the One-Step Binomial Lattice

We derive the option pricing formula from the perspective of a publisher who wants to hedge the revenue risk incurred from price changes, and prove that under the one-step binomial lattice the derived option price is equal to the one that is calculated from the perspective of a risk-neutral advertiser. The derivation here follows the settings proposed by Wang and Chen [2012] and considers the case where an ad option allows its buyer to pay a fixed CPC for display impressions. Therefore, the strike price of the option is the fixed CPC and the underlying price is the uncertain winning payment CPM in online auctions. The total number of future impressions to sell is assumed to be deterministic, denoted by S^M. If the CPM in time 1 goes up, the publisher's revenue can be expressed as

$$R_1^{\{u\}} = \begin{cases} (1-\alpha)S^M/1000M_1^{\{u\}} + \alpha S^M H F^M, & \text{if } M_1^{\{u\}} \geq F^C, \\ (1-\alpha)S^M/1000M_1^{\{u\}} + \alpha S^M/1000M_1^{\{u\}}, & \text{if } M_1^{\{u\}} < F^C, \end{cases} \tag{5.22}$$

where α is the percentage of estimated total impressions to sell via ad options. Eq. (5.22) shows that the publisher's revenue is a combination of guaranteed and non-guaranteed impressions. Eq. (5.22) can be rewritten as $R_1^{\{u\}} = S^M/1000M_1^{\{u\}} - \alpha S^M H \Phi_1^{\{u\}}$, where $\Phi_1^{\{u\}}$ is the option payoff function, defined by $\max\{M_1^{\{u\}}/(1000H) - F^C, 0\}$, and the superscript notation $\{u\}$ represents the upward movement. Similarly, if CPM in time 1 goes down, the publisher's revenue is $R_1^{\{d\}} = S^M/1000M_1^{\{d\}} - \alpha S^M H \Phi_1^{\{d\}}$, where $\Phi_1^{\{d\}} = \max\{M_1^{\{d\}}/(1000H) - F^C, 0\}$.

Since the publisher uses α to control the revenue in bull and bear markets, there exists a value α^* such that $R_1^{\{u\}}(\alpha^*) = R_1^{\{d\}}(\alpha^*)$, then $\alpha^* = (M_1^{\{u\}} - M_1^{\{d\}})/(\Phi_1^{\{u\}} - \Phi_1^{\{d\}})$. As described, the publisher's least requirement on the valuation is that his expected future revenue (including the upfront income in terms of option prices) should be equal to the current revenue level from auctions alone, so the following equation holds:

$$R_0 = \frac{\alpha^* S^M}{1000}\pi_0 + \widetilde{r}^{-1} R_1^{\{u\}}(\alpha^*) = \frac{\alpha^* S^M}{1000}\pi_0 + \widetilde{r}^{-1} R_1^{\{d\}}(\alpha^*).$$

The option price π_0 can then be calculated by

$$\begin{aligned} \pi_0 &= \widetilde{r}^{-1}\left(\frac{\widetilde{r}M_0 - M_1^{\{d\}}}{M_1^{\{u\}} - M_1^{\{d\}}}\Phi_1^{\{u\}} + \frac{M_1^{\{u\}} - \widetilde{r}M_0}{M_1^{\{u\}} - M_1^{\{d\}}}\Phi_1^{\{d\}}\right) \\ &= \widetilde{r}^{-1}\left(\frac{\widetilde{r} - d}{u - d}\Phi_1^{\{u\}} + \frac{u - \widetilde{r}}{u - d}\Phi_1^{\{d\}}\right), \end{aligned} \tag{5.23}$$

where $u = M_1^{\{u\}}/M_0, d = M_1^{\{u\}}/M_0$. Up to this point, we have proved that the calculated option price π_0 is no-arbitrage and hedges the revenue for the publisher.

5.6.2 Convergence of the Binomial Lattice Option Pricing Model to the BSM Model

Hsia [1983] discussed a general proof for the convergence of the binomial lattice option pricing model to the BSM model. His work can be used to derive a continuous-time closed-form option pricing formula for Eq. (5.3). To simplify the discussion, we here adopt the settings proposed by Cox, Ross, and Rubinstein [1979] and derive a BSM-like option pricing formula for Eq. (5.3).

We consider the case where an ad option allows its buyer to pay a fixed CPC for display impressions. Therefore, the strike price of the option is the fixed CPC and the underlying price is the uncertain winning payment CPM in online auctions. Let $\Delta t = T/n, u = e^{\sigma\sqrt{\Delta t}} > 1, d = 1/u < 1$, where σ is the volatility of CPM. Let r be the constant continuous-time risk-less interest rate and let $M(t)$ be the continuous-time CPM at time t. Under the risk-neutral probability measure $\mathbb{Q}$, the GBM underlying can be expressed as

$$dM(t) = rM(t)dt + \sigma M(t)dW^{\mathbb{Q}}(t), \tag{5.24}$$

where $W^{\mathbb{Q}}(t)$ is a standard Brownian motion under $\mathbb{Q}$.

As described, for $j \geq j^*, j = 1, 2, \ldots, n, j^* = 1, 2, \ldots, n$, the advertiser will exercise the option to buy the targeted impressions, so the following inequality holds:

$$\frac{M_0}{1000H} u^{j^*-1} d^{n-j^*+1} < F^C \leq \frac{M_0}{1000H} u^{j^*} d^{n-j^*}, \tag{5.25}$$

then

$$\left(\frac{u}{d}\right)^{j^*-1} \leq \frac{1000HF^C}{M_0 d^n} \leq \left(\frac{u}{d}\right)^{j^*}, \tag{5.26}$$

and, taking logarithms and dividing by $\ln(u/d)$ and subtracting nq from each side gives

$$\frac{j^* - 1 - nq}{\sqrt{n}} \leq \frac{\ln(1000HF^C/M_0) - \ln(u^{nq}d^{n(1-q)})}{\sqrt{n}(\ln(u) - \ln(d))} \leq \frac{j^* - nq}{\sqrt{n}}. \tag{5.27}$$

Since

$$\lim_{n\to\infty} \frac{j^* - nq}{\sqrt{n}} - \frac{j^* - 1 - nq}{\sqrt{n}} = 0,$$

then

$$\frac{\ln(1000HF^C/M_0) - \ln(u^{nq}d^{n(1-q)})}{\sqrt{n}(\ln(u) - \ln(d))} \approx \frac{j^* - nq}{\sqrt{n}}. \tag{5.28}$$

Therefore, we can obtain

$$
\begin{aligned}
\psi(j^*, n, q) &= \mathbb{P}(j \geq j^*, j \in \{1, \dots, n\}) \\
&= \mathbb{P}\left(\frac{j - nq}{\sqrt{nq(1-q)}} \geq \frac{j^* - nq}{\sqrt{nq(1-q)}}\right) = 1 - \mathcal{N}\left(\frac{j^* - nq}{\sqrt{nq(1-q)}}\right) \\
&= \mathcal{N}\left(\frac{nq - j^*}{\sqrt{nq(1-q)}}\right) \\
&= \mathcal{N}\left(\frac{\ln(M_0/(1000HF^c)) + \ln(u^{nq}d^{n(1-q)})}{\sqrt{n}(\ln(u) - \ln(d))\sqrt{q(1-q)}}\right),
\end{aligned} \tag{5.29}
$$

where $\mathcal{N}(\cdot)$ is the cumulative distribution function of a standard normal distribution.

If $n \to \infty$ (or $\Delta t \to 0$), the following convergence results can be obtained:

$$
\begin{aligned}
\sqrt{n}\ln(u) &= \lim_{n \to \infty} \sqrt{n} \ln e^{\sigma\sqrt{\Delta t}} = \sigma\sqrt{T}, \\
\sqrt{n}\ln(d) &= \lim_{n \to \infty} \sqrt{n} \ln e^{-\sigma\sqrt{\Delta t}} = -\sigma\sqrt{T}, \\
q = \frac{\widetilde{r} - d}{u - d} &= \lim_{\Delta t \to 0} \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} \\
&= \lim_{\Delta t \to 0} \frac{\sigma + (r - \frac{1}{2}\sigma^2)\sqrt{\Delta t} + o(\Delta^{3/2})}{2\sigma + o(\Delta t)} = \frac{1}{2}, \\
\lim_{\Delta t \to 0} \frac{2q - 1}{\sqrt{\Delta t}} &= \lim_{\Delta t \to 0} \frac{2}{\sqrt{\Delta t}} \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} = \frac{r}{\sigma} - \frac{\sigma}{2},
\end{aligned}
$$

and then

$$
\begin{aligned}
\ln(u^{nq}d^{n(1-q)}) &= nq\ln(u) + n(1-q)\ln(d) \\
&= nq\sigma\sqrt{\Delta t} - n(1-q)\sigma\sqrt{\Delta t} = 2nq\sigma\sqrt{\Delta t} - n\sigma\sqrt{\Delta t} \\
&\approx 2n\sigma\sqrt{\Delta t} \lim_{\Delta t \to 0} \frac{\sigma + (r - \frac{1}{2}\sigma^2)\sqrt{\Delta t} + o(\Delta t)}{2\sigma + o(\Delta t)} - \lim_{\Delta \to 0} n\sigma\sqrt{\Delta t} \\
&\approx \lim_{\Delta t \to 0} n\sigma\sqrt{\Delta t} + (r - \frac{1}{2}\sigma^2)T - \lim_{\Delta t \to 0} n\sigma\sqrt{\Delta t} = (r - \frac{1}{2}\sigma^2)T.
\end{aligned}
$$

Therefore, we have

$$
\psi(j^*, n, q) = \mathcal{N}\left(\frac{\ln(M_0/(1000HF^C)) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right). \tag{5.30}
$$

Since $\widetilde{q} = qu/\widetilde{r}$, we then have

$$\begin{aligned}
\ln(u^{n\widetilde{q}}d^{n(1-\widetilde{q})}) - \ln(u^{nq}d^{n(1-q)}) &= \ln(u^{n(\widetilde{q}-q)}d^{n(q-\widetilde{q})}) \\
&= n(\widetilde{q}-q)\ln(u) + n(q-\widetilde{q})\ln(d) \\
&= n(\widetilde{q}-q)\sigma\sqrt{\Delta t} - n(q-\widetilde{q})\sigma\sqrt{\Delta t} = 2n(\widetilde{q}-q)\sigma\sqrt{\Delta t} \\
&= 2nq\frac{u-\widetilde{r}}{\widetilde{r}}\sigma\sqrt{\Delta t} = 2q\sigma\frac{T}{\sqrt{\Delta t}}\frac{e^{\sigma\sqrt{\Delta t}} - e^{r\Delta t}}{e^{r\Delta t}} \\
&\approx \lim_{\Delta t\to 0} 2q\sigma T\frac{\sigma - r\Delta t}{1+r\Delta t} = \sigma^2 T.
\end{aligned}$$

Similarly,

$$\widetilde{q} = \lim_{\Delta t\to 0}\frac{1+\sigma\sqrt{\Delta}}{1+\widetilde{\Delta} t}\frac{\sigma + (r-\frac{1}{2}\sigma^2)\sqrt{\Delta t})}{2\sigma} = \frac{1}{2},$$

we then obtain

$$\begin{aligned}
\psi(j^*, n, \widetilde{q}) &= \mathcal{N}\left(\frac{\ln(M_0/(1000HF^C)) + \ln(u^{n\widetilde{q}}d^{n(1-\widetilde{q})})}{\sqrt{n}(\ln(u)-\ln(d))\sqrt{\widetilde{q}(1-\widetilde{q})}}\right) \\
&= \mathcal{N}\left(\frac{\ln(M_0/(1000HF^C)) + \ln(u^{nq}d^{n(1-q)})}{\sqrt{n}(\ln(u)-\ln(d))\sqrt{\widetilde{q}(1-\widetilde{q})}}\right. \\
&\qquad\left. + \frac{\ln(u^{n\widetilde{q}}d^{n(1-\widetilde{q})}) - \ln(u^{nq}d^{n(1-q)})}{\sqrt{n}(\ln(u)-\ln(d))\sqrt{\widetilde{q}(1-\widetilde{q})}}\right) \\
&= \mathcal{N}\left(\frac{\ln(M_0/(1000HF^C)) + (r+\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}\right). \qquad (5.31)
\end{aligned}$$

Finally, the continuous-time option pricing formula can be obtained as follows

$$\pi_0 = \frac{M_0}{1000H}\mathcal{N}(\varsigma_1) - F^C e^{-rT}\mathcal{N}(\varsigma_2), \qquad (5.32)$$

$$\varsigma_1 = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{M_0}{1000HF^C}\right) + (r+\frac{1}{2}\sigma^2)T\right), \qquad (5.33)$$

$$\varsigma_2 = \frac{1}{\sigma\sqrt{T}}\left(\ln\left(\frac{M_0}{1000HF^C}\right) + (r-\frac{1}{2}\sigma^2)T\right). \qquad (5.34)$$

Hence, if the GBM assumption is valid, one can use the closed-form solution to calculate the option price. However, as described, lattice methods provide an alternative way to calculate the option price and, in general, is simpler in terms of implementation. We here use the closed-form pricing formula to examine the convergence error of various lattice methods which are based on the GBM underlying model.

5.6.3 Trinomial Lattice Methods for Pricing Display Ad Options with the GBM Underlying

The calculation of the option price for a display ad option with the GBM underlying model over a trinomial lattice is described in Algorithm 5.2. We here consider the case where an ad option allows its buyer to pay the fixed CPC for display impressions. Therefore, the strike price of the option is the fixed CPC and the underlying price is the uncertain winning payment CPM from online auctions. Algorithm 5.2 can be easily extended to calculate the option price over a binomial lattice (see Table 2.2).

Algorithm 5.2 Trinomial lattice method for pricing an ad option with the GBM underlying. The strike price of the option is the fixed CPC and the underlying price is the uncertain winning payment CPM from online auctions.

---

function $\textsc{OptionPricingTrinomialLattice}(M_0, \sigma, H, T, n, r, F^C)$
    # Initialization:
    $\Delta t \leftarrow T/n; \widetilde{r} \leftarrow e^{r\Delta t}$;
    $u, m, d, q_1, q_2, q_3 \leftarrow$ Boyle-TRIN (or KR-TRIN or Tian-TRIN) in Table 2.2;
    # Build a (recombining) trinomial lattice for CPM
    $\mathbf{\Sigma}_{(n+1)\times(n+1)} \leftarrow \mathbf{0}_{(n+1)\times(n+1)}; \mathbf{\Sigma}_{(1,1)} \leftarrow M_0$;
    for $j \leftarrow 2$ to $n+1$ **do**
        $\mathbf{\Sigma}_{(1,j)} \leftarrow u \times \Sigma_{(1,j-1)}; \mathbf{\Sigma}_{(2,j)} \leftarrow m \times \Sigma_{(1,j-1)}; \mathbf{\Sigma}_{(3,j)} \leftarrow d \times \Sigma_{(1,j-1)}$;
        if $2(j-1)+1 > 3$ **then**
            for $k \leftarrow 4$ to $2(j-1)+1$ **do**
                $\mathbf{\Sigma}_{(k,j)} \leftarrow d \times \mathbf{\Sigma}_{(k-2,j-1)}$;
            end for
        end if
    end for
    # Calculate the terminal payoffs and the option value backward recursively
    $\widetilde{\mathbf{\Sigma}}_{(n+1)\times(n+1)} \leftarrow \mathbf{0}_{(n+1)\times(n+1)}; \widetilde{\mathbf{\Sigma}}_{(:,n+1)} \leftarrow \max\{\mathbf{\Sigma}_{(:,n+1)}/(1000H) - F^C, 0\}$;
    for $j \leftarrow n$ to 1 **do**
        for $k \leftarrow 1$ to $2(j-1)+1$ **do**
            if $k = 1$ **then**
                $\widetilde{\mathbf{\Sigma}}_{(k,j)} \leftarrow \widetilde{r}^{-1}(q_1\widetilde{\mathbf{\Sigma}}_{(k,j+1)} + q_2\widetilde{\mathbf{\Sigma}}_{(k,j+1)} + q_3\widetilde{\mathbf{\Sigma}}_{(k,j+1)})$;
            else if $k \geq 2$ **then**
                $\widetilde{\mathbf{\Sigma}}_{(k,j)} \leftarrow \widetilde{r}^{-1}(q_1\widetilde{\mathbf{\Sigma}}_{(k-1,j+1)} + q_2\widetilde{\mathbf{\Sigma}}_{(k,j+1)} + q_3\widetilde{\mathbf{\Sigma}}_{(k+1,j+1)})$;
            end if
        end for
    end for
    return $\pi_0 \leftarrow \widetilde{\mathbf{\Sigma}}_{(1,1)}$
end function

5.6.4 Binomial Diffusion Approximation

The discussed discrete-time binomial process can be used to approximate a general stochastic diffusion process. Let $M(t)$ be the continuous-time process of CPM at time

$t \in [0, T]$ and let $X(t) = \ln(M(t))$. A general stochastic diffusion process can be expressed as follows:

$$dX(t) = \alpha(t, X(t))dt + \beta(t, X(t))dW(t), \tag{5.35}$$

where $\alpha(t, X(t))$ and $\beta(t, X(t))$ are continuous-time drift and diffusion functions, and $W(t)$ is a standard Brownian motion under the real-world probability measure $\mathbb{P}$. Consider a process $\widetilde{X}(t)$, which is a step function with initial value $X(0)$ and transition movements only at times Δt, $2\Delta t$, ..., $n\Delta t$. For simplicity, we denote the time steps by $t_0 = 0, t_1 = \Delta t, \ldots, t_n = n\Delta t = T$. Then, Eq. (5.35) can be weakly converged by a (discrete-time) binomial process $\widetilde{X}(t)$ if the following conditions hold [Nelson and Ramaswamy, 1990]:

$$\lim_{\Delta t \to 0} \sup_{t_k \le t < t_k + \Delta t, 0 \le k < n} \mid \alpha(t_k, \widetilde{X}(t_k)) - \alpha(t, X(t)) \mid \to 0, \tag{5.36}$$

$$\lim_{\Delta t \to 0} \sup_{t_k \le t < t_k + \Delta t, 0 \le k < n} \mid \beta(t_k, \widetilde{X}(t_k)) - \beta(t, X(t)) \mid \to 0, \tag{5.37}$$

$$\lim_{\Delta t \to 0} \sup_{t_k \le t < t_k + \Delta t, 0 \le k < n} \mid \widetilde{X}^u(t_k + \Delta t) - X(t) \mid \to 0, \tag{5.38}$$

$$\lim_{\Delta t \to 0} \sup_{t_k \le t < t_k + \Delta t, 0 \le k < n} \mid \widetilde{X}^d(t_k + \Delta t) - X(t) \mid \to 0, \tag{5.39}$$

where $\widetilde{X}^u(t_k + \Delta t)$ and $\widetilde{X}^d(t_k + \Delta t)$ are the successors of $X(t)$ and at each time step the process can make one of two possible moves: up to a value $\widetilde{X}^u(t_k + \Delta t)$ or down to a value $\widetilde{X}^d(t_k + \Delta t)$.

5.6.5 Mathematical Results of the SV Model

The mathematical results of the SV model defined in Eqs. (5.7)-(5.8) are provided in this appendix, including the arbitrage-free condition, the estimation of volatility in the constructed binomial lattice, and the estimation of the model parameters κ, θ, δ.

5.6.5.1 Risk-Neutral Probability Measure for Eq. (5.7)

By applying the Itô Lemma to Eq. (5.7), we obtain

$$M(t) = M(0) \exp\left\{ \int_0^t (\mu - \frac{1}{2}\sigma^2(t))ds + \int_0^t \sigma(s)dW(s) \right\}. \tag{5.40}$$

Consider a discount process $D(t) = e^{-rt}$, where r is a constant risk-less interest rate. Then $dD(t) = -rD(t)dt$. Therefore, the discounted CPM process is

$$D(t)M(t) = M(0)\exp\left\{ \int_0^t (\mu - r - \frac{1}{2}\sigma^2(t))ds + \int_0^t \sigma(s)dW(s) \right\}, \tag{5.41}$$

and its differential is

$$\begin{aligned}
d\Big(D(t)M(t)\Big) &= M(t)dD(t) + D(t)dM(t) \\
&= -rD(t)M(t)dt + D(t)(\mu M(t)dt + \sigma(t)M(t)dW(t)) \\
&= \sigma(t)D(t)M(t)\left(\frac{\mu - r}{\sigma(t)}dt + dW(t)\right) \\
&= \sigma(t)D(t)M(t)dW^{\mathbb{Q}}(t),
\end{aligned} \tag{5.42}$$

where $W^{\mathbb{Q}}(t) = W(t) + \int_0^t \frac{\mu-r}{\sigma(s)}ds$. According to the Girsanov Theorem [Wilmott, 2006], if choosing the process $\tau(t) = \frac{\mu-r}{\sigma(t)}$, then $W^{\mathbb{Q}}(t)$ is a standard Brownian motion under a new probability measure $\mathbb{Q}$. This $\mathbb{Q}$ is risk-neutral because it renders $D(t)M(t)$ into a martingale. Therefore, the risk-neutral formulation of Eq. (5.7) is then

$$dM(t) = rM(t)dt + \sigma(t)M(t)dW^{\mathbb{Q}}(t). \tag{5.43}$$

Therefore, $dX(t) = (r - \sigma^2(t)/2)dt + \sigma(t)dW^{\mathbb{Q}}(t)$.

5.6.5.2 Estimation of $\sigma(t_k + \Delta t)$ for the Censored Binomial Lattice

Eq. (5.8) is the Cox-Ingersoll-Ross (CIR) model [Cox et al., 1985]. Several CIR model forecasting results can be used to estimate $\sigma(t_k + \Delta t)$. Since

$$\begin{aligned}
d(e^{\kappa t}\sigma(t)) &= \kappa e^{\kappa t}\sigma(t)dt + e^{\kappa}d\sigma(t) \\
&= \kappa e^{\kappa t}\sigma(t)dt + e^{\kappa}\Big(\kappa(\theta - \sigma(t))dt + \delta\sqrt{\sigma(t)}dZ(t)\Big) \\
&= e^{\kappa t}\kappa\theta dt + e^{\kappa t}\delta\sqrt{\sigma(t)}dZ(t).
\end{aligned} \tag{5.44}$$

Then, we have

$$\sigma(t) = \sigma(0)e^{-\kappa t} + \theta(1 - e^{-\kappa t}) + \delta\int_0^t e^{-\kappa(t-s)}\sqrt{\sigma(s)}dZ(s). \tag{5.45}$$

Recall that the expectation of an Itô integral is zero, we obtain

$$\mathbb{E}[\sigma(t) \mid \mathcal{F}(0)] = \sigma(0)e^{-\kappa t} + \theta(1 - e^{-\kappa t}), \tag{5.46}$$

where $\mathcal{F}(0)$ represents the information up to time 0. Therefore, the conditional stochastic volatility $\sigma(t_k + \Delta t)$ can be obtained by the following formula

$$\sigma(t_k + \Delta t) = \sigma(t_0)e^{-\kappa(t_k+\Delta t)} + \theta(1 - e^{-\kappa(t_k+\Delta t)}).$$

5.6.5.3 Estimation of Parameters κ, θ, δ

Several statistical methods can be used to estimate the values of parameters κ, θ, δ. The simplest method is the ordinary least squares (OLS) method [Kladivko, 2007]. The discreteness form of Eq. (5.8) is

$$\sigma(t_k) - \sigma(t_{k-1}) = \kappa(\theta - \sigma(t_k))\Delta t + \delta\sqrt{\sigma(t_k)}\epsilon(t_k), \tag{5.47}$$

where $\epsilon(t_k) \sim \mathbf{N}(0, \Delta t)$. The equation can be rewritten as follows

$$\frac{\sigma(t_k) - \sigma(t_{k-1})}{\sqrt{\sigma(t_k)}} = \frac{\kappa\theta\Delta t}{\sqrt{\sigma(t_k)}} - \kappa\sqrt{\sigma(t_k)}\Delta t + \delta\epsilon(t_k), \tag{5.48}$$

The sum of square errors $\sum_{k=1}^{\widetilde{n}-1}\Big(\delta\epsilon(t_k)\Big)^2$ can be minimized so that κ and θ can be obtained, where $\widetilde{n}$ is the size of training data. Then, we have

$$(\widehat{\kappa}, \widehat{\theta}) = \text{argmin}_{\kappa,\theta} \sum_{k=1}^{\widetilde{n}-1} \left(\frac{\sigma(t_k) - \sigma(t_{k-1})}{\sqrt{\sigma(t_k)}} - \frac{\kappa\theta\Delta t}{\sqrt{\sigma(t_k)}} + \kappa\sqrt{\sigma(t_k)}\Delta t \right)^2. \tag{5.49}$$

Therefore,

$$\widehat{\kappa} = \frac{(\widetilde{n}+1)^2 + \sum_{k=1}^{\widetilde{n}-1}\sigma(t_{k+1})\sum_{k=1}^{\widetilde{n}-1}\frac{1}{\sigma(t_k)} - \sum_{k=1}^{\widetilde{n}-1}\sigma(t_k)\sum_{k=1}^{\widetilde{n}-1}\frac{1}{\sigma(t_k)} - (n-1)\sum_{k=1}^{\widetilde{n}-1}\frac{\sigma(t_{k+1})}{\sigma(t_k)}}{\left((\widetilde{n}+1)^2 - \sum_{k=1}^{\widetilde{n}-1}\sigma(t_k)\sum_{k=1}^{\widetilde{n}-1}\frac{1}{\sigma(t_k)}\right)\Delta t},$$

$$\widehat{\theta} = \frac{(\widetilde{n}-1)\sum_{k=1}^{\widetilde{n}-1}\sigma(t_{k+1}) - \sum_{k=1}^{\widetilde{n}-1}\frac{\sigma(t_{k+1})}{\sigma(t_k)}\sum_{k=1}^{\widetilde{n}-1}\sigma(t_k)}{(\widetilde{n}+1)^2 + \sum_{k=1}^{\widetilde{n}-1}\sigma(t_{k+1})\sum_{k=1}^{\widetilde{n}-1}\frac{1}{\sigma(t_k)} - \sum_{k=1}^{\widetilde{n}-1}\sigma(t_k)\sum_{k=1}^{\widetilde{n}-1}\frac{1}{\sigma(t_k)} - (n-1)\sum_{k=1}^{\widetilde{n}-1}\frac{\sigma(t_{k+1})}{\sigma(t_k)}}.$$

Then, we obtain the estimation of $\widehat{\delta}$ by the formula

$$\begin{aligned}\widehat{\delta} &= \left(\left(\sum_{k=1}^{\widetilde{n}-1}\Delta t\right)^{-1} \times \sum_{k=1}^{\widetilde{n}-1}\left(\frac{\sigma(t_{k+1})-\sigma(t_k)}{\sqrt{\sigma(t_k)}} - \frac{\widehat{\kappa}\widehat{\theta}\Delta t}{\sqrt{\sigma(t_k)}} + \widehat{\kappa}\sqrt{\sigma(t_k)}\Delta t\right)^2\right)^{1/2} \\ &= \left(\frac{1}{(\widetilde{n}-1)\Delta t}\sum_{k=1}^{\widetilde{n}-1}\left(\frac{\sigma(t_{k+1})-\sigma(t_k)}{\sqrt{\sigma(t_k)}} - \frac{\widehat{\kappa}\widehat{\theta}\Delta t}{\sqrt{\sigma(t_k)}} + \widehat{\kappa}\sqrt{\sigma(t_k)}\Delta t\right)^2\right)^{1/2}. \qquad (5.50)\end{aligned}$$

Kladivko [2007] discussed the Maximum Likelihood (ML) method as an alternative way and compared the results to the OLS method for a specific dataset. In our experiments, we find that the calculated option prices of display ad options are less sensitive to OLS and ML methods; therefore, we adopt the OLS method as it is computationally simpler.

Chapter 6

Conclusion

In this chapter, we summarise the main points made in Chapters 3-5 and point out the future directions that can be carried out to extend the research of this thesis.

6.1 Concluding Remarks

The work presented in this thesis looked at the non-guaranteed delivery problem in online advertising. Three novel solutions were proposed by employing and extending the mathematical models from modern financial theories. In Chapter 3, we studied an optimal dynamic model for a publisher or SSP who engages in RTB to allocate and price the future display impressions into guaranteed contracts. The developed model mimics the advanced booking system in the airline industry, and connects RTB and PG algorithmically in order to maximise the publisher's expected revenue. In Chapter 4, we proposed a multi-keyword multi-click ad option for sponsored search and discussed the corresponding option pricing methods based on the assumption that the underlying winning payment prices of candidate keywords follow a GBM. This option allows an advertiser to target multiple keywords and exercise multiple times in the contract lifetime. Our theoretical and empirical analysis also showed that the search engine can have an increased expected revenue over time. In Chapter 5, we discussed another ad option for display advertising and investigated a lattice framework for option evaluation under a more general situation that the GBM assumption is not valid empirically. This option allows an advertiser to pay a fixed CPM or CPC for an impression or click that is same or different to its underlying measurement model from auctions. We used the SV model to describe the underlying price movement and constructed a censored binomial lattice to approximate the underlying SV model.

Table 6.1: Summary the developments in Chapters 3-5.

	Chapter 3	Chapter 4	Chapter 5
Development	Optimal model	Ad option	Ad option
Ad type/format	Display	Search	Display, search
Objective	Revenue maximisation	No arbitrage	No arbitrage
Expected revenue	Increase	Increase	Increase
Inventory allocation	$\surd$	$\times$	$\times$
Contract pricing	$\surd$	$\surd$	$\surd$
Exercise right	$\times$	$\surd$	$\surd$
Exercise time	Fixed	Flexible	Fixed
Exercise opportunities	Single	Multiple	Single
Underlying inventory	Single	Multiple	Single
Underlying assumption	Probabilistic, empirical	GBM	GBM, SV
Behaviour assumption	Risk-neutral	Risk-neutral	Risk-neutral
Modelling setting	Continuous-time	Continuous-time	Discrete-time

Table 6.1 provides an overview summary and comparison of the developments in Chapters 3-5. These developments have many similarities in solving the non-guaranteed delivery problem. First, they can all be implemented in conjunction with auction mechanisms. Second, advertisers are assumed to be risk-neutral [Wilmott, 2006, Krishna, 2009, Narahari, 2014] in the modelling. Third, these developments can increase the seller's expected revenue. They support the two most popular measurement models (i.e., the CPM and CPC models) in our discussions and can be extended easily to others like the cost-per-action (CPA) model.

In the meantime, these developments are uniquely different from each other. The optimal model discussed in Chapter 3 offers a complete automated system that deals with both optimal allocation and pricing of future inventories. Even though only the guaranteed contracts can be sold, the model is not limited to any presumed bids' distribution (i.e., works for both probabilistic and empirical bids' distributions). The ad option proposed in Chapter 4 is the first research work that discusses the option contract mechanism in the context of sponsored search. Advertisers are able to have greater flexibility and a more personalized delivery. As also described earlier (see Section 4.3.4), the proposed ad option is priced based on a GBM and the GBM assumption is reasonable because several observations have been made in the sponsored search market. However, this GBM assumption is not valid for display advertising. Therefore, Chapter 5 investigated the possibility of using the SV model as the underlying force to describe the winning price movement. The ad option discussed in Chapter 5 has a different contract structure but it is currently limited to the single underlying variable.

6.2 Future Work

Financial methods open up many new possibilities for developing online advertising sales models and markets. The studies discussed in this thesis represent only a small step in using the full potential of financial methods in online advertising. The following research directions are interesting to further explore in the future.

6.2.1 Optimal Stochastic Dynamic Models

Given the static supply and demand, the dynamic model discussed in Chapter 3 maximises the expected revenue of a publisher (or SSP) who wants to sell some of the estimated future display impressions in advance via guaranteed contracts and the remaining estimated supply will be auctioned off in RTB. This research can be further extended with considering the stochastic supply and demand. Simply, the arrival of advertisers and online users can be modelled by two independent stochastic processes. Gallego and van Ryzin [1994] provided some insights on this issue while they only used a Poisson process to represent the arrival of demand. In future research, we can extend the Poisson process to model the supply. In addition, as described in Chapter 3, the expected salvage value of display impressions from RTB is not zero, which is determined by the new levels of remaining future supply and demand. Therefore, the optimal pricing and allocation of display impressions will depend on balancing the filled or unfilled demand with the arriving speed of supply.

6.2.2 Stochastic Processes for Market Price

Stochastic processes that describe the movement of the spot market prices of ad inventories (i.e., the winning payment CPMs or CPCs from online auctions) are the one that have received surprisingly little attention in online advertising. Here the spot market price is similar but slightly different to the one that we see in financial markets because there are no posted winning payment CPMs or CPCs in any ad exchanges or search advertising platforms. However, a seller, such as publisher, SSP and search engine, have the information about the advertisers' bids and winning payment prices for some specific ad slots or keywords. Such information could be used to analyse the market's supply and demand levels.

In Chapters 4-5, we discussed the GBM and SV models for pricing ad options. These stochastic processes can be investigated independently as they can also be used to analyse the statistical properties of market price movements. As described earlier, the studied GBM and SV models still have their limitations. The former is restricted by a constant volatility while the latter is unable to capture the price jumps and spikes. Many other stochastic processes can be further explored in the future. For example, we can study the parametric mean-reverting jump diffusion process [Kou, 2002] to capture the long-term mean-reverting fact as well as the distribution of price jumps.

6.2.3 Game-Theoretical Models for Ad Option Pricing

Game theory can provide an alternative way to evaluate an option written on ad inventories. For an advertiser, two advertising strategies can be considered: (i) bidding in online auctions; (ii) buying and exercising ad options. Advertisers are assumed to be identical and risk-neutral. The interesting and difficult part of this research is to find the advertisers' equilibrium in a mixed strategy game. This equilibrium may not be a simple Nash because most of the current advertising auction mechanisms adopt the GSP auction model where advertisers follow a symmetric Nash (or locally Envy-free) equilibrium [Edelman et al., 2007, Varian, 2007]. However, once a certain type of equilibrium is found in this mixed strategy game, the option can be evaluated accordingly.

Two different objectives can be considered for the game-theoretical option pricing. First, the optimal models which maximise the expected revenue of a publisher or search engine. Second, the efficient models which maximise the social welfare (or surplus) of advertisers. In this research, VCG auctions can be explored. Under a VCG auction, it would be interesting to examine if the incentive compatibility and the individual rationality constraints are satisfied in the option pricing [Narahari et al., 2009, Nadarajah et al., 2012]. The former ensures that an advertiser reveals his true value and the latter ensures the advertiser's payoff to be non-negative. This is a broad research direction and many game-theoretical models can be investigated.

6.2.4 Market Design of Ad Derivatives

Designing an ad derivatives market would be another interesting topic for future research. Most of the current online advertising markets are one-sided auctions [McAfee and Vassilvitskii, 2012], in which an advertiser can only submit his bids to purchase the targeted ad inventories and is limited to real-time transactions. The ad derivatives market will be two-sided auctions, where an advertiser can not only submit bids to buy but also can receive bids to sell. The transactions are for the ad inventories that will be created in the future. This will give advertisers a market place to buy and sell their future ad inventories centrally in order to speculate profits or hedge risks.

One interesting thing for further discussion is whether the prices should be posted in the ad derivatives market or not. Sealed-bid auctions are mostly adopted in online advertising markets and the transaction information is not disclosed to advertisers. If there are posted prices in ad derivatives market, advertisers may use such prices to estimate the spot market prices inversely. They will then adjust their bidding strategies in real-time auctions, which can further influence the prices posted in the ad derivatives market because the values of ad derivatives are calculated based on the corresponding spot market prices. This interesting problem can be examined in two stages. First, we can study this as the implied price problem, similar to the implied volatility problem in option pricing [Wilmott, 2006]. Second, we can move one step further to the system level and discuss the new equilibrium between the spot and guaranteed markets.

Appendix A

Glossary of Technical Terms

The technical terms used throughout the thesis are briefly explained in this glossary. Many term definitions here are drawn from Wilmott [2006], Jansen [2011], IAB [2013] and Wikipedia directly.

Ad: the commercial portion of message content for which an advertiser has paid or will pay when an online user sees his content.

Ad exchange: a technology platform that facilitate the bidded buying and selling of online media advertising inventory from multiple ad networks. The approach is technology-driven as opposed to the historical approach of negotiating price on media inventory.

Ad slot (or ad space): the allocated real estate on a Web page of a site in which an ad can be placed. Each space on a site is uniquely identified; therefore, multiple ad slots can exist on a single page.

Advertiser: also called the marketer, the company who pays for the ad display or click.

Banner ad: an ad embedded on a Web page that is usually intended to drive traffic to a different Web page by linking to the advertiser's site.

Bear market: a market in which market prices are falling and investors, fearing losses, tend to sell. This can create a self-sustaining downward spiral.

Bull market: a market in which the market prices are generally rising.

Click: a click on an ad on a Web page, which takes a user to another site.

Click-through rate (CTR): the rate of clicked ads to total ads displayed.

Cost per click (CPC): also called pay-per-click (PPC), is an online advertising measurement model used to direct traffic to websites, where an advertiser pays a search engine (or a publisher) when his ad is clicked by an online user.

Cost per mille (CPM): also called pay-per-mille (PPM), is an online advertising measurement model used to direct traffic to websites, where an advertiser pays a publisher when his ad is displayed 1000 times to online users.

Demand-side platform (DSP): an automated bidding platform for advertisers to get good impressions at low cost, by participating in multiple auctions among various ad exchanges at the same time.

Impression: a display of an ad to a user on a Web page. Note that a single page view can have more than one impression if there is more than one ad slot on the page.

Interest rate: a percentage used to calculate the cost of borrowing money. In this thesis, we only consider the constant risk-less bank interest rate.

Keyword: a specific word or combination of words that an online searcher might type into a search field. Advertisers can purchase keywords to guarantee that their website information is displayed prominently.

Publisher: an individual or organization that prepares, issues, and disseminates content for public distribution. Simply, a publisher has the space for ads to be displayed.

Query: a series of terms entered by a searcher into a search engine, which initiates a search and results in a search engine result page (SERP) with organic and paid listings.

Real-time bidding (RTB): refers to the means by which ad inventory is bought and sold on a per-impression basis, via programmatic instantaneous auction, similar to financial markets.

Search engine: a program that indexes Web pages and then attempts to match then by relevancy to the users' search requests. Examples of search engines include Google, Bing, Baidu etc.

Search engine results page (SERP): a page that online users see after they have entered their query into the search box. The SERP has two types of result listings in response to the submitted query: organic results and paid results. Organic search results are the Web page listings that most closely match the user's search query based on relevance. Paid results are basically ads – the websites have paid to have their Web pages display for certain keywords, so these listings show up when someone runs a search query containing those keywords.

Supply-side platform (SSP): an automated platform for publishers to sell impressions at an optimal price, by creating multiple auctions for the same impression in different ad exchanges to reach more advertisers who are willing to bid.

User: an individual with access to the Internet and the WWW, and issues ad-hoc topics to express his information needs, such as Web search or surfing.

Web page: the traditional presentation of information online. Web sites are made up of Web pages, analogous to the pages in a book.

Appendix B

Related Publications

The following publications and submissions are related to this thesis:

- Bowei Chen and Jun Wang. A lattice framework for pricing display ad options with the stochastic volatility model. *Working Paper*, 2014. In arXiv: `http://arxiv.org/abs/1409.0697`
- Bowei Chen, Shuai Yuan, and Jun Wang. A dynamic pricing model for unifying programmatic guarantee and real-time bidding in display advertising. In *Proceedings of the 8th International Workshop on Data Mining for Online Advertising (ADKDD'14),* ***Best Paper Award***, pages 1–9, New York, NY, USA, 2014b. ACM
- Bowei Chen, Jun Wang, Ingemar Cox, and Mohan Kankanhalli. Multi-keyword multi-click ad options for sponsored search. *Working Paper*, 2014a. In arXiv: `http://arxiv.org/abs/1307.4980`
- Jun Wang and Bowei Chen. Selling futures online advertising slots via option contracts. In *Proceedings of the 21st International Conference on World Wide Web (WWW '12)*, pages 627–628, Lyon, France, 2012. ACM

There are also other publications that were completed during my PhD study; while relevant to the broader applications of financial methods into information technologies, they are not directly solving the non-guaranteed delivery problem in online advertising:

- Shuai Yuan, Jun Wang, Bowei Chen, Peter Mason, and Sam Seljan. An empirical study of reserve price optimisation in real-time bidding. In *Proceedings of* *the 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'14)*, pages 1897–1906, New York, NY, USA, 2014. ACM

- Weinan Zhang, Jun Wang, Bowei Chen, and Xiaoxue Zhao. To personalize or not: a risk management perspective. In *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys'13)*, pages 229–236, Hong Kong, China, 2013. ACM

- Jamie O'Brien, Chris Coleridge, and Bowei Chen. Enriching domain knowledge of the academic-industrial landscape of an engineering doctorate centre: a multiple-sector industrial R&D survey. In *Triple Helix International Conference 2013 (THA'13)*, London, UK, 2013

Bibliography

Miguel Anjos, Russell Cheng, and Christine Currie. Maximizing revenue in the airline industry under one-way pricing. *Journal of the Operational Research Society*, 55(5): 535–541, 2004.

Miguel Anjos, Russell Cheng, and Christine Currie. Optimal pricing policies for perishable products. *European Journal of Operational Research*, 166(1):246–254, 2005.

Moshe Babaioff, Jason Hartline, and Robert Kleinberg. Selling ad campaigns: online algorithms with cancellations. In *Proceedings of the 10th ACM Conference on Electronic Commerce (EC'09)*, pages 61–70, Stanford, CA, USA, 2009. ACM.

Louis Bachelier. Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, 3(17):21–86, 1900.

Kyle Bagwell. *The Economics Of Advertising*. Edward Elgar Publishing, 2001.

Santiago Balseiro, Jon Feldman, Vahab Mirrokni, and Muthu Muthukrishnan. Yield optimization of display advertising with ad exchange. In *Proceedings of the 12th ACM conference on Electronic commerce (EC'11)*, pages 27–28, 2011.

Anand Bhalgat, Tanmoy Chakraborty, and Sanjeev Khanna. Mechanism design for a risk averse seller. In *Proceedings of the 8th Workshop on Internet and Network Economics (WINE'12)*, pages 198–211, Liverpool, UK, 2012. Springer.

Vijay Bharadwaj, Wenjing Ma, Michael Schwarz, Jayavel Shanmugasundaram, Erik Vee, Jack Xie, and Jian Yang. Pricing guaranteed contracts in online display advertising. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management (CIKM'10)*, pages 399–408, Toronto, ON, Canada, 2010. ACM.

Vijay Bharadwaj, Peiji Chen, Wenjing Ma, Chandrashekhar Nagarajan, John Tomlin, Sergei Vassilvitskii, Erik Vee, and Jian Yang. Shale: an efficient algorithm for allocation of guaranteed display advertising. In *Proceedings of the 18th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'12)*, pages 1195–1203, Beijing, China, 2012. ACM.

Tomas Björk. *Arbitrage Theory in Continuous Time*. Oxford University Press, 3rd edition, 2009.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. ISSN 00223808.

Zvi Bodie, Robert Merton, and David Cleeton. *Financial Economics*. Pearson, 2009.

Peter Boer. *The Real Options Solution Finding Total Value in a High-Risk World*. John Wiley, 2002.

Tilman Börgers, Ingemar Cox, Martin Pesendorfer, and Vaclav Petricek. Equilibrium bids in sponsored search auctions: theory and evidence. *American Economic Journal: Microeconomics*, 5(4):163–187, 2013.

Phelim Boyle. Option valuation using a three-jump process. *International Options Journal*, 3:7–12, 1986.

Phelim Boyle. A lattice framework for option pricing with two state variables. *Journal of Financial and Quantitative Analysis*, 23(1):1–12, 1988.

Phelim Boyle, Jeremy Evnine, and Stephen Gibbs. Numerical evaluation of multivariate contingent claims. *The Review of Finanical Studies*, 2(2):241–250, 1989.

Rick Bruner. The decade in online advertising 1994-2004. *Double Click White Paper*, 2005.

Bowei Chen and Jun Wang. A lattice framework for pricing display ad options with the stochastic volatility model. *Working Paper*, 2014. In arXiv: `http://arxiv.org/abs/1409.0697`.

Bowei Chen, Jun Wang, Ingemar Cox, and Mohan Kankanhalli. Multi-keyword multi-click ad options for sponsored search. *Working Paper*, 2014a. In arXiv: `http://arxiv.org/abs/1307.4980`.

Bowei Chen, Shuai Yuan, and Jun Wang. A dynamic pricing model for unifying programmatic guarantee and real-time bidding in display advertising. In *Proceedings of the 8th International Workshop on Data Mining for Online Advertising (ADKDD'14), **Best Paper Award***, pages 1–9, New York, NY, USA, 2014b. ACM.

Bowei Chen, Shuai Yuan, and Jun Wang. A dynamic pricing model for unifying programmatic guarantee and real-time bidding in display advertising. In *Proceedings of the 8th International Workshop on Data Mining for Online Advertising (ADKDD'14)*, pages 1–9, New York, NY, USA, 2014c. ACM.

Edward Clarke. Multipart pricing of public goods. *Public Choice*, 11(1):17–33, 1971.

William Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74(368):829–836, 1979.

Les Clewlow and Chris Strickland. *Energy Derivatives Pricing and Risk Management*. Lacima, 2000.

Florin Constantin, Jon Feldman, Muthu Muthukrishnan, and Martin Pál. An online mechanism for ad slot reservations with cancellations. In *Proceedings of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'09)*, pages 1265–1274, New York Marriott Downtown, NY, USA, 2009. SIAM.

George Constantinides and Anastasios Malliaris. Options markets. In *Options Markets Vol I Equity Options Markets: Foundations and Pricing*. Edward Elgar Publishing, 2001.

John Cox and Stephen Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3:145–166, 1976.

John Cox, Stephen Ross, and Mark Rubinstein. Option pricing: a simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

John Cox, Jonathan Ingersoll, and Stephen Ross. A theory of the term structure of interest rates. *Economietirca*, 53:385–407, 1985.

Shijie Deng. Stochastic models of energy commodity prices and their applications : mean-reversion with jumps and spikes. In *Working Paper Series of Workable Energy Regulation (POWER)*, Berkeley, CA, USA, 2000. University of California Energy Institute. http://www.ucei.berkeley.edu/PDF/pwp073.pdf.

Shijie Deng and Shmuel Oren. Electricity derivatives and risk management. *Energy*, 31(6):940–953, 2006.

Gavin Dunaway. Paths to programmatic premium. *AdMonsters*, 11 2012. http://www.admonsters.com/blog/programmatic-premium-beyond-rtb.

Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second-price auction: selling billions of dollars worth of keywords. *American Economic Review*, 97(1):242–259, 2007.

eMarketer. Real-time bidding poised to make up quarter of all display spending. http://goo.gl/04Ykx, 2013.

David Evans. The online advertising industry: Economics, evolution, and privacy. *Journal of Economic Perspectives*, 23(3):37–60, 2009.

Jon Feldman, Nitish Korula, Vahab Mirrokni, Muthu Muthukrishnan, and Martin Pál. Online ad assignment with free disposal. In *Proceedings of the 5th International Workshop on Internet and Network Economics (WINE'09)*, pages 374–385, Rome, Italy, 2009. Springer.

Ionut Florescu and Frederi Viens. Stochastic volatility: option pricing using a multinomial recombining tree. *Applied Mathematical Finance*, 15(2):151–181, 2008.

Guillermo Gallego and Garrett van Ryzin. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*, 40(8):999–1020, 1994.

Arpita Ghosh, Preston McAfee, Kishore Papineni, and Sergei Vassilvitskii. Bidding for representative allocations for display advertising. In *Proceedings of the 5th International Workshop on Internet and Network Economics (WINE'09)*, pages 14–18, Rome, Italy, 2009. Springer.

Google. Google adwords traffic estimator. `http://goo.gl/TYdOQ`.

Google. The arrival of real-time bidding and what it means for media buyers. Technical report, White Paper, 2011.

Theodore Groves. Incentives in teams. *Econometrica*, 41(4):617–631, 1973.

Tero Haahtela. Recombining trinomial tree for real option valuation with changing volatility. In *SSRN: `http://dx.doi.org/10.2139/ssrn.1932411`*, 2010.

Jiawei Han, Micheline Kamber, and Jian Pei. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 3rd edition, 2011.

David Hobson. A survey of mathematical finance. *Royal Society Proceedings A*, 460: 3369–3401, 2004.

Chicheng Hsia. On binomial option pricing. *Journal of Financial Research*, 6(1): 41–46, 1983.

John Hull and Alan White. Pricing interest-rate-derivative securities. *The Review of Financial Studies*, 3:573–592, 1990.

IAB. Glossary of interactive advertising terms version 2.0. *Interactive Advertising Bureau (IAB) Report*, 2013.

Wolfgang Jank and Inbal Yahav. E-loyalty networks in online auctions. *Annals of Applied Statistics*, 4(1):151–178, 2010.

Jim Jansen. *Understanding Sponsored Search: Core Elements of Keyword Advertising*. Cambridge University Press, New York, NY, USA, 2011.

Carlos Jarque and Anil Bera. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255–259, 1980.

Herb Johnson. Options on the maximum or the minimum of several assets. *Journal of Financial and Quantitative Analysis*, 22(3):277–283, 1987.

Ana Justel, Daniel Pefia, and Ruben Zamar. A multivariate kolmogorov-smirnov test of goodness of fit. *Statistics & Probability Letters*, 35(3):251–259, 1997.

Bardia Kamrad and Peter Ritchken. Multinomial approximating models for options with k state variables. *Management Science*, 37(23):1640–1652, 1991.

Kamil Kladivko. Maximum likelihood estimation of the cox-ingersoll-ross process: the Matlab implementation. *Technical Computing Prague*, 2007.

Frank Knight. *Risk, Uncertainty and Profit*. Boston MA: Hart, Schaffner and Marx; Houghton Mifflin, 1921.

Steven Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8): 1086–1101, 2002.

Martin Krekel, Johan de Kock, Ralf Korn, and Tin-Kwai Man. An analysis of pricing methods for baskets options. *WILMOTT Magazine*, pages 82–89, 2006.

Vijay Krishna. *Auction Theory*. Academic Press, 2nd edition, 2009.

Sébastien Lahaie and David Pennock. Revenue analysis of a family of ranking rules for keyword auctions. In *Proceedings of the 8th ACM conference on Electronic commerce (EC'07)*, pages 50–60, San Diego, CA, USA, 2007. ACM.

Greta Ljung and George Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

Paolo Malighetti, Stefano Paleari, and Renato Redondi. Pricing strategies of low-cost airlines: the Ryanair case study. *Journal of Air Transport Management*, 15(4):195–203, 2009.

Greg Mankiw. *Principles of Economics*. South-Western College Pub, 4th edition, 2006.

Rahul Marathe and Sarah Ryan. On the validity of the geometric Brownian motion assumption. *Engineering Economist*, 50:159–192, 2005.

James Marshall. *Valuation of Multiple Exercise Options*. PhD thesis, University of Western Ontario, 2012.

James Marshall, Mark Reesor, and Matthew Cox. Simulation valuation of multiple exercise options. In *Proceedings of the 2011 Winter Simulation Conference (WSC'11)*, pages 3772–3783, Phoenix, AZ, USA, 2011. IEEE.

Preston McAfee and Sergei Vassilvitskii. An overview of practical exchange design. *Current Science*, 103(9):1056–1063, 2012.

Thomas Meinl and Benjamin Blau. Web service derivatives. In *Proceedings of the 18th International Conference on World Wide Web (WWW'09)*, pages 271–280, Madrid, Spain, 2009. ACM.

Robert Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

Alexander Mood, Franklin Graybill, and Duane Boes. *Introduction to the Theory of Statistics*. McGraw-Hill, 3rd edition, 1974.

Yongma Moon and Changhyun Kwon. Online advertisement service pricing and an option contract. *Electronic Commerce Research and Applications*, 10(1):38–48, 2010.

Muthu Muthukrishnan. Ad exchanges: research issues. In *Proceedings of the 5th International Workshop on Internet and Network Economics (WINE'09)*, pages 1–12, Rome, Italy, 2009. Springer.

Selvaprabu Nadarajah, Francois Margot, and Nicola Secomandi. Valuation of multiple exercise options with energy applications. *Carnegie Mellon University Working Paper*, 2012.

Yadati Narahari. *Game Theory and Mechanism Design*. World Scientific, 2014.

Yadati Narahari, Dinesh Garg, Ramasuri Narayanam, and Hastagiri Prakash. *Game Theoretic Problems in Network Economics and Mechanism Design Solutions*. Springer, 2009.

Daniel Nelson and Krishna Ramaswamy. Simple binomial processes as diffusion aapproximation in financial models. *The Review of Finanical Studies*, 1990.

Jamie O'Brien, Chris Coleridge, and Bowei Chen. Enriching domain knowledge of the academic-industrial landscape of an engineering doctorate centre: a multiple-sector industrial R&D survey. In *Triple Helix International Conference 2013 (THA'13)*, London, UK, 2013.

OpenX. Ad networks vs ad exchanges: how they stack up. *White Paper*, 2010.

OpenX. Programmatic + premium: current practices and future trends. *White Paper*, 2013.

Michael Ostrovsky and Michael Schwarz. Reserve prices in Internet advertising auctions: a field experiment. In *Proceedings of the 12th ACM Conference on Electronic Commerce (EC'11)*, pages 59–60, San Jose, CA, USA, 2011. ACM.

Peter Ouwehand and Graeme West. Pricing rainbow options. *WILMOTT Magazine*, pages 215–242, 2006.

Furcy Pin and Peter Key. Stochastic variability in sponsored search auctions: observations and models. In *Proceedings of the 12th ACM Conference on Electronic Commerce (EC'11)*, pages 61–70, San Jose, CA, USA, 2011. ACM.

James Primbsa, Muruhan Rathinamb, and Yuji Yamadac. Option pricing with a pentanomial lattice model that incorporates skewness and kurtosis. *Applied Mathematical Finance*, pages 1–17, 2007.

PubMatic. Understanding real-time bidding (RTB) from the publisher perspective. Technical report, White Paper, 2010.

Tao Qin, Wei Chen, and Tie-Yan Liu. Sponsored search auctions: recent advances and future directions. *ACM Transactions on Intelligent Systems and Technology*, 5(4), 2014.

Guillaume Roels and Kristin Fridgeirsdottir. Dynamic revenue management for online display advertising. *Journal of Revenue & Pricing Management*, 8(5):452–466, 2009.

Konstantin Salomatin, Tie-Yan Liu, and Yiming Yang. A unified optimization framework for auction and guaranteed delivery in online advertising. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM'12)*, pages 2005–2009, Maui, HI, USA, 2012. ACM.

Paul Samuelson. Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, pages 41–49, 1965a.

Paul Samuelson. Rational theory of warrant pricing. *Industrial Management Review*, 6:13–31, 1965b.

Frank Schmid. The stock market: beyond risk lies uncertainty. *The Regional Economist*, pages 5–9, 2012.

Samuel Shapiro and Martin Wilk. An analysis of variance test for normality. *Biometrika*, 52(3):591–611, 1965.

William Sharpe. *Investments*. Englewood Cliffs, 1st edition, 1978.

Steven Shreve. *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model*. Springer, 2004.

Nikolai Smirnov. Table for estimating the goodness of fit of empirical distributions. *Annals of Mathematical Statistics*, 19:279–281, 1948.

Suresh Sundaresan. Continuous-time methods in finance: a review and an assessment. *Journal of Finance*, 13(4):1057–1099, 2000.

Kalyan Talluri and Garrett van Ryzin. *The Theory and Practice of Revenue Management*. Springer, 2005.

Brad Templeton. Reflections on the 25th anniversary of spam. 2008. `http://www.templetons.com/brad/spam/spam25.html`.

Yisong Tian. A modified lattice approach to option pricing. *The Journal of Futures Markets*, 1993.

Ruey Tsay. *Analysis of Financial Series*. John Wiley, 2nd edition, 2005.

Hal Varian. The arbitrage principle in financial economics. *Journal of Economic Perspectives*, 1(2):55–72, 1987.

Hal Varian. Position auctions. *International Journal of Industrial Organization*, 25(6): 1163–1178, 2007.

Hal Varian and Christopher Harris. VCG in theory and practice. Google Working Paper, 2013.

William Vickrey. Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*, 16(1):8–37, 1961.

Michele Villinski. Valuing multiple-exercise option contracts: Methodology and application to water markets. *American Agricultural Economics Association*, 2004. Annual Meeting.

Jun Wang and Bowei Chen. Selling futures online advertising slots via option contracts. In *Proceedings of the 21st International Conference on World Wide Web (WWW '12)*, pages 627–628, Lyon, France, 2012. ACM.

Rafal Weron. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*. John Wiley, 2006.

Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1 (6):80–83, 1945.

Paul Wilmott. *Paul Wilmott On Quantitative Finance*. John Wiley, 2nd edition, 2006.

Shuai Yuan and Jun Wang. Sequential selection of correlated ads by POMDPs. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM'12)*, pages 515–524, Maui, HI, USA, 2012. ACM.

Shuai Yuan, Jun Wang, and Xiaoxue Zhao. Real-time bidding for online advertising: measurement and analysis. In *Proceedings of the 7th Workshop on Data Mining for Online Advertising (ADKDD'13)*, Chicago, IL, USA, 2013. ACM.

Shuai Yuan, Jun Wang, Bowei Chen, Peter Mason, and Sam Seljan. An empirical study of reserve price optimisation in real-time bidding. In *Proceedings of the 20th ACM*

SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'14), pages 1897–1906, New York, NY, USA, 2014. ACM.

Peter Zhang. *Exotic Options*. World Scientific, 2nd edition, 1998.

Weinan Zhang, Jun Wang, Bowei Chen, and Xiaoxue Zhao. To personalize or not: a risk management perspective. In *Proceedings of the 7th ACM Conference on Recommender Systems (RecSys'13)*, pages 229–236, Hong Kong, China, 2013. ACM.