HEC MONTRÉAL

Risk-Aware Bid Optimization for Online Display Advertisement

par

Rui Fan

Erick Delage
HEC Montréal
Directeur de recherche

Sciences de la gestion
(Spécialisation M.Sc.)

Mémoire présenté en vue de l'obtention
du grade de maîtrise ès sciences
(M. Sc.)

March 2022

Résumé

Ce projet se concentre sur le problème de l'optimisation des enchères dans le cadre d'enchères en temps réel pour les publicités d'affichage en ligne, où un annonceur ou l'agent de l'annonceur a accès aux caractéristiques des utilisateurs et aux informations sur les emplacements publicitaires et décide des prix d'enchères optimaux.

Dans cette thèse, nous proposons des modèles d'optimisation des enchères qui tient compte du risque de dépassement de budget et qui exploite les données historiques. Dans nos modèles, l'annonceur doit concevoir à l'avance une politique d'enchères, faisant correspondre le type d'opportunité publicitaire à un prix d'enchère dans une période de temps donnée avec un budget prédéterminé. Nous formulons le problème de deux manières, considérant soit la maximisation du profit ou des revenus comme objectif pour l'annonceur. Après avoir employé une relaxation lagrangienne, nous développons de nouvelles stratégies d'enchères optimales à forme fermée en tenant compte du risque.

En revanche, nous proposons également deux modèles d'optimisation des offres neutres vis-à-vis du risque qui maximisent respectivement le revenu et le profit sans tenir compte du risque de violation de la contrainte budgétaire. Les expressions paramétrées à forme fermée pour chaque stratégie d'enchères optimales neutres au risque sont également dérivées dans cette thèse. Nous comparons la performance de nos quatre modèles, ainsi que l'autre stratégie d'enchères tenant compte du risque dans la littérature récente. Notre méthode d'aversion au risque permet de contrôler efficacement le risque de dépassement du budget pour une période de temps donnée en utilisant une mesure de risque entropique. L'efficacité de ces modèles est mesurée sur un jeu de données réelles qui

contient les caractéristiques du trafic d'un site web et les cibles du prix de gain et des informations sur les clics.

Mots-clés

Politique d'enchères, gestion des risques, optimisation, aversion au risque, apprentissage profond, publicité par affichage.

Méthodes de recherche

Apprentissage machine, mesure de risque entropique, relaxation lagrangienne, régression de moyenne et de variance.

Abstract

This thesis focuses on the bid optimization problem in the real-time bidding setting for online display advertisements, where an advertiser or the advertiser's agent has access to the features of the website visitor and type of ad slot, to decide the optimal bidding prices.

We propose two risk-aware data-driven bid optimization models that exploit historical data and where the advertiser needs to design upfront a bidding policy, mapping the type of advertisement opportunity to a bidding price, to be applied in a given period of time with a predetermined budget. Our two models respectively employ expected profit or revenue as the objective to maximize for the advertiser. After employing a Lagrangian relaxation, we derive a parametrized closed-form expression for the optimal bidding strategy.

By contrast, we also propose two risk-neutral bid optimization models that maximize the revenue and profit respectively without considering the risk of violating the budget constraint. Similar parametrized closed-form expressions for each risk-neutral optimal bidding strategy are also derived in this thesis. Using a real world dataset that contains features of website traffic and targets of winning price and click information, we compare the performance of our four models, as well as the other risk-aware bidding strategy in the latest literature. We demonstrate that our risk-averse methods can effectively control the risk of spending over the budget of a period of time while achieving a competitive level of profit or revenue compared with the risk-neutral models and the other risk-aware bidding strategy.

Keywords

Risk-aware Bidding policy, Optimization, Risk Aversion, Deep Learning, Display Advertising

Research methods

Machine learning, entropic risk measure, Lagrangian relaxation, mean and variance regression.

Contents

List of Tables

List of Figures

List of acronyms

RTB Real time bidding

CTR click-through rate

CVaR Conditional value at risk

CDF Cumulative distribution function

DeepFM Deep factorization model

KPI Key performance indicator

RN Risk neutral

RA Risk averse

RNP Risk neutral approach that maximizes profit

RAP Risk averse approach that maximizes profit

RNR Risk neutral approach that maximizes revenue

RAR Risk averse approach that maximizes revenue

ORTB The non-linear bidding strategy model

RMP Risk management over profit approach

Preface

Before enrolling in the Master in Data Science and Business Analytics program at HEC Montreal, I was managing an e-commerce team in China and operating online stores on Taobao.com [1] and JD.com [2], the two largest e-commerce platform in China for a local brand during the year 2016 and 2019. Since the brand was just established and no awareness either online or offline, we had considered many approaches to marketing the brand as well as products. One of the most important channels is directly marketing the products through these e-commerce platforms' online advertisement slots. At that time, we were able to set the highest budget that we would like to pay for the different slots, and the system provided by the platform bids automatically for different advertisers. As a new brand, instead of restricting the risk of profit from this marketing approach, or aiming to lower cost per click (CPC), we were more concerned that we did not gain enough exposure for the brand and products within the given budget in a given time. Therefore, the models in this thesis will focus on modeling the risk in the bid optimization problem for the online advertisement setting. It will use the profit and revenue as the objective function yet put a special emphasis on controlling the risk of incurring large expenses, hence propose two risk-averse bidding strategies b^{rap} and b^{rar} respectively for profit or revenue maximizers.

[1] `https://world.taobao.com/`
[2] `https://global.jd.com/`

Acknowledgements

During this work, I have received a lot of support in many ways.

First of all, I would like to thank my supervisor, professor Erick Delage. During this project, I was drowned in testing new algorithms and distracted by other researchers' codes many times. It was my supervisor, professor Erick Delage has been guiding me throughout the whole time and I cannot complete this research without him. Thanks to Erick, I gradually become a real researcher. He helped me look under the surface, not focus on how to use the algorithm or code but ask why this approach can serve its goal. I become much more concentrated on the theory and learned to regularly take a step back and look at the bigger picture. Surprisingly, it feels even easier to read or write codes when keeping the theory in my mind. And when I have to give up the deep neural network approach for optimization after trying for several months, it was Erick encouraged me to keep going and helped me remodel the problem from the beginning. From this experience, I realized that one of the most important qualities of a good researcher is persistence and being able to recover and learn from failure. And there are so much more improvements that happened to me during the research experience thanks to him. I sincerely appreciate professor Erick Delage for all his support and guidance on my way to becoming a researcher. With a role model like him, I will continue my journey of being a better researcher!

And I would like to thank our research team of the Canada Research Chair in Decision Making under Uncertainty, who have inspired me with their research ideas and helped me use Compute Canada to facilitate my experiments.

Also, I appreciate the financial support from Mitacs Research Grant and the Differential Tuition Fee Exemption offered by HEC Montreal. Especially, I want to thank HEC Montreal for providing me with this great opportunity and resources to do my research.

Finally, I would like to thank my families and friends who have supported me and encouraged me to keep going for my goal. Research life can be sometimes frustrated, but I can always pull myself together for challenges with them on my side.

Chapter 1

Introduction

In real-time bidding (RTB) online display advertisement settings, the advertiser is given website traffic of clients and ad slots with many different features, including logs, regions, ad slots format, etc., in order to decide the optimal bidding prices for ad potentially displayed to a wide range of users. The bidding process in digital advertising (Ostrovsky et al., 2007) is based on a Vickrey auction (also called second-price auction) model (Vickrey, 1961), where advertisers or their agents win the auction of a given opportunity if they offer the highest bidding price among the competitors and pay the second-highest bidder's bidding price. The bidders in the auction activities can be the advertiser themselves or their agents, which we will refer to as the decision makers for the bidding policy.

These bidding policies typically need to be made with incomplete information about the type of traffic and users' profile, the potential click-through rate [1], the winning price (market price)[2], the net value of each customer to the company etc. Instead, they resort to historical data that could be shared by the advertisement platform or accumulated internally by advertisers themselves. The decision makers in this problem need to use these historical data to estimate the click-through rate, winning price and value of customers based on this data and then develop the optimal bidding policy for the forthcoming bid-

[1]The click-through rate of an advertisement is the number of times a click is made on the advertisement, divided by the number of times the ad is shown, which is also called the number of impressions: `https://en.wikipedia.org/wiki/Click-through_rate`

[2]The winning price or the market price is the price that is paid for the ad spot.

ding opportunities.

This project is conducted from the perspective of decision makers in the bid optimization problem. As bid optimization could be described in the form of static, multi-stage or dynamic problems, this thesis will only focus on the static version of the problem. The objective of this project is to explore a risk-aware bid optimization model that exploits historical data, where the advertiser needs to design up-front a greedy bidding policy, mapping the type of advertisement opportunity to a bidding price for a given number of opportunities at one time, while effectively controlling the risk of violating the budget constraint.

The proposed bidding strategy is implemented and evaluated on the open-source real-world iPinyou dataset[3]. The empirical experiments show that our bidding policies effectively control the risk of the expense going over the budget while outperforming other state-of-the-art risk-aware strategies.

1.1 Motivation

Many companies around the world exploit online marketing, and in 2020, the total spending on the digital ad in Canada reached C$ 8.45 billion (Statista, 2021). Yet, effective use of online marketing budgets requires robust statistical methods that identify key interactions among a possibly biased historical dataset and efficient algorithms that efficiently converge to optimal risk-aware bidding policies. Moreover, few of them considered the risk in the bidding activity and none of them have explicitly addressed accounting for risk aversion on total expense.

In reality, the marketing budget is determined ahead and depends on a certain period of time. The budget is given at the beginning of the time and the decision-makers have no knowledge about what opportunities will come next. If they do not consider the risk of budget constraint, it is very likely to spend money on opportunities that appear earlier

[3]"iPinYou Global RTB Bidding Algorithm Competition Dataset", access on `https://contest.ipinyou.com/`

and miss the valuable opportunities that appear later. Therefore, the policy that controls risk on total expense in the given time is of practical concern in the real world to prevent spending over the total budget at the very beginning, and make sure that the budget can cover the whole marketing period of time, which is essential for both marketing objective and finance objective of the company.

1.2 Contributions

In this thesis, we propose a hybrid approach of empirical distribution and parametric distribution to formulate the bid optimization problem. We develop novel closed-form solutions of optimal risk-neutral and risk-aware bidding policies with objectives of either maximizing revenue or profit. Profit maximization is a natural approach generally applicable to mature companies. The revenue-maximizing models are developed for advertisers who prefer to attract as many clicks as possible within the budget constraint. They may consider awareness and exposure are much more important than return on investment (ROI) or profit, especially when the product or company is in the early stage of the life cycle.

The proposed bidding policies are easy to implement and interpret by being functions of the predicted winning price distribution and click-through rate. The issue that arises with such approaches is that in practice, the realized number of clicks and expenses might differ significantly from their respective average, where we employ the expected utility model that accounts for risk aversion.

Our risk-averse methods are able to control the risk of spending over budget by modelling the expected utility of total expense. The proposed bidding strategies are implemented and evaluated on the open-source real-world dataset. The empirical experiments show that our bidding policy effectively controls the risk of the expense going beyond the budget while outperforming other risk-aware strategies. We expect that the results of this research project will motivate the use of new and more robust models, which will help decision-makers who use algorithmic online marketing to make optimal decisions on bidding prices under high frequency, large-scale data, and uncertainty about performance. It will help businesses improve their key performance indicators while increasing the chances of remaining within their allocated budget.

1.3 Outline

The rest of this thesis is organized as follows:

The literature review is presented in Chapter 2. The research methodology and the problem descriptions are detailed in Chapter 3. The closed-form optimal bidding strategies under the risk-neutral approach are derived in Chapter 4, while the risk-averse approach is in Chapter 5. Experiment design is conducted in Chapter 6, followed by the numerical results of experiments presented in Chapter 7. Finally, we conclude our work in the last Chapter 8.

Chapter 2

Literature Review

Our model for bid optimization is built on the estimation of the click-through rate and the winning price of the bidding opportunity. We model the risk based on the expected utility theory. Therefore, we explore the literature in these related aspects respectively.

2.1 Click-through Rate Prediction

The click-through rate (CTR) prediction is a binary classification problem and the most commonly used estimator in computational advertising is the Logistic Regression (Richardson et al., 2007; Zhang et al., 2010; Graepel et al., 2010) based on a linear model. Non-linear models can also be applied, such as tree-based Gradient Boosting Regression Tree (Friedman, 2002), and many models have been developed based on the Factorization Machines (FM) (Rendle, 2010) for the CTR prediction problem (Ta, 2015): Sparse Factorization Machines (SFM) (Pan et al., 2016), Field-aware Factorization Machines (FFM) (Juan et al., 2017), and Field-Weighted Factorization Machines (FwFM) (Pan et al., 2018), which can better fit the feature combination and sparse data often found in the display advertising dataset.

In recent years, with the development of deep learning research and recommendation systems, many researchers have applied deep neural network based models on this CTR prediction task. There are convolution neural network based (Liu et al., 2015), Product-

Based Neural Networks (Qu et al., 2018), Wide&Deep (Cheng et al., 2016), the FM based neural network (DeepFM) (Guo et al., 2017). Built on the top of the DeepFM, the Feature Importance and Bilinear feature Interaction NETwork (Huang et al., 2019) and Field-Embedded Factorization Machines (Pande, 2021) were developed most recently. The idea of combining the deep neural network and factorization machine is widely accepted in the research and represents the leading performance in real-world usage.

2.2 Winning Price Prediction

The winning price prediction is also called the bid landscape problem. Cui et al. (2011) uses the gradient boosting decision trees to model the market price. Wu et al. (2015) proposed the censored regression model to deal with the problem when some historical market prices are unknown to the advertiser.

The recent deep learning advancement also applies to the bid landscape problem. Wu et al. (2018b) proposed the Deep Censored Learning model that uses the deep learning model for click-through rate prediction to boost the prediction quality on the market price and the distribution of market price also influences the learning result. Deep Landscape Forecasting (DLF) (Kan et al., 2019) model combines deep learning for probability distribution forecasting and survival analysis for censorship handling based on a recurrent neural network to model the conditional winning probability with respect to each bid price.

2.3 Bid Optimization

Many models have been proposed to design efficient bid optimization policies based on historical data. Typically, the purpose of bidding algorithms is optimizing advertisers' Key Performance Indicator (KPI), such as the the number of clicks and conversion rate, within a limited budget. In particular, many approaches build their bidding price depending on the click-through rate predictor and optimize a simple parametric bid policy based on this prediction in order to maximize the expected number of clicks while respecting a constraint on the average budget.

Ostrovsky et al. (2007) proposed truthful bidding which is based on the estimation of the value of the click. Continued on the truthful bidding approach, Chen et al. (2011) developed a bidding function with truthful bidding value minus a tuned parameter and solves a bidding problem with multiple campaigns and from the perspective of the publisher using linear programming and duality. Another widely used linear bidding strategy is calculating the bid price via the predicted click-through rate multiplied by a constant parameter tuned according to the campaign budget and performance (Perlich et al., 2012). There is also the bid optimization problem in sponsored search advertising where advertisers bid on queries or keywords (Even-dar et al., 2009; Borgs et al., 2007; Zhang et al., 2012; Yang et al., 2015).

Zhang et al. (2014) introduced a non-linear bidding model: Optimal Real-Time Bidding (ORTB) strategy. Again, with the estimated click-through rate that depends on features as the input of the bidding function which tries to bid more impressions rather than focus on a small set of high-value impressions. In this paper, the bidding policy is cast as a functional optimization problem with maximizing the expected total clicks within the given total budget $\mathscr{B}$ in a total of N opportunities under the dependency assumptions: the winning probability Q_W depends on the bidding price b; the bidding price depends on the observable features X.

Mathematically, the ORTB model can be represented by:

$$\begin{aligned} b(\cdot)_{ORTB} = \quad & \underset{b(\cdot)}{\text{argmax}} \quad N\int_X \theta(X) Q_W(b(X)) p_X(X) dX \\ & \text{subject to} \quad N\int_X \hat{w}(X) Q_W(b(X)) p_X(X) dX \leq \mathscr{B}, \end{aligned} \tag{2.1}$$

where the $p_X(X)$ is the probability density function of features X.

They model the winning probability by:

$$Q_W(b) = \frac{b}{c+b}, \tag{2.2}$$

where c is a constant parameter tuned to fit the winning probability given different value of bidding price compared with the winning price W in the dataset.

To solve the model (2.1), the author introduce λ as a Lagrangian multiplier to approximate the solution to the relaxation function:

$$\mathscr{L}(b(X),\lambda) = \int_X \theta(X)Q_W(b(X))p_X(X)dX - \lambda \int_X \hat{w}(X)Q_W(b(X))p_X(X)dX + \lambda \frac{\mathscr{B}}{N},$$

where the parameter λ is tuned to minimize the difference between the maximum total expense and the given total budget $\mathscr{B}$ for N opportunities.

Since the bidding follows the second-price auction, the advertiser always pay the winning price that is less than the bidding price. The upper bound of expense can be obtained by using bidding price to multiply with the winning probability:

$$N \int_X \hat{w}(X)Q_W(b(X))p_X(X)dX < N \int_X b(X)Q_W(b(X))p_X(X)dX.$$

Therefore, the optimal λ is approximated by:

$$\lambda^* = \text{argmin} \sum_{i=1}^{N} \left(b_\lambda(X)Q_W(b_\lambda(X)) - \frac{\mathscr{B}}{N} \right)^2$$

As a result, an approximately optimal bidding policy with winning probability (2.2) for problem (2.1) is obtained by:

$$b_\lambda(X) = \sqrt{\frac{c}{\lambda}\theta(X) + c^2} - c.$$

They also assume a different winning probability by:

$$Q_W(b(X)) = \frac{b^2(X)}{c^2 + b^2(X)}. \tag{2.3}$$

Therefore, an approximately optimal bidding policy with the winning probability assumption (2.3) can be obtained by:

$$b_\lambda(X) = c[(\frac{\theta(X) + \sqrt{c^2\lambda^2 + \theta(X)^2}}{c\lambda})^{\frac{1}{3}} - (\frac{c\lambda}{\theta(X) + \sqrt{c^2\lambda^2 + \theta(X)^2}})^{\frac{1}{3}})].$$

Fernandez-Tapia et al. (2016) solved the bid problem in cases where impressions generated by homogeneous Poisson processes and winning prices are independent and identically distributed (IID). While Ren et al. (2018) modelled CTR learning and winning price estimation as part of bid optimization for campaign profit maximization as a whole and performed a novel joint optimization.

There are also different objectives in bid optimization. Liu et al. (2017) suggested a dual based bidding framework that is derived from a strict second-price auction assumption and generally applicable to the multiple ads scenario with various objectives and constraints. Besides the profit maximization objective, Yang et al. (2019) studied the common case where advertisers aim to maximize the number of conversions, and set cost-per-click (CPC) as a KPI constraint. They convert such a problem into a linear programming problem and leverage the primal-dual method to derive the optimal bidding strategy.

Researchers also consider the bid optimization in a multi-stage setting as a sequential decision process, where the Reinforcement Learning (RL) approach plays an important role. Typically, the bid optimization problem is formulated by Markov Decision Process (MDP), where the bid prices are the actions to find, and the realized clicks are the rewards to the RL problem. The budget constraint is treated differently in the previous research. Cai et al. (2017) treated the budget as part of the state and derived the solution by dynamic programming while Du et al. (2017) emphasized the budget constraint using the Constrained Markov Decision Process (CMDP) and solving it by linear programming. Wu et al. (2018a) introduced the budget constraint into the reward by adding a linear penalty. Adikari and Dutta (2019) considered the budget also as a decision to make where both budget allocation and bidding price are optimized by dynamic programming. Researchers also formulate it as a multi-agent RL problem. Jin et al. (2018) proposed a cluster-based bidding strategy to achieve overall benefits among advertisers as multiple agents. Zhao et al. (2018) also looked into the cooperative rewards and addressed the problem of the environment changing by modelling the hourly transition probabilities. However, RL approaches are generally more computationally expensive to solve compared with static models.

2.3.1 Risk aware optimization

Since the bidding process can be modelled as a multi-stage problem, many researchers looked into the feedback control problem during the bidding process (an overview can be found in (Karlsson, 2020)), which control the risk of unstable performance and keep the optimization process along the stages using a dynamic system. An approximate solution proposed by Grislain et al. (2019) is based on a Recurrent Neural Network (RNN) architecture, which add the additional penalty to the objective function if the bidding policy falls short of its KPI to improve the robustness of performance under uncertainty. Karlsson (2016) introduced the bid randomization mechanism to help exploration in a partially observed market and control the uncertainty in the auction-based bidding process.

RMP approach

Especially, Haifeng et al. (2017) proposed the risk management on profit (RMP) model that also models the bidding process as a one-shot problem and modeling the risk of profit directly. They adopt a Bayesian logistic regression model to predict the click-through rate distribution because they are modeling the risk of profit, and they assume the risk of profit comes from the uncertainty of the click-through rate. The proposed model introduces the standard deviation of the predicted CTR distribution as a risk penalty added into the bidding function.

For every opportunity, there is a risk of negative return whatever the positive bid price when the advertiser pays for the impression but does not obtain the click. Namely, the expected return $R(b)$ can be modeled as:

$$R(b) = \begin{cases} 0, & b \le W \\ v \cdot \theta - W, & b > W \end{cases}$$

where v is the value of click. θ is the probability of a click for this opportunity, and W is winning price, which are modeled as stochastic random variables with p.d.f. $p_\theta(\theta)$ and $p_W(W)$.

The author defines the risk-controlled return by deducting a risk factor, which is the standard deviation of the return. Therefore, the bidding policy choose the bid price that could maximize the lower bound of return given certain risk. Their RMP strategy is modeled as below:

$$b_{RMP}(R(b)) = \underset{b}{\text{argmax}} \qquad \mathbb{E}[R(b)] - \alpha \text{Std}[R(b)], \tag{2.4}$$

where the risk of profit $R(b)$ is influenced by the hyperparamter α and the standard deviation of profit $\text{Std}[R(b)]$, and this volatility of randomness come from $\theta \sim p_\theta(\theta)$ and $w \sim p_w(w)$.

Another noticeable point is the RMP bidding model obtained without the budget involved. The RMP approach only considers the risk of negative profit but not on the risk of expense going over the budget.

Risk Aversion

Utility is a topic of economics that is typically used to model worth or value.[1] In the display advertising field, the utility of bidding is often defined as the profit of clicks. The papers (Chapelle, 2015; Vasile et al., 2017; Haifeng et al., 2017) define the value of click v as the value of total winning price divided by total clicks and measures utility as the profit of bidding which is the difference between the total value of clicks and the expense paid. In these papers, the expected utility theory (Bernoulli, 1954; John von Neumann, 1944) considers risk in probability and the optimal decisions are the choices that provide higher utility.

In the field of expected utility theory, the exponential utility function accounting for a constant absolute risk aversion is the most commonly used utility function. Expected utility theory with the exponential utility function can also be interprete as employing an entropic risk measure (Rudloff and Wunderlich, 2008), which satisfies the axioms of convex risk measures (Foellmer and Schied, 2010).

[1] https://en.wikipedia.org/wiki/Utility#cite_note-exputility-5

Chapter 3

Research Methodology

In this project, we adopt the common assumptions about dependency relationships between variables from the other literature (Haifeng et al., 2017). For each bidding opportunity with an observable feature vector X that represents both user and ad information, the bidding price depends on the click-through rate, and the realized click C represents if the fact is that the ad gets clicked; the distribution of wining price W; and the net value of this customer to the company V. Also, we assume these variables C, W, V are mutually independent given X as we describe in the Assumption 1:

Assumption 1. *The winning price W, realized C and net value of the customer V are mutually independent given X.*

Therefore, we model C, W, V separately for each given opportunity X.

In this thesis, we propose an approach in the real-time bidding (RTB) setting that optimizes the bid policy over a batch of M opportunities, that are independent and identically distributed drawn. This approach is designed to simulate the situation that the one-shot decision of a bidding policy is made for a given period of time where the number of advertisement display opportunities and the budget are fixed.

In order to evaluate the risk-averse bidding strategies that we propose, we apply the performance metrics that account for risk aversion as well as the common metrics used in the bid optimization problem. To be consistent with the modeling, all the metrics are measured in every batch of opportunities.

3.1 Modeling Conditional Click-through Rate

In reality, the click-through rate is an indicator of the attractiveness of a display advertising opportunity. Therefore, the value of an advertising spot is higher when this spot has a higher probability to be clicked.

We assume that the click-through rate, i.e. the probability of the opportunity getting click, depends on the opportunity's features X, and formally denote:

$$\theta(X) := \mathbb{P}(C|X).$$

In the context of this work, we will employ the DeepFM model (Guo et al., 2017) to estimate $\theta(X)$. Besides, our optimal bidding policy can be easily incorporated with other CTR prediction algorithms, and its performance can be further improved due to more accurate predictions of these dependent variables.

The resource we use for CTR prediction is from the *DeepCTR* library (Shen, 2017) ,[1] which provides a set of easy-to-use packages and models API for feature engineering and deep learning models. By mainly using this *DeepCTR* Models API, we built the predictor for CTR of each opportunity.

Common methods that model unstructured websites' features usually have a strong bias towards interactions between features or require expertise in feature engineering. The DeepFM is an FM-based Neural Network that could emphasize both explicit and implicit feature interactions behind user behaviours. The DeepFM consists of two components, the FM component and the deep component, which share the same input and are jointly trained. The Factorization Machines (FM) model the pairwise feature interactions as the inner product of latent vectors between features. Although the FM model is designed for better revealing proper combinations of basic features, it suffers from expensive compu-

[1] DeepCTR library `https://github.com/shenweichen/deepctr`

tation costs when the order goes above two. Therefore, the DeepFM model is able to combine the advantages of factorization machines for feature learning and a feed-forward neural network for generalization, the model DeepFM has superior performance in dealing with unstructured data and programmatic advertising and recommendation system design.

Remark 1. *Due to the advantages in modeling unstructured data as we explained above, we estimate the click-through rate $\theta(X)$ using the DeepFM model while both ORTB* (2.1) *and RMP* (2.4) *approaches are applying linear-based models: the ORTB method estimates the $\theta(X)$ using the logistic regression, and the RMP approach uses the Bayesian logistic regression.*

3.2 Modeling Conditional Winning Price Distribution

We assume that conditional on observing X, the winning price W follows the normal distribution:

$$W \sim N(\hat{w}(X), \sigma(X)),$$

where the conditional standard deviation $\sigma(X)$ also depends on the given opportunity's features X, so that we have the parametrized probability distribution function of the winning price W modeled as follows:

$$f_{W|X}(w) = \frac{1}{\sigma(X)\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{w-\hat{w}(X)}{\sigma(X)}\right)^2}.$$

Therefore, to model the distribution of winning price, we need estimators for the conditional mean $\hat{w}(X)$ and the conditional standard deviation of the winning price $\sigma(X)$.

Given a dataset containing observed (X,W) pairs, we can obtain an estimator of expected winning price, conditioned on X, by running the regression model:

$$\hat{w}(X) := \arg\min_{w \in \mathscr{W}} \mathbb{E}[(W - w(X))^2],$$

where $\mathscr{W}$ is the set of estimation functions modeled by a certain neural network architecture, and where $\mathbb{E}$ refers to the expected value under the empirical distribution.

The same dataset can also be used to train an estimator of $\sigma(X)$ following a method introduced in the paper (Fan and Yao, 1998). Conceptually, we consider the residual Z defined as:

$$Z := (W - \hat{w}(X))^2,$$

which depends on the estimator $\hat{w}$, as well as the W. We can approximate the conditional variance of the winning price $\sigma^2(X)$ as the expected value of the residual Z.

$$\hat{z}(X) := \arg\min_{z \in \mathscr{Z}} \mathbb{E}[(Z - z(X))^2],$$

where again $\mathscr{Z}$ is the set of estimation functions modeled by a certain neural network architecture. Finally, an estimator of the conditional standard deviation of winning price $\sigma(X)$ can be obtained by:

$$\sigma(X) := \sqrt{\max(\hat{z}(X), \varepsilon)},$$

for some small $\varepsilon > 0$, which ensures that the variance estimate is always positive.

In the experiments, similar with the click-through rate estimator θ, the estimators of mean of winning price $\hat{w}(X)$ and standard deviation of winning price $\sigma(X)$ are predicted based on DeepFM models.

3.3 Modeling Conditional Winning Probability

In the RTB setting, the advertisers follow the second price auction (Ostrovsky et al., 2007) process and win the ad spots only if they bid the highest price among all the advertisers, and the winning price is the second highest price offered among the bidders and is the price that winner needs to pay for the given ad display opportunity. Therefore, we assume that the advertisers or their agents can win the bid if they offer a bid price that larger than the winning price.

Since the expense only happens when the advertiser wins the bids, in order to model uncertainty about expense, we need to model the probability of winning the bid which depends on the bidding price and the winning price. For this purpose, we define a function that indicates whether the bidding price wins the auction:

$$s(b,W) := \mathbf{1}\{b \geq W\} = \begin{cases} 1 & b \geq W \\ 0 & \text{otherwise} \end{cases}$$

where b is the bid price and W is the winning price. Hence the conditional winning probability given X can simply be obtained from $\mathbb{E}[s(b,W)|X]$, which depends on both the bid price b and the conditional winning price distribution given X.

Remark 2. *Both the ORTB approach* (2.1) *and the RMP approach* (2.4) *assume the winning price is a stochastic random variable that is independent with opportunity's features X. For the ORTB approach, the winning probability is obtained by certain reciprocal functions specified in equations* (2.2) (2.3)*, and tune the constant parameter c by fitting the winning probability graph when varying the bid price. As for the RMP approach, the winning probability is estimated using Monte Carlo Simulation.*

However, we assume that the winning price W is conditional on the given opportunity X and follows the Gaussian distribution, which is a more natural and logical assumption as advertisers are willing to bid more for the more valuable opportunity if they can successfully identify them through the observable features X. Based on this assumption, we develop the winning probability conditional on X and further derive the winning probability in the latter section.

3.4 Modeling Conditional Value of Customer

Ideally, if the decision maker identifies a customer who could bring a higher value to the company from the observable features vector X, they will be willing to bid at a higher price to improve the probability of winning the auction. Therefore, we assume the bidding price should depend on the value of customer V and we model this variable conditional on the given opportunity X by:

$$\hat{V}(X) := \mathbb{E}[V|X].$$

Yet, in our experiments, V will be considered independent of X and known for simplicity.

3.5 Evaluation Metrics

Many evaluation metrics are relevant to the RTB setting. In reality, the marketing budget is determined ahead and depends on a certain period of time. The number of ad display opportunities is also fixed given a certain period of time since the number of ad spots is usually fixed for a given period of time unless there is a change in website design.

Therefore, we propose a new evaluation approach in the RTB setting that all the metrics are measured in every batch of opportunities to simulate the number of ad display opportunities that are certain and the budget is fixed for a given period of time. Here, we introduce a batch containing M independent and identical distributed display opportunities:

$$\text{Batch} := \{(X_i, W_i, C_i, V_i)\}_{i=1}^{M}$$

In our research, our objective is to maximize profit for the given number of display opportunities and lower the risk of spending over budget for a certain time. Thus, we not only apply the common advertising metrics for a given time but also we see the batches as experiment samples and apply the risk measures over the distribution of batches.

3.5.1 Batch revenue

We define the revenue of advertiser as the total value of clicks. Therefore, we can formulate the expected revenue by using the CTR and value per click:

$$\text{Batch Revenue} = \sum_{i=1}^{M} V_i C_i s(b(X_i), W_i).$$

3.5.2 Batch expense

We define the expense as total amount of money paid if the bid price is higher than the winning price:

$$\text{Batch Expense} = \sum_{i=1}^{M} W_i s(b(X_i), W_i).$$

3.5.3 Batch profit

We define the profit as revenue deduct profit:

$$\text{Batch Profit} = \text{Batch Revenue} - \text{Batch expense}.$$

3.5.4 Batch number of clicks

The number of clicks here is the realized clicks:

$$\text{Batch number of clicks} = \sum_{i=1}^{M} C_i s(b(X_i), W_i).$$

3.5.5 Batch impression rate

The frequency of time when the bidding price is higher than the winning price among the total opportunities for a given time. The probability that advertiser successfully expose the ad to the customers, which is the realized winning rate:

$$\text{Batch impression rate} = \frac{\sum_{i=1}^{M} s(b(X_i), W_i)}{M}.$$

3.5.6 Sharpe ratio of batch profit or revenue

Sharpe ratio (Sharpe, 1994) is a common metric that measures the return given its risk. We apply this concept to our setting, by the expected value of batch return dividing its standard deviation from sample batches. The return is either batch profit or batch revenue depending on the model:

$$\text{Sharpe ratio of profit} = \frac{\mathbb{E}[\text{Batch Profit}]}{\sigma(\text{Batch Profit})},$$

$$\text{Sharpe ratio of revenue} = \frac{\mathbb{E}[\text{Batch Revenue}]}{\sigma(\text{Batch Revenue})}.$$

3.6 Variable table

In order to better display the problem in the following sections, we summarize the definition of our variables in table 3.1:

Variables	**Description**
V	net value of customer for opportunity
$\hat{V}(X)$	the conditional net value of customer estimator given X
X	observed features of opportunity
C	indicates of click for opportunity
$\theta(X)$	the conditional click-through rate estimator given X
$b(X)$	bidding price for opportunity given X
W	the winning price for opportunity
$\hat{w}(X)$	the conditional expected value of winning price for opportunity given X
$s(b,W)$	indicates of winning the auction with bid b when winning price is W
$\sigma(X)$	the conditional standard deviation of winning price for opportunity given X
B	average budget per opportunity
M	the number of opportunities in a batch
N	total number of opportunities in the dataset

Table 3.1: Variable definitions

Chapter 4

Risk-neutral Approach

The goal of bid optimization is to maximize the return of advertising within a certain budget for a given time.

As the common business problem, the profit of the bid optimization problem is the difference between the revenue and the expense. The revenue here is the value of total clicks, which relates to V and C. But the V and C are unrevealed to the advertisers and we need to estimate both values.

The expense is the amount of expense paid for the winning opportunities. As we introduced in the Chapter 1, the bidding process in the RTB setting follows the second-price auction. Thus, the bidding price here is not the amount that advertiser paid but the winning price, and the winning price W remains unrevealed to the advertisers until the bidding event finishes. Therefore, we develop the estimator for the winning price.

As for the budget, we refer to the budget given for a certain period of time, such as the daily budget or monthly budget that is predetermined by advertisers, which is a common procedure for marketing budgeting practice. Since the number of ad slots during a certain period of time is relatively consistent from time to time, the budget for a given time here is the same as the budget for a given number of bidding opportunities.

We first construct two basic bidding policy models that impose that budget be satisfied in expectation without consideration of risk of spending over budget: b^{rnp} will maximize the expected profit while b^{rnr} will maximize the expected revenue. Chapter 5 will later develop respective risk aware model in terms of the budget constraint.

Variables	**Description**
$b^{\text{rnp}}(\cdot)$	risk-neutral bid policy that maximizes profit problem
$b^{\text{rnr}}(\cdot)$	risk-neutral bid policy that maximizes revenue problem

Table 4.1: Risk-neutral models

4.1 Profit Maximization Problem Formulation

We first construct the b^{rnp} model, where the objective function is to maximize the expected profit subject to a budget constraint for a given batch of opportunities. As we discussed in the Chapter 3, we consider a random batch of M i.d.d. opportunities. We assume that V, C, and W are mutually independent given X (see Assumption 1). We define the predetermined average budget per opportunity is B:

$$B := \frac{\text{Budget for the batch of opportunities}}{M}.$$

Therefore, we describe the b^{rnp} as follows:

$$\begin{aligned} b^{\text{rnp}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[(1/M)\cdot \text{Batch profit}] \\ \text{subject to} \quad & \mathbb{E}[(1/M)\cdot \text{Batch expense}] \leq B, \end{aligned}$$

where $b^{\text{rnp}}(\cdot)$ is the bid price that maximize the expected average profit from a batch of M opportunities, while respect the constraint that the expected expense per opportunity in the batch is within the predetermined budget per opportunity B.

The above model mathematically takes the form:

$$\begin{aligned} b^{\text{rnp}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} V_i C_i s(b(X_i), W_i)] - \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i), W_i)] \\ \text{subject to} \quad & \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i), W_i)] \leq B. \end{aligned}$$

Based on the linearity of expectation, these expressions can be simplified. For example, we can simplify the batch revenue to the expected instantaneous value per opportunity format by:

$$\mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} V_i C_i s(b(X_i),W_i)] = \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}[V_i C_i s(b(X_i),W_i)] = \mathbb{E}[VCs(b(X),W)];$$

Applying the same simplifying method, we can represent the batch expense using the expected instantaneous expense per opportunity:

$$\mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i),W_i] = \mathbb{E}[Ws(b(X),W)];$$

Therefore, the batch profit can be simplified using the difference between expected instantaneous revenue and expense:

$$\begin{aligned}&\mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} V_i C_i s(b(X_i),W_i)] - \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i),W_i)]\\ &= \mathbb{E}[VCs(b(X),W)] - \mathbb{E}[Ws(b(X),W)].\end{aligned}$$

As a result, we obtain an equivalent policy for b^{rnp} using an expected instantaneous profit maximization form:

$$\begin{aligned} b^{\text{rnp}}(\cdot) = \operatorname*{argmax}_{b:\mathscr{X}\to\mathbb{R}^+} \quad & \mathbb{E}[VCs(b(X),W)] - \mathbb{E}[Ws(b(X),W)] \\ \text{subject to} \quad & \mathbb{E}[Ws(b(X),W)] \leq B. \end{aligned} \tag{4.1}$$

In order to obtain the closed-form functions of the objective and the constraint, we need to first transform the model using the estimators defined in Chapter 3 and exploit the conditional independent Assumption 1. We can start with the approximation for revenue and expense. From there, we can easily obtain the profit by taking the difference between revenue and expense.

4.1.1 Reducing expected instantaneous revenue expression

We start by rewriting the expected instantaneous revenue expression as:

$$\begin{aligned}\mathbb{E}[VCs(b(X),W)] &= \mathbb{E}[\mathbb{E}[VCs(b(X),W)|X]] \\ &= \mathbb{E}[\mathbb{E}[V|X]\mathbb{E}[C|X]\mathbb{E}[s(b(X),W)|X]] \\ &= \mathbb{E}[\hat{V}(X)\theta(X)\mathbb{E}[s(b(X),W)|X]],\end{aligned}$$

where we exploited (see Assumption 1) the fact that V, C, and W are mutually independent from each other given X.

The conditional winning probability $\mathbb{E}[s(b,W)|X)]$ can be further reduced since we assume W is normally distributed conditionally on X:

$$\begin{aligned}\mathbb{E}[s(b,W)|X] &= \int_{W\leq b} f_{W|X}(W)dW \\ &= \int_{W\leq b} \frac{1}{\sigma(X)\sqrt{2\pi}}\exp(-\frac{(W-\hat{w}(X))^2}{2\sigma(X)^2})dW \\ &= \int_{W\leq \frac{b-\hat{w}(X)}{\sigma(X)}} \frac{1}{\sqrt{2\pi}}\exp(-\frac{W^2}{2})dW \\ &= \Phi(\frac{b-\hat{w}(X)}{\sigma(X)}),\end{aligned}$$

where $f_{W|X}(w)$ is the probability density function of the winning price W given X, and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

Therefore, we can calculate the expected revenue given X using:

$$\mathbb{E}[\text{Revenue}|X] = \hat{V}(X)\theta(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}). \tag{4.2}$$

4.1.2 Reducing expected instantaneous expense expression

Since advertisers only pay the bidding expense at the winning price W if they win the bid, we exploit the assumption that the winning price follows a normal distribution conditional on X.

In the case of the budget constraint, we have that:

$$
\begin{aligned}
\mathbb{E}[\text{Expense}|X] &= \mathbb{E}[Ws(b,W)|X] \\
&= \int_{W\leq b} W f_{W|X}(W)dW \\
&= \int_{w_s\leq \frac{b-\hat{w}}{\sigma}} (\sigma(X)w_s+\hat{w}(X))\phi(\sigma(X)w_s+\hat{w}(X))\sigma(X)dw_s \\
&= \int_{w_s\leq \frac{b-\hat{w}}{\sigma(X)}} (\sigma(X)w_s+\hat{w}(X))\phi(w_s)dw_s \\
&= \sigma(X)\int_{w_s\leq \frac{b-\hat{w}(X)}{\sigma(X)}} w_s\phi(w_s)dw_s + \hat{w}(X)\int_{w_s\leq \frac{b-\hat{w}}{\sigma(X)}} \phi(w_s)dw_s \\
&= \hat{w}(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) - \sigma(X)\phi(\frac{b-\hat{w}(X)}{\sigma(X)}),
\end{aligned} \tag{4.3}
$$

where w_s follows the standard normal distribution, and $\phi(\cdot)$, $\Phi(\cdot)$ are respectively the density function and cumulative distribution function of a standard normal distribution.

4.1.3 Reducing expected instantaneous profit expression

According to the reduced forms of the expected instantaneous revenue (4.2) and expense (4.3) we developed in the former sections, we can develop the reduced form of expected instantaneous profit as:

$$
\begin{aligned}
\mathbb{E}[\text{Profit}|X] &= \mathbb{E}[VCs(b(X),W)|X] - \mathbb{E}[Ws(b(X),W)|X] \\
&= \hat{V}(X)\theta(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) - [\hat{w}(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) - \sigma(X)\phi(\frac{b-\hat{w}(X)}{\sigma(X)})],
\end{aligned} \tag{4.4}
$$

which is the objective function for profit maximizing model models.

4.2 Closed-form Optimal Solution of Lagrangian Relaxation for RNP Model

After we derive the reduced profit (4.4) and expense (4.3) expressions, we can reformulate the model for b^{rnp} (4.1) using the reduced form:

$$
\begin{aligned}
b^{\text{rnp}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\operatorname{argmax}} \quad & \mathbb{E}[\hat{V}(X)\theta(X)\Phi(\frac{b(X)-\hat{w}(X)}{\sigma(X)}) - g(b(X),X)] \\
\text{subject to} \quad & \mathbb{E}[g(b(X),X)] \leq B,
\end{aligned}
$$

where

$$
\begin{aligned}
g(b,X) :&= \mathbb{E}[Ws(b,W)|X] \\
&= \hat{w}(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) - \sigma(X)\phi(\frac{b-\hat{w}(X)}{\sigma(X)}). \qquad (4.5)
\end{aligned}
$$

In an attempt to solve this problem, one can introduce the Lagrangian coefficient $\lambda \geq 0$ to obtain a relaxation of this risk-neutral profit maximizing problem:

$$
\begin{aligned}
\tilde{b}_\lambda^{\text{rnp}}(\cdot) :&= \underset{b:\mathscr{X}\to\mathbb{R}^+}{\operatorname{argmax}} \quad \mathbb{E}[\hat{V}(X)\theta(X)\Phi(\frac{b(X)-\hat{w}(X)}{\sigma(X)}) - g(b(X),X)] - \lambda\mathbb{E}[g(b(X),X) - B] \\
&= \underset{b:\mathscr{X}\to\mathbb{R}^+}{\operatorname{argmax}} \quad \mathbb{E}[\mathcal{G}_\lambda(b(X),X)],
\end{aligned}
$$

where

$$
\mathcal{G}_\lambda(b,X) := \hat{V}(X)\theta(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) - g(b,X) - \lambda(g(b,X) - B).
$$

The optimal bid price b^{rnp} can be approximated using $\tilde{b}_{\lambda^*}^{\text{rnp}}$ with λ^* as the smallest value of $\lambda \geq 0$ such that:

$$
\mathbb{E}[g(\tilde{b}_{\lambda^*}^{\text{rnp}}(X),X)] \leq B.
$$

We next provide a closed-form solution for $\tilde{b}_\lambda^{\text{rnp}}$ in the form of Lemma 1.

Lemma 1. *For any $\lambda \geq 0$, a maximizer of the Lagrangian relaxation takes the form:*

$$
\tilde{b}_\lambda^{rnp}(X) := \arg \max_{b\in\{0,\, \frac{\hat{V}(X)\theta(X)}{\lambda+1},\, \infty\}} \mathcal{G}_\lambda(b,X),\ \forall X \in \mathscr{X}.
$$

Proof. We start by exploiting the interchangeability property of expected value (see (Shapiro, 2017)), which implies that the optimal bid price for the Lagrangian relaxation can be obtained as the price, for each X, that maximizes the Lagrangian relaxation function $\mathcal{G}$:

$$
\begin{aligned}
\tilde{b}_{\lambda^*}^{\text{rnp}}(X) = \underset{b\in\mathbb{R}^+}{\operatorname{argmax}}\, \mathcal{G}_\lambda(b,X) :&= \hat{V}\theta\Phi(\frac{b-\hat{w}}{\sigma}) - (\lambda+1)[\hat{w}\Phi(\frac{b-\hat{w}}{\sigma}) - \sigma\phi(\frac{b-\hat{w}}{\sigma})] + \lambda B \\
&= -(\hat{w} + \lambda\hat{w} - \hat{V}\theta)\Phi(\frac{b-\hat{w}}{\sigma}) + (1+\lambda)\sigma\phi(\frac{b-\hat{w}}{\sigma}) + \lambda B, \qquad (4.6)
\end{aligned}
$$

where we dropped the relation to X for simplicity of presentation.

Since $\mathcal{G}_\lambda(b,X)$ is twice differentiable with respect to b, the maximizer for b is either $0,\infty$ or at a value where the derivative is 0. For the latter case, we get that:

$$\frac{d\mathcal{G}_\lambda(b,X)}{db} = 0 \Leftrightarrow \frac{\hat{w}+\lambda\hat{w}-v\theta}{\sigma}\phi(\frac{b-\hat{w}}{\sigma})+(1+\lambda)\frac{b-\hat{w}}{\sigma}\phi(\frac{b-\hat{w}}{\sigma}) = 0\,.$$

Hence, we can conclude that the value of b where $\mathcal{G}_\lambda(b,X)$ has a derivative of zero is:

$$b = \frac{\hat{V}(X)\theta(X)}{\lambda+1}.$$

□

4.3 Revenue Maximization Problem Formulation

This b^{rnp} (4.1) model can be easily transformed into a revenue maximizing model by using revenue as the objective function. To be more specific, the objective function here is maximizing the average expected revenue per batch of M opportunities while, once again, satisfying the budget constraint per batch in expectation.

Therefore, the b^{rnr} can be described as:

$$\begin{aligned} b^{\text{rnr}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[(1/M)\cdot \text{Batch revenue}] \\ \text{subject to} \quad & \mathbb{E}[(1/M)\cdot \text{Batch expense}] \leq B, \end{aligned}$$

where $b^{\text{rnr}}(\cdot)$ is the bid price that maximizes the expected average revenue from a batch of M opportunities, while respecting the same constraint as in the expected profit maximization problem.

The above model mathematically takes the form:

$$\begin{aligned} b^{\text{rnr}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} V_iC_is(b(X_i),W_i)] \\ \text{subject to} \quad & \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} W_is(b(X_i),W_i)] \leq B. \end{aligned}$$

Based on linearity of expectation, it is again equivalent to model the expected instantaneous revenue problem:

$$\begin{aligned} b^{\mathrm{rnr}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\operatorname{argmax}} \quad & \mathbb{E}[VCs(b(X),W)] \\ \text{subject to} \quad & \mathbb{E}[Ws(b(X),W)] \leq B. \end{aligned} \tag{4.7}$$

4.4 Closed-form Optimal Solution of Lagrangian Relaxation for RNR Model

According to the reduced form of revenue (4.2) and expense (4.3), we can reformulate the risk-neutral revenue model b^{rnr} using the reduced form:

$$\begin{aligned} b^{\mathrm{rnr}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\operatorname{argmax}} \quad & \mathbb{E}[\hat{V}(X)\theta(X)\Phi(\frac{b(X)-\hat{w}(X)}{\sigma(X)})] \\ \text{subject to} \quad & \mathbb{E}[\hat{w}(X)\Phi(\frac{b(X)-\hat{w}(X)}{\sigma(X)}) - \sigma(X)\phi(\frac{b(X)-\hat{w}(X)}{\sigma(X)})] \leq B. \end{aligned} \tag{4.8}$$

Similarly, in an attempt to solve problem (4.8), one can introduce the Lagrangian coefficient $\lambda \geq 0$ to obtain a relaxation function use the expense representation $g(b,X)$ as defined in (4.5):

$$\begin{aligned} \tilde{b}^{\mathrm{rnr}}_{\lambda}(\cdot) &:= \underset{b:\mathscr{X}\to\mathbb{R}^+}{\operatorname{argmax}} \quad \mathbb{E}[\hat{V}(X)\theta(X)\Phi(\frac{b(X)-\hat{w}(X)}{\sigma(X)})] - \lambda\mathbb{E}[g(b(X),X)-B] \\ &= \underset{b:\mathscr{X}\to\mathbb{R}^+}{\operatorname{argmax}} \quad \mathbb{E}[\mathcal{G}_{\lambda}(b(X),X)], \end{aligned}$$

where

$$\mathcal{G}_{\lambda}(b,X) := \hat{V}(X)\theta(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) - \lambda(g(b,X)-B).$$

The optimal bid price b^{rnr} can be approximated using $\tilde{b}^{\mathrm{rnr}}_{\lambda^*}$ with λ^* as the smallest value of $\lambda \geq 0$ such that:

$$\mathbb{E}[g(\tilde{b}^{\mathrm{rnr}}_{\lambda^*}(X),X)] \leq B.$$

We next provide a closed-form solution for $\tilde{b}^{\mathrm{rnr}}_{\lambda}$ in the form of Lemma 2.

Lemma 2. *For any $\lambda \geq 0$, a maximizer of the Lagrangian relaxation takes the form:*

$$\tilde{b}_{\lambda}^{rnr}(X) := \arg \max_{b \in \{0,\, \frac{\hat{V}(X)\theta(X)}{\lambda},\, \infty\}} \mathcal{G}_{\lambda}(b,X),\ \forall X \in \mathscr{X}\,.$$

Proof. Again, we start by exploiting the interchangeability property of expected value, which implies that the optimal bid price for the Lagrangian relaxation can be obtained as the price, for each X, that maximizes the Lagrangian relaxation function $\mathcal{G}$:

$$\begin{aligned}\tilde{b}_{\lambda^*}^{\mathrm{rnr}}(X) = \operatorname*{argmax}_{b \in \mathbb{R}^+} \mathcal{G}_{\lambda}(b,X) &:= \hat{V}\theta\Phi(\frac{b-\hat{w}}{\sigma}) - \lambda\left(\hat{w}\Phi(\frac{b-\hat{w}}{\sigma}) - \sigma\phi(\frac{b-\hat{w}}{\sigma}) - B\right) \\ &= (\hat{V}\theta - \lambda\hat{w})\Phi(\frac{b-\hat{w}}{\sigma}) + \lambda\sigma\phi(\frac{b-\hat{w}}{\sigma}) + \lambda B\,, \end{aligned} \tag{4.9}$$

where we dropped the relation to X for simplicity of presentation.

Since $\mathcal{G}_{\lambda}(b,X)$ is twice differentiable with respect to b, the maximizer for b is either $0,\infty$ or at a value where the derivative is 0. For the latter case, we get that:

$$\frac{d\mathcal{G}_{\lambda}(b,X)}{db} = 0 \Leftrightarrow \frac{\lambda\hat{w} - \hat{V}\theta}{\sigma}\phi(\frac{b-\hat{w}}{\sigma}) + \lambda\frac{b-\hat{w}}{\sigma}\phi(\frac{b-\hat{w}}{\sigma}) = 0\,.$$

Hence, we can conclude that the value of b where $\mathcal{G}_{\lambda}(b,X)$ has a derivative of zero is:

$$b = \frac{\hat{V}(X)\theta(X)}{\lambda}.$$

□

Chapter 5

Risk-averse Approach

The objective of risk-averse bid optimization model is to maximize the return of bidding while controlling the risk of violating the budget constraint for a given time.

Similarly to the case of the risk-neutral models, we assume the value of customer V, the fact that the ad is clicked C and the winning price W are mutually conditionally independent given X (see Assumption 1). We also consider a similar per opportunity budget B.

The distinguish point of the risk-averse approach is that we take the risk of going over budget into consideration. In this chapter, we develop the risk-averse bid optimization models for profit maximizing b^{rap}, and revenue maximizing b^{rar}, by applying the utility function that considers the risk on the budget constraint.

Variables	**Description**
$b^{\text{rap}}(\cdot)$	risk-averse bid policy that maximizes profit problem
$b^{\text{rar}}(\cdot)$	risk-averse bid policy that maximizes revenue problem

Table 5.1: Risk-averse models

5.1 The Risk-averse Budget Constraint

We introduce the exponential utility function to model risk aversion in the budget constraint, where $u_\alpha(y) := -\exp(\alpha y)$ is a concave utility function, where risk aversion is controlled using the hyperparamter α.

According to (Rudloff and Wunderlich, 2008), this risk-averse approach reduces to the risk-neutral approach (4.1) when $\alpha \to 0$ and can be considered as imposing an upper bound of B on the entropic risk measure ρ of the per opportunity's expense:

$$\rho(\frac{1}{M}\sum_{i=1}^{M} Ws(b(X),W)) \leq B,$$

where the entropic risk measure takes the form of

$$\rho(X) := \frac{1}{\alpha}\ln(\mathbb{E}[\exp(\alpha X)]). \tag{5.1}$$

The constraint in the batch of M opportunities can be represented by:

$$\mathbb{E}[u_\alpha(\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i),W_i))] \geq u_\alpha(B). \tag{5.2}$$

Then we can simplify the constraint based on the fact that the winning price W_i and features X_i are independently and identically distributed variables:

$$\begin{aligned}
\mathbb{E}[u_\alpha(\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i),W_i))] &= -\mathbb{E}[e^{\alpha\left(\frac{1}{M}\sum W_i s(b(X_i),W_i)\right)}] \\
&= -\mathbb{E}[\prod_{i=1}^{M} e^{(\alpha/M)W_i s(b(X_i),W_i)}] \\
&= -\prod_{i=1}^{M}\mathbb{E}[e^{(\alpha/M)W_i s(b(X_i),W_i)}] \\
&= -\mathbb{E}[e^{\alpha/MWs(b(X),W)}]^M \\
&= \mathbb{E}[u_{\alpha/M}(Ws(b(X),W)]^M,
\end{aligned}$$

where the third equality is derived based on the independence assumption and the fourth equality is derived based on the common distribution.

It means that constraint (5.2) can be rewritten as:

$$\mathbb{E}[u_{\alpha/M}(Ws(b(X),W))] \geq u_{\alpha}(B)^{1/M} = u_{\alpha/M}(B).$$

In terms of simplification, we make

$$\alpha' = \frac{\alpha}{M}.$$

Therefore, the final constraint can be written as expected instantaneous format:

$$\mathbb{E}[u_{\alpha'}(Ws(b(X),W))] \geq u_{\alpha'}(B). \tag{5.3}$$

5.2 Reducing the risk-averse budget constraint

We further simplify the constraint to obtain a closed-form representation. In doing so, we start by dividing both side of constraint (5.3) by $-u_{\alpha'}(B) > 0$, given that $B \geq 0$ implies that $-u_{\alpha'}(B) > 0$, in order to normalize this constraint. This leads us to

$$-\mathbb{E}[u_{\alpha'}(Ws(b(X),W))/u_{\alpha'}(B)] \geq -1.$$

We then exploit the Gaussian nature of W when X is known to obtain the following reduction:

$$\begin{aligned}
&-\mathbb{E}[u_{\alpha'}(Ws(b(X),W))/u_{\alpha'}(B)|X] \\
&= -e^{-\alpha' B}\left(\int_{w\leq b(X)} \exp(\alpha' w) f_{W|X}(w)dw + \int_{w>b(X)} \exp(\alpha' \cdot 0) f_{W|X}(w)dw\right) \\
&= -e^{-\alpha' B}\left(\int_{w\leq b} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(w-\hat{w})^2}{2\sigma^2} + \alpha' w)dw + 1 - \Phi(\frac{b-\hat{w}}{\sigma})\right) \\
&= -e^{-\alpha' B}\left(\int_{w\leq b} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{[w-(\hat{w}+\alpha'\sigma^2)]^2}{2\sigma^2} + \frac{(\alpha')^2\sigma^2}{2} + \alpha'\hat{w})dw + 1 - \Phi(\frac{b-\hat{w}}{\sigma})\right) \\
&= -e^{\gamma_1}\int_{w\leq b} \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(w-\hat{w}-\alpha'\sigma^2)^2}{2\sigma^2})dw - e^{\gamma_2} + e^{\gamma_2}\Phi(\frac{b-\hat{w}}{\sigma}) \\
&= -e^{\gamma_1(X)}\Phi(\frac{b(X)-\hat{w}(X)-\alpha'\sigma(X)^2}{\sigma(X)}) - e^{\gamma_2} + e^{\gamma_2}\Phi(\frac{b(X)-\hat{w}(X)}{\sigma(X)})
\end{aligned} \tag{5.4}$$

where we temporarily dropped the relation to X for simplicity and define:

$$\gamma_1(X) := (1/2)(\alpha')^2\sigma(X)^2 + \alpha'\hat{w}(X) - \alpha' B \qquad \& \qquad \gamma_2 := -\alpha' B.$$

Overall, this allows us to capture the risk-averse budget constraint using:

$$\mathbb{E}[h(b(X),X)] \geq -1,$$

where

$$h(b,X) := -e^{\gamma_1(X)}\Phi(\frac{b-\hat{w}(X)-\alpha'\sigma(X)^2}{\sigma(X)}) - e^{\gamma_2} + e^{\gamma_2}\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) \tag{5.5}$$

5.3 Risk-averse Profit Maximization Model

The objective of risk-averse bid optimization b^{rap} model is maximizing the profit of bidding while controlling the risk of violating the budget constraint for a given batch of opportunities. Since we do not consider the risk on profit, the objective function is the same with the model b^{rnp}. We consider the risk of spending over the budget by using the risk-averse budget constraint (5.2):

$$\begin{aligned} b^{\text{rap}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} V_i C_i s(b(X_i),W_i)] - \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i),W_i)] \\ \text{subject to} \quad & \mathbb{E}[u_\alpha(\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i),W_i))] \geq u_\alpha(B). \end{aligned} \tag{5.6}$$

Following the reductions presented in Sections 5.1 and 5.2, we can reduce the problem to the following risk-averse expected instantaneous profit maximization problem:

$$\begin{aligned} b^{\text{rap}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[VCs(b(X),W)] - \mathbb{E}[Ws(b(X),W)] \\ \text{subject to} \quad & \mathbb{E}[h(b(X),X)] \geq -1. \end{aligned} \tag{5.7}$$

In an attempt to solve this problem, one can again introduce the Lagrangian coefficient $\lambda \geq 0$ to obtain a relaxation of this risk-averse profit maximizing problem:

$$\begin{aligned} \tilde{b}_\lambda^{\text{rap}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[VCs(b(X),W)] - \mathbb{E}[Ws(b(X),W)] - \lambda(-1-\mathbb{E}[h(b(X),X)]) \\ = \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[\mathcal{G}_\lambda(b(X),X)], \end{aligned}$$

where

$$\mathcal{G}_\lambda(b,X) := \mathbb{E}[VCs(b(X),W)|X] - \mathbb{E}[Ws(b(X),W)|X] - \lambda(-1-\mathbb{E}[h(b(X),X)|X]) \tag{5.8}$$

$$= \hat{V}(X)\theta(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) - g(b,X) + \lambda(1-h(b,X)),$$

following our definitions of $g(b,X)$ and $h(b,X)$ in (4.5) and (5.5) respectively.

The optimal bid price b^{rap} can be approximated using $\tilde{b}^{\text{rap}}_{\lambda^*}$ with λ^* as the smallest value of $\lambda \geq 0$ such that:

$$\mathbb{E}[h(\tilde{b}^{\text{rap}}_{\lambda^*}(X),X)] \geq -1.$$

We next provide a closed-form solution for $\tilde{b}^{\text{rap}}_{\lambda}$ in the form of Lemma 3.

Lemma 3. *For any $\lambda \geq 0$, a maximizer of the Lagrangian relaxation takes the form:*

$$\tilde{b}^{rap}_{\lambda}(X) := \arg \max_{b \in \{0, -\frac{\mathbf{W}(\lambda\alpha' e^{(\hat{V}(X)\theta(X)+\lambda e^{\gamma_2}-B)\alpha'})}{\alpha'} + \hat{V}(X)\theta(X)+\lambda e^{\gamma_2}, \infty\}} \mathcal{G}_\lambda(b,X),\ \forall X \in \mathscr{X},$$

where $\mathbf{W}$ *is the Lambert W-function, i.e. the inverse function of $f(x) := xe^x$.*

Proof. Similar with the risk-neutral models, we also exploit the interchangeability property of expected value, which implies that the optimal bid price for the Lagrangian relaxation can be obtained as the price, for each X, that maximizes the Lagrangian relaxation function $\mathcal{G}_\lambda(b,X)$:

$$\begin{aligned}
\tilde{b}^{\text{rap}}_{\lambda^*}(X) = \underset{b\in\mathbb{R}^+}{\operatorname{argmax}}\, \mathcal{G}_\lambda(b,X) := {} & (\hat{V}\theta + \lambda e^{\gamma_2})\Phi(\frac{b-\hat{w}}{\sigma}) - \lambda e^{\gamma_1}\Phi(\frac{b-\hat{w}-\alpha'\sigma^2}{\sigma}) \\
& + (1-e^{\gamma_2})\lambda - [\hat{w}\Phi(\frac{b-\hat{w}}{\sigma}) - \sigma\phi(\frac{b-\hat{w}}{\sigma})] \\
& = (\hat{V}\theta + \lambda e^{\gamma_2} - \hat{w})\Phi(\frac{b-\hat{w}}{\sigma}) - \lambda e^{\gamma_1}\Phi(\frac{b-\hat{w}-\alpha'\sigma^2}{\sigma}) \\
& + (1-e^{\gamma_2})\lambda + \sigma\phi(\frac{b-\hat{w}}{\sigma}) \\
& = c_1\Phi(\frac{b-\hat{w}}{\sigma}) - c_2\Phi(\frac{b-c_3}{\sigma}) + (1-e^{\gamma_2})\lambda + \sigma\phi(\frac{b-\hat{w}}{\sigma}),
\end{aligned} \tag{5.9}$$

where we drop the dependence of $\hat{V}$, θ, $\hat{w}$, and σ to X for convenience and where we use c_1, c_2, and c_3 to refer to $c_1(X) := \hat{V}(X)\theta(X) + \lambda e^{\gamma_2} - \hat{w}(X)$, $c_2(X) := \lambda e^{\gamma_1(X)}$, and $c_3(X) := \hat{w}(X) + \alpha'\sigma(X)^2$.

Since $\mathcal{G}_\lambda(b,X)$ is twice differentiable with respect to b, the maximizer for b is either $0, \infty$ or at a value where the derivative is 0. For the latter case, we get that:

$$
\begin{aligned}
&\frac{\mathrm{d}\mathcal{G}_\lambda(b,X)}{\mathrm{d}b} = 0\\
&\Leftrightarrow -\frac{c_1}{\sigma}\phi(\frac{b-\hat{w}}{\sigma}) + \frac{c_2}{\sigma}\phi(\frac{b-c_3}{\sigma}) + \frac{b-\hat{w}}{\sigma}\phi(\frac{b-\hat{w}}{\sigma}) = 0\\
&\Leftrightarrow (c_1 - b + \hat{w})\phi(\frac{b-\hat{w}}{\sigma}) = c_2\phi(\frac{b-c_3}{\sigma})\\
&\Leftrightarrow \ln((\hat{V}\theta + \lambda e^{\gamma_2} - b))\phi(\frac{b-\hat{w}}{\sigma})) = \ln(c_2\phi(\frac{b-c_3}{\sigma}))\\
&\Leftrightarrow \ln((\hat{V}\theta + \lambda e^{\gamma_2} - b)/\sqrt{2\pi}) - \frac{(b-\hat{w})^2}{2\sigma^2} = \ln(c_2/\sqrt{2\pi}) - \frac{(b-c_3)^2}{2\sigma^2}\\
&\Leftrightarrow 2\sigma^2\ln(\hat{V}\theta + \lambda e^{\gamma_2} - b) - (b-\hat{w})^2 = 2\sigma^2[\ln(\lambda) + \gamma_1] - c_3^2 + 2bc_3 - b^2\\
&\Leftrightarrow 2\sigma^2\ln(\hat{V}\theta + \lambda e^{\gamma_2} - b) - (b-\hat{w})^2 = 2\sigma^2[\ln(\lambda) + \gamma_1] - (\hat{w} + \alpha'\sigma^2)^2 + 2b(\hat{w} + \alpha'\sigma) - b^2\\
&\Leftrightarrow 2\ln(\hat{V}\theta + \lambda e^{\gamma_2} - b) = 2[\ln(\lambda) + 0.5(\alpha')^2\sigma^2 + \alpha'\hat{w} - \alpha' B] - (\alpha')^2\sigma^2 - 2w\alpha' + 2b\alpha'\\
&\Leftrightarrow \ln(\frac{\hat{V}\theta + \lambda e^{\gamma_2} - b}{\lambda}) = \alpha'(b-B)\\
&\Leftrightarrow \ln(\frac{\hat{V}\theta + \lambda e^{\gamma_2} - b}{\lambda}) = (V\theta + \lambda e^{\gamma_2} - B)\alpha' - (\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha'\\
&\Leftrightarrow \frac{\hat{V}\theta + \lambda e^{\gamma_2} - b}{\lambda} e^{(V\theta + \lambda e^{\gamma_2} - b)\alpha'} = e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'}\\
&\Leftrightarrow (\hat{V}\theta + \lambda e^{\gamma_2} - b)e^{(V\theta + \lambda e^{\gamma_2} - b)\alpha'} = \lambda e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'}\\
&\Leftrightarrow (\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha' e^{(V\theta + \lambda e^{\gamma_2} - b)\alpha'} = \lambda\alpha' e^{(V\theta + \lambda e^{\gamma_2} - B)\alpha'}\\
&\Leftrightarrow (\hat{V}\theta + \lambda e^{\gamma_2} - b)\alpha' = \mathbf{W}(\lambda\alpha' e^{(V\theta + \lambda e^{\gamma_2} - B)\alpha'})\\
&\Leftrightarrow b = -\frac{\mathbf{W}(\lambda\alpha' e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'})}{\alpha'} + \hat{V}\theta + \lambda e^{\gamma_2},
\end{aligned}
$$

where the $\mathbf{W}$ is the Lambert-$\mathbf{W}$ function. Hence, we can conclude that the value of b where $\mathcal{G}_\lambda(b,X)$ has a derivative of zero is:

$$b = -\frac{\mathbf{W}(\lambda\alpha' e^{(\hat{V}\theta + \lambda e^{\gamma_2} - B)\alpha'})}{\alpha'} + \hat{V}\theta + \lambda e^{\gamma_2}.$$

$\square$

5.4 Risk-averse Revenue Maximization Model

This b^{rap} (5.7) model can be easily modified to a revenue maximizing model while respecting the risk-averse budget constraint, when advertiser is looking for maximizing the expected click revenue while being risk-averse with respect to violating the budget constraint for a given period of time, captured by the batch size M. This is modeled using:

$$\begin{aligned} b^{\text{rar}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[\frac{1}{M}\sum_{i=1}^{M} V_i C_i s(b(X_i), W_i)] \\ \text{subject to} \quad & \mathbb{E}[u_\alpha(\frac{1}{M}\sum_{i=1}^{M} W_i s(b(X_i), W_i))] \geq u_\alpha(B). \end{aligned}$$

Based on the derivation of the batch utility constraint in the earlier Section (5.3), it is again equivalent to model the expected instantaneous value per opportunity, where we denote $\alpha' := \frac{\alpha}{M}$:

$$\begin{aligned} b^{\text{rar}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[VCs(b(X), W)] \\ \text{subject to} \quad & \mathbb{E}[u_{\alpha'}(Ws(b(X), W))] \geq u_{\alpha'}(B). \end{aligned} \tag{5.10}$$

Based on the normalized utility equation (5.4), we can reformulate the risk-averse revenue model b^{rar} as below:

$$\begin{aligned} b^{\text{rar}}(\cdot) := \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad & \mathbb{E}[VCs(b(X), W)] \\ \text{subject to} \quad & \mathbb{E}[h(b(X), X)] \geq -1. \end{aligned}$$

As was done with b^{rap}, we introduce the Lagrangian coefficient $\lambda \geq 0$ to obtain a relaxation of this risk-averse revenue maximizing problem:

$$\begin{aligned} \tilde{b}_\lambda^{\text{rar}} &:= \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad \mathbb{E}[VCs(b(X), W)] - \lambda(-1 - \mathbb{E}[h(b(X), X)]) \\ &= \underset{b:\mathscr{X}\to\mathbb{R}^+}{\text{argmax}} \quad \mathbb{E}[\mathcal{G}_\lambda(b(X), X)], \end{aligned}$$

where

$$\mathcal{G}_\lambda(b,X) := \mathbb{E}[VCs(b(X),W)|X] - \lambda(-1-\mathbb{E}[h(b(X),X)|X]) \tag{5.11}$$
$$= \hat{V}(X)\theta(X)\Phi(\frac{b-\hat{w}(X)}{\sigma(X)}) + \lambda(1+h(b,X)).$$

The optimal bid price b^{rar} can be approximated using $\tilde{b}^{\text{rar}}_{\lambda^*}$ with λ^* as the smallest value of $\lambda \geq 0$ such that:

$$\mathbb{E}[h(\tilde{b}^{\text{rar}}_{\lambda^*}(X),X)] \geq -1.$$

We next provide a closed-form solution for $\tilde{b}^{\text{rar}}_{\lambda}$ in the form of Lemma 4.

Lemma 4. *For any $\lambda \geq 0$, a maximizer of the Lagrangian relaxation takes the form:*

$$\tilde{b}^{rar}_{\lambda}(X) := \arg \max_{b \in \{0,\ \frac{1}{\alpha'}\ln(\frac{\hat{V}(X)\theta(X)}{\lambda}+\exp(-\alpha' B))+B,\ \infty\}} \mathcal{G}_\lambda(b,X),\ \forall X \in \mathscr{X}.$$

Proof. According to interchangeability property of expected value, the optimal bid price for the Lagrangian relaxation can be obtained as the price, for each X, that maximizes the Lagrangian relaxation function $\mathcal{G}_\lambda(b,X)$, namely:

$$\tilde{b}^{\text{rar}}_{\lambda^*}(X) = \underset{b\in\mathbb{R}^+}{\operatorname{argmax}}\, \mathcal{G}_\lambda(b,X) := (\hat{V}\theta + \lambda e^{\gamma_2})\Phi(\frac{b-\hat{w}}{\sigma}) - \lambda e^{\gamma_1}\Phi(\frac{b-\hat{w}-\alpha'\sigma^2}{\sigma}) + (1-e^{\gamma_2})\lambda$$
$$= c_1\Phi(\frac{b-\hat{w}}{\sigma}) - c_2\Phi(\frac{b-c_3}{\sigma}) + (1-e^{\gamma_2})\lambda, \tag{5.12}$$

where we dropped the relation to X for simplicity of presentation, and where $c_1(X) = \hat{V}(X)\theta(X) + \lambda e^{\gamma_2}$, $c_2(X) = \lambda e^{\gamma_1(X)}$, and $c_3(X) := \hat{w}(X) + \alpha'\sigma(X)^2$.

Since $\mathcal{G}_\lambda(b,X)$ is twice differentiable with respect to b, the maximizer for b is either

$0, \infty$ or at a value where the derivative is 0. For the latter case, we get that:

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{G}_\lambda(b,X)}{\mathrm{d}b} = 0 &\Leftrightarrow -\frac{c_1}{\sigma}\phi(\frac{b-\hat{w}}{\sigma}) + \frac{c_2}{\sigma}\phi(\frac{b-c_3}{\sigma}) = 0 \\
&\Leftrightarrow c_1\phi(\frac{b-\hat{w}}{\sigma}) = c_2\phi(\frac{b-c_3}{\sigma}) \\
&\Leftrightarrow \ln(c_1\phi(\frac{b-\hat{w}}{\sigma})) = \ln(c_2\phi(\frac{b-c_3}{\sigma})) \\
&\Leftrightarrow \ln(c_1/\sqrt{2\pi}) - \frac{(b-\hat{w})^2}{2\sigma^2} = \ln(c_2/\sqrt{2\pi}) - \frac{(b-c_3)^2}{2\sigma^2} \\
&\Leftrightarrow (b-c_3)^2 - (b-\hat{w})^2 = 2\sigma^2\ln(c_2/c_1) \\
&\Leftrightarrow 2(\hat{w}-c_3)b = 2\sigma^2\ln(c_2/c_1) + \hat{w}^2 - c_3^2 \\
&\Leftrightarrow b = (1/\alpha')\ln(c_1/c_2) + \frac{\alpha'\sigma^2}{2} + \hat{w} \\
&\Leftrightarrow b = (1/\alpha')\ln(\hat{V}\theta/\lambda + e^{\gamma_2}) + B\ .
\end{aligned}$$

Hence, we can conclude that the value of b where $\mathcal{G}_\lambda(b,X)$ has a derivative of zero is:

$$b = (1/\alpha')\ln(\hat{V}\theta/\lambda + e^{\gamma_2}) + B.$$

$\square$

Chapter 6

Experiment Design

In the previous chapters, we have modelled the bid optimization problems that maximize the expected profit or revenue while respecting the budget constraint with a risk-neutral approach and the risk-averse approach. We obtained closed-form expressions of approximately optimal bidding policies under each circumstances. In this Chapter, we design the experiments to test the effectiveness of our bidding policies using the real-life iPinyou Dataset. Besides, we present empirical experiments assumptions, implementation of the batch evaluation approach, and parameters tuning algorithms.

6.1 iPinyou Dataset

The iPinYou dataset was released by iPinYou Information Technologies Co., Ltd (iPinYou)[1] for the global RTB algorithm competition in 2013. The iPinyou Dataset includes logs of ad biddings, impressions, clicks, and final conversions, and data are collected from different industries of three seasons. Since our study requires the click, winning price information for certain advertisers, we use the second and third seasons of data where *Paying Price* and *Advertiser ID* are available. Even though the conversion may seem more directly related to the advertiser's goal, it is defined by each advertiser and different between companies. For example, for an e-commerce company, the conversion has hap-

[1] `http://www.ipinyou.com`

pened when the user makes purchases but for a software company, they are looking for the number of downloads. Therefore, we use the common target click C as an objective to build our model.

Zhang et al. (2015) have analyzed the distributions of data from different industries in this dataset, which are very different from every industry and relatively stable inside of one industry. Since the bidding models that we designed are based on the assumptions that the characteristics of each opportunity is independent and identically distributed, we implement the experiments separately for the different industries.

In this thesis, we use one industry data: the Chinese vertical e-commerce industry collected from *Advertiser ID* 1458, to demonstrate our experiment results.

6.1.1 Data format

The log data format can be seen from the table 6.1 where each row of this table represents a column/variable in the dataset (column with * means that the data in the column is hashed or modified before the log is released.) and a detailed description for each variable can be found in (Liao et al., 2014). Here, we only generally introduce the relevant information.

For each observation in the dataset, the auction and ad slot features (all the columns except the columns # 3, 20, and 21) are sent to the bidding engine, where the advertiser propose the according to bidding price (column 21). If their bidding price is the highest among the competitors, they win the auction for this ad slot and pay the second-highest bidding price among the competition, which is the *Paying Price* (column 20). Once the advertiser wins, their ad will be presented to the user captured by the observation at the specific ad slot that was won. The *Log Type* (column 3) indicates the user response(impression:1, click:2, conversion:3), which determines the performance that is achieved for this observation if the bid is won.

Column #	Column Name	Example
*1	Bid ID	01530000008a77e7ac18823f5a4f5121
2	Timestamp	20130218134701883
3	Log Type	1
*4	iPinyou ID	35605620124122340227135
5	User-Agent	Mozilla/5.0 (compatible; MSIE 9.0; Windows NT 6.1; WOW64; Trident/5.0)
*6	IP	118.81.189.*
7	Region ID	15
8	City ID	16
9	ID Exchange	2
*10	Domain	e80f4ec7f5bfbc9ca416a8c01cd1a049
*11	URL	hz55b000008e5a94ac18823d6f275121
12	Anonymous URL	null
13	Ad Slot ID	2147689_8764813
14	Ad Slot Width	300
15	Ad Slot Height	250
16	Ad Slot Visibility	SecondView
17	Ad Slot Format	Fixed
18	Ad Slot Floor Price	0
19	Creative ID	e39e178ffdf366606f8cab791ee56bcd
*20	Bidding Price	753
*21	Paying Price	15
*22	Landing Page URL	a8be178ffdf366606f8cab791ee56bcd
23	Advertiser ID	1458
*24	User Profile IDs	123,5678,3456

Table 6.1: Data Format

In our research, the auction and ad slot features are the observable features X (i.e. all the columns except the column # 3, 20, and 21). In particular, columns # 14, 15, 18 are numerical variables and all the others are categorical variables. Column # 21 and column # 3, i.e. the *Paying Price* and *Log Type*, remain unobserved before the bidding and are the dependent variables defined as W and C. Column # 20, *Bidding Price* is the decision variable that we actually want to optimize. Therefore, all the estimators and bidding policies that we designed are built on features X which are available to the advertisers before the auction to be applied to real-world circumstances.

6.1.2 Data split

The dataset is separated into **Train, Valid** and **Test** sets. The **Train** set contains 3,083,056 observations and are used to train estimators for $\theta, \hat{w}, \sigma$ and optimize the Lagrangian relaxation parameter λ. Both **Valid** and **Test** set contains 307,319 observations. We explore the influence of hyperparameters and select the risk controlling factor α in the **Valid** for the risk-averse models we proposed. The **Test** set is used to evaluate and compare different bidding policies with performance metrics defined in the Section 3.5.

6.1.3 Descriptive statistics of the dataset

Winning price

Figure 6.1 presents the winning price distribution based on the train set (the test set applies a similar distribution). More proportion of winning prices are distributed at the lower price range than in the higher price range, with 83.4% percentage distributed within the price of 100 and the mean of winning price W is 68.9 (all prices are in Chinese Yuan, CNY).

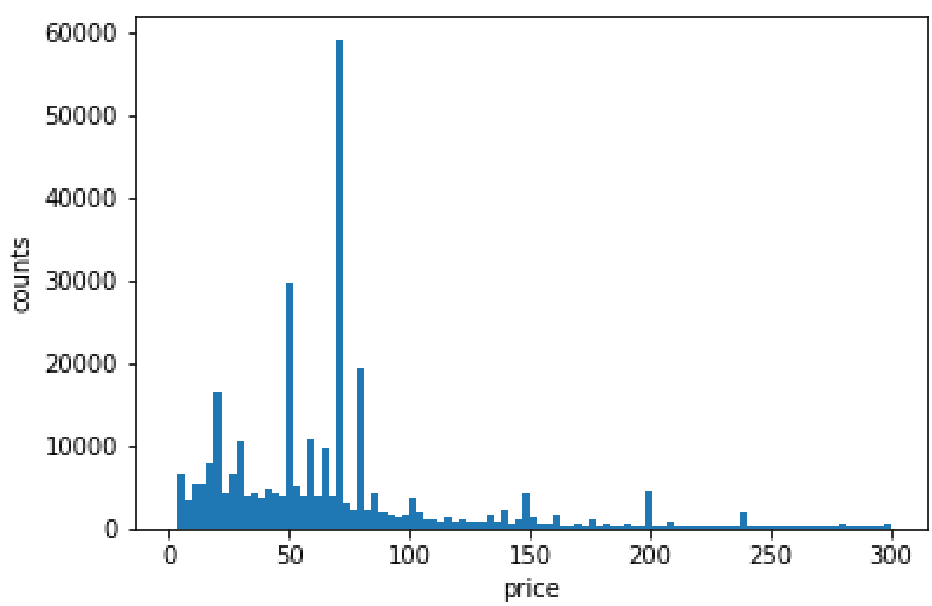

Figure 6.1: Winning price distribution

Clicks

Click events, i.e. $C = 1$, are very sparse in the dataset. Only 2,454 clicks in a total of 3,083,056 observations in the **Train** set, which is 0.080% of the dataset. Similar to the test set, there are only 0.078% of observations are clicked.

Considering the sparsity of the click event, we design the batch size relatively larger to make sure every batch contains the realized clicks. Since in the **Train** set, there are only 0.080% observations are clicked, we use the batch size $M = 10,000$ with an average of 8 clicks per batch. We assume that the distribution of the number of clicks in each batch follows the normal distribution which could be identified from the Figure 6.2.

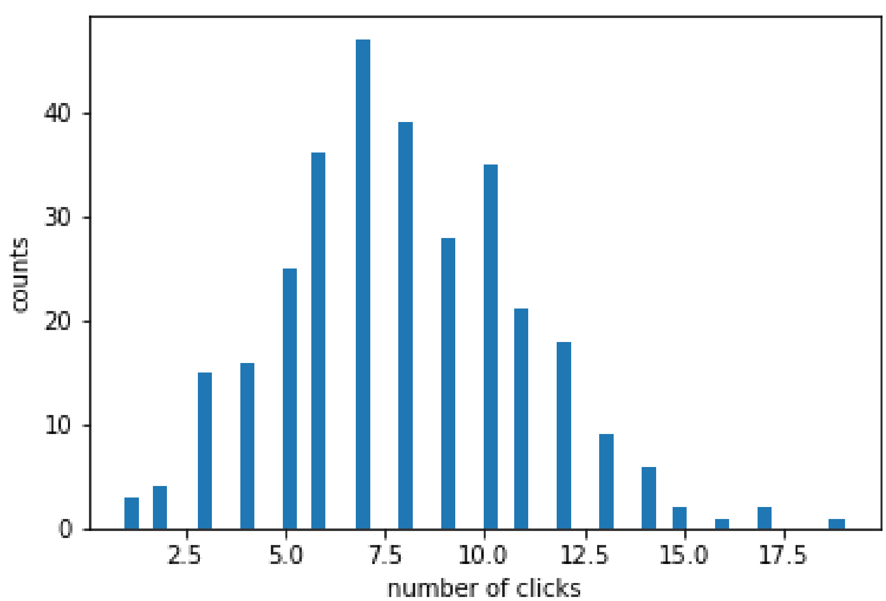

Figure 6.2: Click distribution

6.2 Experiment Assumptions

6.2.1 Budget setting

If the budget is large enough, a simple strategy of bidding at a very high price can win all the bidding opportunities, all the clicks and achieve the maximum revenue. However, it is more common that advertisers only have access to a limited budget and our models are developed for this situation. Therefore, in the experiments, we set our predetermined budget B lower than the historical total cost of 23,428,808 for 307,319 opportunities in the **Valid** set. To be comparable with the RMP approach, we run the experiments under 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, and 1 of the historical average cost, which is 1.19, 2.38, 4.76, 9.53, 19.06, 38.12, 76.24, respectively.

6.2.2 Value per click

Since the distribution of valuation is not available we could not model value depending on features X. We assume the value of a click is equivalent to the empirical value per click v:

$$\hat{V}(X) = v = \frac{\sum_{j=0}^{N} W_j}{\sum_{j=0}^{N} C_j}.$$

However, our approach could easily integrate an estimator of $\hat{V}(X)$ if the right type of data was available. This could generate more efficient bidding price with consideration of different value for different observations.

6.2.3 Early stop frequency

Bidding activity stops when the advertiser or their agents use up the budget. In other words, when the advertisers or their agents try to bid on ad opportunities, they might have to stop bidding before seeing all the ad opportunities for a given time because of running out of budget. It is unwanted because there might be more valuable web traffic later on. Therefore, it is considered as a risk measure, and we define it as Early Stopping Frequency in our setting:

$$\text{Early Stop Frequency} = \mathbb{P}(\frac{\text{Batch expense}}{M} > B).$$

Empirical early stop frequency

In the experiment, we simulate the situation of early stop as it happens when the algorithm has an insufficient fund for the next bidding opportunity. We use the empirical early stop event e to calculate the empirical early stop frequency on the samples of **Valid** set to search the risk hyperparameter α and use samples of **Test** set to present the model performance. To be more specific, the probability $\mathbb{P}$ is estimated using observations among the sample batches:

$$e := \text{the number of early stop events},$$

$$\text{Empirical early stop frequency} := \frac{e}{\text{the number of sample batches}},$$

where the number of sample batches of a dataset that contains N observations and uses batch size of M can be obtained by $\lfloor \frac{N}{M} \rfloor$.

Batch performances when early stop happens

When early stop happens in the middle of a batch of opportunities, the algorithm will stop bidding for any other left opportunities in the batch.

Mathematically, when the early stop happens, for a given batch of M opportunities $\{(X_i, W_i, C_i, V_i)\}_{i=1}^{M}$:

$$\eta := \min\{\eta = 1, \ldots, M | (1/M) \sum_{i=1}^{\eta} W_i s(b(X_i), W_i) > B\}$$

Since the batch budget BM has been used up, the decision makers will stop bidding by simply submitting the $\varepsilon' \geq 0$ which is a small value that left after paying all the expenses before the η-th opportunity in the batch, as the bidding price for the left opportunities in the batch:

$$\tilde{b}_i(X) := \begin{cases} b(X_i) & \text{if } i < \eta \\ \varepsilon' & \text{otherwise} \end{cases}$$

where $\varepsilon' = BM - \sum_{i=1}^{\eta-1} W_i s(b(X_i), W_i)$.

In the experiments, once $\varepsilon' < b(X_\eta)$, it means the budget left is insufficient for bidding and we define the early stop event happens. The algorithm will continue bidding for the following opportunities at the price of ε' until it is spent or sees the whole batch of M opportunities.

As a result, with a small bid price ε', the conditional winning probability $s(\tilde{b}_i(X_i), W_i) \approx 0$ for the left opportunities unless there is no other bidder in auction and the ad exchange platform accept the lowest payment as 0. So we have the empirical batch revenue as:

$$\begin{aligned}
\text{Early Stop Batch Revenue} &= \sum_{i=1}^{M} V_i C_i s(\tilde{b}_i(X_i), W_i) \\
&= \sum_{i=1}^{\eta-1} V_i C_i s(b(X_i), W_i) + \sum_{i=\eta}^{M} V_i C_i s(0, W_i) \\
&\approx \sum_{i=1}^{\eta-1} V_i C_i s(b(X_i), W_i)
\end{aligned}$$

Similarly, the empirical batch performance for expense, profit, clicks, impression rate with early stop can be defined as follows:

$$\begin{aligned}
\text{Early Stop Batch expense} &= \sum_{i=1}^{M} W_i C_i s(\tilde{b}_i(X_i), W_i) \\
&\approx \sum_{i=1}^{\eta-1} W_i s(b(X_i), W_i)
\end{aligned}$$

The batch profit is still the difference between batch revenue and batch expense.

$$\begin{aligned}
\text{Early Stop Batch number of clicks} &= \sum_{i=1}^{M} C_i s(\tilde{b}_i(X_i), W_i) \\
&\approx \sum_{i=1}^{\eta-1} C_i s(b(X_i), W_i)
\end{aligned}$$

$$\text{Early Stop Batch impression rate} = \frac{\sum_{i=1}^{M} s(\tilde{b}_i(X_i), W_i)}{M} \approx \frac{\sum_{i=1}^{\eta-1} s(b(X_i), W_i)}{M}$$

6.2.4 Implementation Detail

During the implementation, there is a numerical imprecision problem for the same term of the Lagrangian relaxation function (5.9) and (5.12) that model the risk-averse models RAP and RAR when $\frac{b-\hat{w}-\alpha\sigma^2}{\sigma}$ is small. This term is:

$$c_2\Phi(\frac{b-c_3}{\sigma}) = \lambda e^{\gamma_1}\Phi(\frac{b-\hat{w}-\alpha\sigma^2}{\sigma}) \tag{6.1}$$

The L'Hôpital's rule (Alaya, 2017) can be applied to improve the numerical precision of this term. Where $\ln(\Phi(y))$ is obtained, we use an approximation:

$$\ln(\Phi(y)) \approx -\ln(\sqrt{2\pi}) - y^2/2 - \ln(-y).$$

Applying this L'Hôpital's rule to the term (6.1), we define:

$$y := \frac{b-\hat{w}-\alpha\sigma^2}{\sigma},$$

therefore, the term (6.1) can be approximated by:

$$\begin{aligned}
\lambda e^{\gamma_1}\Phi(\frac{b-\hat{w}-\alpha\sigma^2}{\sigma}) &= \lambda e^{\gamma_1}\Phi(y) \\
&= \lambda \exp\left(\gamma_1 + \ln(\Phi(y))\right) \\
&\approx \lambda \exp\Big(\gamma_1 - \ln(\sqrt{2\pi}) - y^2/2 - \ln(-y)\Big).
\end{aligned}$$

6.3 Experiment Steps

We explore the bid optimization problem by applying the following steps:

6.3.1 Training estimators

The estimators are trained and cross-validated on the **Train** set. First, we obtain $\theta(X)$, i.e. the estimators for the click-through rate. After we obtain the estimator of winning price $\hat{w}(X)$, we then train the estimator of conditional standard deviation of winning price $\sigma(X)$ using the approach we described in the Section 3.2.

6.3.2 Searching for optimal parameters

To better model the effect of expense and risk, we introduce two parameters as introduced in the earlier Chapters and summarised in the following table 6.2:

Parameter	**Description**
λ	Lagrangian relaxation parameter
α	risk controlling parameter

Table 6.2: Hyperparameter for $b^{\text{rnp}}, b^{\text{rnr}}, b^{\text{rap}}$ and b^{rar}

These two parameters are developed for the bid optimization models to solve two different problems.

The parameter λ is needed for both risk-neutral and risk-aware approaches. To obtain the closed-form of bidding price, we introduce the Lagrange multiplier $\lambda \geq 0$. It represents the trade-off between maximizing the objective function, either revenue (4.2), or profit (4.4) depending on the model; and respect to the budget constraint that expense remain under the predetermined budget (4.3).

On the other hand, the parameter $\alpha \geq 0$ indicates the extent of being risk-averse with respect to the budget constraint, which applies in the risk-averse models. The α is larger when the decision-makers tend to be more risk-averse.

For both risk-neutral and risk-averse approaches, Lagrangian relaxation parameter λ is selected from the train set. As for risk-averse models, for each α, we have a optimal λ obtained from the **Train** set; then, the advertiser can select the α level using the performance observed on the **Valid** set. In our experiments, we will select the α that led to the best performance in terms of the Sharpe ratio of the early stop batch revenue or batch profit depending on the model. One can also use the α as a subjective hyperparameter selected by the decision makers depending on the their risk preference.

Explore the Lagrangian relaxation parameter λ

The λ is the Lagrange multiplier that we introduced into the model to integrate the constraint in the objective function of our model and make it amenable to optimization using first order conditions.

Specifically, in our research as defined in objective function (4.6) and (5.9), the λ can be seen as a trade-off parameter that indicates the strength of constraint. When the λ is close to zero, the algorithm will maximize the objective function, either revenue or profit and ignore the budget constraint, which leads to high bidding price and expenses; on the other hand, when the λ is large enough, the algorithm will only predict zero bidding price to respect the strong constraint of budget, which causes the unwanted result of zero bid, zero expense and zero value of the objective function. Therefore, the ideal λ should generate a non-zero bidding price that maximizes the objective function while respecting the constraint.

In addition, we use the expectation approach to explore the optimal λ. Under this expectation approach, the λ is conditional on winning price prediction $\hat{w}$ and the click-through rate prediction θ. No matter the objective function is either revenue or profit, the process of searching the optimal λ is the same. This is because either the objective function is maximizing the expected revenue or profit, the model is equivalent to capturing as many click opportunities as possible, which means spending all the predetermined budget to win every possible bid.

As the relationship between λ and expected expense is monotonous, when the λ increases, the budget constraint will be stronger and leads to the bidding price decreases, as in the Figure 6.3) and lower the expected expense as in the Figure 6.4. Both figures are produced with $B = 2.38$ in the train set as an example, and the relationship between λ and the expected expense or bidding price remains the same monotonous given other hyperparameters.

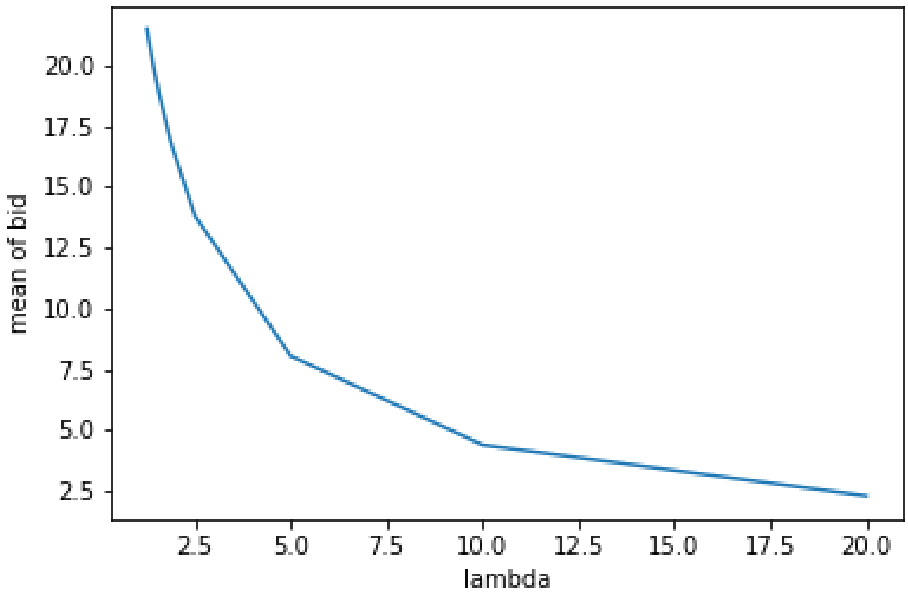

Figure 6.3: Bidding price under different λ

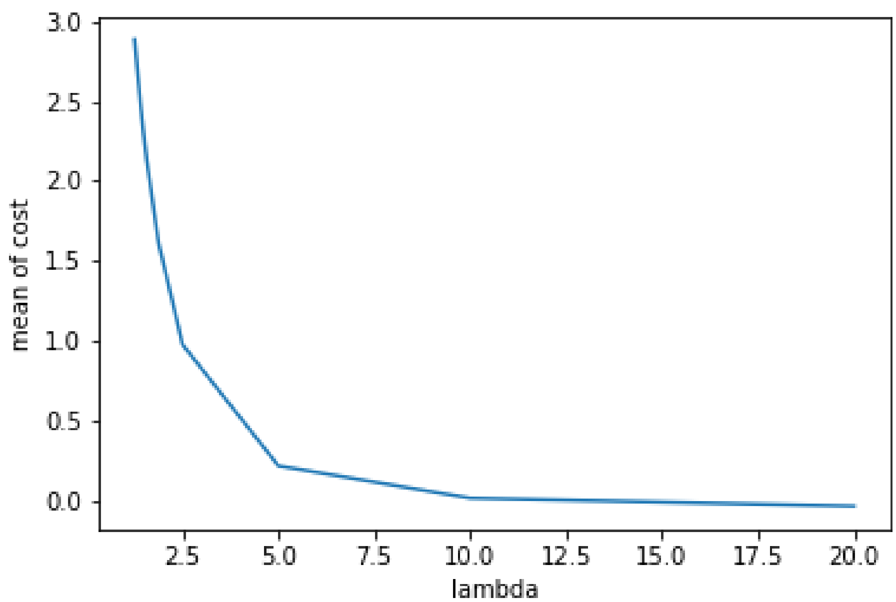

Figure 6.4: Expected expense under different λ

Therefore, we use the bisection search method to find the optimal λ that could spend the expected expense for risk-neutral models, or the certainty equivalent CE measured using the entropic risk measure (5.1) for risk-averse models, that is close to predetermined budget B:

Algorithme 1 : Bisection search for the best Lagrangian relaxation parameter λ

Input : tolerance $\varepsilon = 1e-2$, max $\lambda = 20$ (the max λ is larger when α is smaller), Predetermined average budget B, trained $\hat{w}(X), \theta(X), \sigma(X)$

Output : Best value of λ^*

Data : Training set

$\lambda_{min} \leftarrow 0$;

$\lambda_{max} \leftarrow 1$;

while $\lambda_{max} - \lambda_{min} > \varepsilon$ **do**

  $\lambda = (\lambda_{max} + \lambda_{min})/2$;

  Evaluate model in the train set, calculate the expected expenses $\mathbb{E}[\text{expense}]$

  (or the certainty equivalent CE for the risk-averse approaches)

  Then, compared with budget B;

  if $B > \mathbb{E}[expense](or\ CE\ for\ risk\text{-}averse\ approaches)$ **then**

    $\lambda_{max} = \lambda$;

  end

  if $B < \mathbb{E}[expense](or\ CE\ for\ risk\text{-}averse\ approaches)$ **then**

    $\lambda_{min} = \lambda$;

  end

end

$\lambda^* = \lambda_{max}$;

return λ^*;

Explore the risk controlling parameter α

The risk-neutral models without α, or $\alpha \to 0$, are simply controlling the expected expense without consideration of risk. The risk-averse models are using α as the risk controlling hyperparameter and controlling the entropic risk measure of expense. The extent of risk aversion gets higher when the α is increasing.

In reality, when the decision-maker is more risk-averse, they tend to reduce the expense by bidding at the lower price so that they are certain that they will not pay over the bidding price. Our model also demonstrates this monotonic relationship between risk aversion level and bidding price. The Figure 6.5 is produced using the RAP approach on the **Valid** set while the λ is obtained using the Algorithm 1 from the **Train** set. From there, we can clearly identify the bidding price is under more control when the α is higher.

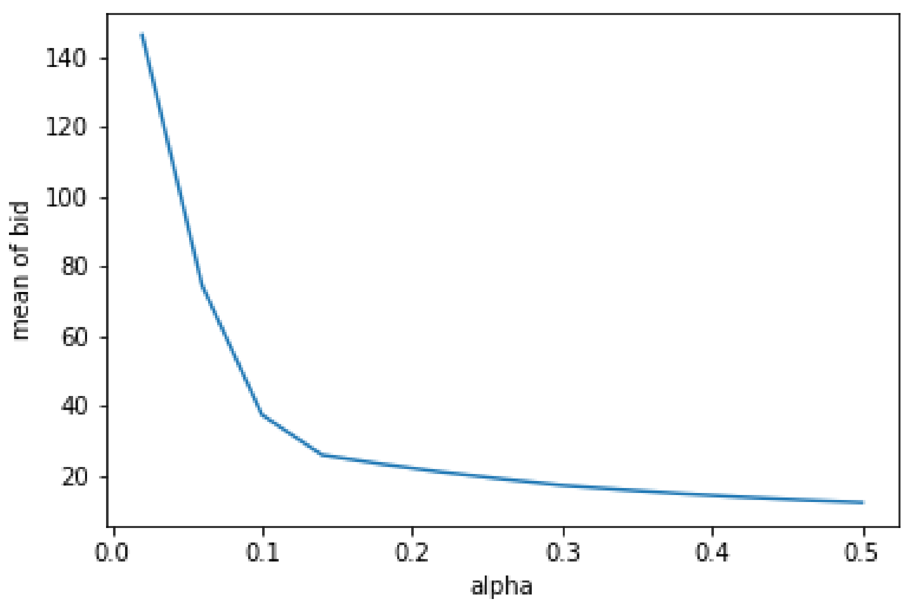

Figure 6.5: Bidding price under different α

For the RAP model, if we select the hyperparameter α only based on the Sharpe ratio, very often we will select a small alpha, which is very similar to directly use the RNP model that maximize the profit without consideration of risk of running out of the budget. The risk exposure is very similar. To better demonstrate our model's advantage in risk control on the budget constraint, we select the lowest α that more than certain probability of batches remain under the budget, and we define this probability as confidence level

p. In the experiment, we use the number of early stop events e to represent whether the algorithm goes over the budget or not.

In the experiment, the optimal α we use here is the one that produces the highest Sharpe ratio, using the objective value as the return, and remains under the budget for at least p =95% batches of **Valid** set. The detailed process for the RAP model is described in the Algorithm 2:

Algorithme 2 : Exhaustive search for the optimal α of RAP model

---

Input : Predetermined average budget B, a list of α, batch size M, a confidence level of p samples remain under the budget, trained $\hat{w}(X), \theta(X), \sigma(X)$, calibrated optimal λ_α^*
Output : Optimal α^*
Data : Valid set;
the **Valid** set contains N observations, $\lfloor \frac{N}{M} \rfloor$ batches

The best Sharpe ratio $s^* = 0$;
for *each* α **do**
    Initial empirical Early Stop frequency given α is $e_\alpha = 0$;
    for *each batch j in the **Valid** set* **do**
        Let objective function value for this batch obj_j be the calculated **Batch profit** ;
        if *batch budget run out* **then**
            $e_\alpha = e_\alpha + 1$;
        end
    end
    Mean of obj in batches given α: $\mu_\alpha = \overline{obj}$;
    Standard deviation of obj in batches given α: $\sigma_\alpha = \sigma_{obj}$;
    Sharpe ratio given α: $s_\alpha = \frac{\mu_\alpha}{\sigma_\alpha}$;
    if $e_\alpha < (1-p) * \lfloor \frac{N}{M} \rfloor$ *and* $s_\alpha > s^*$ **then**
        the optimal $s^* = s_\alpha$;
        the optimal $\alpha^* = \alpha$
    end
end
return α^*;

This Algorithm 2 can be easily adapted to the RAR model. When searching for optimal α of RAR model, the only difference compared to the Algorithm 2 is that we calculate the **Batch revenue** as the objective function value obj instead of **Batch profit**.

6.3.3 Model Evaluation and Comparison

The models' performance is evaluated using the **Test** set and the detailed metrics are described in the Section 3.5.

Among the bidding strategies in the latest literature, the RMP bidding strategy has a similar setting by modeling as one-stage problem, and similar goal which is to control the risk in the bidding while maximizing the profit. Our difference is that the RMP approach targets to control the risk of negative profit while our RAP approach is controlling the risk of violating the budget constraint. Therefore, we will also compare our bidding strategy with the RMP approach in the **Test** set.

In order to better compare the strategy itself, we tried to use the same setting for both approaches. All the comparisons we made using the same dataset, under the same budget constraint; we both use the empirical value per click v in the experiments to represent V_i and assume the valuation of click is constant, independent from the distribution of features X in the experiments.

Plus, when we replicate their approach on the dataset, we use a larger range of risk control and bid scale ratio hyperparameters to make sure the result is comparable and does not decline due to an insufficient search of hyperparameters. In addition, we apply the same criteria to select hyperparameters. Instead of using the profit obtained from the **Valid** set as RMP approach, we changed it to select the hyperparameters combinations that more than the confidence level of p, remain under the predetermined budget in every batch and gives the highest Sharpe ratio in the batches of **Valid** set. We also generate the RMP_neutral model by removing its risk factor (when $\alpha = 0$) to compare with our RNP model.

However, the RMP model considers the risk in profit coming from the uncertainty of clicks and estimates their mean and standard deviation by Bayesian regression on **Train** set, and W is a stochastic random variable independent given X. In comparison, we will use the conditional mean estimators on the click-through rate and winning price, since we mainly focus on the risk of expense going beyond the budget where the risk comes from the winning price distribution.

After we obtain all the bidding strategies, we compare the metrics performance in batches of M opportunities using the empirical cumulative distribution of sample batches.

Chapter 7

Numerical Results

In this chapter, we explore the behaviour of risk controlling parameter α with different extents of risk aversions and the influence of different budget levels on the **Valid** set. Then, we select the optimal α based on the Sharpe ratio for the RAP and RAR models and compare the performance of different approaches on the **Test** set.

7.1 Level of Risk Aversion

In our risk-averse bidding models RAP and RAR, the level of risk aversion is controlled by the α. The bidding price b^{rap} and b^{rar} are lower when increasing the value of α as we discussed in the last Chapter 6 about the Figure 6.5. As a result, the model is less likely to win the bid opportunities with less bidding price when competing with others, so the risk of spending over the budget before seeing all the opportunities could be reduced by only trying to win cheap bids that cost less. However, when the $\alpha \to 0$, which represents the decision maker is risk-neutral towards the expense, our risk-neutral bidding models RNP and RNR can be applied.

7.1.1 Profit maximization models

First, we present the performance of profit maximizing models, the risk-averse model RAP and risk-neutral RNP bidding strategies. The performance of the RNP model is denoted as $\alpha = 0$ in the figures.

We compare the model performance **Batch profit, Batch expense**, which represents the model's objective function value and constraint, using the empirical CDF graphs under different levels of risk aversion, given the same average budget. For example, given the budget level $B = 38.12$, we have:

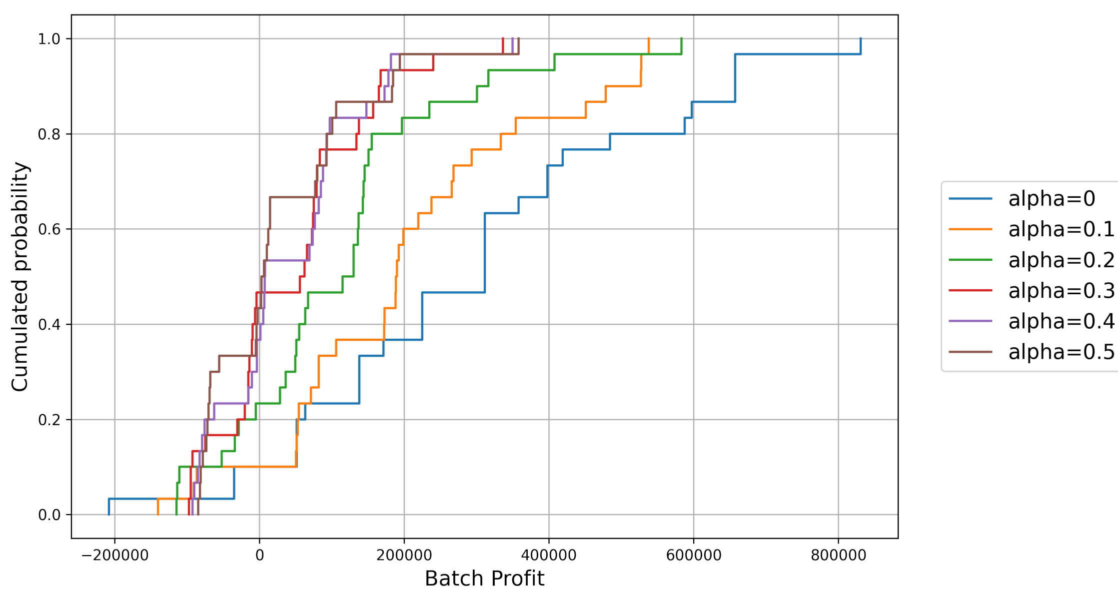

Figure 7.1: Empirical distribution of Batch profit under different risk level for the profit model with B=38.12

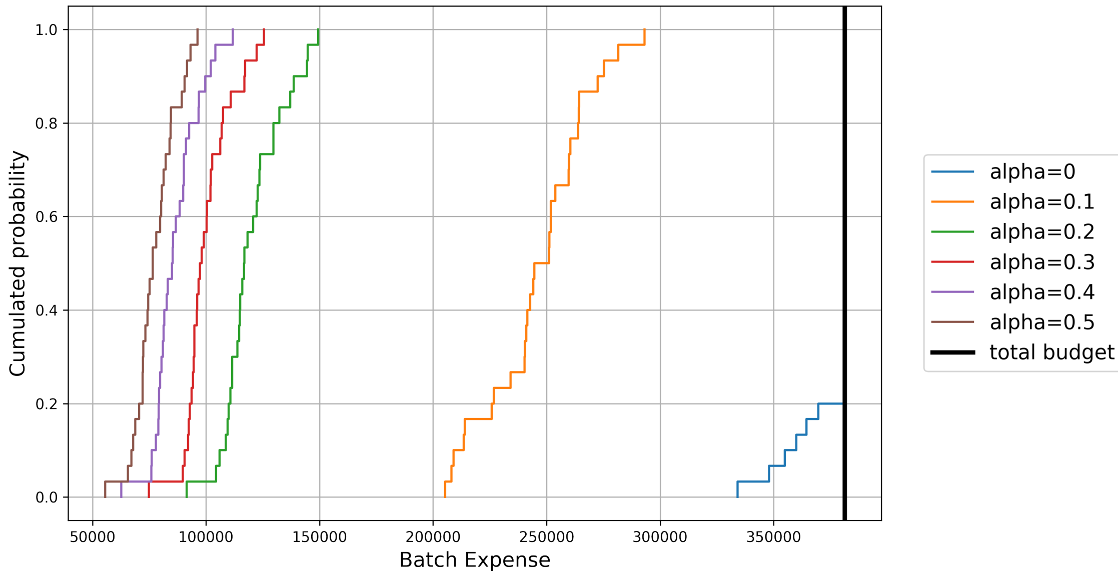

Figure 7.2: Empirical distribution of Batch expense under different risk level for the profit model with B=38.12

The empirical CDF graphs of batch profit in Figure 7.1 and batch expense in Figure 7.2 demonstrate the similar behaviour given different levels of budget.

Our risk-neutral approach RNP outperforms the risk-averse RAP approach in terms of the profit of each $M = 10,000$ opportunities batch. In Figure 7.1, the blue line that represents the RNP where $\alpha = 0$ is on the right side, and dominates all the other RAP models with $\alpha > 0$. Other than that, there is no clear dominance relationship between lines of different $\alpha > 0.3$ using the RAP approach with this budget level, which means when the α is relatively larger, the difference in profit has been reduced and even generate similar profit that overlapping on the empirical CDF graph. If we look at the left tail of the graph 7.1, the risk-averse policy with a larger α, or more risk-averse, have a lower probability of getting negative profit, and we can identify the RNP model has a higher probability of extreme negative profit.

On the other hand, the risk-averse approach RAP demonstrates effective control of spending over budget. The bold black line in the right part of the Figure 7.2 represent the predetermined batch budget. Since the algorithm stops when using up the budget in the batch, there is no CDF line passing the budget black line. As we can see from the expense Figure 7.2, all the RAP models with positive α never use up the batch budget. But the risk-neutral model RNP represented by the blue line uses up the budget 80% of time, which can be identified by only 20% of RNP's CDF line is shown on the left of the budget line.

The Empirical early stop frequency is another main metric that demonstrates the performance of risk control on the budget constraint:

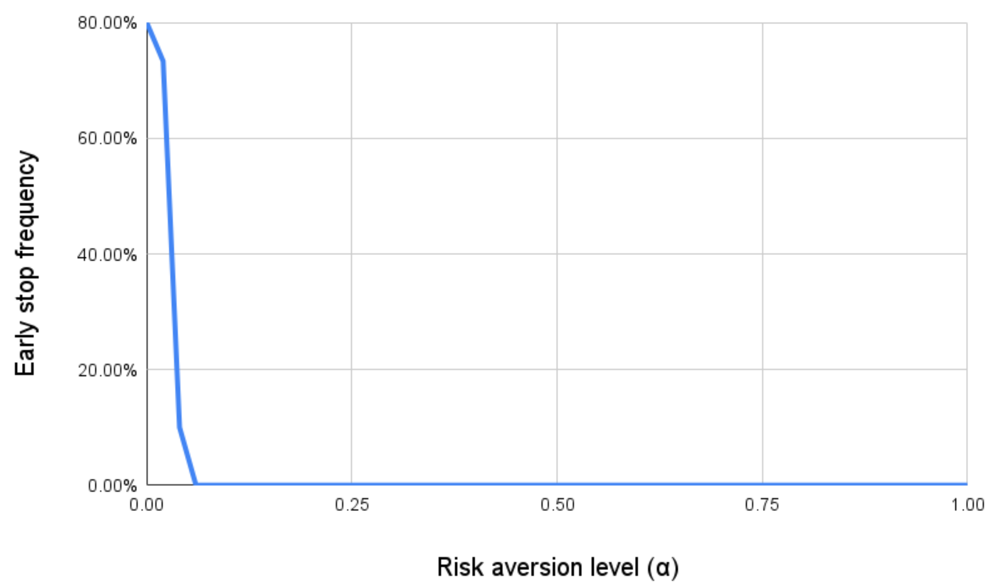

Figure 7.3: Empirical Early stop frequency under different risk level for the profit model with B=38.12

The Figure 7.3 shows a steep decline in the Empirical early stop frequency when we increase the value of α. If we have a much less budget, $B = 2.38$, we can have a more clear view about how the α influences on Empirical early stop frequency, since the smaller budget requires more risk control.

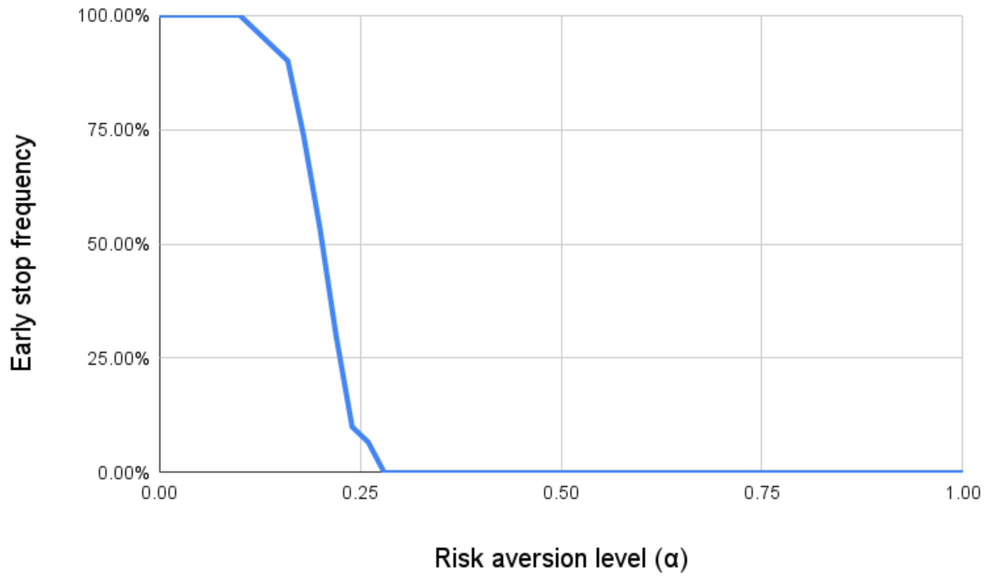

Figure 7.4: Empirical Early Stop Frequency under different risk level for the profit model with B=2.38

Combined the Figure 7.3 and Figure 7.4, we can conclude that a larger value of α does effectively control the Early stop frequency from majority of batches to none. Especially with a smaller budget $B = 2.38$, when $\alpha \to 0$, the budget runs out before the last bidding opportunity for every sample batch. Then, the early stop frequency drops quickly as α increases and ends up with zero early stop frequency when α is greater than a certain value, approximately 0.08 when $B = 38.12$ and 0.28 when $B = 2.38$. Therefore, by adjusting the value of α, the RAP model can control the risk of spending over the budget given a certain period of time.

One may notice that even though the model is more risk-averse when α is higher, it still generates negative profit in the Figure 7.1. Again, our model controls the risk of going over the budget, not directly on the risk of negative profit. Therefore the effect of control negative profit is only marginal. When we experiment on batches, the model is trying to capture all the opportunities by spending all the certainty equivalent amount of the budget given the risk level α while maximizing the objective function but do not account for it might end up with a negative profit.

7.1.2 Revenue maximization models

The RAR and RNR models are aiming for maximizing the revenue of bidding, where the risk-averse model RAR is presented by lines with $\alpha > 0$ and risk-neutral model RNR is denoted as $\alpha = 0$. We can draw a similar conclusion as we conclude for the profit maximizing models in terms of performance with different risk levels.

Given the same average budget $B = 38.12$, we compare the model performance **Batch revenue, Batch expense**, which represent the model's objective function value and constraint under different levels of risk aversion:

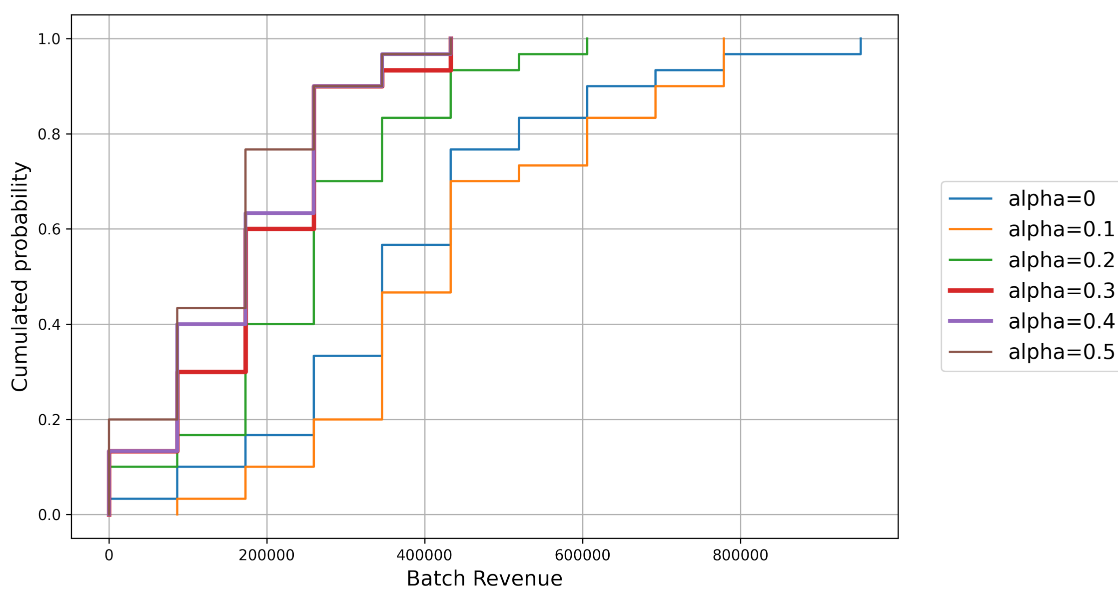

Figure 7.5: Empirical distribution of Batch revenue under different risk level for the revenue model with B=38.12

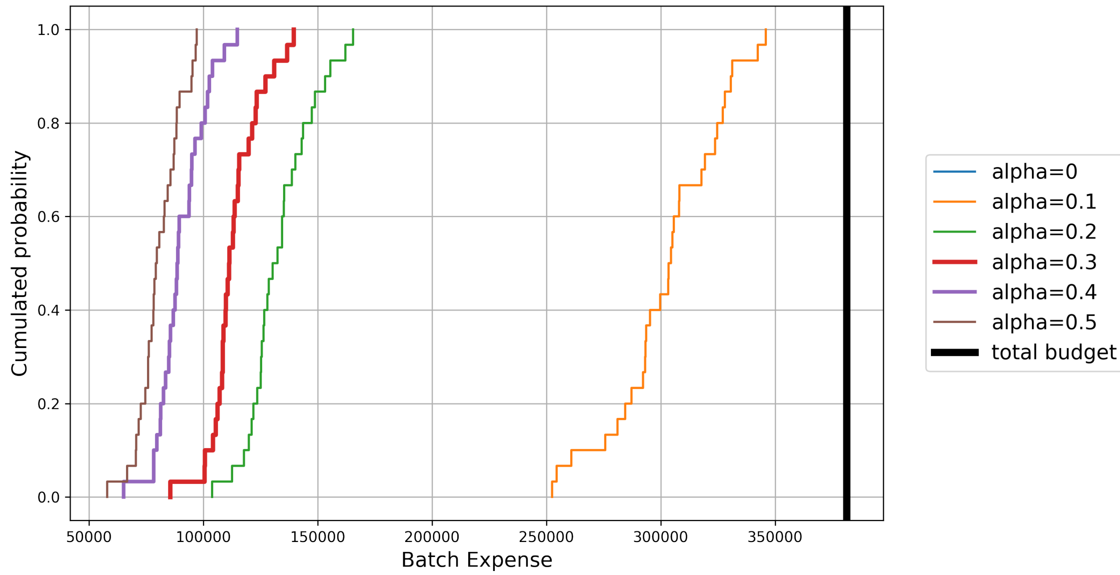

Figure 7.6: Empirical distribution of Batch expense under different risk level for the revenue model with B=38.12

Again, the empirical CDF graphs of batch revenue in Figure 7.5 and batch expense in Figure 7.6 demonstrate the similar behaviour given different levels of budget.

Both empirical CDF figures 7.5 and 7.6 show that there is a clear dominance relationship between different α. The batch revenue and expense move in the same direction and both of them are higher when the α is lower. Surprisingly, the risk-neutral model RNR is stochastic dominated by the RAR model when $\alpha = 0.1$. This is due to when the model is bidding with consideration of the risk of violating the budget constraint, the b^{rar} is higher than b^{rnr} for opportunities that have a higher click-through rate θ but bid less on the opportunities with a lower click-through rate. Therefore, when the α is within a certain range, the RAR outperforms the RNR by spending more on valuable opportunities.

Similarly, the risk-averse approach RAR demonstrates effective control on the risk of spending over budget. First, the Figure 7.6 shows that all of the samples with our risk-averse model do not run out of the batch budget and the more risk-averse, the expense is more restricted.

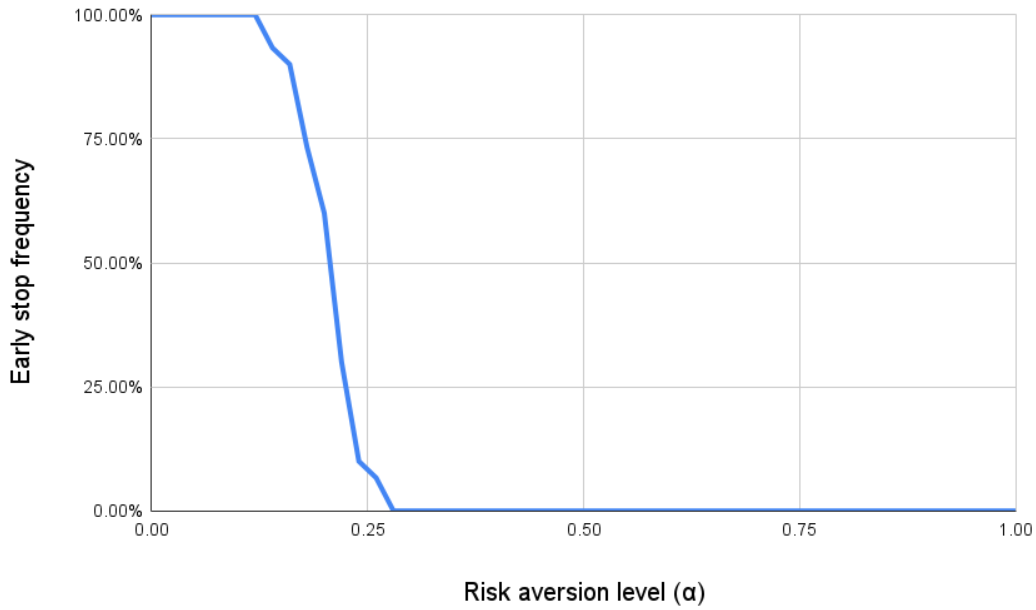

Figure 7.7: Empirical distribution of Early stop frequency under different risk level for the revenue model with B=2.38

We use a smaller budget of $B = 2.38$ to better see the influence on the Empirical early stop frequency. As seen from the Figure 7.7, a larger value of α does control the Early stop frequency from running out of the budget for every sample batch to none of them. Therefore, we have a similar conclusion that by adjusting the value of α, the model RAR can control the risk of spending over the budget given a certain period of time.

7.2 Level of Budget

All of our models depend on the Lagrangian relaxation parameter λ, which is optimized using the algorithm 1 by restricting the expected expense, or certainty equivalent of expense within but close to the predetermined budget. Therefore, our models depend on the budget level and we explore the influence of different levels of budget B on the model performance.

Given the same $\alpha = 0.1$, the model produces different bidding prices for the same set of opportunities according to the different budget levels, from a large average budget when $B = 76.24$ to a small average budget $B = 1.19$. Here, we use the risk-averse with profit objective model RAP as an example to present the influence, the same conclusion can be drawn from other models.

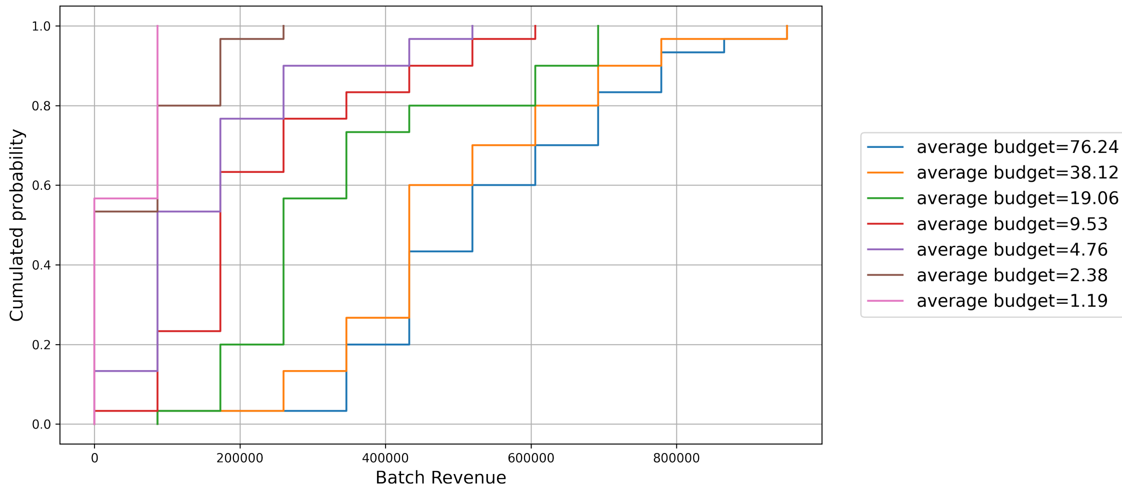

Figure 7.8: Empirical distribution of Batch revenue under different budget level for the profit model

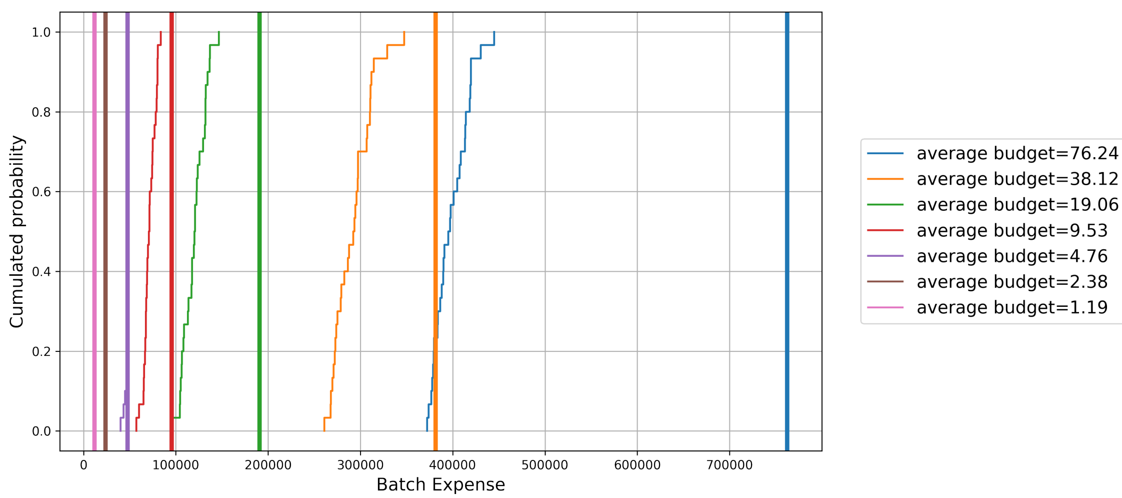

Figure 7.9: Empirical distribution of Batch expense under different budget level for the profit model

From the Figure 7.8 and 7.9, there are clear dominance relationship between the empirical CDF of different budget levels. In the Figure 7.9, the total budget lines are represented by the bold vertical lines with the same color of their batch expense lines, where the CDF of batch expense is on the left side of its total budget line if the model remains under the budget in the experiment batches. The $\alpha = 0.1$ provides enough control on the budget constraint for $B \geq 9.53$ but requires a larger α, or say more risk-averse control, on the smaller budgets, as the batch expense lines of $B \leq 4.76$ largely overlapping with their total budget bold lines. We can conclude that given the same risk level when we have more budget, our model can win more opportunities, therefore producing more revenue but at the same time more expense.

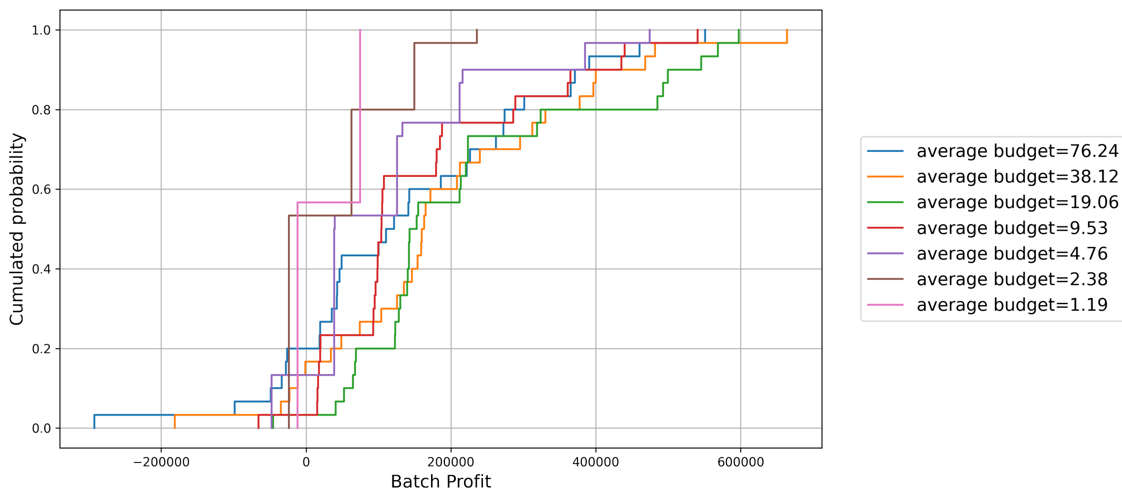

Figure 7.10: Empirical distribution of Batch profit under different budget level for the profit model

Since both revenue and expense move in the same direction, the change in profit is uncertain. If we look at the profit performance Figure 7.10, many empirical CDF graphs are overlapping when $B \geq 19.06$. Part of the reason is that if we see the distribution of the winning price from the Figure 6.1, the distribution is skewed with a long tail on the right. Therefore, when we have a relatively sufficient budget, the budget is very likely to cover most of the valuable opportunities, which leads to an overlapping performance. Furthermore, the click distribution is very sparse and one can win many display opportunities but none contains the realized clicks. Every incremental of clicks can cost a lot more expenses, which also causes the overlapping on profit. Also, one could interpret the budget level as a regularization parameter that concentrates on the most productive opportunities. Thus, the batch profit given a smaller budget with a larger regularization can have comparable performance.

However, the empirical CDFs of $B = 19.06, 9.53, 4.76$ stochastically dominates the profit CDF of smaller budgets, as profit is higher when the budget is higher. The model with a lower budget produces lower bidding prices, therefore it loses more distribution of the opportunities and leads to a lower batch profit. But overlapping happens again given $B = 2.38$ and $B = 1.19$, as the approach seems to have difficulty identifying cheap profitable opportunities given the risk aversion level $\alpha = 0.1$.

7.3 Model Comparison

In this section, we compare the different models in terms of metrics defined in 3.5 on the out-of-sample **Test** set.

7.3.1 The profit models and revenue models

In this thesis, we propose both profit maximization models, RAP and RNP, and revenue maximization models, RAR and RNR. Given the budget $B = 38.12$, we obtain the optimal pair of λ and α according to the Algorithm 2 for the risk-averse models RAP and RAR; and applying the Algorithm 1 to find the optimal λ for the risk-neutral models, RNP and RNR. Then, we use the optimal bidding prices obtained for each model to compare their performance of **Batch profit, Batch revenue, Batch expenses** using the empirical CDF graphs:

From the batch profit performance in the Figure 7.11, the empirical distributions of the two profit-maximizing models dominate the two revenue-maximizing models. For the two profit models, RNP and RAP are overlapping most of the time so there is no dominance relationship, which is mainly due to the RNP model costs more while achieving higher

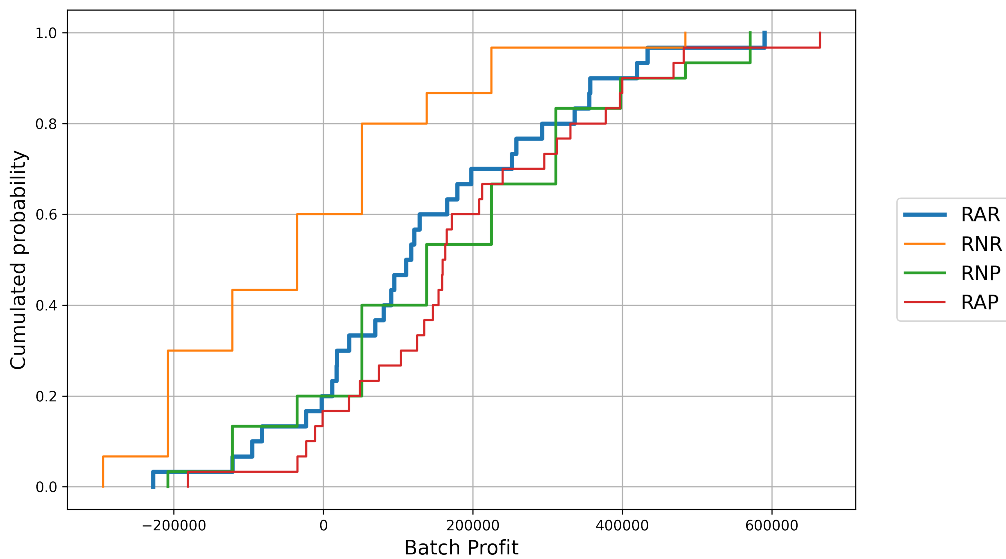

Figure 7.11: Empirical distribution of Batch profit under different methods when B=38.12

revenue than the RAP model. One could interpret that the RNP model uses the in-sample estimators to find the λ that controls the expected expense according to the Algorithm 1, and there is a difference between empirical results and expected values.

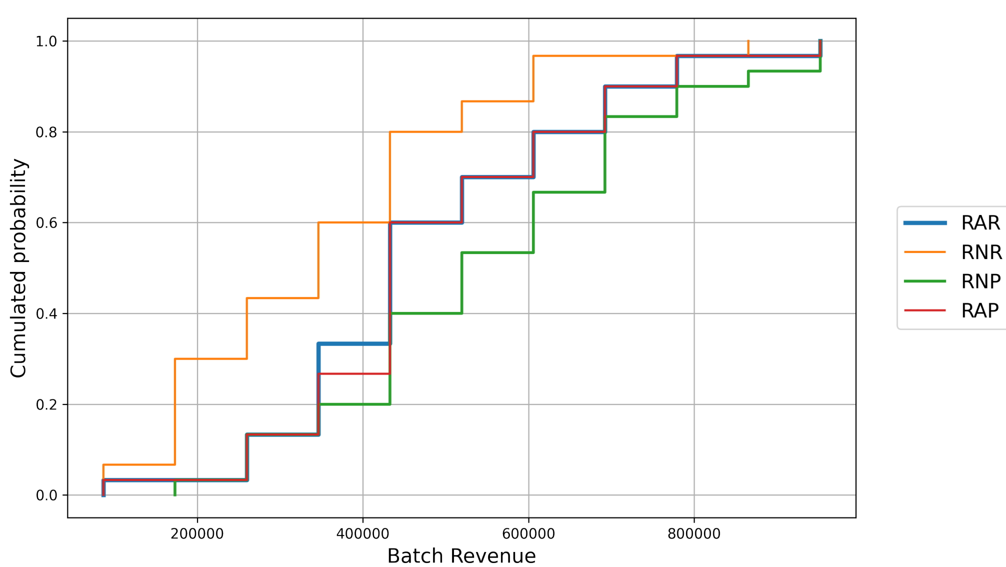

Figure 7.12: Empirical distribution of Batch profit under different methods when B=38.12

When we look at the batch revenue performance in the Figure 7.12, we find out that the

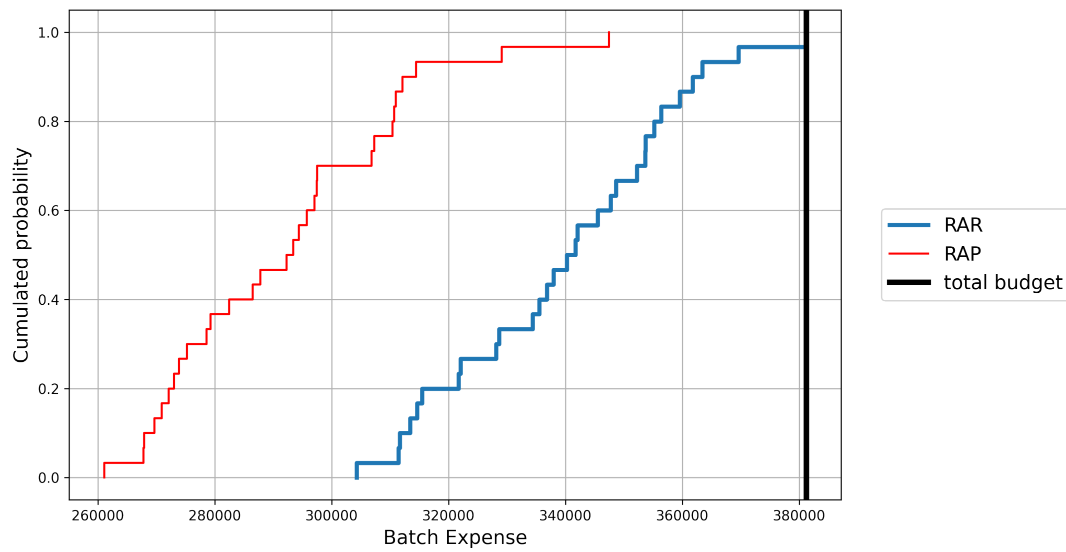

Figure 7.13: Empirical distribution of Batch expense under different methods when B=38.12

RAR and RAP models have very similar performance. However, the RNR unexpectedly produce the worst revenue compared with other methods. This is because, with the RNR approach, the model was trying to maximize the expected value of clicks by a higher bidding price b^{rnr} without considering of risk of running over the budget. Therefore, the RNR model runs out of the budget in every batch and loses the valuable opportunities that happen after its early stop point. Even though the RNP approach also happens early stop for every batch, aiming for maximizing profit is actually equivalent to adding an additional penalty for the expense. Therefore, the RNP approach wins more opportunities than the RNR model which can be identified by a higher impression rate per batch and will be discussed later.

Both risk-averse approaches demonstrate effective control of the expense going beyond the budget. As we can identify from the Figure 7.13, Only RAR and RAP models' expense CDF graphs can be found in the figure and more than 95% of probability that the batch expense remains under the batch budget. On the contrary, the CDF graphs of both risk-neutral approaches combine completely with the bold black line for the total batch budget, which means the risk-neutral models run out of the budget for every batch.

Performance on other metrics

Advertisers also look into metrics like the number of clicks and impression rate as their KPI. In this section, we present the results given different levels of budget for other metrics as stated in Section 3.5, as well as the Empirical early stop frequency as an indicator of the probability of violating the budget constraint among the **Test** set.

As shown in the table 7.1, we present the results of **Batch number of clicks, Batch impression rate, Sharpe ratio of profit, Sharpe ratio of revenue, Early stop frequency** by keeping 4 significant figures:

First, all the risk-averse models demonstrate better control on the risk of violating the budget constraint compared with the risk-neutral models, which can be concluded from the number of **Early stop frequency**.

When the budget is relatively large $B \geq 9.53$, the risk-averse models generally outperform the risk-neutral models in terms of **Sharpe ratio of profit, Sharpe ratio of revenue**. The profit models outperform the revenue models on **Sharpe ratio of profit, Batch number of clicks**. Even though the revenue model is targeted at maximizing revenue, which is equivalent to maximizing the number of clicks, the algorithm does not always better on **Batch number of clicks, Batch impression rate, Sharpe ratio of revenue**. We have discussed earlier that this phenomenon is caused by revenue models tending to be less conservative on expense and more likely to bid higher for the opportunity that has a higher click-through rate, which leads to a higher expense and less capability of capturing the profitable opportunities.

When the budget is relatively small $B \leq 4.76$, the risk-neutral models generally outperform the risk-averse models in terms of **Sharpe ratio of profit, Sharpe ratio of revenue, Batch number of clicks**, whereas the risk-averse models have a better performance in the **Batch impression rate**. This is because the bidding price is lower by risk-averse models, it is less likely for the risk-averse models to win costly opportunities that could generate more profit, but with the risk control on the budget constraint, the risk-averse models can certainly win more opportunities overall which leads to a higher **Batch impression rate**.

Budget	Measures	RAP	RNP	RAR	RNR
$B = 76.24$	Average batch number of clicks	7.233	7.467	7.467	7.700
	Average batch impression rate	79.93%	80.55%	96.93%	99.90%
	Sharpe ratio of profit	1.005	1.024	0.02149	-0.1987
	Sharpe ratio of revenue	3.119	3.199	3.391	3.251
	Early stop frequency	0	0	0	90%
$B = 38.12$	Average batch number of clicks	5.600	6.367	5.533	4.033
	Average batch impression rate	64.51%	69.64%	71.18%	54.55%
	Sharpe ratio of profit	1.083	0.8473	0.7699	-0.1802
	Sharpe ratio of revenue	2.678	2.749	2.604	1.961
	Early stop frequency	0	100%	3.333%	100%
$B = 19.06$	Average batch number of clicks	5.067	3.300	3.500	2.567
	Average batch impression rate	47.08%	34.81%	42.26%	30.02%
	Sharpe ratio of profit	1.381	0.5946	0.9866	0.2364
	Sharpe ratio of revenue	2.314	1.787	1.884	1.664
	Early stop frequency	13.33%	100%	0	100%
$B = 9.53$	Average batch number of clicks	2.633	1.700	1.833	1.700
	Average batch impression rate	28.85%	18.35%	26.52%	18.34%
	Sharpe ratio of profit	1.077	0.4297	0.8461	0.4297
	Sharpe ratio of revenue	1.565	1.219	1.364	1.219
	Early stop frequency	0	100%	0	100%
$B = 4.76$	Average batch number of clicks	0.9667	1.433	0.9667	1.433
	Average batch impression rate	19.36%	11.06%	19.11%	11.02%
	Sharpe ratio of profit	0.4970	0.7026	0.5452	0.7026
	Sharpe ratio of revenue	0.8272	1.141	0.8954	1.141
	Early stop frequency	0	100%	0	100%
$B = 2.38$	Average batch number of clicks	0.3333	1.267	0.4000	1.267
	Average batch impression rate	7.739%	6.928%	7.162%	6.883%
	Sharpe ratio of profit	0.4500	0.8018	0.5236	0.8018
	Sharpe ratio of revenue	0.6202	1.024	0.6547	1.024
	Early stop frequency	0	100%	0	100%

Table 7.1: Metrics results

If we compare the performance between RAR and RAP, RNR and RNP, the metrics results are similar in **Batch impression rate, Batch number of clicks**, and risk-neutral models RNP and RNR have almost the same metrics. Under the limited budget, risk-averse models bid lower than the risk-neutral models with the same objective, which leads to a higher impression rate and better control on the budget constraint but also lowers the probability of winning the valuable opportunities that has a higher probability to get clicks. As for the two risk-neutral models, the RNP model wins slightly more opportunities than the RNR model thanks to more consideration of cost, but still, both b^{rnp} and b^{rnr} are too low to win more valuable opportunities.

7.3.2 Model comparison with the RMP approach

To compare with the RMP approach, we use the same budget amount and value of click setting to make the result more comparable and apply the procedure specified in Section 6.3.3 to optimize the RMP hyperparameters.

Here, we present the out-of-sample performance for our risk-averse profit maximizing model RAP and the RMP approach. Since the budget level could influence the model performance as we discussed earlier, we present the results with two budget levels separately.

Model performance when budget is relatively large

With a relatively large average budget, $B = 38.12$, the batch profit empirical CDF of different models are overlapping. When looking at the left tail of the batch profit graph, our b^{rnp} generates more loss than the other models and the RMP and RMP_neutral models show a good performance on controlling the risk of negative profit.

From the batch expense's empirical CDF graph 7.15, both the RMP and our RAP models effectively control the risk of running out of budget. The RMP approaches cost less than our approaches, even though they have no control over the risk of expense going beyond the budget. However, our RNP model uses up the budget for every batch since its CDF completely combines with the budget line.

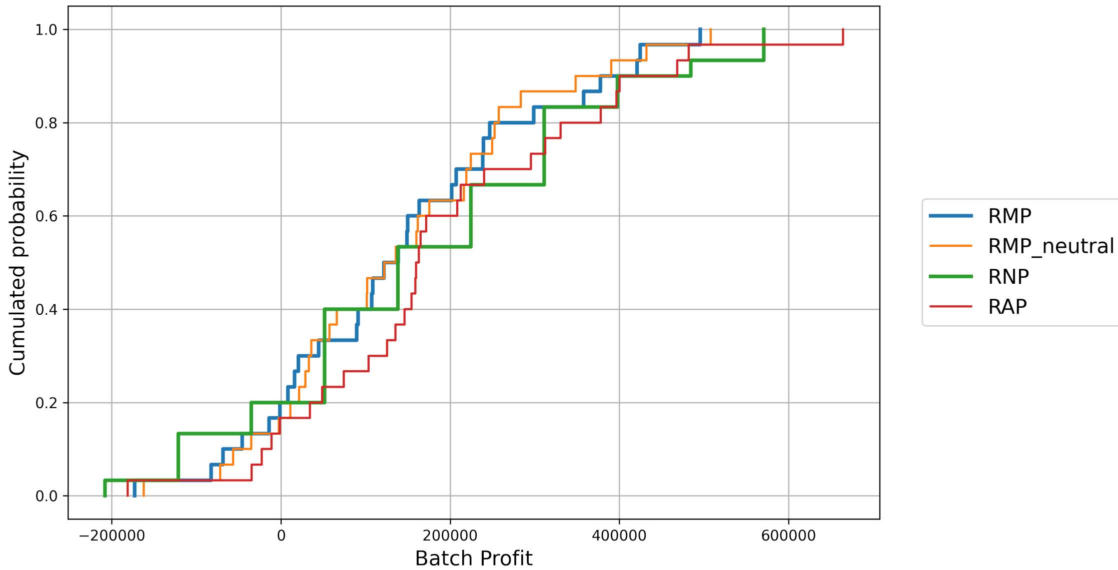

Figure 7.14: Empirical distribution of Batch profit compared with the RMP when B=38.12

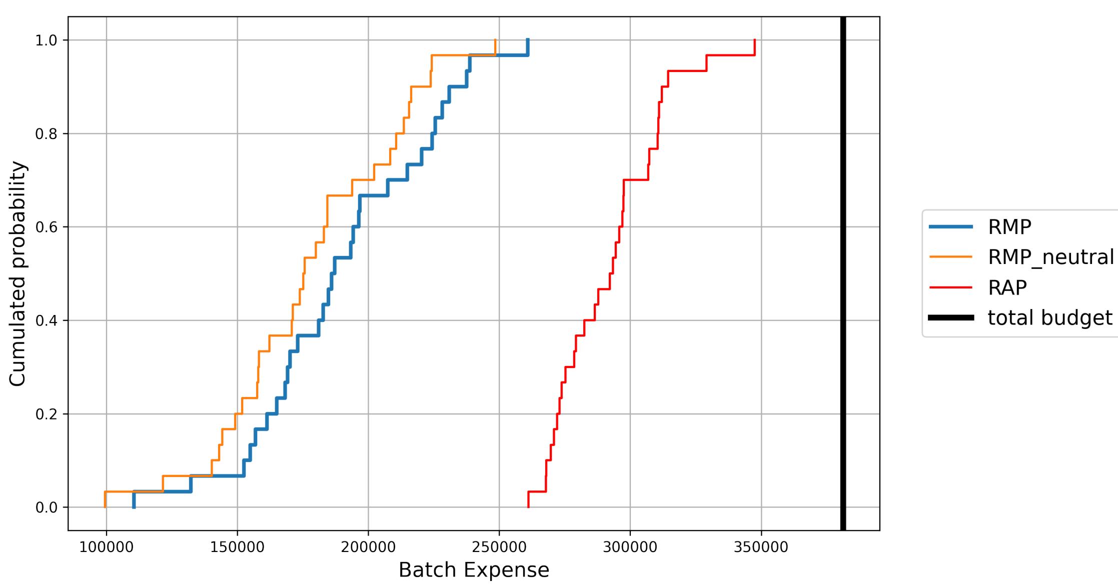

Figure 7.15: Empirical distribution of Batch expense compared with the RMP when B=38.12

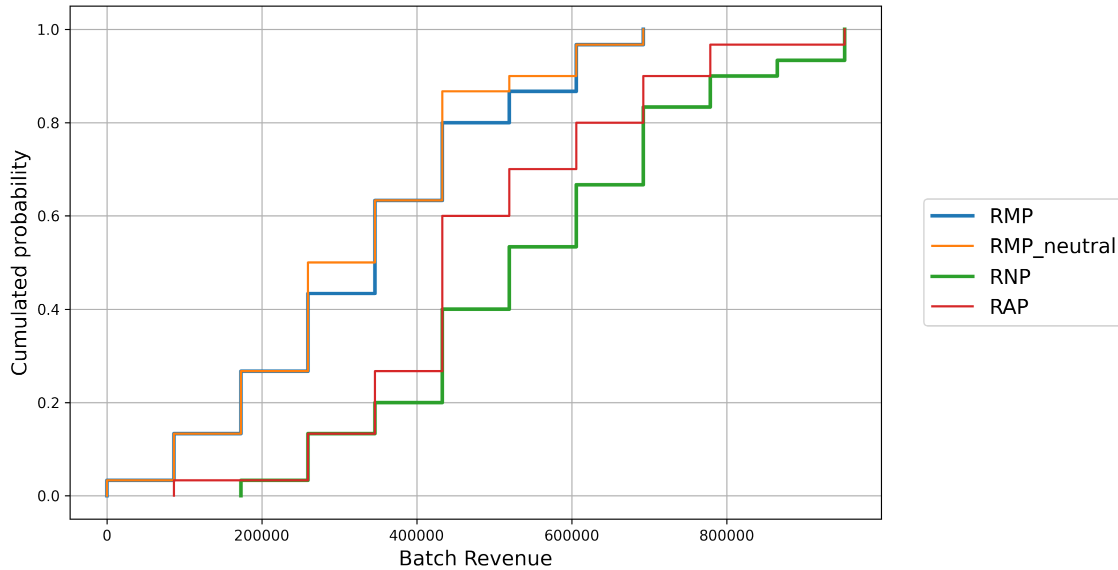

Figure 7.16: Empirical distribution of Batch revenue compared with the RMP when B=38.12

However, our RAP and RNP models dominate the RMP approaches on the batch revenue performance, which means our model produced more clicks than the RMP approaches.

Model performance when budget is relatively small

The results are different when using a relatively small average budget, $B = 2.38$. If we look at the empirical CDF on batch profit graph 7.17, our risk-neutral RNP model dominates all the other models; our risk-averse RAP model has similar results with the RMP model.

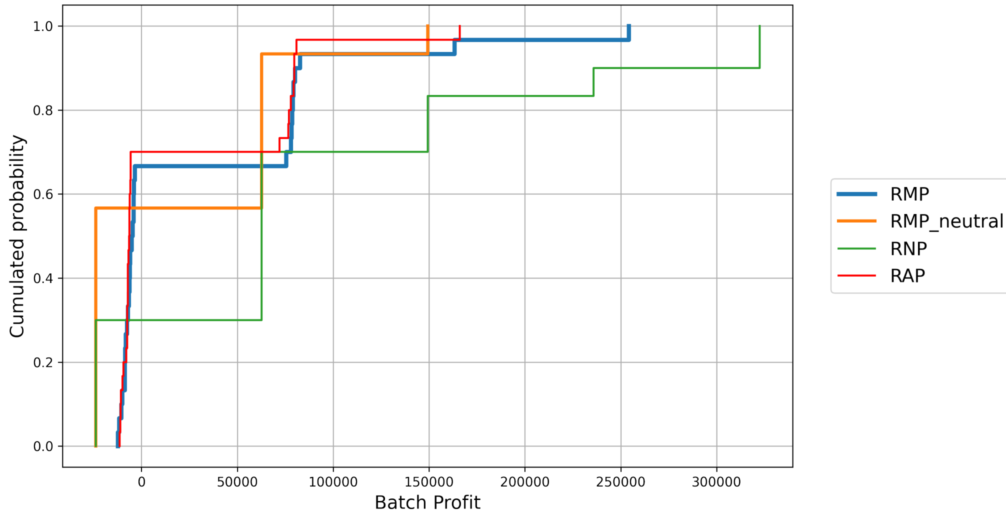

Figure 7.17: Empirical distribution of Batch profit compared with the RMP when B=2.38

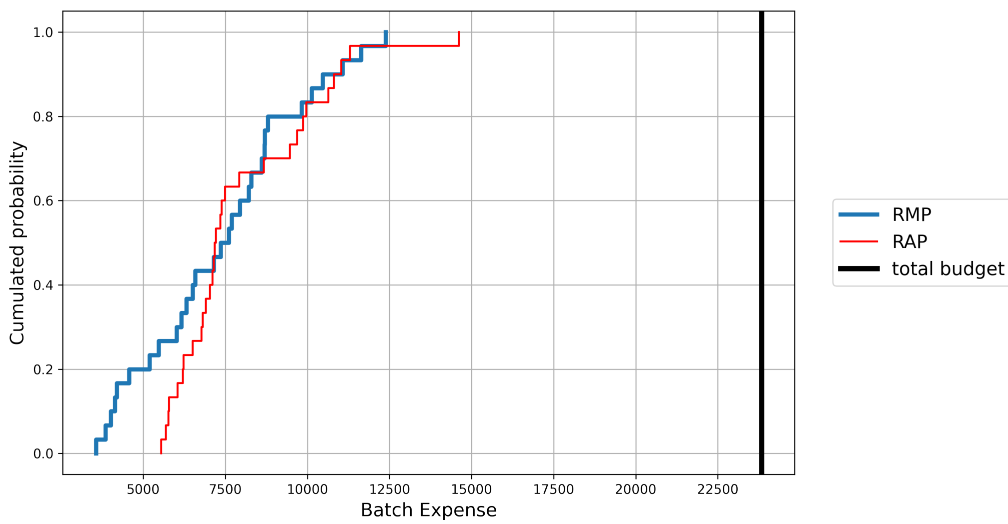

Figure 7.18: Empirical distribution of Batch expense compared with the RMP when B=2.38

Both risk-neutral models fail to remain under the budget by completely integrating with the black budget line on the right and both risk-averse models successfully control the risk of violating the budget constraint, as seen from Figure 7.18. Two risk-averse models have similar results batch expense, by overlapping with each other.

Performance on other metrics

As for other metrics, our approaches provide competitive results compared with the RMP approaches. Especially when the budget is relatively higher, our RAP model significantly dominates the RMP and RMP_neutral models in terms of **Batch number of clicks, Batch impression rate, Sharpe ratios**. When the budget is lower, the RAP model outperforms the RMP on **Batch impression rate** by 2.7% and has comparable results to the best of RMP and RMP_neutral models in the other three metrics. Perhaps the high profit opportunities are all quite homogeneous since we use the same empirical value of click v to represent $V(X)$. The RNP model performs the best among these four models on a low budget but seems unstable in terms of performance with a high Empirical early stop frequency.

Budget	**Measures**	**RAP**	**RNP**	**RMP**	**RMP neutral**
$B = 38.12$	Average batch number of clicks	5.600	6.367	3.867	3.700
	Average batch impression rate	64.51%	69.64%	41.5%	38.78%
	Sharpe ratio of profit	1.083	0.8473	0.8915	0.9218
	Sharpe ratio of revenue	2.678	2.749	1.938	1.947
	Early stop frequency	0	100%	0	0
$B = 2.38$	Average batch number of clicks	0.3333	1.267	0.4333	0.5000
	Average batch impression rate	7.739%	6.928%	4.984%	7.285%
	Sharpe ratio of profit	0.4500	0.8018	0.4867	0.3630
	Sharpe ratio of revenue	0.6202	1.024	0.6056	0.8076
	Early stop frequency	0	100%	0	100%

Table 7.2: Compare metrics results with RMP approach

Chapter 8

Conclusion

In this thesis, we model the risk of spending over the budget given the number of bidding opportunities by using the exponential utility function and entropic risk measures on expense. We not only propose the risk-averse and risk-neutral model that maximizes the profit, the model RAP and RNP, but also expand our approach to revenue maximization problem, the model RAR and RNR. We derive closed-form solutions of optimal bidding policies for each model, which are very easy to implement in reality. After training the estimators and finding the optimal α using the algorithm 2, the decision maker only needs to input their predetermined average budget for a given period of time and risk preference, and the model can predict the bidding price that controls the risk on expense going beyond the budget while maximizing the objective function value, either profit or revenue.

The two risk-averse bidding policies, RAP and RAR can effectively control the risk of spending over the budget for a given time. By adjusting the risk controlling parameter α, the decision maker can control the model's level of risk aversion according to their risk preference. Larger α leads to more risk aversion and less expense. We also propose the algorithm of selecting the optimal α based on the Sharpe ratio of the objective function value detailed in Algorithm 2. Or the decision maker can input their risk preference of α and obtain a bidding policy with a customized level of risk aversion on the expense control.

We propose the batch experiment scheme that is more applicable in reality. Advertisers usually make their budget plan based on the period of time, like a monthly budget or daily budget. The opportunities for online ad slots are a certain number until there is a change in website design. Other bidding strategies do not consider the batch performance but look at the overall opportunities in the dataset so that their bidding policies simply bid until running out the budget. However, there is the risk that the bidding policy will miss the valuable clicks that happened in the following opportunities when they have already stopped bidding. Even if they lower the bidding price by adding a bid scaler hyperparameter to reduce this risk, this bid scaler will largely depend on the number of opportunities in the dataset. On the contrary, our batch experiment scheme considers the budget on batches as subsets of the dataset and the risk hyperparameter selected over the batch performances. Therefore, even though the hyperparameter performance might vary due to the sparsity of clicked opportunities, the empirical result will converge when the dataset is large enough and the number of opportunities in the dataset will have no influence on the bidding price.

The proposed bidding policies are developed based on the static setting, which is less computationally expensive to train compared with sequential models with the multi-stage setting. The sequential models are often formed by complicated architectures such as neural networks or reinforcement learning as we discussed in Chapter 2, which requires the decision makers with a high level of expertise to understand and implement. However, our models demonstrate superior interpretability and feasibility as closed-form solutions depend on predicted winning price distribution and click-through rate. Decision makers can easily take advantage of continuously advancing models developed for these conditional statistics. Also, if these predictions are available, our policies can be faster in solving the problem on a large scale.

Our approach outperforms the RMP strategy when the budget is relatively large in our batch experiment setting. The RMP approach controls risk over profit, where the risk is considered using the mean-standard deviation approach on the click-through rate, while our methods control risk over the expense and minimize the entropic risk by adjusting the exponential utility of the expense with a hyperparameter α. Also, when the budget is relatively ample, our bidding policies generate more revenue/clicks given the budget constraint for the advertiser, which is meaningful, especially for the brand or product in their early life cycle.

Finally, there are many directions that we could continue to explore within the bid optimization topic:

- In this thesis, we propose the risk control on expense going over the budget by modeling the expected utility of expense and identifying there is less control on the negative profit. We can consider risk aversion in objective by modeling the risk-averse profit $\mathbb{E}[u(\text{profit})]$.
- Other than the exponential utility, there are other methods available for risk modeling, for example, the CVaR and VaR;
- It is possible to consider modeling the risk over other objectives, like ROI, using the utility approach with risk consideration;
- Our approach can be continuously improved by applying the dynamic system and reinforcement learning models

Bibliography

Adikari, S. and Dutta, K. (2019). A new approach to real-time bidding in online advertisements: Auto pricing strategy. *INFORMS Journal on Computing*, 31(1):66–82.

Alaya (2017). Approximation of logarithm of standard normal cdf. Cross Validated. URL:https://stats.stackexchange.com/q/288160 (version: 2017-06-30).

Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1):23–36.

Borgs, C., Chayes, J., Immorlica, N., Jain, K., Etesami, O., and Mahdian, M. (2007). Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, page 531–540, New York, NY, USA. Association for Computing Machinery.

Cai, H., Ren, K., Zhang, W., Malialis, K., Wang, J., Yu, Y., and Guo, D. (2017). Real-time bidding by reinforcement learning in display advertising. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*. ACM.

Chapelle, O. (2015). Offline evaluation of response prediction in online advertising auctions. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, page 919–922, New York, NY, USA. Association for Computing Machinery.

Chen, Y., Berkhin, P., Anderson, B., and Devanur, N. R. (2011). Real-time bidding algorithms for performance-based display ad allocation. In *Proceedings of the 17th ACM*

SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11, page 1307–1315, New York, NY, USA. Association for Computing Machinery.

Cheng, H.-T., Koc, L., Harmsen, J., Shaked, T., Chandra, T., Aradhye, H., Anderson, G., Corrado, G., Chai, W., Ispir, M., Anil, R., Haque, Z., Hong, L., Jain, V., Liu, X., and Shah, H. (2016). Wide & deep learning for recommender systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, DLRS 2016, page 7–10, New York, NY, USA. Association for Computing Machinery.

Cui, Y., Zhang, R., Li, W., and Mao, J. (2011). Bid landscape forecasting in online ad exchange marketplace. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, page 265–273, New York, NY, USA. Association for Computing Machinery.

Du, M., Sassioui, R., Varisteas, G., State, R., Brorsson, M., and Cherkaoui, O. (2017). Improving real-time bidding using a constrained markov decision process.

Even-dar, E., Mansour, Y., Mirrokni, V., Muthukrishnan, S., and Nadav, U. (2009). Bid optimization in broad-match ad auctions.

Fan, J. and Yao, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika*, 85(3):645–660.

Fernandez-Tapia, J., Guéant, O., and Lasry, J.-M. (2016). Optimal real-time bidding strategies.

Foellmer, H. and Schied, A. (2010). Convex and coherent risk measures. *Encyclopedia of Quantitative Finance*.

Friedman, J. H. (2002). Stochastic gradient boosting. *Comput. Stat. Data Anal.*, 38(4):367–378.

Graepel, T., Candela, J. Q. n., Borchert, T., and Herbrich, R. (2010). Web-scale bayesian click-through rate prediction for sponsored search advertising in microsoft's bing search engine. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, page 13–20, Madison, WI, USA. Omnipress.

Grislain, N., Perrin, N., and Thabault, A. (2019). Recurrent neural networks for stochastic control in real-time bidding. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '19, page 2801–2809, New York, NY, USA. Association for Computing Machinery.

Guo, H., Tang, R., Ye, Y., Li, Z., and He, X. (2017). Deepfm: A factorization-machine based neural network for ctr prediction.

Haifeng, Z., Weinan, Z., Yifei, R., Kan, R., Wenxin, L., and Jun, W. (2017). Managing risk of bidding in display advertising. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*.

Huang, T., Zhang, Z., and Zhang, J. (2019). Fibinet: Combining feature importance and bilinear feature interaction for click-through rate prediction. In *Proceedings of the 13th ACM Conference on Recommender Systems*, RecSys '19, page 169–177, New York, NY, USA. Association for Computing Machinery.

Jin, J., Song, C., Li, H., Gai, K., Wang, J., and Zhang, W. (2018). Real-time bidding with multi-agent reinforcement learning in display advertising. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 2193–2201, New York, NY, USA. Association for Computing Machinery.

John von Neumann, O. M. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Juan, Y., Lefortier, D., and Chapelle, O. (2017). Field-aware factorization machines in a real-world online advertising system. In *Proceedings of the 26th International Conference on World Wide Web Companion*, WWW '17 Companion, page 680–688, Republic and Canton of Geneva, CHE. International World Wide Web Conferences Steering Committee.

Kan, R., Jiarui, Q., Lei, Z., Weinan, Z., and Yong, Y. (2019). Deep landscape forecasting for real-time bidding advertising. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM.

Karlsson, N. (2016). Control problems in online advertising and benefits of randomized bidding strategies. *European Journal of Control*, 30(C):31–49.

Karlsson, N. (2020). Feedback control in programmatic advertising: The frontier of optimization in real-time bidding. *IEEE Control Systems Magazine*, 40(5):40–77.

Liao, H., Peng, L., Liu, Z., and Shen, X. (2014). Ipinyou global rtb bidding algorithm competition dataset. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, ADKDD'14, page 1–6, New York, NY, USA. Association for Computing Machinery.

Liu, H., Zhu, M., Meng, X., Hu, Y., and Wang, H. (2017). Dual based dsp bidding strategy and its application.

Liu, Q., Yu, F., Wu, S., and Wang, L. (2015). A convolutional click prediction model. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, CIKM '15, page 1743–1746, New York, NY, USA. Association for Computing Machinery.

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

Pan, J., Xu, J., Ruiz, A. L., Zhao, W., Pan, S., Sun, Y., and Lu, Q. (2018). Field-weighted factorization machines for click-through rate prediction in display advertising. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, page 1349–1357, Republic and Canton of Geneva, CHE. International World Wide Web Conferences Steering Committee.

Pan, Z., Chen, E., Liu, Q., Xu, T., Ma, H., and Lin, H. (2016). Sparse factorization machines for click-through rate prediction. In *2016 IEEE 16th International Conference on Data Mining (ICDM)*, pages 400–409.

Pande, H. (2021). Field-embedded factorization machines for click-through rate prediction.

Perlich, C., Dalessandro, B., Hook, R., Stitelman, O., Raeder, T., and Provost, F. (2012). Bid optimizing and inventory scoring in targeted online advertising. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, page 804–812, New York, NY, USA. Association for Computing Machinery.

Qu, Y., Fang, B., Zhang, W., Tang, R., Niu, M., Guo, H., Yu, Y., and He, X. (2018). Product-based neural networks for user response prediction over multi-field categorical data. *ACM Trans. Inf. Syst.*, 37(1).

Ren, K., Zhang, W., Chang, K., Rong, Y., Yu, Y., and Wang, J. (2018). Bidding machine: Learning to bid for directly optimizing profits in display advertising. *IEEE Transactions on Knowledge and Data Engineering*, 30(4):645–659.

Rendle, S. (2010). Factorization machines. In *2010 IEEE International Conference on Data Mining*, pages 995–1000. IEEE.

Richardson, M., Dominowska, E., and Ragno, R. (2007). Predicting clicks: Estimating the click-through rate for new ads. In *Proceedings of the 16th International Conference on World Wide Web*, WWW '07, page 521–530, New York, NY, USA. Association for Computing Machinery.

Rudloff, B. and Wunderlich, R. (2008). Entropic risk constraints for utility.

Shapiro, A. (2017). Interchangeability principle and dynamic equations in risk averse stochastic programming. *Operations Research Letters*, 45(4):377–381.

Sharpe, W. F. (1994). The sharpe ratio. *The Journal of Portfolio Management*, 21(1):49–58.

Shen, W. (2017). Deepctr: Easy-to-use,modular and extendible package of deep-learning based ctr models. `https://github.com/shenweichen/deepctr`.

Statista (2021). Digital advertising in canada - statistics & facts. https://www.statista.com/topics/3048/digital-advertising-in-canada/#dossierKeyfigures.

Ta, A.-P. (2015). Factorization machines with follow-the-regularized-leader for ctr prediction in display advertising. In *2015 IEEE International Conference on Big Data (Big Data)*, pages 2889–2891.

Vasile, F., Lefortier, D., and Chapelle, O. (2017). Cost-sensitive learning for utility optimization in online advertising auctions. In *Proceedings of the ADKDD'17*, ADKDD'17, New York, NY, USA. Association for Computing Machinery.

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

Wu, D., Chen, X., Yang, X., Wang, H., Tan, Q., Zhang, X., Xu, J., and Gai, K. (2018a). Budget constrained bidding by model-free reinforcement learning in display advertising. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, page 1443–1451, New York, NY, USA. Association for Computing Machinery.

Wu, W., Yeh, M.-Y., and Chen, M.-S. (2018b). Deep censored learning of the winning price in the real time bidding. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, page 2526–2535, New York, NY, USA. Association for Computing Machinery.

Wu, W. C.-H., Yeh, M.-Y., and Chen, M.-S. (2015). Predicting winning price in real time bidding with censored data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, page 1305–1314, New York, NY, USA. Association for Computing Machinery.

Yang, X., Li, Y., Wang, H., Wu, D., Tan, Q., Xu, J., and Gai, K. (2019). Bid optimization by multivariable control in display advertising. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.

Yang, Y., Zeng, D., Yang, Y., and Zhang, J. (2015). Optimal budget allocation across search advertising markets. *INFORMS Journal on Computing*, 27(2):285–300.

Zhang, W., Yuan, S., and Wang, J. (2014). Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, page 1077–1086, New York, NY, USA. Association for Computing Machinery.

Zhang, W., Yuan, S., Wang, J., and Shen, X. (2015). Real-time bidding benchmarking with ipinyou dataset.

Zhang, W., Zhang, Y., Gao, B., Yu, Y., Yuan, X., and Liu, T.-Y. (2012). Joint optimization of bid and budget allocation in sponsored search. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, page 1177–1185, New York, NY, USA. Association for Computing Machinery.

Zhang, Y., Wang, D., Wang, G., Chen, W., Zhang, Z., Hu, B., and Zhang, L. (2010). Learning click models via probit bayesian inference. In *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, CIKM '10, page 439–448, New York, NY, USA. Association for Computing Machinery.

Zhao, J., Qiu, G., Guan, Z., Zhao, W., and He, X. (2018). Deep reinforcement learning for sponsored search real-time bidding. In *Proceedings of the 24th ACM SIGKDD*

International Conference on Knowledge Discovery & Data Mining, KDD '18, page 1021–1030, New York, NY, USA. Association for Computing Machinery.