Learning in Search Advertising

W. Jason Choi        Amin Sayedi *

March 12, 2018

*W. Jason Choi (wc2452@columbia.edu) is a Ph.D. candidate in Marketing at the Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027. Amin Sayedi (aminsa@uw.edu) is Assistant Professor of Marketing at the Foster School of Business, University of Washington, Seattle, WA 98195. The authors thank Kinshuk Jerath, Miklos Sarvary, Jiwoong Shin, Woochoel Shin, Andrey Simonov, Hema Yoganarasimhan, and participants at the 2017 NYU Conference on Digital, Mobile Marketing, and Social Media Analytics, Yale seminar, and 2018 UTD Dallas FORMS Conference for their valuable feedback.
Learning in Search Advertising

Abstract

Prior literature on search advertising primarily assumes that search engines know advertisers’ click-through rates, the probability that a consumer clicks on an advertiser’s ad. This information, however, is not available when a new advertiser starts search advertising for the first time. In particular, a new advertiser’s click-through rate can be learned only if the advertiser’s ad is shown to enough consumers, i.e., the advertiser wins enough auctions. Since search engines use advertisers’ expected click-through rates when calculating payments and allocations, the lack of information about a new advertiser can affect new and existing advertisers’ bidding strategies. In this paper, we use a game theory model to analyze advertisers’ strategies, their payoffs, and the search engine’s revenue when a new advertiser joins the market. Our results indicate that a new advertiser should always bid higher (sometimes above its valuation) when it starts search advertising. However, the strategy of an existing advertiser, i.e., an incumbent, depends on its valuation and click-through rate. A strong incumbent increases its bid to prevent the search engine from learning the new advertiser’s click-through rate, whereas a weak incumbent decreases its bid to facilitate the learning process. Interestingly, we find that, under certain conditions, the search engine benefits from not knowing the new advertiser’s click-through rate because its ignorance could induce the advertisers to bid more aggressively. Nonetheless, the search engine’s revenue sometimes decreases because of this lack of information, particularly, when the incumbent is sufficiently strong. We show that the search engine can mitigate this loss, and improve its total profit, by offering free advertising credit to new advertisers.

1 Introduction

Theoretical literature on search advertising has focused on advertisers’ and search engines’ long-term equilibrium strategies, and has often assumed that the advertisers know each others’ types (e.g., valuations, click-through rates (CTR), and quality scores) when submitting their bids (e.g., Edelman et al. 2007; Katona and Sarvary 2010; Jerath et al. 2011). In the short-run, however, this assumption is often violated due to the lack of available information, especially regarding new advertisers who have had no prior interactions with the search engine. Surprisingly, limited attention has been paid to search engines’ and advertisers’ strategies when a new advertiser of an unknown type joins the platform.
When a new advertiser starts search advertising for the first time, the CTRs of its ads are typically unknown to the search engine, other advertisers, and the new advertiser itself. The search engine could at best have an expectation of the CTR based on a few observable characteristics of the advertisers.\footnote{For instance, Google assigns an average Quality Score to new advertisers based on the performances of other advertisers using the same keyword. See \url{https://searchengineland.com/didnt-know-recent-quality-score-changes-259559}.} The actual CTR becomes known only when the ads are displayed to consumers enough number of times such that sufficient impression and click data become available. In other words, the new advertiser has to win advertising auctions sufficiently many times before the search engine and the advertisers could learn its CTR.

The initial lack of information about a new advertiser’s CTR has important implications for existing and new advertisers, as well as the search engine. In particular, because search engines take advertisers’ expected CTRs into account when calculating payments and allocations, lack of information about a new advertiser’s CTR could affect advertisers’ bidding strategies. Furthermore, since learning takes place only if the new advertiser’s ad is displayed to consumers, advertisers may increase or decrease their bids (from what they would bid if the CTR was known) to hinder or facilitate the learning process.

To illustrate the advertisers’ trade-offs when a new advertiser joins the platform, consider the following example.

**Example.** Suppose an advertiser, $A$, is the only advertiser on a keyword $K$ on Google, and assume that there is only one ad slot available on this keyword. $A$’s bid is $1 and its CTR is 15%, which means that for 100 impressions, on average its ad will be clicked 15 times. Assume that $B$ is a new advertiser who wants to start advertising on Google for the first time. $B$’s bid for the same keyword is also $1, but its CTR is not known to anybody at the time of entry. Therefore, for the initial auctions, Google assigns an average CTR estimate (e.g., based on the performance of advertisers with similar characteristics) of, say, 10%.

However, Google can eventually learn the new advertiser’s CTR after sufficient impression
and click data for the new advertiser become available. And note that \( B \) can facilitate this learning process by bidding aggressively and thereby winning in the early rounds. Doing so allows Google to observe more click data for \( B \)'s ad which would in turn allow Google to more accurately estimate \( B \)'s true CTR. Note that Google’s estimate of \( B \)'s CTR directly affects the payment and allocation of the advertisers. This is because Google uses effective bids, computed as advertisers’ submitted bids multiplied by their expected CTRs, to calculate payment and allocation.\(^2\) Given this, would the new advertiser \( B \) prefer to have its CTR learned by the Google quickly or not?

If the new advertiser privately knew its true CTR, then the answer is evident. For example, if it knew that its true CTR is 20% (i.e., twice that of Google’s estimated average), then \( B \) would unambiguously like Google to quickly update its CTR from the 10% average to the true 20% CTR. The reason is that updating its CTR to a higher value would not only make \( B \)'s future effective bid more competitive against the existing advertiser \( A \), but also lower \( B \)'s cost-per-click when it wins. In particular, with its $1 bid and 20% CTR, \( B \) will outrank \( A \)'s effective bid of $1 \times 15\% and win the auction for a cost-per-click of $0.75, whereas it would have lost the auction had its CTR remained at the average of 10% (see Table 1). Conversely, \( B \)'s incentive to facilitate Google’s learning its CTR would diminish if \( B \) knew its true CTR is lower than the average estimate. In this case, \( B \)'s long-term payoff would suffer if its low CTR is learned quickly.

\[
\begin{array}{ccc}
\text{B’s true CTR} & \text{Not known (CTR=10\%)} & \text{Known (CTR=20\%)} & \text{Known (CTR=5\%)} \\
B & \text{Has to bid (and pay)} & \text{Wins at cost-per-click} & \text{Has to bid (and pay)} \\
& \$1 \times 15\%/10\% = \$1.5 \text{ to win} & \$1 \times 15\%/20\% = \$0.75 & \$1 \times 15\%/5\% = \$3 \text{ to win} \\
A & \text{Wins at cost-per-click} & \text{Has to bid (and pay)} & \text{Wins at cost-per-click} \\
& \$1 \times 10\%/15\% = \$0.66 & \$1 \times 20\%/15\% = \$1.33 \text{ to win} & \$1 \times 5\%/15\% = \$0.33 \\
\end{array}
\]

*Table 1: When Search Engine Knows vs. Does Not Know New Advertiser’s CTR*

\(^2\)In practice, effective bids also include other elements such as landing page experience; however, for the purpose of this example, we only consider the expected CTR and the submitted bid that are the two most important elements of effective bids.
In reality, however, when $B$ first joins the search platform, it does not know whether its true CTR is lower or higher than an average advertiser with similar characteristics. Therefore, it is not clear whether the new advertiser $B$ should increase or decrease its bid to accelerate or slow down the learning process if it wants to maximize its profit.

Similarly, for the existing advertiser $A$, Google’s learning the new advertiser $B$’s CTR can be a double-edged sword. If $B$’s CTR turns out to be higher than the estimated average, then $A$ may lose the ad position; if it turns out to be lower, $A$ can win the auction at half the cost than when $B$’s CTR is not known to Google (from $0.66$ to $0.33$ in Table 1). Again, given that the existing advertiser $A$ can facilitate (hinder) Google’s learning process by decreasing (increasing) its bids when $B$ joins, it is not clear which bidding strategy would maximize its profit.

In this paper, we study advertisers’ strategic responses to a search engine’s learning process. To that end, we use a game-theoretic model to analyze advertisers’ and the search engine’s strategies when a new advertiser, whose CTR is unknown, joins the platform. We are interested in answering the following research questions.

1. Does a new advertiser (entrant) benefit from its CTR being learned? How does this affect the entrant’s bidding strategy?

2. How should an existing advertiser adjust its bid when a new advertiser joins?

3. How does the lack of information about a new advertiser’s CTR affect the search engine’s revenue and strategy?

In answering the first set of questions, we show that a new advertiser’s expected payoff when its CTR is learned by the search engine is higher than when it is not. The higher payoff incentivizes the new advertiser to bid aggressively to accelerate the learning process. As a result, the entrant should always bid higher (sometimes even above its valuation) in the beginning when its CTR is unknown to the search engine, than in the long run after its CTR
becomes known. This finding is in line with what industry experts commonly recommend with regard to new advertisers setting starting bids—namely, bid aggressively “into high positions” and “make adjustments after [accumulating] data.” Despite the risk of paying a high initial cost, the experts explain that bidding high and thereby attaining top positions early on could help improve the advertisers’ long-run profits.

Our result indicate that even for advertisers whose long-run equilibrium cost-per-click is low, the initial cost-per-click (at the time of joining the platform) may be above their valuation. In other words, advertisers should be prepared to lose money in the beginning when they start search advertising for the first time. Moreover, they should not be discouraged from using search advertising even if the initial cost-per-clicks, upon joining the platform, are higher than their willingness to pay.

In answering the second question, we find that an incumbent’s response to an entrant joining the auction depends on the incumbent’s CTR. If the incumbent’s CTR is high, the incumbent bids aggressively to impede the entrant’s CTR from being learned by the search engine. This is because an incumbent with high CTR does not want to risk compromising its margin (or worse, losing its ad position) in the event the entrant’s CTR turns out to be high.

This “preemptive” strategy, however, is too expensive for an incumbent with a low CTR. As we show, an incumbent with a low CTR lowers its bid when an entrant joins, so that the entrant’s CTR is learned more quickly. In this case, competing with an advertiser whose CTR is unknown is too costly for the incumbent; by accelerating the learning process, the incumbent hopes that the entrant’s CTR will turn out to be lower than expectation.

In answering the third question, interestingly, we find that the search engine may benefit from not knowing the new advertiser’s CTR. The intuition is that the entrant, and sometimes the incumbent as well, bids more aggressively when the entrant’s CTR is not known, which increases the search engine’s revenue. Under certain conditions, however, the search engine’s

---

3[https://searchengineland.com/4-ways-to-determine-your-your-starting-bids-144616]
ignorance could also hurt its revenue. For instance, if the entrant’s CTR is high, the search engine misses clicks (and hence opportunities for earning higher revenue) by not displaying the entrant’s ad in the beginning. The negative effect becomes more pronounced when the incumbent’s CTR is high because a strong incumbent bids aggressively to mask the entrant’s CTR. This deters the search engine from learning the entrant’s potentially high CTR.

We find that the search engine can mitigate the downside risk of not knowing the entrant’s CTR by offering free ad credit to new advertisers. For example, Google provides $75 ad credit to new advertisers when they spend $25 on AdWords[^1] Microsoft’s Bing also offers similar promotions to newcomers. While these programs have traditionally been viewed as promotions to attract new advertisers, our research reveals new strategic incentives beyond new customer acquisition that motivate search engines to offer ad credit.

First, offering ad credit could help the search engine learn the entrant’s CTR, thereby reducing the revenue loss the search engine incurs for not knowing the CTR. Furthermore, when the entrant is *ex ante* weaker than the incumbent, which is likely for many keywords in practice, the ad credit increases the search engine’s revenue by “helping the weaker player.” These strategic incentives affect the conditions under which it is optimal for the search engine to offer ad credit, as well as the optimal amount of ad credit.

*Theoretical Contribution.* While, from a managerial point of view, our work sheds light on advertisers’ and search engines’ strategies regarding new entries, we also want to highlight two unique aspects of our model from a theoretical point of view. First, in the context of search advertising, we study the *transition* of a game from a partial information game to a full information one. While the previous literature on search advertising assumes that the game is either always full information (e.g., [Edelman et al., 2007]) or always partial information (e.g., [Edelman and Schwarz, 2010]), in practice, the level of information is constantly changing. Our paper takes a first step towards bridging this gap by analyzing this transition[^5] We show

[^1]: https://www.google.com/ads/adwords-coupon.html
[^5]: In fact, since new advertisers constantly join this market, and even existing advertisers often revamp
that the advertisers’ and the search engine’s strategies regarding the transition are different from those in full information and partial information games. In this paper, we focus on the learning dynamics associated with new advertisers’ CTRs. However, the scope of the insights and implications obtained herein is not limited to CTRs and can be applied to a broader range of related parameters such as valuations and conversion rates.

Second, our analysis demonstrates how some of the standard results from learning theory may be reversed when the subjects of learning are not as “passive” as has been commonly assumed in the literature (e.g., Gittins and Jones, 1979; Katehakis and Veinott, 1987). For example, exploration-exploitation trade-off from standard learning theory suggests that knowing less about new advertisers would only hurt the search engine’s revenue because the search engine must then learn about new advertisers through costly exploration. In contrast, our model shows that the search engine may be better off knowing less about the new advertiser due to the advertisers’ strategic reactions during the search engine’s learning process. In other words, when the subjects are strategic agents, exploration could be profitable for the learner.

The rest of this paper is structured as follows. In Section 2 we discuss related literature. In Section 3 we present the model. We analyze the model and discuss advertisers’ strategies in Section 4. The search engine’s strategy and the role of free advertising credit are examined in Section 5. We explore extensions of the main model in Section 6 to establish the robustness of our main results, and conclude in Section 7. All proofs are relegated to Appendix A.

2 Related Literature

The increasing prevalence of search advertising has motivated a growing body of empirical (e.g., see Rutz and Bucklin, 2011; Yao and Mela, 2011; Haruvy and Jap, 2018) and theoretical papers in the marketing literature. Katona and Sarvary (2010) and Jerath et al. (2011) their campaigns, change their ad copies and landing pages, or change their ad agencies altogether, one could argue that this market is always in transition.
study advertisers’ incentives in obtaining lower vs. higher positions in search advertising auctions. Sayedi et al. (2014) investigate advertisers’ poaching behavior on trademarked keywords, and their budget allocation across traditional media and search advertising. Desai et al. (2014) analyze the competition between brand owners and their competitors on brand keywords. Lu et al. (2015) and Shin (2015) study budget constraints, and budget allocation across keywords. Zia and Rao (2017) look at the budget allocation problem across search engines. Zhu and Wilbur (2011) and Hu et al. (2016) study the trade-offs involved in choosing between “cost-per-click” and “cost-per-action” contracts. Wilbur and Zhu (2009) find the conditions under which it is in a search engine’s interest to allow some click fraud. Cao and Ke (2017) model a manufacturer and retailers’ cooperation in search advertising and show how it affects intra- and inter-brand competition. Amaldoss et al. (2015a) show how a search engine can increase its profits and also improve advertisers’ welfare by providing first-page bid estimates. Berman and Katona (2013) study the impact of search engine optimization, and Amaldoss et al. (2015b) analyze the effect of keyword management costs on advertisers’ strategies. Berman (2016) explores the effects of advertisers’ attribution models on their bidding behavior and their profits. Sayedi et al. (2017) study exclusive placement auctions in search and display advertising. Katona and Zhu (2017) show how quality scores can incentivize advertisers to invest in their landing pages and to improve their conversion rates. Following Edelman et al. (2007), by arguing that players learn each others’ types after playing the game repeatedly, the vast majority of this literature uses a full information setup to model search advertising auctions. There are a few papers (e.g., Amaldoss et al. 2015a,b; Edelman and Schwarz 2010) that use an incomplete information setting for modeling search advertising. In these papers, however, the game remains an incomplete information game; i.e., players do not learn each others’ types. To the best of our knowledge, our paper is the first on search advertising to model the learning process, wherein the game starts as an incomplete information game and, if a new advertiser’s type is learned, transitions to a full information game.
Parts of our model may resemble the literature on games with asymmetric information. For instance, in Jiang et al. (2011), a seller may want to hide its type from a platform by mimicking another type in a pooling equilibrium, thereby limiting the platform’s information. Despite some similarities, the key distinction between our model and games with asymmetric information is that there is no information asymmetry in our model. Although players take certain actions to facilitate or hinder the revelation of information, in our model, those actions do not signal their types. Furthermore, in signaling games, players mimic other players’ strategies to hide or reveal information, whereas in our model, advertisers interfere in the search engine’s learning process in order to do so.

There is a growing body of research in Computer Science and Operations Research that addresses dynamic learning in repeated auction settings. Li et al. (2010) solve for an advertiser’s optimal bidding strategy when it is uncertain about its CTR and faces an exogenous distribution of competing bids. Similarly, Iyer et al. (2014) adopt a mean field approximation to demonstrate how bidders’ strategies are affected by their learning of uncertain valuations. Hummel and McAfee (2016) characterize the search engine’s optimal bid on behalf of advertisers under uncertain CTRs in a repeated game, and Balseiro and Gur (2017) introduce adaptive bidding strategies for budget-constrained advertisers in repeated auctions of incomplete information. While closely related in spirit, the research focus and the modeling approach of these previous works are sharply distinct from this paper. In particular, many of the aforementioned works invoke mean field equilibrium concepts under the assumption of an infinitely large market size for a given ad slot. More importantly, these papers do not explicitly model competitors’ responses, which may generate important strategic nuances as demonstrated in this paper.
3 Model

Our model consists of a search engine and two advertisers, the incumbent and the entrant, indexed by $S$, $I$ and $E$, respectively. The search engine sells one ad slot in a second-price auction with reserve price $R$. Each advertiser has an advertiser-specific CTR — $c_I$ for the incumbent and $c_E$ for the entrant — that represents the average CTR of the advertiser if placed in the ad slot. In other words, when an ad is displayed to a consumer, the consumer clicks on the incumbent’s (entrant’s) ad with probability $c_I (c_E)$. Parameters $c_I$ and $c_E$ depend on the advertisers’ ad copies, as well as the relevance and strength of their brands with respect to the consumer’s search query. We assume that both advertisers have the same valuation per click, which we normalize to 1. This assumption allows us to focus on the effects of learning CTRs; however, we relax this assumption in Section 6. The incumbent (entrant) submits a bid $b_{It}$ ($b_{Et}$), where $t$ indexes the game stage. The bids indicate how much advertisers are willing to pay per click.

Search engines use quality scores to determine payment and allocation of search advertising auctions. Specifically, an advertiser’s effective bid (commonly referred to as the ad rank) is the product of the advertiser’s submitted bid and its quality score. The key determinant of the advertiser’s quality score is its expected CTR, i.e., the probability that its ad will be clicked if displayed to consumers. This is captured by parameters $c_I$ and $c_E$ in our model. To focus on the role of CTRs, we assume that the two advertisers are the same along other dimensions of the quality score (e.g., landing page experience). Therefore, the ad ranks of the incumbent and the entrant at stage $t$ are $c_I b_{It}$ and $c_E b_{Et}$, respectively. The advertiser with the higher ad rank wins the auction, provided its ad rank is greater than or equal to the reserve price. The winner pays (per-click) the minimum bid required to win the auction; i.e., if the incumbent wins, it pays $\max[c_E b_{Et}, R]/c_I$ and if the entrant wins, it pays $\max[c_E b_{Et}, R]/c_I$.

---

6We consider a multiple-slot Generalized Second-Price (GSP) auction in Section 6.

7For example, Google claims that “the biggest [component of a quality score] by far is click-through rate.”

8For a detailed discussion of how quality scores affect advertisers’ strategies, see Katona and Zhu (2017).
We assume that $c_I$ and $c_E$ are drawn from $F_I$ and $F_E$, respectively, where $F_I$ and $F_E$ are differentiable cumulative distribution functions. Since the incumbent has been advertising with the search engine for an extended period of time, following Edelman et al. (2007) (and many other papers in the literature), we assume that its CTR, $c_I$, is common knowledge. On the other hand, the entrant’s CTR is not known at the time of entry because the entrant is new to the platform. When the entrant joins, the search engine, the incumbent, and the entrant only know the distribution of the entrant’s CTR. We assume that the expected value of $c_E$, denoted by $\mu_E$, is greater than the reserve price (i.e., $\mu_E > R$), so that the entrant could beat the reserve price in expectation. Similarly, we assume that $c_I \geq R$ so that the incumbent is in the game when the entrant enters.\footnote{This assumption is to facilitate exposition. When $c_I < R$, we can show that the advertisers’ strategies remain the same as when $c_I = R$.}

Before we proceed further, we should elaborate on the meaning of the CTR parameters $c_I$ and $c_E$. In our model, these parameters represent the advertiser-specific CTRs which, as explained above, depend on the advertisers’ ad copies and brand strengths among others. Advertiser-specific CTRs are independent of position effects where higher ad slot position increases ad’s click propensity. Indeed, search engines only take into account advertiser-specific CTRs, controlling for position effects, when computing advertisers’ ad ranks. Position-specific CTRs will be incorporated in the the multi-slot extension model in Section 6.2. The CTR that is initially unknown and is the central subject of learning throughout the paper is the new advertisers’ advertiser-specific CTR.

Next, we describe the timing of the game (see Figure 1).

**Stage 1**: The entrant joins the market. The entrant’s CTR is initially unknown, and is therefore set to its expected value $\mu_E$.\footnote{In Google AdWords, new advertisers received an average Quality Score of 6. See https://searchengineland.com/minimum-quality-score-can-save-money-adwords-226757} The incumbent and the entrant simultaneously
submit their bids $b_{I1}$ and $b_{E1}$ to the search engine. The incumbent’s ad rank is $c_I b_{I1}$ whereas the entrant’s is $\mu_E b_{E1}$, since the search engine does not know the entrant’s CTR yet. If the incumbent wins, it pays (per-click) $\max[\mu_E b_{E1}, R]/c_I$, and if the entrant wins, it pays $\max[c_I b_{I1}, R]/\mu_E$. If the entrant wins, its CTR becomes known to the search engine by the next stage; otherwise, it remains unknown.

To simplify the analysis, we assume that if the entrant wins a single auction, i.e., the auction in Stage 1, then the search engine learns its CTR. In practice, the entrant would have to win sufficiently many times for the search engine to accurately learn its CTR. Stage 1 in our model corresponds to as many auctions as the entrant needs to win for the search engine to learn its CTR. Furthermore, in practice, learning is continuous and gradual; i.e., the search engine’s estimate of the entrant’s CTR improves every time the entrant wins. Our model can be viewed as a discrete approximation of this learning process: the search engine either knows or does not know the entrant’s CTR.

**Stage 2:** The advertisers submit their bids $b_{I2}$ and $b_{E2}$. The incumbent’s ad rank is $c_I b_{I2}$. The entrant’s ad rank depends on the outcome of the Stage 1 auction. If the entrant had won in Stage 1, then its CTR is known to the search engine by Stage 2, and therefore, its ad rank is $c_E b_{E2}$. Otherwise, as in Stage 1, its CTR is not learned and its ad rank is $\mu_E b_{E2}$.

---

11 If the entrant wins the auction in Stage 1, the search engine learns $c_E$; however, we do not need to make any assumptions on whether the incumbent also learns $c_E$ or not. Specifically, as we show in Lemma 1, the incumbent bids truthfully in Stage 2 regardless of the outcome of Stage 1.
We capture the relative weight of Stage 2 compared to Stage 1 with parameter $\delta > 0$. Note that since the advertisers’ decisions in Stage 1 affects their payoffs in Stage 2, $\delta$ affects how the advertisers trade off short-term revenue (in Stage 1) for long-term revenue (in Stage 2).\footnote{One might argue that the search engine eventually learns the entrant’s CTR, even if the entrant does not win in Stage 1. For instance, its CTR may be learned if the entrant’s ad is displayed on the second page of the search results for a sufficiently long period of time. In this case, we could assume that the game has a Stage 3 in which, regardless of the outcomes of Stages 1-2, $c_E$ becomes learned by the search engine. It is easy to show that both advertisers bid truthfully in Stage 3, and that the existence of Stage 3 does not affect the advertisers’ strategies in Stages 1-2. In this model, $\delta$ could be interpreted as the length of time required for the search engine to learn the entrant’s CTR if the entrant does not win in Stage 1 (compared to when it wins in Stage 1).}

The advertisers’ strategies are a pair of first and second stage bids $(b_{j1}, b_{j2}) \in \mathbb{R}_+^2$ for $j = I, E$. The incumbent’s expected profit is the sum of its first and second stage payoffs. That is, $E[\pi_I] = \pi_{I1} + \delta E[\pi_{I2}]$ where $\pi_{I1}$ denotes the incumbent’s first stage payoff, and $\pi_{I2}$ its second stage payoff contingent on the realization of $c_E$, over which expectation is taken. Specifically,

$$\pi_{I1} = \begin{cases} c_I \left(1 - \frac{\max[c_E b_{E1}, R]}{\mu_E}\right) & \text{if } c_I b_{I1} \geq \max[c_E b_{E1}, R]; \\ 0 & \text{otherwise,} \end{cases} \quad \pi_{I2} = \begin{cases} c_I \left(1 - \frac{\max[\tilde{c}_E b_{E2}, R]}{c_I}\right) & \text{if } c_I b_{I2} \geq \max[\tilde{c}_E b_{E2}, R]; \\ 0 & \text{otherwise,} \end{cases}$$

where

$$\tilde{c}_E = \begin{cases} c_E & \text{if } c_E \text{ is learned (i.e., entrant won Stage 1 auction),} \\ \mu_E & \text{if } c_E \text{ is not learned (i.e., entrant lost Stage 1 auction).} \end{cases} \quad (3.1)$$

Similarly, the entrant’s expected profit is $E[\pi_E] = E[\pi_{E1}] + \delta E[\pi_{E2}]$, where

$$\pi_{E1} = \begin{cases} c_E \left(1 - \frac{\max[c_E b_{E1}, R]}{\mu_E}\right) & \text{if } \mu_E b_{E1} \geq \max[c_E b_{E1}, R]; \\ 0 & \text{otherwise,} \end{cases} \quad \pi_{E2} = \begin{cases} c_E \left(1 - \frac{\max[\tilde{c}_E b_{E2}, R]}{c_E}\right) & \text{if } \tilde{c}_E b_{E2} \geq \max[c_I b_{I2}, R]; \\ 0 & \text{otherwise.} \end{cases}$$

Finally, the search engine’s expected profit is $E[\pi_S] = \pi_{S1} + \delta E[\pi_{S2}]$, where

$$\pi_{S1} = \begin{cases} \max[\mu_E b_{E1}, R] & \text{if } c_I b_{I1} \geq \max[\mu_E b_{E1}, R]; \\ \max[c_I b_{I1}, R] & \text{if } \mu_E b_{E1} > \max[c_I b_{I1}, R], \end{cases} \quad \pi_{S2} = \begin{cases} \max[\tilde{c}_E b_{E2}, R] & \text{if } c_I b_{I2} \geq \max[\tilde{c}_E b_{E2}, R]; \\ \max[c_I b_{I2}, R] & \text{if } \tilde{c}_E b_{E2} > \max[c_I b_{I2}, R]; \\ 0 & \text{otherwise.} \end{cases}$$
We use subgame perfect Nash equilibrium as the solution concept and solve by backward induction.

4 Analysis

4.1 Full Information Setting

As a benchmark, we first consider the case where the entrant’s CTR is common knowledge. This corresponds to what most of the previous theoretical papers in search advertising literature assume. Even though the auction is not a standard second-price auction because advertisers’ bids are multiplied by their CTRs, it is easy to show that truthful bidding (i.e., bidding the per-click valuation) is still a weakly dominant strategy for both advertisers. The advertisers’ equilibrium strategies and their payoffs under full information are summarized in the following proposition.

Proposition 1 (Bids and Payoffs Under Full Information). Under full information, truthful bidding is a weakly dominant strategy for both advertisers. The payoffs of the incumbent, the entrant, and the search engine, respectively, are \( \pi_I^F = (1 + \delta) \max[c_I - \max[c_E, R], 0] \), \( \pi_E^F = (1 + \delta) \max[c_E - c_I, 0] \), and \( \pi_S^F = (1 + \delta) \max[\min[c_I, c_E], R] \).

Proposition shows that when the search engine knows the entrant’s CTR, both advertisers always bid truthfully. This finding is not new to the literature and is presented here for the sake of completeness. Interestingly, in the next section, we show that truthful bidding is no longer an equilibrium strategy when the search engine does not know the entrant’s CTR.
### 4.2 Partial Information Setting

In practice, there is little information regarding the entrant’s CTR that is available to the search engine. Therefore, unlike the case for the incumbent’s CTR, the advertisers and the search engine have at best only partial information about the entrant’s CTR.

**Advertisers’ Strategies**

We begin our analysis under partial information with the advertisers’ bids. We focus on dominant strategy equilibrium where advertisers play weakly dominant strategies. As we show in Lemma 1, the second stage auction is trivial: advertisers bid truthfully. Intuitively, this is because there are no strategic considerations of future payoffs in the last stage, and thus the truthfulness property of standard second-price auctions holds.

**Lemma 1 (Bids in Stage 2 Under Partial Information).** Regardless of the outcome of Stage 1, bidding truthfully is a weakly dominant strategy for both advertisers in Stage 2.

In contrast, the advertisers’ bidding strategies in Stage 1 are not always truthful. Their bids can be either below or above valuation depending on their Stage 2 expected payoffs. The following lemma characterizes the advertisers’ first stage equilibrium bids.

**Lemma 2 (Bids in Stage 1 Under Partial Information).** In Stage 1, it is weakly dominant for the incumbent and the entrant, respectively, to bid

\[
\begin{align*}
    b_{I1}^* &= 1 + \delta \left( 1 - \frac{\mu_E}{c_I} \right) I_{\{c_I \geq \mu_E\}} - \int_0^{c_I} \left( 1 - \frac{\max[c_E, R]}{c_I} \right) dF_E, \\
    b_{E1}^* &= 1 + \delta \left( \int_{c_I}^{1} (c_E - c_I) dF_E - (\mu_E - c_I) I_{\{\mu_E > c_I\}} \right),
\end{align*}
\]

where \( I_{\{\cdot\}} \) is the indicator function.

We see from expressions (4.1) and (4.2) that the advertisers’ bids are no longer truthful. In
general, truthful bidding is a weakly dominant strategy in a second-price auction even under incomplete information assumption. What drives the change in advertisers’ strategies from truthful bidding in our model is the advertisers’ incentive (or lack thereof) to help the search engine learn the entrant’s CTR. In other words, the advertisers’ preference to play a game in Stage 2 in which the entrant’s CTR is $\mu_E$ versus $c_E$ (where $c_E$ is randomly drawn from $F_E$) affects their Stage 1 bids. For example if the entrant’s expected payoff in Stage 2 is higher when its CTR is $c_E$ (i.e., the CTR is learned), compared to when it is $\mu_E$ (i.e., the CTR is not learned), the entrant increases its Stage 1 bid.

Does the entrant want the search engine to learn its CTR? We find the answer to be affirmative. For the entrant, the benefits of revealing its CTR are two-fold. First, it allows the entrant to outrank the incumbent in Stage 2 with some probability even when $c_I \geq \mu_E$, a situation in which the entrant would have surely lost in Stage 2 if its CTR was unknown. Secondly, it provides an opportunity for the entrant to pay lower cost-per-click in the event that its CTR turns out to be high, compared to the case when its CTR is assigned the mean estimate $\mu_E$. Evidently, there is also the risk of its CTR turning out to be low, in which case the entrant would have been better off being assigned $\mu_E$. The reward of a high CTR realization, however, is disproportionately larger than the loss the entrant incurs for a low realization. The reason is that while the gains for the entrant increase proportionally with high realizations of $c_E$, the loss of a low $c_E$ is bounded from below by zero. This can also be seen from the following expressions of the entrant’s expected Stage 2 profit when $\mu_E > c_I$:

<table>
<thead>
<tr>
<th>Search engine does not know $c_E$</th>
<th>Search engine knows $c_E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\int_0^1 c_E(1 - c_I/\mu_E)dF_E$</td>
<td>$\int_0^1 c_E(1 - c_I/c_E)I{c_E &gt; c_I}dF_E$</td>
</tr>
<tr>
<td>$= (1 - c_I/\mu_E)\int_0^1 c_EdF_E = (1 - c_I/\mu_E)\mu_E$</td>
<td>$= \int_{c_I}^1 c_E(1 - c_I/c_E)dF_E$</td>
</tr>
<tr>
<td>$\mu_E - c_I = \int_0^1(c_E - c_I)dF_E$</td>
<td>$= \int_{c_I}^1(c_E - c_I)dF_E$</td>
</tr>
</tbody>
</table>

Table 2: Entrant’s Stage 2 Profit

From Table 2, we see that the entrant’s Stage 2 profit when the search engine does not know the entrant’s CTR (left-hand side) is integrated over negative values as well (in the
range $c_E \in (0, c_I)$). This integral value is lower than that when the search engine knows $c_E$ (right-hand side) where only positive values are integrated. In sum, for any entrant CTR distribution $F_E$, the entrant’s Stage 2 profit is higher in expectation if the search engine learns its CTR. Therefore, the entrant bids aggressively in Stage 1 in order to facilitate the learning process.

The incumbent’s bidding strategy is slightly more nuanced: the incumbent underbids for low $c_I$ and overbids for high $c_I$. Suppose $c_E$ is not learned by the search engine in Stage 2. If $c_I$ is close to $\mu_E$, then the incumbent either loses the Stage 2 auction or, even if it wins the auction, receives a low Stage 2 payoff because the cost-per-click $\mu_E/c_I$ is high. In this case, the incumbent is better off shading its Stage 1 bid below valuation, thereby helping the entrant win the first stage auction. Why would the incumbent want to help its competitor win the ad auction? The intuition is that by facilitating the revelation of the entrant’s CTR, the incumbent foregoes its first stage payoff, but creates an opportunity to reap a large second stage payoff. Thus, the incumbent has a strategic incentive to underbid.

On the other hand, if $c_I$ is significantly greater than $\mu_E$, then the incumbent’s Stage 1 strategy switches from underbidding to overbidding. To illustrate, suppose $c_I$ is high and compare the incumbent’s Stage 2 payoff when $c_E$ is concealed vs. revealed. Had $c_E$ been concealed, the incumbent would win in Stage 2 at a low cost-per-click of $\mu_E/c_I$, since $c_I \gg \mu_E$. Conversely, had $c_E$ been revealed, there are two possibilities: if $c_E$ turns out to be low, the incumbent will pay an even lower cost; if $c_E$ turns out to be high, the incumbent will pay a high cost (if not lose the ad position). The latter risk of a high $c_E$ realization outweighs the reward of a low realization because the incumbent’s potential to reap larger margins for a low $c_E$ realization is limited by the reserve price. Therefore, the incumbent has strong incentives to conceal $c_E$ when its CTR is high, and thus bids above valuation in Stage 1. This can also be seen from the following expressions of the incumbent’s Stage 2 profit when $c_I > \mu_E$:
Table 3: Incumbent’s Stage 2 Profit when $c_I > \mu_E$

<table>
<thead>
<tr>
<th>Search engine does not know $c_E$</th>
<th>Search engine knows $c_E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_I(1 - \mu_E/c_I)$</td>
<td>$\int_0^1 c_I(1 - \max(c_E, R)/c_I)\mathbb{I}_{c_I &gt; c_E} dF_E$</td>
</tr>
<tr>
<td>$= c_I - \mu_E$</td>
<td>$= \int_0^R c_I(1 - R/c_I)dF_E + \int_{c_I}^{c_E} c_I(1 - c_E/c_I)dF_E$</td>
</tr>
<tr>
<td>$= \int_0^1 (c_I - c_E) dF_E$</td>
<td>$= \int_0^R (c_I - R)dF_E + \int_{c_I}^{c_E} (c_I - c_E)dF_E$</td>
</tr>
</tbody>
</table>

From Table 3, we see that the incumbent’s Stage 2 profit when the search engine does not know the entrant’s CTR (left-hand side) is $c_I - c_E$ integrated over all values of $c_E$. When $c_E$ is known (right-hand side), for values of $c_E \in (0, R)$, we have $c_I - R$ integrated; since $R > c_E$, the incumbent is better off when the search engine does not know $c_E$ for this integration range. Within the integration range of $c_E \in (R, c_I)$, the expressions on both sides are equal to $c_I - c_E$. Finally, within the range $c_E \in (c_I, 1)$, negative values are integrated on the left-hand side expression whereas the right-hand side expression is zero. For this integration range, the incumbent is better off when the search engine knows $c_E$. Overall, the negative effect of learning $c_E$ on the incumbent’s profit (which happens for $c_E \in (0, R)$) is constant as $c_I$ increases, but the positive effect (which happens for $c_E \in (c_I, 1)$) shrinks as $c_I$ increases. Therefore, a weak incumbent with low $c_I$ is better off in Stage 2 when $c_E$ is learned, whereas a strong incumbent with high $c_I$ is better off when $c_E$ is not learned. This incentivizes a weak (strong) incumbent to underbid (overbid) in Stage 1. We summarize these results in the following proposition.

**Proposition 2** (Advertisers’ Strategies in Stage 1 Under Partial Information). In Stage 1, the entrant always bids above its valuation. The incumbent bids below its valuation if $c_I$ is low, and bids above its valuation if $c_I$ is high.

Figure 2 illustrates the results of Proposition 2 when $c_E$ is drawn from a uniform distribution. The entrant bids above its valuation (represented by the large dashed line) for all values of $c_I$, whereas the incumbent underbids for low values of $c_I$ and overbids for high values. An interesting feature of Figure 2 is that the entrant’s bid is non-monotonic in the incumbent’s CTR. Initially, its bid increases with $c_I$ because as the incumbent’s CTR approaches the
Entrant’s mean CTR, the entrant’s Stage 2 profit when its CTR is not learned, $\mu_E - c_I$, decreases to zero; therefore, revealing its CTR and hoping the realization is higher than expectation becomes a more attractive strategy. As the incumbent’s CTR increases beyond $\mu_E$, however, the chances of winning in Stage 2 when its CTR is learned become slimmer. This implies that the expected benefit of revealing its CTR in Stage 2 decreases, which in turn dampens the entrant’s incentive to overbid.

**Search Engine’s Revenue**

Next, we turn to the implications of partial information (i.e., unknown CTR of the entrant) on the search engine’s revenue. Is the search engine unequivocally better off knowing $c_E$ before the auctions take place? One may conjecture that being more informed about the bidders can only benefit the search engine as the search engine would be able to allocate the ad slot more efficiently to the advertiser that yields higher revenue. Surprisingly, we find that this is not always the case. Under certain conditions, not knowing the entrant’s CTR increases the search engine’s revenue.\(^{13}\)

The intuition revolves around two effects. First, the search engine’s ignorance of the entrant’s CTR induces the entrant to bid more aggressively in Stage 1. As explained above, the incentive to bid higher arises from the fact that the entrant’s expected payoff in Stage 2

\(^{13}\text{To be more precise, the common knowledge that the search engine does not know the entrant’s CTR may increase its revenue.}\)
is higher if the search engine learns its CTR. This forces the incumbent to pay a higher cost-per-click if it wins, which results in higher Stage 1 revenue for the search engine. Note that this effect is most prominent for moderate values of $c_I$.

The second effect is subtler. Consider the search engine’s Stage 2 revenue when $c_I > \mu_E$, and recall that the advertisers bid truthfully in Stage 2. If $c_E$ is not known, the search engine’s expected revenue in Stage 2 is $(\mu_E/c_I)c_I = \mu_E$. Using the definition of $\mu_E$, this can be rewritten as

$$\int_0^1 c_E \, dF_E.$$  \hspace{1cm} (4.3)

If $c_E$ is known, the search engine’s Stage 2 revenue depends on the realization of $c_E$ and can be written as

$$\int_0^R R \, dF_E + \int_{c_I}^{c_E} c_E \, dF_E + \int_{c_I}^1 c_I \, dF_E.$$ \hspace{1cm} (4.4)

Comparing the two integral expressions (4.3) and (4.4), we see that within the integration range $c_E \in (0, R)$, (4.4) is larger; within the range $c_E \in (R, c_I)$, the two expressions are equal, and within the range $c_E \in (c_I, 1)$, (4.3) is larger. Thus, if $c_I$ is not too high, then the search engine’s revenue when it does not know $c_E$ (i.e., Expression (4.3)) is larger than when it does (i.e., Expression (4.4)). Intuitively, since the benefit of a high realization of $c_E$ is bounded from above by $c_I$, i.e., the search engine cannot fully reap the benefits of a high $c_E$, the search engine’s Stage 2 revenue may be higher when $c_E$ is not known than when it is.

Together, these two effects show us that the search engine’s ignorance of the entrant’s CTR is blissful for moderate values of $c_I$. This result is formalized in the following proposition.

**Proposition 3** (Search Engine Revenue: Ignorance is Bliss). The search engine’s revenue is higher not knowing the entrant’s CTR than knowing it if and only if (i) $\underline{c} < c_I < \overline{c}$, or (ii) $c_I \leq \mu_E$ and $\delta < 1$, where $\underline{c}$ and $\overline{c}$ are defined in the appendix.

What does this mean for the search engine? Proposition 3 suggests that search engines do not always have to be concerned about not knowing the new advertisers’ types. In fact,
not knowing the new advertisers’ CTRs can sometimes increase the search engine’s revenue because ignorance induces advertisers to bid aggressively. However, Proposition 3 also reveals conditions under which the search engine’s ignorance can be a curse. For example, if the incumbent is strong (e.g., sufficiently high $c_I$ in Figure 3), then not knowing the entrant’s CTR decreases the search engine’s revenue. This is because, when $c_I$ is sufficiently high, the entrant, who is the “price setter” in the auction, bids less aggressively as $c_I$ increases. Furthermore, when $c_I$ is high, since the search engine does not learn the entrant’s CTR in equilibrium due to the incumbent’s aggressive bidding, it suffers from suboptimal allocation of the ad slot (i.e., missing out on a potentially high $c_E$).

Given that the search engine could sometimes incur a revenue loss for not knowing $c_E$, one may wonder what strategies a search engine could deploy to mitigate this loss. In the next section, we show that offering free ad credit to new advertisers may be a solution. While offering such promotions can be costly for the search engine, our findings suggest that their costs are minimized and their benefits maximized when the incumbent happens to be strong, i.e., the situation in which the ignorance-induced revenue loss is most severe.
Figure 4: Ad Credit by Google and Bing

5 Free Ad Credit for New Advertisers

Search engines run promotions that provide free ad credit to new advertisers. For example, Google offers ad coupons to new advertisers worth up to $75 which can be redeemed within 30 days of spending the first $25 in advertising.14 Similarly, Microsoft’s search engine Bing sends promotional codes to new advertisers, which allow new advertisers to redeem credit toward free advertising (see Figure 4). In this section, we seek to answer the following questions: (i) how does ad credit affect the new and existing advertisers’ bidding strategies? and (ii) what are its implications on the search engine’s revenue?

To that end, we augment the main model by adding the search engine’s ad credit decision prior to Stage 1. That is, the search engine sets the ad credit, denoted by $\alpha \geq 0$, and then the incumbent’s CTR is drawn from $F_I$. The advertisers observe $\alpha$ and the rest of the game proceeds identically as described in the main model. The effect of the search engine’s ad credit is to transfer free ad credit $\alpha$ to the entrant if it wins the first stage auction. Therefore, the entrant’s Stage 1 payoff when the search engine offers ad credit $\alpha$ is

$$\pi_{E1}(\alpha) = \begin{cases} c_E \left( 1 - \frac{\max[c_I b_{I1}, R]}{\mu_E} \right) + \alpha & \text{if } \mu_E b_{E1} > \max[c_I b_{I1}, R], \\ 0 & \text{if } \mu_E b_{E1} \leq \max[c_I b_{I1}, R]. \end{cases}$$

The reason that we assume $\alpha$ is set before $c_I$ is realized is that search engines use the same

14 https://www.google.com/ads/adwords-coupon.html
amount of ad credit across many keywords for which incumbents have different CTRs; that is, in practice, \( \alpha \) is not a function of \( c_I \). We should also note that if \( \alpha \) is decided after \( c_I \) is realized, the truthfulness nature of these second-price auctions will break down. This is because the incumbent would benefit from lowering its bid to induce the search engine to set a lower \( \alpha \). Finally, the assumption that the ad credit is only available in the first stage reflects the fact that these promotions typically expire after a short period of time.

5.1 Advertisers’ Strategies with Ad Credit

We first describe the effect of the ad credit on the advertisers’ bidding strategies. The ad credit does not affect the incumbent’s weakly dominant bidding strategy in (4.1) because it does not change its underlying payment mechanism. The ad credit does, however, increase the entrant’s payoff when it wins in Stage 1, and thus incentivizes the entrant to bid more aggressively in the first stage. We present the advertisers’ bidding strategies with ad credit in the following proposition.

**Proposition 4** (Ad Credit and Advertisers’ Bids). *If the search engine offers ad credit \( \alpha \geq 0 \) to the entrant, then compared to the benchmark bids (4.1) and (4.2), the incumbent’s first stage bid remains unchanged, whereas the entrant’s bid increases by \( \frac{\alpha}{\mu_E} \) to

\[
b^*_E(\alpha) = 1 + \frac{\alpha}{\mu_E} + \frac{\delta}{\mu_E} \left( \int_{c_I}^1 (c_E - c_I) dF_E - (\mu_E - c_I) \mathbb{1}_{(\mu_E > c_I)} \right). \tag{5.1}
\]

5.2 Search Engine’s Revenue with Ad Credit

We turn to the impact of ad credit on the search engine’s revenue. Given the first stage bids of the incumbent and the entrant in (4.1) and (5.1), respectively, the search engine’s
expected revenue as a function of ad credit $\alpha$ is

$$\mathbb{E}[\pi_S(\alpha)] = \int_0^1 \Pi_S(\alpha, c_I) \, dF_I, \quad (5.2)$$

where

$$\Pi_S(\alpha, c_I) = \begin{cases} 
\mu_E b_{E1}^* (\alpha) + \delta \min[c_I, \mu_E] & \text{if } c_I b_{I1}^* \geq \mu_E b_{E1}^* (\alpha), \\
\max[c_I b_{I1}^*, R] - \alpha + \delta \left( \int_0^{c_I} \max[c_E, R] dF_E + (1 - F_E(c_I)) c_I \right) & \text{if } c_I b_{I1}^* < \mu_E b_{E1}^* (\alpha).
\end{cases} \quad (5.3)$$

Note that we express the entrant’s bid as $b_{E1}^* (\alpha)$ to make its dependence on the ad credit explicit. The expression in (5.2) is simply an expectation of the search engine’s revenue over the incumbent’s CTR realization. The top expression in (5.3) represents the sum of the search engine’s first and second stage payoffs when the incumbent wins in Stage 1, such that the entrant’s CTR remains unknown in Stage 2. The bottom expression in (5.3) represents the sum of the first and second stage payoffs when the entrant wins in Stage 1 and its CTR becomes known. In this case, the search engine transfers ad credit $\alpha$ to the entrant.

To help develop intuition for the impact of ad credit on the search engine’s revenue, consider the search engine’s revenue after the incumbent’s CTR has been realized (see Figure 5). If $c_I$ is low such that the entrant wins, then offering ad credit can be costly, as shown by the lower revenue curve with ad credit (large dashed curve). On the other hand, if $c_I$ is high, the benefits of ad credit begin to emerge, as shown by the higher revenue curve. This
suggests that the effect of ad credit on the search engine’s revenue will depend crucially on the strength of the incumbent. Next, we analyze the three effects associated with ad credit, namely, the extraction effect, the ad credit cost, and the learning effect.

**Incumbent Surplus Extraction**

Recall that the entrant’s bid \( b_{E1}^*(\alpha) \) in (5.1) increases in proportion to the ad credit \( \alpha \) (see Proposition 4). This implies that if the incumbent wins in Stage 1, then its cost-per-click, \( \mu_{E} b_{E1}^*(\alpha) \), will increase with \( \alpha \); the search engine can thus extract additional surplus from the incumbent by inflating its cost-per-click using ad credit. We call this the extraction effect. It is defined as the expected increase in the search engine’s Stage 1 payoff due to the ad credit when the incumbent wins:

\[
\int_0^1 (\mu_{E} b_{E1}^*(\alpha) - \mu_{E} b_{E1}^*(0)) \mathbb{1}_{\{c_Ib_{I1}\geq \mu_{E} b_{E1}^*(\alpha)\}} dF_I = \alpha \mathbb{P}\{c_Ib_{I1}\geq \mu_{E} b_{E1}^*(\alpha)\}.
\]

**Ad Credit Cost**

If the entrant wins in Stage 1, the search engine incurs a cost of \( \alpha \), corresponding to the ad credit the search engine transfers to the entrant.

**Learning Entrant’s CTR**

The third and subtler effect concerns the change in the search engine’s Stage 2 payoff due to the ad credit. To illustrate, suppose the incumbent’s CTR is high. In this case, as we saw in Proposition 3 and Figure 3, knowing the entrant’s CTR leads to a higher search engine revenue than that under ignorance. This is due to the more efficient allocation of the ad slot as well as a higher expected payment of the winner. Since offering ad credit could help the entrant win and thus facilitate the search engine learning its CTR, it could also improve the search engine’s Stage 2 revenue. We call this the learning effect. It is defined as the additional Stage 2 revenue the search engine gains from offering ad credit, compared to the
Stage 2 payoff when it does not offer any; i.e., $\delta(\pi_{S2}(\alpha) - \pi_{S2}(0))$ where

$$
\pi_{S2}(\alpha) = \int_{cIb_{1}\geq \mu Eb_{E1}(\alpha)} \min[cI, \mu_E]dF_I \\
+ \int_{cIb_{1}< \mu Eb_{E1}(\alpha)} \left( \int_{0}^{cI} \max[cE, R]dF_E + (1 - F_E(cI))cI \right) dF_I.
$$

These three effects summarize all the pros and cons of offering ad credit in our model. Note that we are not discussing the customer acquisition effect of offering free ad credit. Promotional incentives for attracting new customers have been extensively studied in the literature (e.g., Jedidi et al., 1999; Nijs et al., 2001; van Heerde et al., 2003). Instead, we focus on the extraction and learning effects of ad credit that are new to the literature. In other words, we show that even if the free ad credit does not attract new advertisers, the search engine may still benefit from offering it because of these two positive effects.

In order to assess the net effect of ad credit on the search engine’s revenue, we need to examine how the three forces described above change with respect to (i) the ad credit amount $\alpha$ (Figure 6a) and (ii) the strength incumbent (Figure 6b).\footnote{The plots in Figure 6 are not to scale, and are displayed together for ease of exposition.}

**How the Amount of Ad Credit Affects the Three Forces**

Consider the extraction effect, represented by the solid line in Figure 6a. When $\alpha$ is small, the extraction value increases with the ad credit. The reason is if the incumbent wins, it pays in proportion to the entrant’s bid, which increases with $\alpha$. As $\alpha$ increases further, however, the expected value extracted from the incumbent decreases. This is because as the large ad credit motivates the entrant to bid more aggressively, the probability of the incumbent winning diminishes.

Secondly, consider the learning effect, represented by the small dashed line in Figure 6a. The value of learning the entrant’s CTR increases with ad credit. Intuitively, the higher the ad credit offered to the entrant, the more likely the entrant will win the first stage auction, and thereby reveal its CTR. And for sufficiently high reserve price, knowing the entrant’s CTR
increases the search engine’s second stage revenue since the search engine can capitalize on the event that the entrant’s CTR turns out to be high.\footnote{Proposition 8 in Section A.12 in the appendix provides an analytical characterization of how the three forces change with respect to the ad credit amount.}

**How the Strength of the Incumbent Affects the Three Forces**

Next, we look at how the three forces change as the incumbent becomes stronger (shown in Figure 6b). As \( n \) increases, the incumbent’s CTR, \( c_{i} \), will be high with a higher probability; i.e., the incumbent becomes stronger.\footnote{Using distribution \( F_{I}(c) = c^{n} \) in this example could also be interpreted as the incumbent being the winner (and the survivor) of \( n \) players with uniform distributions who have played this game in the past.} Figure 6b reveals interesting relationships between the three effects and the incumbent’s strength. First, the extraction effect becomes stronger as \( n \) increase because the extraction effect is proportional to the probability of the incumbent winning in Stage 1, which increases with \( n \). The learning effect is inverted U-shaped: it initially increases due to the more efficient allocation of the ad slot as well as a higher expected payment of the winner. For larger \( n \), however, the learning effect decreases because the efficiency gains is offset by the diminished probability of the entrant winning against a strong incumbent. The cost decreases monotonically with \( n \) because the search engine has to pay the ad credit only if the entrant wins, but the probability of the entrant winning decreases as the incumbent becomes stronger. In sum, the three curves in Figure 6b suggest that when the incumbent is sufficiently strong, a likely situation for many keywords in

\footnote{The reason that we use \( n \), instead of \( c_{i} \), to measure the incumbent’s strength is that the ad credit \( \alpha \) is set before \( c_{i} \) is realized. Therefore, it is more intuitive to analyze the effect of \( \alpha \) as a function of \( F_{I} \), the distribution of \( c_{i} \), rather than the realization of \( c_{i} \).}
search advertising, the search engine benefits from offering ad credit. This result is formally presented in Proposition 5.

**Proposition 5.** If the probability of the incumbent’s CTR being high, \( P\{c_I > \bar{c}\} \) (where \( \bar{c} \) is defined in the appendix), is sufficiently high, then offering ad credit increases the search engine’s revenue.

Proposition 5 shows that the search engine should offer positive ad credit if the incumbent’s CTR is likely to be high. The reason is that if the incumbent is strong, the positive effects associated with extraction and learning effects are large while the negative cost effect is small. Put together, the three forces create strong incentives for the search engine to provide ad credit to new advertisers.

Recall from Section 4.2 that the search engine’s ignorance of the entrant’s CTR acted as a double-edged sword. On the one hand, when the incumbent’s CTR was low, the search engine’s ignorance induced advertisers to bid more competitively, resulting in higher revenues compared to the full information case. On the other hand, when the incumbent’s CTR was high, the ignorance led to high opportunity costs for the search engine in terms of missed clicks for the entrant with a potentially high CTR. Thus, an important implication of Proposition 5 is that, if the incumbent’s CTR is likely to be high, then offering ad credit to the entrant may serve as an effective tool for the search engine to mitigate its loss from not knowing the entrant’s CTR. Figure 7 demonstrates the marked reduction in the region for which not knowing \( c_E \) decreases the search engine’s revenue.\(^{19}\)

Note that Proposition 5 is not contingent on any specific functional forms of the CTR distributions \( F_I \) and \( F_E \). The finding that the search engine benefits from offering positive ad credit when the incumbent is likely to be strong holds in general. In particular, it is interesting to note that the result holds in the absence of customer acquisition effects. Thus, Proposition 5 suggests that even in settings where customer acquisition is not a first-order

\(^{19}\)Optimal ad credit was used to generate Figure 7b, i.e., \( \alpha^* (n, \delta) = \operatorname{argmax}_{\alpha \geq 0} \mathbb{E}[\pi_S(\alpha)] \)
managerial concern, search engines may still benefit from offering ad credit.

Optimal Ad Credit. Since we are working with general distribution functions $F_E$ and $F_I$, we cannot calculate a closed-form solution for the optimal ad credit that maximizes the search engine’s revenue. Furthermore, since we are not modeling the customer acquisition effect of ad credit, calculating a closed-form expression may not be managerially relevant. Our goal in this paper is to demonstrate other positive effects of offering free ad credit, particularly the learning effect, that have been overlooked in the literature. Nonetheless, in Appendix B, we provide examples of specific functional forms of $F_E$ and $F_I$, and discuss how the optimal ad credit is affected by exogenous parameters $\delta$ and $n$.

6 Extensions

In developing our model, we made two important simplifying assumptions. First, we assumed that both advertisers have the same valuation per click. Because of this assumption, variation in advertisers’ ad ranks came only from variation in their CTRs. While this assumption accentuates the effect of unknown vs. known CTRs on advertisers’ strategies, it is not a desirable assumption from a practical point of view. In Section 6.1, we relax this assumption and consider advertisers’ valuations drawn from uniform distributions. The insights remain
the same as those in the main model except that we work with ad ranks instead of CTRs (e.g., when defining a weak vs. strong incumbent).

Second, in the main model, we assumed that there is only one ad slot available for each keyword. This assumption leads to the existence of *weakly dominant* bidding strategies. Thus, any deviation from truthful bidding in advertisers’ strategies is caused by whether they want $c_E$ to be learned or not. This assumption allowed us to isolate the effect of learning $c_E$ on advertisers’ bidding strategies from other confounding factors such as advertisers’ strategic responses to each others’ bids. Furthermore, it allowed us to avoid using non-standard equilibrium refinements. Nonetheless, we know that, in practice, each keyword has multiple ad slots. In Section 6.2 we establish the robustness of our results under the Generalized Second Price auction with two ad slots and three advertisers.

### 6.1 Advertisers with Different Valuations

An assumption underlying the main model is that the advertisers share a common per-click valuation of the ad slot, which is normalized to 1. However, it is possible that due to various advertiser-specific characteristics (e.g., conversion rates and profit margins), the advertisers’ valuations may be not only different, but also only privately known. In this section, we extend our main model to allow for independent, private per-click valuations of the ad slot. Thus, we assess the robustness of the previous insights by relaxing the common valuation assumption.

To begin, suppose the advertisers’ valuations are drawn independently from a uniform distribution and are privately observed. In terms of game timing, valuations for both advertisers are realized before Stage 1 and the rest of the game proceeds identically as before.

Our results show that the same forces that drive the advertisers’ bidding patterns under common valuation remain operative under different valuations. In particular, the entrant
always bids above its valuation in Stage 1, because revealing its CTR to the search engine increases its Stage 2 payoff. The incumbent’s bidding pattern is similar, except that the pattern is now dependent on its ad rank, $a_I \triangleq c_I v_I$, instead of its CTR as in the main model. Specifically, if $a_I$ is low, then the incumbent lowers its bid in hopes to reveal a low $c_E$, thereby creating an opportunity to increase its second stage payoff. In contrast, if $a_I$ is high, then the incumbent bids preemptively high to mask the entrant’s CTR. The intuition is analogous to that of the main model. The following proposition summarizes the findings.

**Proposition 6** (Heterogeneous Private Valuations). Suppose the advertisers’ per-click valuations are i.i.d. $U(0, 1)$ and privately observed. In Stage 1, the entrant always bids above its valuation. If $a_I \leq R$, the incumbent bids truthfully. If $a_I > R$, the incumbent bids below its valuation if $a_I$ is low, and bids above its valuation if $a_I$ is high.

### 6.2 Multiple Advertising Slots

Another assumption in the main model is that the search engine offers a single ad slot. In practice, however, search engines sell more than one ad slots, which are allocated via the Generalized Second-Price (GSP) auction. In this section, we test whether the main insights derived from the base model carry over to the multiple-slot GSP setting. We consider a two-slot, three-player game where two incumbents face an entry from a new advertiser. To simplify the analysis, we assume that the reserve price is zero, and that the incumbents also learn $c_E$ if the entrant wins in Stage 1. All advertisers share a common per-click valuation of 1. We normalize position-specific CTR of the first ad slot to 1 and denote that of the second slot as $\theta \in (0, 1)$. We index by $i$ and $I$ the incumbent with the smaller and larger CTR, respectively (i.e., $c_i < c_I$), and normalize $c_I$ to 1. We use the lowest-revenue

---

20 The assumption that the incumbents learn $c_E$ can be justified by the fact that advertisers can estimate the ad rank of the advertisers below them by observing the amount they are charged.

21 The normalization can also be interpreted as assuming that the CTR of the average entrant does not exceed that of the strong incumbent. This assumption simplifies expressions but is not necessary. The analysis without this assumption is provided in the appendix.
envy-free equilibrium of the game (Edelman et al., 2007) for equilibrium selection.

Our analysis shows that the main results are robust to multiple-slot settings under GSP auction. Although the bidding patterns do not carry over perfectly from the single-slot—chiefly due to the change in the equilibrium concept—the insight that the entrant receives a higher payoff in Stage 2 when its CTR is learned by the search engine still holds. Moreover, the findings that (i) a weak incumbent prefers to reveal the entrant’s CTR, and that (ii) a strong incumbent has incentives to mask the entrant’s CTR are preserved in the multiple-slot extension. The GSP extension results are summarized in the proposition below.

**Proposition 7** (Multiple-Slot GSP Auction). The entrant and the incumbent with the lower CTR are always better off in Stage 2 if the entrant’s CTR is learned. The incumbent with the higher CTR is better off in Stage 2 if the entrant’s CTR is learned if and only if (i) $\mu_E < c_i$ and $\theta > \hat{\theta}$; or (ii) $c_i \leq \mu_E$ and $\theta > \frac{1}{2}$, where $\hat{\theta}$ is defined in the appendix.

Proposition 7 reveals interesting nuances related to the position-specific CTR of the second ad slot, $\theta$. For example, if $\theta$ is low, then the incumbent with the higher CTR is better off masking the entrant’s CTR and securing the top position. The intuition is as follows. Had the search engine learned the entrant’s CTR, the strong incumbent becomes exposed to the risk of losing the top position and being allocated to the second slot with low CTR. On the other hand, if $\theta$ is sufficiently high, then the second ad slot is almost as profitable as the first slot. In this case, the strong incumbent benefits from the search engine learning the entrant’s CTR because the incumbent can capitalize on a low $c_E$ realization, while the risk of being driven down to the second ad slot against a high $c_E$ is mitigated by the high $\theta$.

Next, we examine how these advertisers’ incentives are reflected in their Stage 1 bids. Figures 8a and 8b depict the entrant’s bid with respect to the weak incumbent’s CTR, $c_i$. Observe that for high $c_i$, the entrant’s bid follows a similar pattern as the single-slot case (see Figure 2). Again, the intuition is the same as before: if the competing incumbent’s CTR is high, then the entrant can earn positive payoffs if and only if (i) its CTR is learned
by the search engine in Stage 2 and (ii) the realized CTR turns out to be higher than $c_i$. Therefore, when facing a strong incumbent, the entrant bids aggressively in Stage 1 in order to create the opportunity to receive a positive payoff in Stage 2.

Next, consider the entrant’s bidding strategy against a low $c_i$. Recall from Proposition 2 that in the single-slot case, the entrant always bids above valuation. In contrast, we find that under the multiple-slot GSP auction, the entrant may lower its bid when the incumbent’s CTR is low. Although seemingly disparate, the underlying intuition is the same in that the weak incumbent has strong incentives to help the search engine learn the entrant’s CTR. Indeed, Figures 8c and 8d show that a weak incumbent shades its bid in order to help the entrant secure the second ad slot, thereby facilitating the search engine learning $c_E$. And since in the lowest-revenue envy-free equilibrium the advertisers’ bids change in proportion to their competitors’, the entrant shades its bid for low $c_i$.

In short, the discrepancy in the entrant’s bidding strategy between the single- and multiple-slot scenarios arises from the fact that while the advertisers play weakly dominant strategies (i.e., independent of opponents’ bids) in the single-slot auction, the entrant shades its bid in
response to a lower bid from its competitor in the multiple-slot GSP auction.

7 Conclusion

In this paper, we study market entry in search advertising. We investigate how a search engine’s lack of information about a new advertiser’s click-through rate affects the strategies of new and existing advertisers, as well as the search engine. Our theoretical analysis offers useful insights for several issues of managerial importance.

Implications for New Advertisers. We show that when a new advertiser starts search advertising, it should bid aggressively in the beginning, sometimes even above its valuation. The reason is that the new advertiser earns a higher expected future payoff when its CTR is learned by the search engine than when it is not. The fact that the new advertiser’s CTR can only be learned when the advertiser wins sufficiently many auctions provides strong incentives for the new advertiser to bid aggressively until its CTR is learned.

Our results also indicate that a new advertiser should be prepared to, temporarily, pay more than its valuation per click in the beginning. If the advertiser’s CTR turns out to be high, the average cost-per-click will decline over time. In other words, a new advertiser should not leave the market even if the initial cost of advertising is high.

Implications for Existing Advertisers. The entry of a new advertiser has two negative effects for existing advertisers. First, if the new advertiser’s CTR turns out to be high, the existing advertiser risks losing its ad position to the new advertiser. Second, since the search advertising slots are sold in auctions, entry of a new advertiser increases the payment of the existing advertiser. We demonstrate that, in response to these entry effects, an existing advertiser with a high CTR—e.g., branded keyword advertisers—should bid more aggressively to make it harder for the new advertiser to reveal its CTR. On the other hand, an existing advertiser with a low CTR—e.g., lowest-slot advertisers—should lower its bid to
make the revelation process easier. By doing so, the existing advertiser foregoes its short-term profit, but creates an opportunity to earn a larger long-term profit in the event that the new advertiser’s CTR turns out to be low.

**Implications for the Search Engine.** When a new advertiser enters the market, the search engine does not know its CTR; the CTR can only be learned if the new advertiser’s ad is displayed to consumers sufficiently many times. On the surface, it appears that this lack of information about the new advertiser would lead to a suboptimal allocation of the ad slot, and thus lower the search engine’s expected revenue. Surprisingly, our result shows that the ignorance may be a boon to the search engine: its ignorance may incentivize the advertisers to bid more aggressively, which in turn may increase the search engine’s revenue compared to the full information benchmark.

The search engine’s ignorance, however, is not always blissful. In particular, if the existing advertiser’s CTR is high, the lack of information about the new advertiser may hurt the search engine’s long-term revenue. We show the search engine can mitigate this loss by offering free ad credit to new advertisers. Intuitively, the ad credit encourages the new advertisers to bid higher, which increases their chances of winning the auctions. This allows the search engine to learn the new advertisers’ CTR more quickly, which in turn increases the search engine’s long-term revenue.

**Future Research.** Our work is a first step towards understanding how agents strategically respond to a platform’s learning process. Future research could explore other scenarios where agents and platforms interact in a learning environment. For instance, a platform may want to learn sellers’ qualities of products for ranking purposes, or customers’ willingness to pay for pricing purposes. In addition, while we allow the transition from an incomplete to a full information game, we make several simplifying assumptions in doing so. For example, the transition is discrete and binary in our model. Analyzing the advertisers’ strategies in a model with gradual, continuous learning process could lead to interesting additional insights.
References


37
A Proofs

A.1 Proof of Proposition 1

Proof. Consider the incumbent’s second stage payoff as a function of its second stage bid $b_{I2}$, given the entrant’s second stage bid $b_{E2}$.

$$
\pi_{I2}(b_{I2}|b_{E2}) = \begin{cases} 
    c_I \left( 1 - \frac{\max[c_E b_{E2}, R]}{c_I} \right) & \text{if } c_I b_{I2} \geq \max[c_E b_{E2}, R], \\
    0 & \text{if } c_I b_{I2} < \max[c_E b_{E2}, R].
\end{cases}
$$

We will show that truthful bidding weakly dominates bidding below and above valuation. To that end, suppose the incumbent bids below valuation such that $c_I b_{I2} < c_I$. If $\max[c_E b_{E2}, R] \leq c_I$, then truthful bidding ensures a positive payoff of $c_I - \max[c_E b_{E2}, R]$ whereas bidding below valuation yields either the same payoff (if $\max[c_E b_{E2}, R] \leq c_I b_{I2} < c_I$), or a lower payoff of zero (if $c_I b_{I2} < \max[c_E b_{E2}, R] < c_I$). And both strategies yield zero payoff if $c_I < \max[c_E b_{E2}, R]$. Therefore, truthful bidding weakly dominates underbidding.

Next, suppose the incumbent bids above valuation such that $c_I b_{I2} > c_I$. If $\max[c_E b_{E2}, R] \leq c_I$, then both strategies yield the same positive payoff of $c_I - \max[c_E b_{E2}, R]$, and if $c_I b_{I2} < \max[c_E b_{E2}, R]$, then both strategies yield zero payoff as the incumbent loses the auction. On the other hand, if $c_I < \max[c_E b_{E2}, R] \leq c_I b_{I2}$, then truthful bidding yields zero payoff whereas overbidding yields a negative payoff of $c_I - \max[c_E b_{E2}, R]$. Therefore, truthful bidding weakly dominates overbidding.

Given the second stage auction outcome, we analyze the first stage auction. The incumbent’s first stage payoff as a function of its first stage bid, given the entrant’s first and second stage
bids $b_{E1}$ and $b_{E2}$ is

$$
\pi_{I1}(b_{I1}|b_{E1}) = \begin{cases} 
    c_I \left(1 - \frac{\max[c_E b_{E1}, R]}{c_I}\right) + \delta \pi_{I2}(b_{I2} = 1|b_{E2}) & \text{if } c_I b_{I1} \geq \max[c_E b_{E1}, R], \\
    \delta \pi_{I2}(b_{I2} = 1|b_{E2}) & \text{if } c_I b_{I1} < \max[c_E b_{E1}, R],
\end{cases}
$$

where the second stage payoff $\pi_{I2}(b_{I2} = 1|b_{E2} = 1)$ reflects the advertisers’ rational anticipation of truthful bidding equilibrium in the second stage auction.

Since the incumbent’s second stage payoff is the same regardless of the outcome of the first stage auction, it is, in effect, immaterial when the incumbent determines its first stage bidding strategy. Therefore, by the same reasoning as above, we can show that a weakly dominant strategy in the first stage is also truthful bidding. The weak dominance of truthful bidding strategy for the entrant can be shown in a similar manner and is omitted.

Finally, the search engine’s revenue can be computed based on the advertisers’ weakly dominant bids above. In any stage, the search engine receives $c_I \left(\max[c_E, R]/c_I\right)$ if the incumbent wins, and $c_E \left(\max[c_I, R]/c_E\right)$ if the entrant wins. The search engine receives nothing if both advertisers’ ad ranks are below the reserve price. The result follows.

\section*{A.2 Proof of Lemma 1}

\textit{Proof.} Since the weak dominance argument of truthful bidding in the second stage auction under full information in Proposition 1 does not rest on any assumption regarding an advertiser’s knowledge of its competitor’s CTR, the truthful bidding result is preserved under partial information.

\section*{A.3 Proof of Lemma 2}

\textit{Proof.} Consider the incumbent’s payoff. If the incumbent wins the first stage auction, then it pays $\max[\mu_E b_{E1}, R]/c_I$ and the entrant’s CTR remains unknown. Thus, in the second stage
On the other hand, if the incumbent does not win the first stage auction, then its first stage payoff is zero; however, it has a chance of winning the second stage auction in the event that \( c_E \) turns out to be very low. Specifically, its second stage payoff conditional on losing the first stage auction is \( \int_0^1 c_I \left(1 - \max[c_E, R]/c_I\right) \mathbb{I}_{c_I \geq \max[c_E, R]} dF_E \). In sum, the incumbent’s payoff as a function of its bid is

\[
\mathbb{E}[\pi_I] = \begin{cases} 
    c_I \left(1 - \frac{\max[c_E, R]}{c_I}\right) + \delta c_I \left(1 - \frac{\mu_E}{c_I}\right) \mathbb{I}_{c_I \geq \mu_E} & \text{if } b_{I1} \geq \frac{\max[c_E, R]}{c_I}, \\
    \delta \int_0^1 c_I \left(1 - \frac{\max[c_E, R]}{c_I}\right) \mathbb{I}_{c_I \geq \max[c_E, R]} dF_E & \text{if } b_{I1} < \frac{\max[c_E, R]}{c_I}.
\end{cases}
\]

Therefore, a weakly dominant first stage bid for the incumbent is

\[
b_{I1}^* = \frac{1}{c_I} \left( c_I + \delta c_I \left(1 - \frac{\mu_E}{c_I}\right) \mathbb{I}_{c_I \geq \mu_E} \right) - \delta \int_0^1 c_I \left(1 - \frac{\max[c_E, R]}{c_I}\right) \mathbb{I}_{c_I \geq \max[c_E, R]} dF_E.
\]

Next, consider the entrant’s payoff. If the entrant wins the first stage auction, then it pays \( \max[c_I b_{I1}, R]/\mu_E \) and the entrant’s CTR becomes known. Therefore, in the second stage auction, the entrant’s expected payoff is \( \int_0^1 c_E \left(1 - \max[c_I, R]/c_E\right) \mathbb{I}_{c_E \geq \max[c_I, R]} dF_E \). On the other hand, if the entrant does not win the first stage auction, then its first stage payoff is zero and its CTR remains unknown. However, it still has a chance of winning the second stage auction when the incumbent’s CTR is sufficiently low. Specifically, the entrant’s second stage payoff conditional on losing the first stage auction is \( c_E \left(1 - \max[c_I, R]/\mu_E\right) \mathbb{I}_{\mu_E > c_I} \).

In sum, the entrant’s payoff as a function of its bid is

\[
\mathbb{E}[\pi_E] = \begin{cases} 
    \mu_E \left(1 - \frac{\max[c_I b_{I1}, R]}{\mu_E}\right) + \delta \int_{\max[c_I, R]}^1 c_E \left(1 - \frac{\max[c_I, R]}{c_E}\right) dF_E & \text{if } b_{E1} > \frac{\max[c_I b_{I1}, R]}{\mu_E}, \\
    \delta \mu_E \left(1 - \frac{\max[c_I, R]}{\mu_E}\right) \mathbb{I}_{\mu_E > c_I} & \text{if } b_{E1} \leq \frac{\max[c_I b_{I1}, R]}{\mu_E}.
\end{cases}
\]

(A.1)
Therefore, a weakly dominant first stage bid for the entrant is

\[ b^*_{E_1} = 1 + \frac{\delta}{\mu_E} \left( \int_{\max[c_l,R]}^{1} (c_E - \max[c_l,R]) dF_E - (\mu_E - \max[c_l,R]) \mathbb{I}_{(\mu_E > c_l)} \right). \]

The expressions simplify with the assumption that \( c_l > R \). ■

### A.4 Proof of Proposition 2

We first prove two intermediary results which will be used for the proof.

**Claim 1.** Suppose a differentiable function \( f(x) \) is single-peaked on the interval \([a,b]\) (i.e., there exists some \( \xi \in (a,b) \) such that \( f'(x) \geq 0 \) for all \( x \leq \xi \) and \( f'(x) \leq 0 \) for all \( x \geq \xi \) and \( f(a) < 0 < f(b) \). Then there exists a pair \( \tilde{x}_1 \leq \tilde{x}_2 \) in \((a,b)\) such that (i) \( f(x) < 0 \) for all \( x \in [a, \tilde{x}_1] \), (ii) \( f(x) = 0 \) for all \( x \in [\tilde{x}_1, \tilde{x}_2] \), and (iii) \( f(x) > 0 \) for all \( x \in (\tilde{x}_2, b] \).

**Proof of Claim.** By the Intermediate Value Theorem (IVT), there must exist at least one root in the interval \((a,b)\). From the set of roots (which could be a singleton), let \( \tilde{x}_1 \) be the smallest; i.e., \( \tilde{x}_1 \triangleq \min\{x \in (a,b) : f(x) = 0\} \). By definition of \( \tilde{x}_1 \), we have \( f(x) < 0 \) for all \( x \in [a, \tilde{x}_1] \). Next, let \( \tilde{x}_2 \triangleq \min\{x \in [\tilde{x}_1, b) : f'(x) > 0\} \). The existence of \( \tilde{x}_2 \) is guaranteed by \( f(b) > 0 \) and continuity of \( f(x) \). Then \( f(x) = 0 \) for all \( x \in [\tilde{x}_1, \tilde{x}_2] \). To see this, suppose to the contrary that either \( f(x) < 0 \) or \( f(x) > 0 \) for some \( x' \in [\tilde{x}_1, \tilde{x}_2] \). In the first case, single-peakedness implies \( f(x) < 0 \) for all \( x \geq x' \), which contradicts \( f(b) > 0 \). In the second case, continuity of \( f(x) \) implies that there exists some \( x'' \in [\tilde{x}_1, x') \) such that \( f'(x) > 0 \). This contradicts the definition of \( \tilde{x}_2 \).

Finally, \( f(x) > 0 \) for all \( x \in (\tilde{x}_2, b] \). For otherwise, if \( f(x) \leq 0 \) for some \( x \in (\tilde{x}_2, b] \), then it must be that \( f(x) \) had crossed the \( x \)-axis from above for some \( x'' \in (\tilde{x}_2, b] \); but then, single-peakedness implies \( f(x) < 0 \) for all \( x > x'' \), which again contradicts \( f(b) > 0 \). This completes the proof. ■
Claim 2. If $F_E(c_E)$ is continuous on the interval $(0,1)$, then

$$\frac{\partial}{\partial c_I} \int_{c_I}^{1} 1 - \frac{c_E}{c_I} dF_E = \frac{1}{c_I^3} \int_{c_I}^{1} c_E dF_E.$$ 

Proof of Claim 2. At $c_I = 1$, we have $\frac{\partial}{\partial c_I} \int_{c_I}^{1} 1 - \frac{c_E}{c_I} dF_E = \frac{\partial}{\partial c_I} 0 = 0 = \frac{1}{(1)^2} \int_{1}^{1} c_E dF_E$. Next, suppose $c_I \in [0,1)$. We have

$$\frac{\partial}{\partial c_I} \int_{c_I}^{1} 1 - \frac{c_E}{c_I} dF_E = \lim_{\delta \to 0} \frac{1}{\delta} \left( \int_{c_I+\delta}^{1} 1 - \frac{c_E}{c_I + \delta} dF_E - \int_{c_I}^{1} 1 - \frac{c_E}{c_I} dF_E \right)$$

$$= \lim_{\delta \to 0} \left( \int_{c_I+\delta}^{1} \frac{c_E}{c_I(c_I+\delta)} dF_E - \frac{1}{\delta} \int_{c_I}^{c_I+\delta} 1 - \frac{c_E}{c_I} dF_E \right). \quad (A.2)$$

The limit operator can be distributed inside the bracket because the limit values of the summands converge. For the first summand, $\lim_{\delta \to 0} \int_{c_I+\delta}^{1} \frac{c_E}{c_I(c_I+\delta)} dF_E = \frac{1}{c_I} \int_{c_I}^{1} c_E dF_E$. On the other hand, the second summand converges to zero. To see this, note that the right limit (i.e., $\delta \to 0^+$) of the second summand is “squeezed” from below and above by values which converge to zero:

$$\frac{1}{\delta} \int_{c_I}^{c_I+\delta} 1 - \frac{c_I + \delta}{c_I} dF_E \leq \frac{1}{\delta} \int_{c_I}^{c_I+\delta} 1 - \frac{c_E}{c_I} dF_E \leq \frac{1}{\delta} \int_{c_I}^{c_I+\delta} 1 - \frac{c_I}{c_I} dF_E.$$ 

The left-most term simplifies to $-\int_{c_I}^{c_I+\delta} \frac{1}{c_I} dF_E = -\frac{1}{c_I} (F(c_I + \delta) - F(c_I))$, which vanishes to zero as $\delta \to 0^+$ due to right-continuity of $F_E$. The right-most term also vanishes to zero because the term inside the integral is zero.

Similarly, the left limit (i.e., $\delta \to 0^-$) can be shown to converge to zero by bounding the integral from below and above and then invoking the continuity assumption.
Therefore,

\[
(A.2) = \lim_{\delta \to 0} \left( \int_{c_1 + \delta}^{1} \frac{c_E}{c_I (c_1 + \delta)} \, dF_E - \frac{1}{\delta} \int_{c_1}^{c_1 + \delta} 1 - \frac{c_E}{c_I} \, dF_E \right)
\]

\[
= \lim_{\delta \to 0} \int_{c_1 + \delta}^{1} \frac{c_E}{c_I (c_1 + \delta)} \, dF_E - \lim_{\delta \to 0} \frac{1}{\delta} \int_{c_1}^{c_1 + \delta} 1 - \frac{c_E}{c_I} \, dF_E
\]

\[
= \frac{1}{c_I^2} \int_{c_1}^{1} c_E \, dF_E.
\]

\[\blacksquare\]

**Proof of Proposition 2.** Consider the incumbent’s bid. Whether the bid is below or above valuation depends on the sign of

\[g(c_I) \triangleq \left( 1 - \frac{\mu_E}{c_I} \right) I_{\{c_I \geq \mu_E\}} - F_E(R) \left( 1 - \frac{R}{c_I} I_{\{c_I \geq R\}} \right) - \int_{R}^{1} \left( 1 - \frac{c_E}{c_I} \right) I_{\{c_I \geq c_E\}} \, dF_E.\]

If \(0 \leq c_I \leq R\), then \(g(c_I) = 0\), so the incumbent bids truthfully. If \(R < c_I \leq \mu_E\), then \(g(c_I) = -F_E(R) \left( 1 - \frac{R}{c_I} \right) - \int_{R}^{c_I} 1 - \frac{c_E}{c_I} \, dF_E\). And since \(0 < \mathbb{P}\{c_E \leq R\}\), we have \(g(c_I) < 0\), which means that the incumbent bids below valuation. Finally, if \(\mu_E < c_I \leq 1\), then

\[g(c_I) = \left( 1 - \frac{\mu_E}{c_I} \right) - F_E(R) \left( 1 - \frac{R}{c_I} \right) - \int_{R}^{c_I} 1 - \frac{c_E}{c_I} \, dF_E.\]  \hspace{1cm} (A.3)

We will show that \((A.3)\) satisfies the properties of Claim 1, thereby proving that there exists a pair of thresholds \(\tilde{c}_1 \leq \tilde{c}_2\) in \((\mu_E, 1)\) that satisfies the properties stated in the proposition.
(i) Differentiability

\[ g'(c_I) = \frac{\partial}{\partial c_I} \left( 1 - \frac{\mu_E}{c_I} - \int_0^R 1 - \frac{R}{c_I} dF_E - \int_{c_I}^R 1 - \frac{c_E}{c_I} dF_E \right) \]

\[ = \frac{\partial}{\partial c_I} \left( \int_0^1 1 - \frac{c_E}{c_I} dF_E - \int_0^1 1 - \frac{\max[R, c_E]}{c_I} dF_E + \int_{c_I}^1 1 - \frac{\max[R, c_E]}{c_I} dF_E \right) \]

\[ = \frac{\partial}{\partial c_I} \left( \frac{1}{c_I} \int_0^1 \left( \max[R, c_E] - c_E \right) dF_E + \int_{c_I}^1 1 - \frac{c_E}{c_I} dF_E \right) \]

\[ = -\frac{1}{c_I^2} \int_0^1 \max[R - c_E, 0] dF_E + \frac{\partial}{\partial c_I} \int_{c_I}^1 1 - \frac{c_E}{c_I} dF_E \quad (\because R < c_I) \]

\[ = -\frac{1}{c_I^2} \int_0^1 \max[R - c_E, 0] dF_E + \frac{1}{c_I} \int_{c_I}^1 c_E dF_E, \quad (A.4) \]

where the last equality follows from Claim 2. Since the derivative is well-defined for all \( c_I \in (0, 1) \), we conclude that \( g(c_I) \) is differentiable.

(ii) Single-peakedness

From \( (A.4) \), it follows that the sign of \( g'(c_I) \) is equal to the sign of \( h(c_I) \triangleq \int_{c_I}^1 c_E dF_E - \int_0^1 \max[R - c_E, 0] dF_E \). At \( c_I = 0^+ \)\(^{22}\) \( h \) is positive because

\[ \int_{c_I}^1 c_E dF_E - \int_0^1 \max[R - c_E, 0] dF_E = \int_1^1 c_E dF_E - \int_0^1 \max[R - c_E, 0] dF_E \]

\[ = \int_0^1 c_E - \max[R - c_E, 0] dF_E \]

\[ \geq \int_0^1 c_E - RdF_E \]

\[ = \mu_E - R > 0. \]

At \( c_I = 1 \), \( h \) is negative because

\[ \int_{c_I}^1 c_E dF_E - \int_0^1 \max[R - c_E, 0] dF_E = \int_1^1 c_E dF_E - \int_0^1 \max[R - c_E, 0] dF_E \]

\[ = -\int_0^1 \max[R - c_E, 0] dF_E < 0. \]

\(^{22}\)We evaluate at the right-limit \( 0^+ \) because \( g'(c_I) \) is undefined at \( c_I = 0 \).
Finally, \( h(c_I) \) is non-increasing because for any \( \delta > 0 \),

\[
\begin{align*}
  h(c_I + \delta) - h(c_I) &= \int_{c_I + \delta}^{1} c_E dF_E - \int_{c_I}^{1} c_E dF_E \\
  &= \int_{c_I + \delta}^{1} c_E dF_E - \int_{c_I + \delta}^{1} c_E dF_E - \int_{c_I}^{c_I + \delta} c_E dF_E \\
  &= -\int_{c_I}^{c_I + \delta} c_E dF_E \leq 0.
\end{align*}
\]

In total, since \( h(c_I) \) is non-increasing in \([0, 1]\), \( h(0) \geq 0 \), and \( h(1) \leq 0 \), by the IVT, there exists a \( \xi \in (0, 1) \) such that \( h(c_I) \geq 0 \) for all \( c_I \leq \xi \) and \( h(c_I) \leq 0 \) for all \( c_I \geq \xi \). By the sign equivalence, we have \( g'(c_I) \geq 0 \) for all \( c_I \leq \xi \) and \( g'(c_I) \leq 0 \) for all \( c_I \geq \xi \).

(iii) Endpoint values

\[
\begin{align*}
g(\mu_E) &= 1 - \frac{\mu_E}{\mu_E} - F_E(R) \left( 1 - \frac{R}{\mu_E} \right) - \int_{\mu_E}^{R} \left( 1 - \frac{c_E}{\mu_E} \right) dF_E \\
  &= -F_E(R) \left( 1 - \frac{R}{\mu_E} \right) - \int_{\mu_E}^{R} \left( 1 - \frac{c_E}{\mu_E} \right) dF_E < 0 \\
g(1) &= 1 - \mu_E - F_E(R) (1 - R) - \int_{R}^{1} 1 - c_E dF_E \\
  &= 1 - \mu_E - \int_{0}^{R} 1 - RdF_E - \int_{R}^{1} 1 - c_E dF_E \\
  &= 1 - \mu_E - \int_{0}^{1} 1 - \max[R, c_E] dF_E \\
  &= \int_{0}^{1} \max[R, c_E] - c_E dF_E \\
  &= \int_{0}^{R} R - c_E dF_E > 0
\end{align*}
\]

Therefore, \( g(c_I) \) satisfies the properties of Claim 1 which implies that there exists a pair \( \bar{c}_1 \leq \bar{c}_2 \) in \((\mu_E, 1)\) such that \( g(c_I) < 0 \) for all \( c_I \in (\mu_E, \bar{c}_1) \), \( g(c_I) = 0 \) for all \( c_I \in [\bar{c}_1, \bar{c}_2] \), and \( g(c_I) > 0 \) for all \( c_I \in (\bar{c}_2, 1] \). This, in turn, implies that the incumbent bids below valuation, truthfully, and above valuation for \( c_I \in (\mu_E, \bar{c}_1) \), \( c_I \in [\bar{c}_1, \bar{c}_2] \), and \( c_I \in (\bar{c}_2, 1] \), respectively.
Second, consider the entrant’s bid

\[ b^*_E(c_I) = 1 + \frac{\delta}{\mu_E} \left( \int_{\max[c_I, R]}^{1} (c_E - \max[c_I, R]) \, dF_E - (\mu_E - \max[c_I, R]) \mathbb{I}_{\{\mu_E > \max[c_I, R]\}} \right) \]

Whether the entrant bids below or above valuation depends on the sign of

\[ k(c_I) \triangleq \int_{\max[c_I, R]}^{1} (c_E - \max[c_I, R]) \, dF_E - (\mu_E - \max[c_I, R]) \mathbb{I}_{\{\mu_E > \max[c_I, R]\}}. \]

If \( \mu_E \leq \max[c_I, R] \), then \( k(c_I) = \int_{\max[c_I, R]}^{1} c_E - \max[c_I, R] \, dF_E \geq 0 \), where the inequality holds with equality only if \( c_I = 1 \). If \( \mu_E > \max[c_I, R] \), then

\[
\begin{align*}
    k(c_I) &= \int_{\max[c_I, R]}^{1} c_E - \max[c_I, R] \, dF_E - (\mu_E - \max[c_I, R]) \mathbb{I}_{\{\mu_E > \max[c_I, R]\}} \\
    &= -\int_{\max[c_I, R]}^{1} c_E - \max[c_I, R] \, dF_E > 0.
\end{align*}
\]

This completes the proof. \[\blacksquare\]

### A.5 Proof of Proposition 3

**Proof.** First, we will show that there exists a unique \( \hat{c} \in (\mu_E, 1) \) such that the entrant wins the first stage auction for \( c_I < \hat{c} \), and the incumbent wins otherwise. To that end, consider the difference in ad ranks \( D_A(c_I) \triangleq c_I b^*_I(c_I) - \mu_E b^*_E(c_I) \). If \( c_I \leq R \), then

\[
D_A(c_I) = c_I - \mu_E + \delta \left( \mu_E - R - \int_{R}^{1} c_E - RdF_E \right) \\
= c_I - \mu_E + \delta \left( \int_{0}^{1} c_E - RdF_E - \int_{R}^{1} c_E - RdF_E \right) \\
= c_I - \mu_E + \delta \left( \int_{0}^{R} c_E - RdF_E + \int_{R}^{1} c_E - RdF_E - \int_{R}^{1} c_E - RdF_E \right) \\
= -(\mu_E - c_I) - \delta \int_{0}^{R} R - c_E dF_E < 0.
\]

46
Therefore, the entrant wins the first stage auction for all \( c_I \leq R \). Note that the entrant also beats the reserve price because \( b^*_E(c_I) \geq 1 \) (cf. Lemma 2) and \( \mu_E > R \).

If \( R < c_I < \mu_E \), then

\[
D_A(c_I) = c_I - \mu_E - \delta \int_{c_I}^{1} c_E - c_I dF_E + \delta \left( c_I - \mu_E - \int_{0}^{1} (c_I - \max[c_E, R]) \mathbb{I}_{(c_I \geq \max[c_E, R])} dF_E \right) < 0.
\]

Therefore, the entrant wins the first stage auction in this interval as well.

Finally, if \( \mu_E \leq c_I \leq 1 \), then

\[
D_A(c_I) = (1 + \delta)(c_I - \mu_E) - \delta \int_{c_I}^{1} c_E - c_I dF_E - \delta \int_{0}^{1} (c_I - \max[c_E, R]) \mathbb{I}_{(c_I \geq \max[c_E, R])} dF_E.
\]

The derivative of \( D_A(c_I) \) with respect to \( c_I \) is

\[
D'_A(c_I) = 1 + \delta(1 - F_E(c_I)) + \delta \left( 1 - \frac{\partial}{\partial c_I} \int_{0}^{1} (c_I - \max[c_E, R]) \mathbb{I}_{(c_I \geq \max[c_E, R])} dF_E \right)
\]

\[
= \begin{cases} 
1 + \delta(1 - F_E(c_I)) + \delta (1 - 1) & \text{if } c_I \geq \max[c_E, R], \\
1 + \delta(1 - F_E(c_I)) + \delta (1 - 0) & \text{if } c_I < \max[c_E, R],
\end{cases}
\]

Therefore, \( D_A(c_I) \) is strictly increasing in the interval \([\mu_E, 1]\). Combined with the fact that

\[
D_A(\mu_E) = -\delta \int_{\mu_E}^{1} c_E - \mu_E dF_E - \delta \int_{0}^{1} (\mu_E - \max[c_E, R]) \mathbb{I}_{(\mu_E \geq \max[c_E, R])} dF_E < 0
\]

\[
D_A(1) = 1 - \mu_E + \delta \left( 1 - \mu_E - \int_{0}^{1} 1 - \max[c_E, R] dF_E \right)
\]

\[
= 1 - \mu_E + \delta \left( \int_{0}^{1} 1 - c_E dF_E - \int_{0}^{1} 1 - \max[c_E, R] dF_E \right)
\]

\[
= 1 - \mu_E + \delta \int_{0}^{1} \max[c_E, R] - c_E dF_E > 0,
\]

we have, by the IVT, a unique \( \hat{c} \in (\mu_E, 1) \) such that \( D_A(c_I) < 0 \) for all \( c_I < \hat{c} \) and \( D_A(c_I) > 0 \) for all \( c_I > \hat{c} \). More generally, combining the results from the intervals above yield that the entrant wins the first stage bid for all \( c_I < \hat{c} \) and the incumbent wins for all \( c_I \geq \hat{c} \).

Next, we characterize the search engine’s expected payoff when it knows the entrant’s CTR
before the auctions take place. If the entrant’s CTR is known, then advertisers bid truthfully in each round. Therefore, the search engine’s total expected revenue under full information is 
\[ \mathbb{E}[\pi_S^F] = (1 + \delta) \left( \int_m^1 m dF_E + \int_0^1 \max[c_E, R] \mathbb{I}_{c_l \geq \max[c_E, R]} dF_E \right) \] , where \( m \triangleq \max[c_l, R] \). On the other hand, if \( c_E \) is \emph{a priori} unknown to all players and is revealed only if the entrant wins an auction, then the search engine’s expected revenue is 
\[
\mathbb{E}[\pi_S] = \begin{cases} 
\mu_E b_{E1}^* + \delta \max[\min[c_l, \mu_E], R] & \text{if } c_l b_{I1}^* \geq \mu_E b_{E1}^*, \\
\max[c_l b_{I1}^*, R] + \delta \left( \int_m^1 m dF_E + \int_0^1 \max[c_E, R] \mathbb{I}_{c_l \geq \max[c_E, R]} dF_E \right) & \text{if } c_l b_{I1}^* < \mu_E b_{E1}^*.
\end{cases}
\] (A.5)

Next, define the difference \( D_\pi(c_l) \triangleq \mathbb{E}[\pi_S] - \mathbb{E}[\pi_S^F] \). The set of \( c_l \) for which \( D_\pi(c_l) > 0 \) corresponds to the region where the search engine’s revenue is higher when it does not know \( c_E \) than when it does. To formally characterize this set, consider first the case when \( 0 \leq c_l \leq R \). In this case, the difference simplifies to 
\[ D_\pi(c_l) = \max[c_l, R] - \int_R^1 R dF_E \geq R - (1 - F_E(R)) R = RF_E(R) > 0. \]
Therefore, ignorance yields higher revenue than full information for \( c_l \leq R \).

If \( R < c_l \leq \mu_E \), then 
\[ D_\pi(c_l) = \max[c_l - \delta \int_0^{c_l} c_l - \max[c_E, R], R] - \left( \int_0^{c_l} \max[c_E, R] dF_E + \int_{c_l}^1 c_l dF_E \right), \]
which is positive if and only if
\[ c_l - \int_R^{c_l} c_E dF_E - \int_{c_l}^1 c_I dF_E - \int_0^R R dF_E - \delta \left( \int_R^{c_l} c_l - c_E dF_E + \int_0^R c_l - R dF_E \right) > 0. \]
But the expression on the left-hand side is 0 at \( c_l = R \), and its derivative with respect to \( c_l \) is \( (1 - \delta) F_E(c_l) \). This implies that if \( \delta < 1 \), then \( D_\pi(c_l) > 0 \) for all \( R < c_l \leq \mu_E \), and if \( \delta > 1 \), then \( D_\pi(c_l) < 0 \) for all \( R < c_l \leq \mu_E \).

If \( \mu_E < c_l < \hat{c} \), then
\[
D_\pi(c_l) = \max \left[ c_l + \delta \left( c_l - \mu_E - \int_0^{c_l} c_l - \max[c_E, R] dF_E \right), R \right] - \left( \int_0^{c_l} \max[c_E, R] dF_E + \int_{c_l}^1 c_l dF_E \right),
\]
which is positive if and only if

\[ c_I \left( \int_R^{c_I} c_E dF_E - \int_0^{c_I} c_I dF_E - \int_0^R RdF_E + \delta \left( c_I - \mu_E - \int_R^{c_I} c_I c_E dF_E - \int_0^R c_I - RdF_E \right) \right) > 0. \]

Now, the expression on the left-hand side is strictly increasing in \( c_I \) because \( \frac{\partial}{\partial c_I} LHS = F_E(c_I) + \delta(1 - F_E(c_I)) > 0 \). Furthermore, the difference is negative and positive at \( c_I = R \) and \( c_I = 1 \), respectively:

\[
D_\pi(R) = \delta(R - \mu_E) < 0
\]

\[
D_\pi(1) = 1 - \int_R^1 c_E dF_E - \int_0^R RdF_E + \delta \left( 1 - \mu_E - \int_R^1 1 - c_E dF_E - \int_0^R 1 - RdF_E \right)
= 1 - \int_0^1 \max[c_E, R] dF_E + \delta \left( \int_0^1 1 - c_E dF_E - \int_0^1 1 - \max[c_E, R] dF_E \right)
= 1 - \int_0^1 \max[c_E, R] dF_E + \delta \int_0^1 \max[c_E, R] - c_E dF_E > 0.
\]

Therefore, by the IVT, there exists a unique \( \tilde{c}_1 \in (r, 1) \) such that \( D_\pi(c_I) < 0 \) for all \( c_I \in (R, \tilde{c}_1) \) and \( D_\pi(c_I) > 0 \) for all \( c_I \in (\tilde{c}_1, 1) \). However, the interval in question here is \( (\mu_E, \hat{c}) \), so we re-define the threshold as \( \bar{c} \triangleq \max[\mu_E, \min[\hat{c}, \tilde{c}_1]] \).

Finally, if \( \hat{c} \leq c_I \leq 1 \), then the incumbent wins the first stage auction and the difference in payoffs between the uncertain and full information cases is

\[
D_\pi(c_I) = (1 + \delta) \left( \mu_E - \left( \int_0^R RdF_E + \int_R^{c_I} c_E dF_E + \int_0^{c_I} c_I dF_E \right) \right) + \delta \int_{c_I}^1 c_E - c_I dF_E.
\]
Similarly as above, we invoke the IVT to prove the unique existence of a root. We have

\[ D'(c_I) = -(1 + 2\delta)(1 - F_E(c_I)) < 0 \]
\[ D_E(\mu_E) = (1 + \delta) \left( \mu_E - \left( \int_R^R \mu_E \ c_E \ dF_E + \int_{\mu_E}^1 \mu_E \ c_E \ dF_E \right) \right) + \delta \int_{\mu_E}^1 c_E - \mu_E dF_E \]
\[ = (1 + \delta) \left( \int_R^R \mu_E - c_E dF_E + \int_{\mu_E}^R \mu_E - R dF_E \right) + \delta \int_{\mu_E}^1 c_E - \mu_E dF_E > 0 \]
\[ D_E(1) = (1 + \delta) \left( \mu_E - \int_0^1 c_E dF_E - \int_R^R R dF_E \right) \]
\[ = (1 + \delta) \int_0^R c_E - R dF_E < 0. \]

Therefore, by the IVT, there exists a unique \( \tilde{c}_2 \in (\mu_E, 1) \) such that \( D_E(c_I) > 0 \) for all \( c_I \in (\mu_E, \tilde{c}_2) \) and \( D_E(c_I) < 0 \) for all \( c_I \in (\tilde{c}_2, 1) \). However, the interval in question here is \([\hat{c}, 1]\), so we bound the threshold as \( \overline{c} \triangleq \max[\hat{c}, \tilde{c}_2] \).

Putting together all the sets for which \( D_E > 0 \) yields the result.

\[ \square \]

A.6 Proof of Proposition 4

Proof. From (A.1), we have

\[ \mathbb{E}[\pi_E] = \begin{cases} \mu_E \left( 1 - \frac{\max[c_Ib_{11}, R]}{\mu_E} \right) + \delta \int_{\max[c_I, R]}^1 \ c_E \left( 1 - \frac{\max[c_I, R]}{c_E} \right) dF_E + \alpha & \text{if } b_{E1} > \frac{\max[c_Ib_{11}, R]}{\mu_E}, \\
\delta \mu_E \left( 1 - \frac{\max[c_I, R]}{\mu_E} \right) \mathbb{I}_{\{\mu_E > c_I\}} & \text{if } b_{E1} \leq \frac{\max[c_Ib_{11}, R]}{\mu_E}. \end{cases} \]

Therefore, a weakly dominant first stage bid for the entrant is

\[ b_{E1}^*(\alpha) = 1 + \frac{\delta}{\mu_E} \left( \int_{\max[c_I, R]}^1 (c_E - \max[c_I, R]) dF_E - (\mu_E - \max[c_I, R]) \mathbb{I}_{\{\mu_E > c_I\}} \right) + \frac{\alpha}{\mu_E}. \]

\[ \square \]
A.7 Proof of Proposition \textsuperscript{5}

Proof. We will derive a condition equivalent to $\frac{\partial}{\partial \alpha} \mathbb{E}[\pi_S(\alpha)]|_{\alpha=0} > 0$, which is a sufficient condition for the search engine to provide positive ad credit.

First, to simplify notation, let $\tilde{b}_E \triangleq \mu_E + \delta \left( \int_{\max[c_I, R]}^{1} (c_E - \max[c_I, R]) dF_E - (\mu_E - \max[c_I, R]) \mathbb{I}_{\{\mu_E > c_I\}} \right)$, so that the entrant’s first stage ad rank is $\mu_E b_{E1}^* = \alpha + \tilde{b}_E$. Based on this, denote the terms in the search engine’s revenue expression that are independent of $\alpha$ by

$$\xi_1 \triangleq \tilde{b}_E + \delta \max[\min[c_I, \mu_E], R],$$

$$\xi_2 \triangleq \max[c_I b_{I1}^*, R] + \delta \left( \int_{0}^{c_I} \max[c_E, R] dF_E + (1 - F_E(\max[c_I, R])) \max[c_I, R] \right).$$

Then, we have

$$\Pi_S(\alpha, c_I) = \begin{cases} 
\alpha + \xi_1 & \text{if } c_I \geq \min[\hat{c}(\alpha), 1], \\
-\alpha + \xi_2 & \text{if } c_I < \min[\hat{c}(\alpha), 1], 
\end{cases}$$

where $\hat{c}(\alpha)$ is the value such that $c_I b_{I1}^* \geq \mu_E b_{E1}^*(\alpha)$ for all $c_I \geq \hat{c}(\alpha)$, and $c_I b_{I1}^* < \mu_E b_{E1}^*(\alpha)$ for all $c_I < \hat{c}(\alpha)$. Note that we can show the unique existence of such threshold following a similar reasoning as in the first part of the proof of Proposition \textsuperscript{3}. Using these notations, we can re-write the search engine’s revenue as follows: $\mathbb{E}[\pi_S(\alpha)] = \int_{0}^{1} \Pi_S(\alpha, c_I) dF_I = \int_{0}^{\min[\hat{c}(\alpha), 1]} -\alpha + \xi_2 dF_I + \int_{\min[\hat{c}(\alpha), 1]}^{1} \alpha + \xi_1 dF_I$. Therefore, if $\hat{c}(\alpha) > 1$, which occurs if and only if $\alpha > 1 - \mu_E + \delta \int_{0}^{1} \max[c_E, R] - c_E dF_E$, then the derivative with respect to $\alpha$ is $\frac{\partial}{\partial \alpha} \mathbb{E}[\pi_S(\alpha)] = \frac{\partial}{\partial \alpha} \int_{0}^{1} -\alpha + \xi_2 dF_I = \frac{\partial}{\partial \alpha} (-\alpha) = -1$. On the other hand, if $\hat{c}(\alpha) \leq 1$, then

$$\frac{\partial}{\partial \alpha} \mathbb{E}[\pi_S(\alpha)] = \frac{\partial}{\partial \alpha} \int_{0}^{\hat{c}(\alpha)} -\alpha + \xi_2 dF_I + \frac{\partial}{\partial \alpha} \int_{\hat{c}(\alpha)}^{1} \alpha + \xi_1 dF_I$$

$$= f_I(\hat{c}(\alpha)) \hat{c}'(\alpha)(-2\alpha - \xi_1 + \xi_2) + 1 - 2F_I(\hat{c}(\alpha)).$$

Rearranging terms, we obtain that $\frac{\partial}{\partial \alpha} \mathbb{E}[\pi_S(\alpha)] > 0$ if and only if

$$\alpha \leq 1 - \mu_E + \delta \int_{0}^{1} \max[c_E, R] - c_E dF_E \quad (A.6)$$
and
\[ \mathbb{P}\{c_I > \hat{c}(\alpha)\} > \frac{1 - f_I(\hat{c}(\alpha)) - \hat{c}'(\alpha)(-2\alpha - \xi_1 + \xi_2)}{2}. \] (A.7)

Substituting \( \alpha = 0 \), condition [A.6] holds and condition [A.7] simplifies as stated in the proposition with \( \hat{c} \triangleq \hat{c}(0) \).

A.8 Proof of Proposition 6

We first prove a preliminary result.

Claim 3. Let \( a, b, c \in \mathbb{R} \) such that \( a < b < c \). Suppose a function \( f(x) \) is convex on \((a, b)\), strictly increasing on \((b, c)\), and satisfies \( f(a) \leq 0 \). (i) If \( f(c) > 0 \), then there exists a unique \( \tilde{x} \in (a, c) \) such that \( f(\tilde{x}) = 0 \), and (ii) if \( f(c) \leq 0 \), then \( f(x) < 0 \) for all \( x \in (a, c) \).

Proof of Claim 3. For the first part, we consider three distinct cases (i) \( f(b) < 0 \), (ii) \( f(b) = 0 \), and (iii) \( f(b) > 0 \). If \( f(b) < 0 \), then since \( f(c) > 0 \) by hypothesis, the IVT implies the existence of a root \( \tilde{x} \in (b, c) \), which is unique since \( f \) is strictly increasing in \((b, c)\). If \( f(b) = 0 \), then \( \tilde{x} = b \) is the unique root. Finally, if \( f(b) > 0 \), then there exists a unique root \( \tilde{x} \in (a, b) \), again by the IVT and the convexity of \( f \).

For the second part, if \( f(c) \leq 0 \), then by the strictly increasing property, it must be that \( f(x) < 0 \) for all \( x \in [b, c] \). In particular, since \( f(b) < 0 \), it follows by convexity and \( f(a) \leq 0 \) that \( f(x) < 0 \) for all \( x \in (a, b) \).

Proof of Proposition 6. Consider the incumbent’s payoff given its valuation \( v_I \).

\[
\mathbb{E}[\pi_I(v_I)] = \begin{cases} 
  c_I v_I - \max[\mu_E b_E, R] + \delta \int_0^1 (c_I v_I - \max[\mu_E v_E, R]) \mathbb{I}_{\{c_I v_I \geq \max[\mu_E v_E, R]\}} dF_{v_E} & \text{if } c_I b_I \geq \max[\mu_E b_E, R], \\
  \delta \int_0^1 \int_0^1 (c_I v_I - \max[c_E v_E, R]) \mathbb{I}_{\{c_I v_I \geq \max[c_E v_E, R]\}} dF_{v_E} dF_{v_E} & \text{if } c_I b_I < \max[\mu_E b_E, R].
\end{cases}
\] (A.8)
Therefore, a weakly dominant first stage bid for the incumbent is

\[ b_I^*(c_I, v_I) = v_I + \frac{\delta}{c_I} \left( \int_0^1 (c_I v_I - \max[\mu_E v_E, R]) \mathbb{I}_{\{c_I v_I \geq \max[\mu_E v_E, R]\}} dv_E \right) \]

\[ \quad - \int_0^1 \int_0^1 (c_I v_I - \max[c_E v_E, R]) \mathbb{I}_{\{c_I v_I \geq \max[c_E v_E, R]\}} dF_E dv_E. \]

(A.9)

It follows that whether the incumbent bids above or below valuation depends on the sign of the second component of (A.9):

\[ g(a_I) \triangleq \int_0^1 (a_I - \max[\mu_E v_E, R]) \mathbb{I}_{\{a_I \geq \max[\mu_E v_E, R]\}} dv_E \]

\[ \quad - \int_0^1 \int_0^1 (a_I - \max[c_E v_E, R]) \mathbb{I}_{\{a_I \geq \max[c_E v_E, R]\}} dF_E dv_E, \]

where we have substituted \( c_I v_I \) with the ad rank parameter \( a_I \). We immediately obtain that if \( a_I < R \), then \( g(a_I) = 0 \) and thus the incumbent bids truthfully. If \( a_I \geq R \), then the expressions in the indicator brackets simplify such that

\[ g(a_I) = \int_0^1 (a_I - \max[\mu_E v_E, R]) \mathbb{I}_{\{a_I \geq \max[\mu_E v_E, R]\}} dv_E - \int_0^1 \int_0^1 (a_I - \max[c_E v_E, R]) \mathbb{I}_{\{a_I \geq \max[c_E v_E, R]\}} dv_E dF_E \]

\[ = \int_0^{\min\left[\frac{a_I}{\sigma_E}, 1\right]} a_I - \max[\mu_E v_E, R] dv_E - \int_0^1 \int_0^{\min\left[\frac{a_I}{\sigma_E}, 1\right]} a_I - \max[c_E v_E, R] dv_E dF. \]

Next, we establish five properties related to \( g(a_I) \) which will be used to prove the proposition.

1. \( g(1) \geq 0 \): This holds because \( g(1) = \int_0^1 \int_0^1 \max[c_E v_E, R] - \max[\mu_E v_E, R] dF_E dv_E \), and for any \( v_E \in [0, 1] \), the inner integrand \( \int_0^1 \max[c_E v_E, R] - \max[\mu_E v_E, R] dF_E \geq 0 \) by Jensen’s inequality.

2. \( g(a_I) \) is strictly increasing in \( a_I \) for \( a_I \in (\mu_E, 1) \): It suffices to show that \( g'(1) \geq 0 \) and \( g''(a_I) < 0 \). The first inequality holds because for \( a_I \in (\mu_E, 1) \), we have

\[ g'(a_I) = 1 - F_E(a_I) - a_I \int_{a_I}^{1} \frac{1}{c_E} dF_E, \]

(A.11)

from which it follows that \( g'(1) = 1 - F_E(1) - 1 \cdot \int_{a_I}^{1} \frac{1}{c_E} dF_E = 0 \). The second inequality holds because we obtain from (A.11) that \( g''(a_I) = -f_E(a_I) - \int_{a_I}^{1} \frac{1}{c_E} dF_E - \)

53
\[
a_I \left( -\frac{1}{a_I} f_E(a_I) \right) = - \int_{a_I}^{1} \frac{1}{c_E} dF_E < 0.
\]

3. \( g''(\mu_E) > 0 \): We have

\[
g'(a_I) = \frac{a_I}{\mu_E} - \mathbb{P}\{ \max[c_E v_E, R] \leq a_I \} = \frac{a_I}{\mu_E} - F_E(a_I) - a_I \int_{a_I}^{1} \frac{1}{c_E} dF_E,
\]

\[
g''(a_I) = \frac{1}{\mu_E} - \int_{a_I}^{1} \frac{1}{c_E} dF_E,
\]

from which it follows \( g''(\mu_E) = \frac{1}{\mu_E} - \int_{\mu_E}^{1} \frac{1}{c_E} dF_E \geq \frac{1}{\mu_E} - \frac{1}{\mu_E} \int_{\mu_E}^{1} dF_E = \frac{1}{\mu_E} F_E(\mu_E) > 0. \)

4. There exists \( \tilde{r} \in (0, \mu_E) \) such that \( g''(R) < 0 \) for \( R < \tilde{r} \), and \( g''(R) \geq 0 \) for \( R \geq \tilde{r} \):

The second derivative at \( a_I = R \) is \( g''(R) = \frac{1}{\mu_E} - \int_{R}^{1} \frac{1}{c_E} dF_E \), which is strictly increasing in \( R \) because \( g''(R) = \frac{1}{R} f_E(R) > 0. \) For \( R = \mu_E \), we have \( g''(R) > 0 \) (see previous property); for \( R = 0 \), we have \( g''(R) = \frac{1}{\mu_E} - \int_{0}^{1} \frac{1}{c_E} dF_E < 0 \) by Jensen’s inequality. The unique existence of \( \tilde{r} \) is guaranteed by the IVT.

5. There exists \( \hat{r} \in (\tilde{r}, \mu_E) \) such that \( g'(R) < 0 \) for \( R \in (0, \hat{r}) \) and \( g'(R) \geq 0 \) for \( R \in [\hat{r}, \mu_E) \): We have \( g'(R) = \frac{R}{\mu_E} - F_E(R) - R \int_{R}^{1} \frac{1}{c_E} dF_E \), which implies that \( g' \) is convex: \( g''(R) = \frac{1}{R} f_E(R) > 0. \) Thus, \( g' \) attains its minimum (which is negative because the minimum should be smaller than \( g'(0) = 0 \)) at \( R = \tilde{r} \), where \( g' \) is flat; i.e., \( g''(\tilde{r}) = \frac{1}{\mu_E} - \int_{\tilde{r}}^{1} \frac{1}{c_E} dF_E = 0. \) Next, using the fact that \( g'(\mu_E) = 1 - F_E(\mu_E) - \mu_E \int_{\mu_E}^{1} \frac{1}{c_E} dF_E > 1 - F_E(\mu_E) - \mu_E \frac{1}{\mu_E} (1 - F_E(\mu_E)) = 0 \), we obtain, by the IVT, the unique existence of \( \hat{r} \in (\tilde{r}, \mu_E) \).

In summary, if \( R < \hat{r} \), then since \( g(R) = 0 \) and \( g'(a_I) < 0 \) for all \( a_I < \hat{r} \), we have \( g(\hat{r}) < 0 \).

Combined with the fact that \( g'(a_I) > 0 \) for \( a_I > \hat{r} \) and that \( g(1) \geq 0 \), we obtain by the IVT a unique \( \tilde{a}_I \in (\hat{r}, 1) \) such that \( g(a_I) < 0 \) for all \( a_I < \tilde{a}_I \) and \( g(a_I) > 0 \) for all \( a_I > \tilde{a}_I \). On the other hand, if \( R > \hat{r} \), then \( g(R) = 0 \) and \( g \) increases thereafter; thus, \( g(a_I) > 0 \) for all \( a_I > R \).

Next, we prove that the entrant always bids above valuation. For ease of exposition, let \( m \triangleq \max[c_I v_I, R] \). The entrant bids above valuation if and only if \( \int_{0}^{1} \int_{0}^{1} (c_E v_E - m) I_{c_E v_E > m} - \).
\[(\mu_Ev_E - m)\mathbb{I}_{(\mu_Ev_E > m)}dF_EdG(m) \geq 0,\] where \(G(m)\) is the cdf of \(m\). It suffices to show that the inner integral

\[
\int_0^1 (c_Ev_E - m)\mathbb{I}_{(c_Ev_E > m)} - (\mu_Ev_E > m)\mathbb{I}_{(\mu_Ev_E > m)}dF_E
\]

is non-negative for any \(m \geq 0\). To that end, suppose \(m \geq \mu_Ev_E\). Then (A.12) = \(\int_0^1 (c_Ev_E - m)\mathbb{I}_{(c_Ev_E > m)}dF_E = \int_{\min\left[\frac{m}{v_E}, 1\right]}^1 v_E\left(c_E - \frac{m}{v_E}\right)dF_E \geq 0\). And if \(m < \mu_Ev_E\), then

\[
(A.12) = \int_0^1 c_Ev_E - mdF_E - (\mu_Ev_E - m)
\]

\[
= \int_0^1 c_Ev_E - mdF_E - \int_0^{\min\left[\frac{m}{v_E}, 1\right]} c_Ev_E - mdF_E - \int_0^1 c_Ev_E - mdF_E
\]

\[
= v_E\int_0^{\min\left[\frac{m}{v_E}, 1\right]} m - c_EdF_E \geq 0.
\]

This completes the proof. ■

A.9 Lemmas for GSP Extension

**Lemma 3** (Assortative Matching). *In any locally envy-free equilibrium, the resulting ad slot allocation is assortative, such that the advertiser assigned to position \(i\) has higher CTR than the advertiser in position \(i + 1\).*

**Proof of Lemma 3**. Let \(c(i), b(i),\) and \(p(i)\), respectively, denote the CTR, bid, and payment of the advertiser in position \(i\). The envy-free conditions are

\[
\alpha_i c(i) - p(i) \geq \alpha_{i+1} c(i) - p(i+1),
\]

\[
\alpha_{i+1} c(i+1) - p(i+1) \geq \alpha_i c(i+1) - p(i).
\]

Rearranging and combining the inequalities yield \((\alpha_i - \alpha_{i+1})c(i) \geq p(i) - p(i+1) \geq (\alpha_i - \alpha_{i+1})c(i+1)\), which together with the fact that \(\alpha_i > \alpha_{i+1}\) (i.e., higher positions are associated
Lemma 4 (LREF Stage 2 Payoffs). Let I(i) denote the incumbent with the higher (lower) CTR. If the entrant’s CTR is learned, then the advertisers’ expected Stage 2 payoffs are

\[ E[\pi^I] = F_E(c_i)(c_I - c_i) + \int_0^{c_i} \theta(c_i - c_E) dF_E + \int_{c_i}^{c_I} (c_I - c_E) + \theta(c_E - c_i) dF_E + (1 - F_E(c_i)) \theta(c_I - c_i). \]

\[ E[\pi^i] = \int_0^{c_i} \theta(c_i - c_E) dF_E, \]

\[ E[\pi^E] = \int_{c_i}^{c_I} \theta(c_E - c_i) dF_E + \int_{c_i}^{1} (c_E - c_I) + \theta(c_I - c_i) dF_E. \]

If the entrant’s CTR is not learned, then the expected payoffs are

\[
\begin{array}{ccc}
\mu_E \leq c_i < c_I & c_i < \mu_E \leq c_I & c_i < c_I < \mu_E \\
E[\pi^I] & (c_I - c_i) + \theta(c_i - \mu_E) & (c_i - \mu_E) + \theta(\mu_E - c_i) & \theta(c_I - c_i) \\
E[\pi^i] & \theta(c_i - \mu_E) & 0 & 0 \\
E[\pi^E] & 0 & \theta(\mu_E - c_i) & (\mu_E - c_I) + \theta(c_I - c_i) \\
\end{array}
\]

Proof of Lemma 4. First, consider the advertisers in positions 2 and 3, where the “third position” is the no-display slot with a position-CTR of 0. The envy-free conditions imply

\[ \alpha_2 c_{(2)} \left( 1 - \frac{c_{(3)} b_{(3)}}{c_{(2)}} \right) \geq 0 \quad \text{and} \quad 0 \geq \alpha_2 c_{(3)} \left( 1 - \frac{c_{(3)} b_{(3)}}{c_{(3)}} \right), \]

which simplify to 1 \leq b_{(3)} \leq c_{(2)}/c_{(3)}. Note that there exist values of b_{(3)} that satisfy this inequality because Lemma 3 implies that c_{(2)}/c_{(3)} \geq 1. And since we choose the lowest revenue envy-free (LREF) equilibrium, we have that b_{(3)} = 1; i.e., the losing advertiser bids truthfully.

Next, we find the LREF bid of advertiser (2) by recursion. Again, the envy-free conditions between advertisers (1) and (2) imply

\[ \alpha_1 c_{(1)} \left( 1 - \frac{c_{(2)} b_{(2)}}{c_{(1)}} \right) \geq \alpha_2 c_{(1)} \left( 1 - \frac{c_{(3)} b_{(3)}}{c_{(1)}} \right), \]

\[ \alpha_2 c_{(2)} \left( 1 - \frac{c_{(3)} b_{(3)}}{c_{(2)}} \right) \geq \alpha_1 c_{(2)} \left( 1 - \frac{c_{(2)} b_{(2)}}{c_{(2)}} \right). \]
Substituting $b(3) = 1$ and simplifying yields

$$c(2) - \frac{\alpha_2}{\alpha_1}(c(2) - c(3)) \leq c(2)b(2) \leq c(1) - \frac{\alpha_2}{\alpha_1}(c(1) - c(3)). \quad (A.14)$$

Using Lemma 3 and $\alpha_1 > \alpha_2$, it can be easily shown that the lower bound of (A.14) is indeed smaller than the upper bound; i.e., $c(2) - \frac{\alpha_2}{\alpha_1}(c(2) - c(3)) < c(1) - \frac{\alpha_2}{\alpha_1}(c(1) - c(3))$; that is, there exist values of $b(2)$ that satisfy the envy-free conditions. In LREF equilibrium, we have $b(2) = 1 - \frac{\alpha_2}{\alpha_1}(1 - c(3)/c(2))$.

The profits for advertisers (1), (2), and (3), respectively, are

$$\pi(1) = \alpha_1 c(1) \left( 1 - \frac{c(2) - \frac{\alpha_2}{\alpha_1}(c(2) - c(3))}{c(1)} \right),$$

$$\pi(2) = \alpha_2 c(2) \left( 1 - \frac{c(3)}{c(2)} \right),$$

$$\pi(3) = 0.$$

Using the expressions above, combined with the assumptions that $\alpha_1 = 1, \alpha_2 = \theta$, we can write the advertisers’ payoffs when $c_E$ is learned (i.e., entrant won in Stage 1):

<table>
<thead>
<tr>
<th></th>
<th>$c_E \leq c_i &lt; c_I$</th>
<th>$c_i &lt; c_E \leq c_I$</th>
<th>$c_i &lt; c_I &lt; c_E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\pi_I$</td>
<td>$(c_I - c_i) + \theta(c_i - c_E)$</td>
<td>$(c_I - c_E) + \theta(c_i - c_E)$</td>
<td>$\theta(c_i - c_I)$</td>
</tr>
<tr>
<td>$\pi_i$</td>
<td>$\theta(c_i - c_E)$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\pi_E$</td>
<td>0</td>
<td>$\theta(c_E - c_i)$</td>
<td>$(c_E - c_I) + \theta(c_I - c_i)$</td>
</tr>
</tbody>
</table>

The payoffs when the $c_E$ is not learned and is thus assigned $\mu_E$ follow immediately.

**A.10 Proof of Proposition 7**

To ensure subgame perfection, we use the expressions for the second stage equilibrium given in Lemma 4.
Entrant

Proof. If $\mu_E \leq c_i$, then the entrant clearly wants to reveal its CTR, for otherwise, it would have no chance of winning in the Stage 2. If $c_i < \mu_E \leq c_I$, then the difference between its payoffs when its CTR is revealed and concealed is

$$\Delta_1 \triangleq \int_{c_i}^{c_I} \theta(c_E - c_i) dF_E + \int_{c_i}^{c_I} (c_E - c) + \theta(c_I - c_i) dF_E - \theta(\mu_E - c_i)$$

$$= - \int_0^{c_i} \theta(c_E - c_i) dF_E + \int_0^{c_i} \theta(c_E - c_i) dF_E + \int_{c_i}^{c_I} \theta(c_E - c_i) - \theta(c_E - c_i) dF_E + \int_{c_i}^{c_I} c_E - c + \theta(c_I - c_i) dF_E$$

$$- \theta(\mu_E - c_i)$$

$$= - \int_0^{c_i} \theta(c_E - c_i) dF_E + \int_{c_i}^{c_I} c_E - c + \theta(c_I - c_i) dF_E$$

$$= - \int_0^{c_i} \theta(c_i - c_E) dF_E + \int_{c_i}^{c_I} (1 - \theta)(c_E - c_i) dF_E > 0. \quad (A.15)$$

Therefore, the entrant is better off revealing its CTR when $c_i < \mu_E \leq c_I$.

Finally, if $c_I < \mu_E$, then the difference between its payoffs when its CTR is revealed and concealed is $\Delta_2 \triangleq \int_{c_i}^{c_I} \theta(c_E - c_i) dF_E + \int_{c_I}^{c_i} (c_E - c) + \theta(c_I - c_i) dF_E - ((\mu_E - c_I) + \theta(c_I - c_i))$, and differentiating $\Delta_2$ with respect to $\theta$ yields

$$\frac{\partial \Delta_2}{\partial \theta} = c_i - c_I + \int_{c_i}^{c_I} c_E - c_i dF_E - \int_{c_i}^{c_I} c_i - c_I dF_E$$

$$= \int_0^{c_I} c_i - c_I dF_E + \int_{c_i}^{c_I} c_E - c_i dF_E$$

$$= \int_0^{c_i} c_i - c_I dF_E + \int_{c_i}^{c_I} c_E - c_I dF_E < 0.$$
for all \( \theta \in (0, 1) \). But the positivity holds at \( \theta = 1 \) because

\[
\Delta_2|_{\theta=1} = \int_{c_i}^{c_I} c_E - c_i dF_E + \int_{c_i}^{1} c_E - c_I + c_I - c_i dF_E - \int_0^{1} c_E - c_I + c_I - c_i dF_E \\
= \int_{c_i}^{1} c_E - c_i dF_E - \int_0^{1} c_E - c_i dF_E \\
= \int_0^{c_i} c_i - c_E dF_E > 0,
\]

This completes the proof. \(\blacksquare\)

**Incumbent with Lower CTR**

*Proof.* First, if \( \mu_E > c_i \), then it is clearly better for the incumbent with the lower CTR if the entrant’s CTR is learned because *only then* does it have a chance of winning the second ad slot. The proof is trivial and is omitted. Second, suppose \( \mu_E \leq c_i \). The difference between its payoffs when the entrant’s CTR is revealed and concealed is \( \Delta_3 \Delta_3 \triangleq \int_0^{c_i} \theta(c_i - c_E)dF_E - \theta(c_i - \mu_E) = -\int_{c_i}^{1} \theta(c_i - c_E)dF_E \geq 0 \). This completes the proof. \(\blacksquare\)

**Incumbent with Higher CTR**

*Proof.* Suppose \( \mu_E \leq c_i \). Then the difference between the payoffs of the incumbent with the higher CTR when the entrant’s CTR is revealed and when it is not is

\[
\Delta_4 \triangleq F_E(c_i)(c_I - c_i) + \int_0^{c_i} \theta(c_i - c_E)dF_E + \int_{c_i}^{c_I} (c_I - c_E) \\
+ \theta(c_E - c_i)dF_E + (1 - F_E(c_I))\theta(c_I - c_i) - ((c_I - c_i) + \theta(c_i - \mu_E)).
\]
Differentiating $\Delta_4$ with respect to $\theta$ yields

$$
\frac{\partial \Delta_4}{\partial \theta} = -c_i + \int_0^{c_i} c_i - c_E dF_E + \int_{c_i}^{c_I} c_I - c_i dF_E + \int_{c_i}^{1} c_I - c_i dF_E + \mu_E
$$

$$
= \int_0^{c_i} c_i - c_E - c_i + c_E dF_E + \int_{c_i}^{c_I} c_E - c_i - c_E + c_I dF_E + \int_{c_i}^{1} c_I - c_i - c_i + c_E dF_E
$$

$$
= 2 \int_{c_i}^{c_I} c_E - c_i dF_E + \int_{c_i}^{1} c_I - c_i dF_E + \int_{c_i}^{1} c_E - c_i dF_E > 0,
$$

where the last positivity follows from the positivity of each term in the sum in the line before.

Therefore, $\Delta_4$ is strictly increasing in $\theta$. At the end points of $\theta$, we have $\Delta_4 < 0$ and $\Delta_4 > 0$, respectively:

$$
\Delta_4|_{\theta=0} = c_I - c_I + \int_0^{c_i} c_I - c_i dF_E + \int_{c_i}^{c_I} c_I - c_i dF_E
$$

$$
= \int_0^{c_i} c_I - c_i + c_i - c_E dF_E + \int_{c_i}^{c_I} c_I - c_E + c_i - c_I dF_E + \int_{c_i}^{1} c_I - c_i dF_E
$$

$$
= \int_{c_i}^{c_I} c_I - c_E dF_E + \int_{c_i}^{1} c_i - c_I dF_E < 0,
$$

and

$$
\Delta_4|_{\theta=1} = \mu_E - c_I + \int_0^{c_i} c_I - c_i dF_E + \int_{c_i}^{1} c_I - c_i dF_E + \int_{c_i}^{c_I} c_I - c_i dF_E + \int_{c_i}^{1} c_I - c_i dF_E
$$

$$
= \mu_E - c_I + \int_0^{c_i} c_I - c_E dF_E + \int_{c_i}^{1} c_I - c_i dF_E + \int_{c_i}^{c_I} c_I - c_E + c_I dF_E
$$

$$
= \int_0^{c_i} c_I - c_E + c_E - c_I dF_E + \int_{c_i}^{1} c_I - c_i + c_E - c_I dF_E + \int_{c_i}^{c_I} c_I - c_i + c_E - c_I dF_E
$$

$$
= \int_{c_i}^{1} c_E - c_i dF_E + \int_{c_i}^{c_I} c_E - c_i dF_E > 0.
$$

Therefore, by the IVT, there exists a unique $\hat{\theta} \in (0, 1)$ that solves $\Delta_4(\theta) = 0$ such that $\Delta_4 < 0$ for $\theta < \hat{\theta}$ and $\Delta_4 > 0$ for $\theta > \hat{\theta}$. This means that the incumbent with the higher CTR wants the entrant’s CTR to be learned if and only if the position-CTR of the second slot is sufficiently large.

Next, suppose $c_i < \mu_E \leq c_I$. Then the difference between the payoffs of the incumbent with
the higher CTR when the entrant’s CTR is revealed and when it is not is

$$\Delta_5 \triangleq F_E(c_i)(c_I - c_i) + \int_0^{c_i} \theta(c_i - c_E)dF_E + \int_{c_i}^{c_I} (c_I - c_E)dF_E + \theta(c_E - c_i)dF_E + \theta(c_I - c_i) - ((c_I - \mu_E) + \theta(\mu_E - c_i)).$$

The difference $\Delta_5$ strictly decreases with $c_I$ because

$$\frac{\partial \Delta_5}{\partial c_I} = -1 + F_E(c_I) + \theta(1 - F_E(c_I)) = -1 + F_E(c_I) + (1 - F_E(c_I)) + \theta(1 - F_E(c_I)) = -(1 - \theta)(1 - F_E(c_I)) < 0.$$ 

Next, we examine the sign of $\Delta_5$ at the endpoints $c_I = \mu_E$ and $c_I = 1$. At $c_I = 1$, we have

$$\Delta_5|_{c_I=1} = \int_0^{c_i} 1 - c_i + \theta(c_i - c_E)dF_E + \int_{c_i}^{1} 1 - c_E + \theta(c_E - c_i)dF_E - ((1 - \mu_E) + \theta(\mu_E - c_i))$$

$$= \int_0^{c_i} 1 - c_i + \theta(c_i - c_E) - (1 - c_E) - \theta(c_E - c_i)dF_E$$

$$= \int_0^{c_i} c_E - c_i + 2\theta(c_i - c_E)dF_E$$

$$= (2\theta - 1) \int_0^{c_i} c_i - c_EdF_E.$$

Therefore,

$$\Delta_5|_{c_I=1} = \begin{cases} + & \text{if } \theta \geq \frac{1}{2}, \\ - & \text{if } \theta < \frac{1}{2}. \end{cases}$$ (A.16)

And since $\Delta_5$ is decreasing in $c_I$, we obtain $\Delta_5 \geq 0$ for all $c_I$ if $\theta \geq \frac{1}{2}$.

The sign of $\Delta_5$ is only inconclusive for $\theta < \frac{1}{2}$, so we examine this case next. If $\Delta_5|_{c_I=\mu_E} \leq 0$ under $\theta < \frac{1}{2}$, then by the decreasing property of $\Delta_5$ we would obtain $\Delta_5 \leq 0$ for all $c_I$. On the other hand, if $\Delta_5|_{c_I=\mu_E} > 0$, then by the IVT, there would exist a unique zero $\tilde{c}_I$ such that $\Delta_5 > 0$ for all $c_I < \tilde{c}_I$, and $\Delta_5 < 0$ for all $c_I > \tilde{c}_I$. We will show that the latter holds; i.e., $\Delta_5|_{c_I=\mu_E} > 0$: 

\[\text{61}\]
\[ \Delta_5 |_{c_i = \mu_E} = \int_0^{c_i} \mu_E - c_i + \theta(c_i - c_E) dF_E + \int_{c_i}^\mu \mu_E - c_E + \theta(c_E - c_i) dF_E + \int_0^1 \theta(\mu_E - c_i) dF_E \]
\[ - \theta(\mu_E - c_i) \]
\[ = \int_0^{c_i} \mu_E - c_i + \theta(c_i - c_E) - \theta(c_E - c_i) dF_E + \int_{c_i}^\mu \mu_E - c_E + \theta(c_E - c_i) - \theta(c_E - c_i) dF_E \]
\[ + \int_\mu^1 \theta(\mu_E - c_i) - \theta(c_E - c_i) dF_E \]
\[ = \int_0^{c_i} \mu_E - c_i + 2\theta(c_i - c_E) dF_E + \int_{c_i}^\mu \mu_E - c_E dF_E + \int_\mu^1 \theta(\mu_E - c_i) dF_E \]
\[ - \int_0^{c_i} \mu_E - c_i + 2\theta(c_i - c_E) dF_E + \int_{c_i}^\mu \mu_E - c_E dF_E + \int_\mu^1 \theta(\mu_E - c_i) dF_E \]
\[ = \int_0^{c_i} c_E - c_i + 2\theta(c_i - c_E) dF_E + \int_{c_i}^\mu c_E - c_E dF_E + \int_\mu^1 c_E - c_E dF_E \]
\[ + \int_\mu^1 \theta(\mu_E - c_E)dF_E - \int_\mu^1 \mu_E - c_E dF_E \]
\[ = \int_0^{c_i} (2\theta - 1)(c_i - c_E) dF_E + \int_\mu^1 (1 - \theta)(c_E - \mu_E) dF_E. \]

Since this is a linear function of \( \theta \), to show that \( \Delta_5 |_{c_i = \mu_E} > 0 \) for all \( \theta \in (0, \frac{1}{2}) \), it suffices to show that \( \Delta_5 |_{c_i = \mu_E} > 0 \) at the endpoints \( \theta = 0 \) and \( \theta = \frac{1}{2} \). At \( \theta = \frac{1}{2} \), we have \( \Delta_5 |_{c_i = \mu_E} = \frac{1}{2} \int_\mu^1 c_E - \mu_E dF_E > 0 \). At \( \theta = 0 \), we have

\[ \Delta_5 |_{c_i = \mu_E} (\theta = 0) = \int_0^1 c_E - \mu_E dF_E - \int_0^{c_i} c_i - c_E dF_E \]
\[ = \int_0^\mu c_E - \mu_E dF_E + \int_0^1 c_E - \mu_E dF_E - \int_\mu^1 \mu_E - \mu_E dF_E - \int_0^{c_i} c_i - c_E dF_E \]
\[ = - \int_0^\mu c_E - \mu_E dF_E - \int_0^{c_i} c_i - c_E dF_E \]
\[ = - \int_0^c c_E - \mu_E dF_E - \int_\mu^c c_E - \mu_E dF_E - \int_0^{c_i} c_i - c_E dF_E \]
\[ = \int_0^c \mu_E - c_i dF_E + \int_\mu^c \mu_E - c_E dF_E > 0, \]

where the last positivity follows from \( \mu_E > c_i \). Thus, we have shown that \( \Delta_5 |_{c_i = \mu_E} > 0 \) for
all $\theta < \frac{1}{2}$. Combined with the fact that $\Delta_5|_{c_I=1} < 0$ for $\theta < \frac{1}{2}$, the IVT implies the unique existence of $\tilde{c}_I \in (\mu_E, 1)$ such that $\Delta_5 > 0$ for all $c_I < \tilde{c}_I$, and $\Delta_5 < 0$ for all $c_I > \tilde{c}_I$.

In summary, if $c_i < \mu_E \leq c_I$, then

$$\Delta_5 = \begin{cases} - & \text{if } \theta < \frac{1}{2} \text{ and } c_I > \tilde{c}_I, \\ + & \text{otherwise.} \end{cases}$$

(A.17)

We consider the last case of $c_I < \mu_E$. The difference between the payoffs of the incumbent with the higher CTR when the entrant’s CTR is revealed and when it is not is

$$\Delta_6 \triangleq F_E(c_i)(c_I - c_i) + \int_0^{c_i} \theta(c_i - c_E)dF_E + \int_{c_i}^{c_I} (c_I - c_E) \theta(c_I - c_i) - \theta(c_I - c_i).$$

Differentiating $\Delta_6$ with respect to $c_I$ yields $\frac{\partial \Delta_6}{\partial c_I} = F_E(c_I) - \theta F_E(c_I) = (1 - \theta)F_E(c_I) > 0$. Since $\Delta_6$ is increasing in $c_I$, to show $\Delta_6 > 0$ for all $c_I$, it suffices to show positivity only at the lowest value $c_I = c_i$. But if $c_I = c_i$, then $\Delta_6$ simplifies to $\int_0^{c_i} \theta(c_i - c_E)dF_E$, which is clearly positive. Therefore, $\Delta_6 > 0$ if $c_I < \mu_E$. ■

A.11 LREF Stage 1 Bids

We denote by $[a_1, a_2, a_3]$ the locally envy-free equilibrium candidate where advertiser $a_1$ gets the first slot, advertiser $a_2$ the second, and advertiser $a_3$ the third (null) slot.

Case 1: Suppose $\mu_E \leq c_i < c_I$.

(i) $[I, i, E]$

EF condition for $E$:

$$\theta \mu_E (1 - \mu_E b_E/\mu_E) + \delta \mathbb{E}[\pi_i^E] \leq 0 \iff b_E \geq 1 + \frac{\delta}{\theta \mu_E} \mathbb{E}[\pi_i^E]$$
EF condition for $i$:

$$
\theta c_i (1 - \mu_E b_i / c_i) + \delta \mathbb{E}[\pi^0_i] \geq \delta \mathbb{E}[\pi^1_i] \iff b_i \leq \frac{c_i}{\mu_E} + \frac{\delta}{\theta \mu_E} (\mathbb{E}[\pi^0_i] - \mathbb{E}[\pi^1_i]),
$$

$$
\theta c_i (1 - \mu_E b_E / c_i) + \delta \mathbb{E}[\pi^0_i] \geq c_i (1 - c_i b_i / c_i) + \delta \mathbb{E}[\pi^0_i] \iff c_i b_i \geq (1 - \theta) c_i + \theta \mu_E b_E
$$

EF condition for $I$:

$$
c_I (1 - c_i b_i / c_i) + \delta \mathbb{E}[\pi^0_I] \geq \theta c_I (1 - \mu_E b_E / c_I) + \delta \mathbb{E}[\pi^0_I] \iff c_i b_i \leq (1 - \theta) c_I + \theta \mu_E b_E
$$

Given $b_E$, there always exist non-negative $b_i$ such that EF conditions constraining $b_i$ are satisfied, because we need $(1 - \theta) c_i + \theta \mu_E b_E \leq c_i b_i \leq (1 - \theta) c_I + \theta \mu_E b_E$ but $c_i < c_I$. On the other hand, there does not always exist non-negative $b_E$ that satisfies EF conditions constraining $b_E$. To see this, note that we must have

$$
1 + \frac{\delta}{\theta \mu_E} \mathbb{E}[\pi^1_E] \leq \frac{c_i}{\mu_E} + \frac{\delta}{\theta \mu_E} (\mathbb{E}[\pi^0_i] - \mathbb{E}[\pi^1_i]) \quad (A.18)
$$

for $E$ to set envy-free bids. Using the fact that $\mathbb{E}[\pi^0_I] - \mathbb{E}[\pi^1_I] = \theta(c_i - \mu_E) - \int_0^{c_i} \theta(c_i - c_E) dF_E = \int_{c_i}^1 \theta(c_i - c_E) dF_E < 0$, it can be shown that the inequality \((A.18)\) holds if and only if

$$
\delta \leq \frac{\theta(c_i - \mu_E)}{\mathbb{E}[\pi^1_I] - (\mathbb{E}[\pi^0_I] - \mathbb{E}[\pi^1_i])}. \quad (A.19)
$$

In summary, $[I, i, E]$ is an envy-free equilibrium if and only if $\delta$ is sufficiently small \((A.19)\); the corresponding lowest revenue equilibrium bids are $b^*_E = 1 + \frac{\delta}{\theta \mu_E} \mathbb{E}[\pi^1_E]$ and $b^*_i = 1 - \theta + \frac{\theta \mu_E}{c_i} b^*_E$.

(ii) $[I, E, i]$

EF condition for $i$:

$$
\theta c_i (1 - c_i b_i / c_i) + \delta \mathbb{E}[\pi^0_i] \leq \delta \mathbb{E}[\pi^1_i] \iff 1 + \frac{\delta}{\theta c_i} (\mathbb{E}[\pi^0_i] - \mathbb{E}[\pi^1_i]) \leq b_i
$$
EF condition for $E$:

$$\theta \mu_E (1 - c_i b_i / \mu_E) + \delta \mathbb{E}[\pi^I_E] \geq 0 \iff b_i \leq \frac{\mu_E}{c_i} + \frac{\delta}{\theta c_i} \mathbb{E}[\pi^I_E],$$

$$\theta \mu_E (1 - c_i b_i / \mu_E) + \delta \mathbb{E}[\pi^I_E] \geq \mu_E (1 - \mu_E b_E / \mu_E) + \delta \mathbb{E}[\pi^I_E] \iff \mu_E b_E \geq (1 - \theta) \mu_E + \theta c_i b_i$$

EF condition for $I$:

$$c_I (1 - \mu_E b_E / c_I) + \delta \mathbb{E}[\pi^I] \geq \theta c_I (1 - c_i b_i / c_I) + \delta \mathbb{E}[\pi^I] \iff \mu_E b_E \leq (1 - \theta) c_I + \theta c_i b_i$$

Given $b_i$, there always exist non-negative $b_E$ such that EF conditions constraining $b_E$ are satisfied, because we need $(1 - \theta) \mu_E + \theta c_i b_i \leq \mu_E b_E \leq (1 - \theta) c_I + \theta c_i b_i$ but $\mu_E < c_I$. On the other hand, there does not always exist non-negative $b_i$ that satisfies EF conditions constraining $b_i$. To see this, note that we must have

$$1 + \frac{\delta}{\theta c_i} (\mathbb{E}[\pi^0_i] - \mathbb{E}[\pi^I_i]) \leq \frac{\mu_E}{c_i} + \frac{\delta}{\theta c_i} \mathbb{E}[\pi^I_E]$$

(A.20)

for $i$ to set envy-free bids. Again, using the fact that $\mathbb{E}[\pi^0_i] - \mathbb{E}[\pi^I_i] < 0$, it can be shown that the inequality (A.20) holds if and only if

$$\delta \geq \frac{\theta (c_i - \mu_E)}{\mathbb{E}[\pi^I_E] - (\mathbb{E}[\pi^0_i] - \mathbb{E}[\pi^I_i])}$$

(A.21)

In summary, $[I, E, i]$ is an envy-free equilibrium if and only if $\delta$ is sufficiently large (A.21); the corresponding lowest revenue equilibrium bids are $b_i^* = 1 + \frac{\delta}{\theta c_i} (\mathbb{E}[\pi^0_i] - \mathbb{E}[\pi^I_i])$ and $b_E^* = 1 - \theta + \frac{\theta c_i}{\mu_E} b_i^*$.

(iii) $[E, I, i]$

This cannot be an envy-free equilibrium because the EF conditions for $I$ and $E$ are

$$\theta c_I (1 - c_i b_i / c_I) + \delta \mathbb{E}[\pi^I_i] \geq c_I (1 - c_i b_I / c_I) + \delta \mathbb{E}[\pi^I_i] \iff c_I b_I \geq (1 - \theta) c_I + \theta c_i b_i,$$

$$\mu_E (1 - c_i b_I / \mu_E) + \delta \mathbb{E}[\pi^I_E] \geq \theta \mu_E (1 - c_i b_i / \mu_E) + \delta \mathbb{E}[\pi^I_E] \iff c_I b_I \leq (1 - \theta) \mu_E + \theta c_i b_i.$$

And since $\mu_E < c_I$, there does not exist any $b_I$ for any $b_i$ such that $I$ and $E$ are
envy-free.

Case 2: Suppose $c_i < \mu_E \leq c_I$.

(i) $[I, i, E]$

This cannot be an envy-free equilibrium because the EF conditions for $E$ and $i$ are

$$
\theta \mu_E (1 - \mu_E b_E / \mu_E) + \delta \mathbb{E}[\pi_I^i] \leq \delta \mathbb{E}[\pi_E^0] \iff \theta \mu_E b_E \geq \theta \mu_E + \delta (\mathbb{E}[\pi_E^i] - \mathbb{E}[\pi_E^0]),
$$

$$
\theta c_i (1 - \mu_E b_E / c_i) + \delta \mathbb{E}[\pi_i^0] \geq \delta \mathbb{E}[\pi_i^i] \iff \theta \mu_E b_E \leq \theta c_i - \delta \mathbb{E}[\pi_i^i],
$$

where we have used the fact that $\mathbb{E}[\pi_i^0] = 0$ if $c_i < \mu_E$. Since $\mathbb{E}[\pi_E^i] - \mathbb{E}[\pi_E^0] \geq 0$ from (A.15), it follows that $\theta \mu_E + \delta (\mathbb{E}[\pi_E^i] - \mathbb{E}[\pi_E^0]) > \theta c_i - \delta \mathbb{E}[\pi_i^i]$. This implies that there does not exist any $b_E$ such that $E$ and $i$ are envy-free.

(ii) $[I, E, i]$

EF conditions for $i$ and $E$:

$$
\theta c_i (1 - c_i b_i / c_i) + \delta \mathbb{E}[\pi_i^0] \leq \delta \mathbb{E}[\pi_i^i] \iff \theta c_i - \delta \mathbb{E}[\pi_i^i] \leq \theta c_i b_i
$$

$$
\theta \mu_E (1 - c_i b_i / \mu_E) + \delta \mathbb{E}[\pi_E^i] \geq \delta \mathbb{E}[\pi_E^0] \iff \theta c_i b_i \leq \theta \mu_E + \delta (\mathbb{E}[\pi_E^i] - \mathbb{E}[\pi_E^0])
$$

EF conditions for $E$ and $I$:

$$
c_I (1 - \mu_E b_E / c_I) + \delta \mathbb{E}[\pi_I^i] \geq \theta c_I (1 - c_i b_i / c_I) + \delta \mathbb{E}[\pi_I^i] \iff \mu_E b_E \leq (1 - \theta) c_I + \theta c_i b_i
$$

$$
\mu_E (1 - \mu_E b_E / \mu_E) + \delta \mathbb{E}[\pi_E^i] \leq \theta \mu_E (1 - c_i b_i / \mu_E) + \delta \mathbb{E}[\pi_E^0] \iff \mu_E b_E \geq (1 - \theta) \mu_E + \theta c_i b_i
$$

Therefore, the lowest revenue envy-free equilibrium bids are $b_i^* = \max \left[ 1 - \delta \int_0^{c_i} 1 \frac{\mu_E}{c_i} dF_E, 0 \right]$ and $b_E^* = 1 - \theta + \frac{\theta c_i b_i^*}{\mu_E}$.

(iii) $[E, I, i]$
This cannot be an envy-free equilibrium because the EF conditions for $I$ and $E$ are

\[ c_I (1 - c_I b_I / c_I) + \delta E[I] \leq \theta c_I (1 - c_I b_I / c_I) + \delta E[I] \iff c_I b_I \geq (1 - \theta) c_I + \theta c_i b_i, \]

and $\mu E(1 - c_I b_I / \mu E) + \delta E[I] \geq \theta \mu E(1 - c_i b_i / \mu E) + \delta E[I] \iff c_I b_I \leq (1 - \theta) \mu E + \theta c_i b_i$.

And since $\mu E \leq c_I$, there does not exist any $b_I$ for any $b_i$ such that $I$ and $E$ are envy-free.

### A.12 Statement and Proof of Proposition 8

**Proposition 8 (Three Forces Generated by Ad Credit).** Suppose $c_I$ and $c_E$ are distributed according to $F_I(c) = c^n$ and $F_E(c) = c$, respectively, for $c \in [0, 1]$, and $R \geq \sqrt{1+\delta - \frac{1}{\delta}}$. Then the extraction effect is concave in $\alpha$, and the expected cost of ad credit increases monotonically in $\alpha$. The learning effect increases convexly in $\alpha$ for $\alpha \leq \frac{1+R^2\delta}{2}$, and plateaus thereafter.

**Proof.** We prove the results in turn for the extraction effect, cost, and learning effect.

**Extraction effect:** From the proof of Proposition 3, we obtained the unique existence of $\hat{c} > \mu E = \frac{1}{2}$ such that the entrant wins the first stage auction for all $c_I < \hat{c}$ and the incumbent wins for all $c_I > \hat{c}$. Since this result holds without ad credit (i.e., $\alpha = 0$), it must be that when the search engine offers positive ad credit to the entrant, the threshold $c_I$ after which the incumbent wins must be at least as large as $\hat{c}$. For otherwise, it would imply that the incumbent wins for a larger interval of $c_I$ even if the ad credit makes the entrant’s ad score more competitive (see Proposition 4), which cannot be true. Therefore, we obtain that if the search engine offers non-zero ad credit, the entrant wins the first stage auction for all $c_I < \mu E = \frac{1}{2}$. This property helps simplify the expression for the two positive forces
generated by the search engine’s ad credit. The extraction effect can be written as

\[ \int_0^1 (\mu_E b_{E1}^*(\alpha) - \mu_E b_{E1}^*(0)) \mathbb{I}_{\{c_I b_{I1} \geq \mu_E b_{E1}^*(\alpha)\}} dF_I = \alpha \int_{c_I b_{I1} \geq \mu_E b_{E1}^*(\alpha)} dF_I = \alpha \int_{c_I \geq \frac{1}{2}, c_I b_{I1} \geq \mu_E b_{E1}^*(\alpha)} n c_I^{n-1} dF_I \quad (A.22) \]

Next, we will express the set \( \{c_I : c_I \geq \frac{1}{2}, c_I b_{I1} \geq \mu_E b_{E1}^*(\alpha)\} \) in terms of \( c_I \). To that end, we have

\[
c_I b_{I1} = c_I \left(1 + \delta \left(1 - \frac{\mu_E}{c_I} \right) \mathbb{I}_{\{c_I \geq \mu_E\}} - \int_0^1 \left(1 - \frac{\max[c_E, R]}{c_I} \right) dc_E \right)
= c_I \left(1 + \delta \left(1 - \frac{1}{2c_I} \right) - \int_0^{c_I} \left(1 - \frac{\max[c_E, R]}{c_I} \right) dc_E \right)
= c_I \left(1 + \delta \left(1 - \frac{1}{2c_I} \right) - \int_0^R \left(1 - \frac{R}{c_I} \right) dc_E - \int_R^{c_I} \left(1 - \frac{c_E}{c_I} \right) dc_E \right)
= c_I + \delta \left(c_I - \frac{1}{2} + \frac{R^2 - c_I^2}{2} \right)
\]

\[
\mu_E b_{E1}^*(\alpha) = \mu_E \left(1 + \frac{\alpha}{\mu_E} + \frac{\delta}{\mu_E} \left(\int_{\max[c_I, R]}^1 (c_E - c_I) df_E - (\mu_E - \max[c_I, R]) \mathbb{I}_{\{\mu_E > c_I\}}\right)\right)
= \frac{1}{2} + \alpha + \delta \int_{c_I}^1 (c_E - c_I) df_E
= \frac{1}{2} + \alpha + \frac{\delta}{2} (1 - c_I)^2,
\]

which implies that the condition “\( c_I \geq \frac{1}{2} \) and \( c_I b_{I1} \geq \mu_E b_{E1}^*(\alpha)\)” is equivalent to

\[
\alpha \leq \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta} \quad \text{and} \quad c_I \geq \max \left(\frac{1}{2}, \frac{2\delta - \sqrt{-4\alpha\delta + 2\delta + 2\delta^2 R^2 + 1} + 1}{2\delta}\right) = \tilde{c}_3.
\]

Therefore, we have

\[(A.22) = \begin{cases} 
\alpha(1 - F_I(\tilde{c}_3)) & \text{if } \alpha \leq \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta}, \\
0 & \text{if } \alpha > \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta}.
\end{cases}\]

Finally, to show that \( (A.22) \) is concave in \( \alpha \) for \( \alpha \leq \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta} \), it suffices to establish that
"1 − F_I(\tilde{c}_3)" is decreasing and concave in \( \alpha \). For then,

\[
\frac{\partial^2 \alpha(1 - F_I(\tilde{c}_3))}{\partial \alpha^2} = 2(1 - F_I(\tilde{c}_3))' + \alpha(1 - F_I(\tilde{c}_3))'' < 0.
\]

But the decreasing property and concavity hold because

\[
(1 - F_I(\tilde{c}_3))' = (1 - (\tilde{c}_3)^n)' = -\frac{n \left( \frac{2\delta - \sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1} + 1}{2\delta} \right)^{n-1}}{\sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1}} < 0
\]

and \((1 - F_I(\tilde{c}_3))'' = (1 - (\tilde{c}_3)^n)''\), which in turn equals to

\[
-\frac{n 4\delta^2 \left( \frac{q_1(\alpha, \delta, n, R)}{2\delta} \right)^n q_2(\alpha, \delta, n, R)}{(-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1)^{3/2} q_1(\alpha, \delta, n, R)^2} \\
\propto -q_1(\alpha, \delta, n, R)^n q_2(\alpha, \delta, n, R) \\
= -q_1(\alpha, \delta, n, R)^{n+1} - q_1(\alpha, \delta, n, R)^n (n - 1) \sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1} \\
\leq -q_1(\alpha, \delta, n, R)^{n+1} < 0,
\]

where

\[
q_1(\alpha, \delta, n, R) = 2\delta - \sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1} + 1,
\]

\[
q_2(\alpha, \delta, n, R) = 2\delta + n \sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1} - 2\sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1} + 1.
\]

The last inequality \(-q_1(\alpha, \delta, n, R)^{n+1} < 0\) follows from

\[
-(2\delta - \sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1} + 1)^{n+1} < 0 \iff 2\delta - \sqrt{-4\alpha\delta + 2\delta + 2\delta^2R^2 + 1} + 1 > 0 \\
\iff \alpha > -\frac{1}{2} - \delta \left( 1 - \frac{R^2}{2} \right) \\
\iff 0 > -\frac{1}{2} - \delta \left( 1 - \frac{R^2}{2} \right) \\
\iff \delta \geq 0, R \leq 1.
\]

*Expected cost:* Recall that the entrant wins the first stage auction for \( c_I \leq \frac{1}{2} \), so the expected
ad cost is increasing in $\alpha$ (at a directly proportional rate) for this region. We will show that the expected cost for $c_I > 1/2$ is also increasing in $\alpha$. To that end, it suffices to show that

$$\frac{\partial}{\partial \alpha} \int_{c_I \geq \frac{1}{2}, c_I b_I^* \geq \mu_E b_{E_1}^*(\alpha)} dF_I \leq 0. \quad (A.23)$$

For then,

$$\frac{\partial}{\partial \alpha} \int_{c_I \geq \frac{1}{2}, c_I b_I^* \geq \mu_E b_{E_1}^*(\alpha)} dF_I = \frac{\partial}{\partial \alpha} \alpha \left( 1 - \int_{c_I \geq \frac{1}{2}, c_I b_I^* \geq \mu_E b_{E_1}^*(\alpha)} dF_I \right)$$

$$= 1 - \int_{c_I \geq \frac{1}{2}, c_I b_I^* \geq \mu_E b_{E_1}^*(\alpha)} dF_I - \frac{\partial}{\partial \alpha} \int_{c_I \geq \frac{1}{2}, c_I b_I^* \geq \mu_E b_{E_1}^*(\alpha)} dF_I \geq 0.$$

But the desired inequality $[A.23]$ holds because

$$\frac{\partial}{\partial \alpha} \int_{c_I \geq \frac{1}{2}, c_I b_I^* \geq \mu_E b_{E_1}^*(\alpha)} dF_I = \frac{\partial}{\partial \alpha} \left\{ \begin{array}{ll} 1 - F_I(\tilde{c}_3) & \text{if } \alpha \leq \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta}, \\ 0 & \text{if } \alpha > \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta} \end{array} \right\}$$

$$= \begin{cases} -f_I(\tilde{c}_3) \frac{\partial \tilde{c}_3}{\partial \alpha} & \text{if } \alpha \leq \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta}, \\ 0 & \text{if } \alpha > \frac{2\delta + 2\delta^2 R^2 + 1}{4\delta} \end{cases} \leq 0,$$

where the last inequality holds because $f_I(\cdot) \geq 0$ and

$$\frac{\partial \tilde{c}_3}{\partial \alpha} = \begin{cases} (-4\alpha \delta + 2\delta + 2\delta^2 R^2 + 1)^{-\frac{1}{2}} & \text{if } \delta + 1 \geq \sqrt{-4\alpha \delta + 2\delta + 2\delta^2 R^2 + 1}, \\ 0 & \text{otherwise,} \end{cases} \geq 0.$$

**Learning effect:** For comparative statics, we need only consider terms in the learning value expression which are dependent on $\alpha$. Since the entrant wins the first stage auction for all $\alpha \geq 0$ if $c_I < \frac{1}{2}$, the search engine’s second stage payoff depends on $\alpha$ only for $c_I \geq \frac{1}{2}$. 70
Therefore, \( \frac{\partial (\pi_{S2}(\alpha) - \pi_{S2}(0))}{\partial \alpha} = \frac{\partial \pi_{S2}(\alpha)}{\partial \alpha} \), which in turn is equal to

\[
\frac{\partial}{\partial \alpha} \left( \int_{c_l \geq \frac{1}{2}, c_l b^*_1 \geq \mu_E b^*_E_1(\alpha)} \frac{1}{2} dF_I \right) + \int_{c_l \geq \frac{1}{2}, c_l b^*_1 < \mu_E b^*_E_1(\alpha)} \left( \int_0^{c_l} \max[c_E, R] dF_E + (1 - F_E(c_l)) c_l \right) dF_I,
\]

which simplifies to

\[
\frac{\partial}{\partial \alpha} \left( \int_{c_l \geq \frac{1}{2}, c_l b^*_1 \geq \mu_E b^*_E_1(\alpha)} \frac{1}{2} dF_I + \int_{c_l \geq \frac{1}{2}, c_l b^*_1 < \mu_E b^*_E_1(\alpha)} \left( \frac{R^2 - c_l^2}{2} + c_l \right) dF_I \right). \tag{A.25}
\]

Again, recall from the proof of Proposition 3 that there exists a unique \( \tau > \mu_E = \frac{1}{2} \) such that the entrant wins the first stage auction for all \( c_I < \tau \) and the incumbent wins for all \( c_I \geq \tau \). Therefore, we can find a threshold ad credit level \( \alpha \) such that for all \( \alpha > \alpha \), the entrant wins the first stage auction for all \( c_I \). For this range of \( \alpha \), the learning value would be zero because the entrant’s CTR would have been revealed regardless of the incumbent’s CTR. By monotonicity, the threshold \( \alpha \) must occur at the \( \alpha \) such that the ad ranks of the two advertisers coincide for \( c_I = 1 \). Simple algebra yields \( \alpha = \frac{1+\delta R^2}{2} \). Therefore,

\[(A.25) = \begin{cases} 
\frac{\partial}{\partial \alpha} \left( \frac{1}{2} c_n c_I^{n-1} dc_I + \int_{c_I}^{c_n} \left( \frac{R^2 - c_I^2}{2} + c_I \right) n c_I^{n-1} dc_I \right) & \text{if } \alpha \leq \frac{1+\delta R^2}{2}, \\
0 & \text{if } \alpha > \frac{1+\delta R^2}{2}, \\
\frac{\partial}{\partial \alpha} \left( -\frac{1}{2} \tilde{c}_4 + \left( \frac{R^2 - \tilde{c}_4}{2} - \frac{n}{2} \tilde{c}_4^{n+2} + \frac{n}{n+1} \tilde{c}_4^{n+1} \right) \right) & \text{if } \alpha \leq \frac{1+\delta R^2}{2}, \\
0 & \text{if } \alpha > \frac{1+\delta R^2}{2}. 
\end{cases}
\]

where \( \tilde{c}_4 \triangleq \frac{2\delta - \sqrt{-4\alpha \delta + 2\delta + 2\delta^2 R^2 + 1} + 1 + \delta^2 R^2 - 1}{2\delta} \). By explicitly solving the derivative in the top branch, it can be shown that the derivative is proportional to

\[
2 \alpha \delta - \delta + \sqrt{-4\alpha \delta + 2\delta + 2\delta^2 R^2 + 1} + \delta^2 R^2 - 1. \tag{A.26}
\]

Therefore, using the fact that the \((A.26)\) is increasing in \( R \), we obtain the following learning
effect pattern for general range of $R$:

$$\text{(A.26)} = \begin{cases} + & \text{if } \frac{1}{2} < \alpha \leq \frac{1+R^2\delta}{2} \text{ or } \alpha \leq \frac{1}{2}, R > r', \\ - & \text{if } \alpha \leq \frac{1}{2}, R \leq r', \end{cases}$$

where $r' = \sqrt{1 + (1 - 2\alpha)\delta - 1 + \frac{1}{2}}$. With the added assumption that $R \geq \sqrt{1 + \delta - 1}$, we obtain that the derivative is always positive for $\alpha \leq \frac{1+R^2\delta}{2}$ and zero for $\alpha > \frac{1+R^2\delta}{2}$. To establish convexity, we need only differentiate (A.26) with respect to $\alpha$ once more. This yields that the second derivative of the learning effect is

$$\frac{\partial^2 \text{(A.26)}}{\partial \alpha^2} = 2\delta + \frac{1}{2} (-4\alpha\delta + 2\delta + 2\delta^2 R^2 + 1)^{-\frac{1}{2}} (-4\delta) \propto 1 - (-4\alpha\delta + 2\delta + 2\delta^2 R^2 + 1)^{-\frac{1}{2}} \geq 0,$$

where the last inequality follows from the fact that $\alpha \leq \frac{1+R^2\delta}{2} \implies -4\alpha\delta + 2\delta + 2\delta^2 R^2 + 1 \geq 1 \implies 1 - (-4\alpha\delta + 2\delta + 2\delta^2 R^2 + 1)^{-\frac{1}{2}} \geq 0$.

**B Optimal Ad Credit**

The search engine determines the ad credit rate $\alpha$ prior to the realization of $c_I$ to maximize its expected profit. A closed-form analytical expression for optimal $\alpha$ is not tractable with general CTR functional forms $F_I$ and $F_E$. In this section, we provide numerical plots for specific functional forms and parameter values to better understand the forces that shape the search engine’s optimal ad credit level.

Figure 9 plots the optimal ad credit with respect to the Stage 2 weight parameter $\delta$ and the incumbent CTR strength parameter $n$. Figure 9a reveals an interesting relationship between the optimal ad credit level and $\delta$. When $n$ is small such that the incumbent is likely to be weak, then the ad credit decreases with $\delta$ (solid line in Figure 9a). On the other hand, when $n$ is large such that the incumbent is likely to be strong, then the ad credit increases with $\delta$ (dashed line in Figure 9a). This pattern can be understood based on the insights obtained
in the main model.

Intuitively, when the incumbent is likely to be weak, the entrant is more likely to win in Stage 1; therefore, offering ad credit would be costly for the search engine. And as $\delta$ increases, the entrant has stronger incentives to overbid to capitalize on the second stage benefits of revealing its CTR. Combined together, these two effects lower $\alpha^*$. On the other hand, when the incumbent is likely to be strong, then the entrant bids close to valuation, whereas the incumbent overbids by a larger magnitude as $\delta$ increases (see Lemma 2). The search engine rationally anticipates that under such bidding patterns, the entrant’s CTR will be increasingly masked by the incumbent for higher $\delta$. To address this, the search engine offers larger ad credit to the entrant in order to both learn the entrant’s CTR more quickly and also extract more surplus from the incumbent.

Finally, Figure 9b shows that the ad credit level is increasing with the incumbent strength parameter $n$. This pattern resonates with the discussion of Proposition 5 which can be roughly interpreted as ‘the stronger is the incumbent, the greater the benefits of ad credit.’ Intuitively, if the incumbent is likely to be strong, then the entrant’s CTR is likely to be masked by the incumbent’s aggressive bid. This induces the “harmful ignorance” where not knowing the entrant’s CTR lowers the search engine’s revenue compared to the full information benchmark. Therefore, the higher the $n$, the larger the ad credit the search engine offers to benefit from the learning effect and incumbent surplus extraction effect.