Essays on Advertising

Woohyun (Jason) Choi

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

Abstract

Essays on Advertising

Woohyun (Jason) Choi

According to eMarketer, the total advertising spend in US alone was estimated to be over $238 billion. Firms invest large amounts of money in advertising to promote and inform consumers about their products and services, as well as to persuade them to purchase. The broad theme of advertising has been examined from many different angles in the marketing literature, ranging from empirically measuring effects of TV ads on sales to analytically characterizing the key economic forces stemming from enhanced targetability in online advertising. The purpose of my dissertation is to study some of the key questions which remain unaddressed in the advertising literature. In the first essay, I examine firms' choices of advertising content in a competitive setting. I demonstrate that competitive forces sometimes induces firms to choose advertising content that shifts consumers' perception of product quality. While this strategy hurts firms in a monopoly setting, it increases their profits under competition because it may increase the utility of their offering in comparison with the competing offering. In the second essay, I investigate the optimal mechanism for selling online ads in a learning environment. Specifically, I show that when ad sellers, such as Google, design their ad auctions, it is optimal for them to favor new advertisers in the auction in order to expedite learning their ad performance. In the third essay, I study the impact of tracking consumers' Internet activities on the online advertising ecosystem in the presence of regulations that, motivated by privacy concerns, endow consumers with the choice to have their online activity be tracked or not. I find that when ad effectiveness is intermediate,

fewer ads are shown to opt-in consumers, who can be tracked and have their funnel stages inferred by advertisers, than to opt-out consumers, who cannot be tracked. In this case, consumers trade-off the benefit of seeing fewer ads by opting-in to tracking (positive instrumental value of privacy) with the disutility they feel from giving up their privacy (intrinsic cost of privacy). Overall, these findings shed light on novel strategic forces that provide guidance for marketers' advertising decisions in three distinct contexts.

Table of Contents

List of Tables

vi

List of Figures

viii

Acknowledgments

First and foremost, I would like to thank God for giving me the opportunity to be a part of the Columbia community, where I had the unique privilege of befriending and learning from some of the brightest intellectuals. Everyday I am grateful for Your subtle interventions in my life. I pray that this degree will be used for nothing less than the glorification of Your name, and the fulfillment of Your will.

Second, I would like to extend a sincere note of gratitude to my advisors Kinshuk and Miklos. Thank you for being patient with me throughout this long journey, and for infusing confidence in me when I most needed it. To say that I have learned valuable lessons from my interactions with you would be an understatement. It has been an immense pleasure and privilege learning from you, and I greatly look forward to continuing our collaboration in the coming years.

Last but not least, I would like to express my deepest love and appreciation to my wife, Jessica. Without your unwavering support and love, I would not have made it through the arduous process, especially the last year of the program. I am overwhelmed with joy and gratitude as we close together this chapter of our lives and start the next as a family of three.

Chapter 1: Introduction

According to eMarketer, the total US ad spend in 2019 increased 6.84% year-over-year to $238 billion.[1] This growth has been partially driven by advances in information technology which generated (i) a host of new media platforms beyond traditional TV and newspapers where firms and consumers can interact (e.g., social media, search platforms, and streaming platforms), and (ii) sophisticated mechanisms for matching the two parties with extraordinary efficiency (e.g., online tracking, pay-per-click pricing, and real-time auctions) (Economist, 2006, 2015).

The burgeoning industry, however, is not without its challenges. Many issues deter marketers' advertising efforts; for instance, marketers are concerned about (i) how to display relevant advertising content[2], (ii) how to identify profitable advertising opportunities and which bidding strategies to deploy[3], and (iii) how to adjust their business to shifting regulatory landscapes (e.g., the European Union's General Data Protection Regulation that empowers consumers to data privacy rights).

In my dissertation, I attempt to address important questions — motivated by these managerial issues — pertaining to firms' advertising strategies, which have received scant attention in the literature. I develop game-theoretic models to shed light on key economic forces that affect the players' incentives and strategies, thereby providing guidance on marketers' advertising strategies. In Chapter 2, I describe a model of firms' advertising content choice strategies in a competitive setting and present, among others, the key finding that when consumers exhibit reference-dependent preferences that yield concave utility functions (e.g., due to loss aversion), then the high-quality firm sometimes displays reference-shifting advertising content that, surprisingly, reduces consumers'

---

[1] https://forecasts-na1.emarketer.com/584b26021403070290f93ac0/5851918c0626310a2c186c0c

[2] https://marketingland.com/37-consumers-say-ads-placed-next-offensive-content-impact-brand-perception-217504

[3] https://searchengineland.com/4-ways-to-determine-your-your-starting-bids-144616.

1

absolute valuation of its offering. While this strategy lowers a monopolist's profit, it can increase the high-quality firm's profit in a competitive setting because it increases the *relative* utility of its offerings compared with the low-quality firm's offerings. In Chapter 3, I show how learning incentives in online advertising significantly impact the advertisers' bidding strategies and the seller's auction mechanism. While previous literature suggests that advertisers should bid their truthful valuation in a second-price auction, I demonstrate that this is not necessarily true for new advertisers. That a new advertiser's ad performance information is revealed to the search engine only if it wins the auction incentivizes the new advertiser to bid above its valuation in early rounds of the auction. In Chapter 4, I discuss the implications of tracking consumers' online activities on the advertising ecosystem, particularly in the context of recent privacy regulations. I find that endowed with new privacy rights, consumers sometimes choose to opt-in to online tracking in order to trade-off their privacy cost with the benefit of seeing fewer ads. Interestingly, my analysis suggest that endogenous tracking choices may result in scenarios where displaying more effective ads decreases the ad network's profit.

Chapter 2: Persuasive Advertising in a Vertically Differentiated Market

2.1 Introduction

One of the key roles of advertisements besides conveying product information is to influence consumers' product evaluation in a way that is more favorable to the advertising firm. Advertisements can aid in creating stronger preference, and thus enhanced market power for the advertising firm. If ads indeed influence consumers, how should competing firms strategically choose their advertising messages? In this paper I study this question in the context of a vertically differentiated market.

Consider the examples in Figure 2.1. An ad for Mercedes-Benz, a high-end automobile manufacturer, reads "The best or nothing" (Figure 2.1a). Similarly, an ad for Verizon, a telecommunications company, boasts that it has the nation's "largest 3G network" (Figure 2.1c). These two ads seek to persuade buyers of cars and telecommunications services, respectively, that they should not compromise on quality when making a purchase. In contrast, an ad for a car produced by Kia, a car brand commonly associated with affordability, states that the car "was named Best Compact Car for the Money" (Figure 2.1b). This statement emphasizes value for money rather than quality-related features. Within the car category, while the ad by Mercedes-Benz succinctly underscores the importance of quality, thereby inducing consumers to care less about price, the focus on "value for money" in the Kia ad creates the opposite effect whereby consumers are called to trade off quality against price. This example illustrates a typical difference in advertising content between competing firms — high-quality firms typically emphasize quality while lower-quality firms often focus on price-related features in their ads. Therefore, ads can affect consumers' sensitivity for quality.

Another effect that ads can have is that of shifting the reference point against which quality is

(a)

(b)
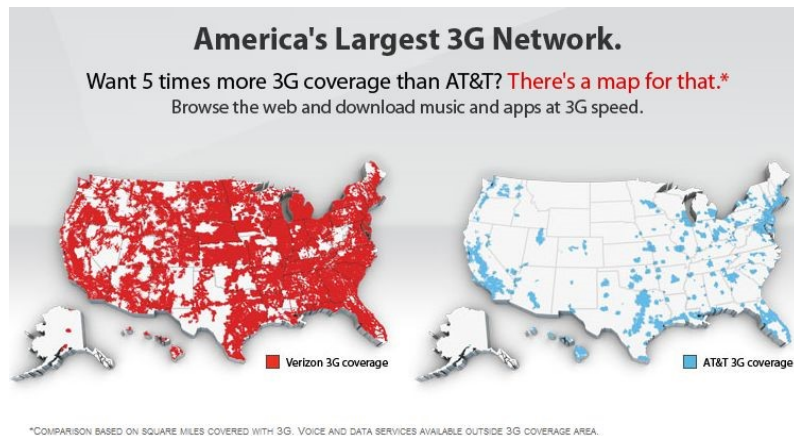
(c)

Figure 2.1: Ads with Different Persuasive Effects

evaluated; i.e., ads can influence the internal benchmark quality level that consumers invoke when evaluating quality by providing anchor points to consumers. For instance, the Verizon ad displays a map with nearly complete network coverage (and explicitly compares this with its competitor's significantly lower network coverage), which can shift the standard against which potential phone service buyers evaluate network coverage. The other two ads can also be argued to influence the quality anchor albeit to a lesser extent (the Mercedes-Benz ad because it does not show or describe a car, and the Kia ad because its message focuses primarily on the quality-price trade-off). Past research on anchoring has demonstrated that across a wide range of domains, consumers' evaluations of products and experiences are heavily influenced by anchor points (Epley and Gilovich, 2001; Mussweiler and Strack, 2001). These anchor points have been shown to be malleable to even subtle interventions that are unrelated to the task (e.g., Ariely et al. (2003) show that subjects' evaluations of an experience were significantly affected by the recall of their social security numbers). Furthermore, the literature on context-dependent preferences suggests that quality-message advertisements may shift consumers' reference quality (Simonson and Tversky, 1992). Based on these ideas, I posit that quality messages in ads can shift the "quality anchor" or reference quality against which consumers evaluate quality.

The two types of advertising content effects discussed above (influencing the customers' sensitivity to quality, and influencing their reference point for quality) are characteristic of vertically differentiated markets where firms offer products of different quality levels. My goal in this paper is to understand the strategic impact of these different ad effects in a competitive setting. Also, while a particular ad can be expected to have both effects simultaneously, ads can be designed that have more of one effect than the other. I investigate and provide guidance on how firms should, under different market conditions, optimally choose ad content corresponding to each type of effect. Overall, my work falls under the umbrella of persuasive advertising in the context of vertically differentiated markets.

The paper presents a game theory model in which two competing firms offering products of different qualities decide their ad strategies — whether to show ads and, if yes, what their content

should be — as well as prices. I assume that ad content can be of two types with different effects — shifting a consumer's valuation of quality such that she values quality either more or less relative to price (valuation shifting), and shifting a consumer's perception of quality by altering the reference point against which she evaluates quality (reference shifting). A representative consumer has private knowledge about her marginal valuation of quality, has diminishing marginal utility for quality, and evaluates quality relative to a reference point. Using this framework, I ask the following two main questions: (1) If firms show ads, what type of content should they include in their ads? (2) How does this affect firms' pricing, profits and consumer surplus?

I start my analysis by considering a monopolist selling two products. I find that the monopolist only chooses either ad content that increases consumer valuation of quality, or ad content that lowers the reference quality; both of these strategies increase the consumer's perceived surplus from the product. Specifically, the monopolist never chooses ad content that increases the reference quality because doing so would only reduce the consumers' perceived value of quality.

However, the insights are quite different in a competitive setting. For instance, the low-quality firm may choose valuation-shifting ad content that decreases consumers' valuation of quality — while this reduces the extractable surplus, it helps the lower quality firm attract more customers. Furthermore, the high-quality firm may choose reference-shifting ad content to raise the reference quality, because even though this undermines consumers' perceived value of quality, it increases perceived product differentiation. Therefore, under competition, firms may counterintuitively choose ad strategies with apprently negative effects such as reducing the average valuation of quality in the market and reducing the perceived quality of the product. These are strategies that a monopolist would never choose.

I also derive implications for (perceived) consumer surplus. Standard economic theory suggests that as competition intensifies, prices fall and consumer surplus increases. This reasoning, however, rests on the premise that consumers' quality preferences remain constant. In the context of persuasive advertising, firms' advertising strategies change consumers' perceptions and product valuations. I show that even if stronger competition lowers firms' prices, a switch in advertis-

ing regimes may cause consumers to value quality less such that the overall perceived consumer surplus decreases.

A key assumption on which my results depend is that consumers have diminishing marginal utility for quality, i.e., utility from quality is concave in quality. This assumption is supported by Tversky and Kahneman (1979), who argue that Prospect Theory, which features a concave value function in the gains domain, is applicable not only to monetary outcomes, but also "to choices involving other attributes; e.g., quality." In fact, in my main model, I operationalize the concavity in the utility function by invoking components of Prospect Theory, specifically, reference dependence and loss aversion. Furthermore, to the extent that quality can be viewed as a form of stimulus experienced by the consumer (i.e., the perceived level of happiness the consumer feels for a given level of quality), this property is also supported by the psychology literature, which suggests that the marginal change in hedonic response to a stimuli "decreases with the distance from the reference point" (Tversky and Kahneman, 1991; Frederick and Loewenstein, 1999).

My work is related to the literature on persuasive advertising, which is advertising that operates by changing the consumers' perceived utility function (Dixit and Norman, 1978; Bagwell, 2007). Within this broad classification, my focus is on a competitive setting. A large body of research has studied the "combative" role of advertising where horizontally differentiated firms decide the level of spending on ads that shift consumer preferences (Von der Fehr and Stevik, 1998; Bloch and Manceau, 1999; Tremblay and Polasky, 2002; Chen et al., 2009). Surprisingly, little attention has been paid to the persuasive role of advertising in a vertically differentiated market. A couple of notable exceptions are Tremblay and Martins-Filho (2001) and Tremblay and Polasky (2002). Tremblay and Martins-Filho (2001) study demand-expansion effects of persuasive ads that change consumer's taste for quality, whereas I keep the market size fixed and allow for more flexibility in the firms' ad content choices. Tremblay and Polasky (2002) assume that "perceived quality equals the amount of advertising," whereas I study firms endowed with distinct quality levels advertising to influence not only consumers' perceptions of quality, but also their valuation for quality. Furthermore, while the main object of interest in Tremblay and Martins-Filho (2001) and Tremblay

and Polasky (2002) is advertising intensity, my focus is primarily on the types of ad content firms choose, namely, valuation-shifting content and reference-shifting content.

My work is also related to Kopalle and Lehmann (1995), Kopalle and Assunção (2000) and Kopalle and Lehmann (2006), who study the impact of setting reference quality on firm strategy, albeit in a monopoly setting. The research on reference-dependent quality evaluations is also related to my research; specifically, Hardie et al. (1993) and Kopalle et al. (1996) use the concept of reference-dependent quality evaluations in the contexts of choice models and dynamic product quality decisions, respectively.[1]

Finally, I note that a large literature studies the informative role of advertising through signaling when consumers are not perfectly informed about the product (Nelson, 1974; Kihlstrom and Riordan, 1984; Milgrom and Roberts, 1986; Mayzlin and Shin, 2011; Gardete, 2013). In my framework, consumers know absolute quality and thus there is no information asymmetry. Previous work that has examined the content of advertising also falls under the umbrella of the informative role of advertising (e.g., Anderson and Renault (2006)).

The rest of the paper is organized as follows. In Section 2.2, I develop the basic model used in my analysis. In Section 2.3, I examine the benchmark model in which a monopolist firm decides its ad strategy — namely, it chooses whether or not to show an ad and what the ad content should be — and price. In Section 2.4, I investigate the main model in which competing firms choose their ad strategies and prices. In Section 2.5, I assess the robustness of the main results by examining various extensions. Finally, in Section 2.6, I discuss managerial implications and conclude. For ease of exposition, I relegate all proofs and lengthy expressions to the appendix.

2.2 Model

Firms

I assume that there are two firms, each producing one good of distinct quality level. I denote the qualities of the high- and low-quality products by q_H and q_L, respectively, where $q_H > q_L > 0$.

---

[1]I note that this paper is distinct from research on the effects of reference prices on firm behavior (Kopalle and Winer, 1996; DellaVigna, 2009; Baucells et al., 2011).

I refer to the firm that produces the high-quality product as Firm H, and its low-quality counterpart as Firm L. To focus on the role of advertising, I abstract from the quality-production decision and assume that the quality of the firms' products is exogenously determined. I normalize the marginal costs of producing both quality levels to zero.[2]

Firms decide whether to show an ad or not. If a firm advertises, it also chooses its ad content. The cost of showing an ad is $k > 0$. Firms then set prices in competition. I discuss all decisions and their timings in detail as I develop the model.

Consumer

I assume that there is one representative consumer. The consumer's utility from consuming product quality q_j from Firm $j \in \{H, L\}$ is

$$u(q_j; \theta) = A(q_j; \theta) + \beta C(q_j; \chi) + \varepsilon_j. \tag{2.2.1}$$

The above function has three components; $A(q_j; \theta)$ and $\beta C(q_j; \chi)$ are deterministic components and ε_j is a random component. The deterministic components are "the absolute pleasure of consumption" and "the sensation of gain and loss" relative to some reference point (**?**); I call these components the *absolute* and *comparative* utilities, respectively.

$A(q_j; \theta)$ denotes the absolute utility, which measures the utility associated with the product's *absolute* quality level q_j. To model this, I invoke the standard linear utility function:

$$A(q_j; \theta) = V + \theta \, q_j.$$

$V > 0$ denotes the quality-independent, intrinsic utility, and $\theta \, q_j$ deotes quality-dependent utility. The parameter θ denotes the consumer's marginal valuation or taste for quality; it is equal to θ_h with probability α, and equal to θ_l with probability $1 - \alpha$, where $\alpha \in (0, 1)$ and $\theta_h > \theta_l > 0$. For simplicity, I normalize θ_h to 1. While α is common knowledge, only the consumer knows

---

[2]This assumption has been commonly used in models of vertical differentiation in the literature. See, for example, Gardete (2013) and Zhu and Dukes (2017).

her realization of $\theta \in \{\theta_h, \theta_l\}$ and knows this before making a product choice (in that sense, this component of utility is deterministic for the consumer). Note that parameter α can also be interpreted as a preference heterogeneity parameter in a unit-mass consumer model.

$C(q_j; \chi)$ denotes the comparative utility with respect to some reference quality χ (which will be defined shortly). I assume that $C(q_j; \chi)$ is concave in q_j. The concavity can be modeled in various ways. For example, it may be derived from the property of diminishing marginal utility of quality (e.g., $C(q_j; \chi) = v(q_j - \chi)$ with $v''(\cdot) < 0$), or from a linear loss-aversion model (LAM) where the utility in the loss domain is steeper than in the gains domain. While the main insights of the paper hold regardless of the modeling choice,[3] I choose to work with the LAM formulation as it is (a) more convenient for exposition purposes, and (b) more widely used in the literature to model reference-dependent utilities and loss aversion (e.g., Hardie et al., 1993; Bodner and Prelec, 1994; Kivetz et al., 2004a,b; Orhun, 2009; Chen and Turut, 2013; Amaldoss and He, 2018). Therefore, I specify the comparative utility as follows:

$$
C(q_j; \chi) = \begin{cases} q_j - \chi & \text{if } q_j \geq \chi, \\ -\lambda(\chi - q_j) & \text{if } q_j < \chi. \end{cases}
$$

The comparative utility, $C(q_j; \chi)$, represents the utility the consumer derives from comparative evaluation of q_j with respect to the reference point χ. Three key features of $C(q_j, \chi)$ are: (i) the reference quality lies between the attribute levels of the choice-set: $\chi \in [q_L, q_H]$; (ii) the comparative utility is linear in the attribute level; and (iii) the consumer is loss-averse: $\lambda > 1$. These properties constitute the linear LAM as discussed above.

The last component of the consumer's utility, ε_j, denotes the brand-specific preference shock experienced by the consumer on buying Firm j's product. This can arise from various situational factors unobserved by the firms, such as the subtle, behavioral influences of the unique purchase environment associated with each firm (e.g., Schwartz and Clore, 1983; Payne and Johnson, 1993).

---

[3]In Section A.3 of the appendix, I provide a detailed analysis of an alternative formulation of concave utility with $C(q_j; \chi) = v(q_j - \chi)$ where $v''(\cdot) < 0$.

I assume that $\varepsilon_j - \varepsilon_k$ ($j \neq k$) is distributed uniformly over $[-\sigma, \sigma]$ for some large $\sigma > 0$ (e.g., Subramanian et al., 2014; Gardete, 2013). This yields a tractable linear demand model that simplifies the analysis of firms' pricing strategies while preserving the essence of the model. In Section A.3 of the appendix, I present a model without the error structure and show that the qualitative insights hold.

If a consumer purchases the product from Firm j, she pays price p_j which leads to a disutility of p_j (i.e., utility of $-p_j$). Finally, there exists an outside option that yields utility $u_0 + \varepsilon_0$, where u_0 is normalized to zero; this would correspond to, say, not making any purchase. The consumer compares the net utilities from the two products and the outside option, and chooses whichever yields the highest.

Types of Advertising

As discussed in the introduction, I consider two effects of ads — valuation shifting and reference shifting. While an ad can have both types of effects, a particular ad can have content to emphasize one type of effect more than the other. Based on this argument, I make a stylized assumption for my main model that a particular ad can have only one type of effect. Therefore, for my main model, I consider two qualitatively distinct types of persuasive ads: valuation-shifting ads and reference-shifting ads. Making this assumption enables us to communicate the key insights cleanly. In later analysis, I show that allowing ads to have both effects simultaneously does not change the insights qualitatively.

Next, I discuss how I operationalize valuation-shifting and reference-shifting. A *valuation-shifting ad* influences consumer's marginal valuation of quality. Depending on the firm's ad content, a valuation-shifting ad could either increase or decrease the likelihood of a consumer having high valuation for quality, α. For instance, ads that emphasize the importance of quality (e.g., the Mercedes-Benz ad in Figure 2.1a) would increase α by persuading the consumer to become more concerned about (absolute) quality. I denote this type of valuation-shifting ad by $V^\uparrow$; its effect is to increase α by a magnitude $v^\uparrow$.

On the other hand, ads that highlight value (e.g., the Kia ad in Figure 2.1b) have the effect of

11

making consumers less concerned about absolute quality and more concerned about the quality-price trade-off. I denote this type of valuation-shifting ad by $V^\downarrow$ and assume that it increases the likelihood of a consumer having low valuation for quality by a magnitude $v^\downarrow$. In what follows, I assume that $v^\uparrow = v^\downarrow = v$; this assumption is for analytical simplicity and does not alter the key trade-offs captured and insights obtained by the model. If one firm chooses $V^\uparrow$ and the other $V^\downarrow$, the advertising effects negate one another and the consumer's valuation distribution does not change.

The second type of ad is a *reference-shifting ad*. This form of advertising alters consumers' perception of quality by changing the reference point against which consumers evaluate quality (e.g., the Verizon ad in Figure 2.1c); i.e., it influences the internal benchmark quality level that consumers invoke when evaluating quality. As discussed earlier, this is motivated by the streams of behavioral research on anchoring effects (Epley and Gilovich, 2001; Mussweiler and Strack, 2001; Ariely et al., 2003) and context-dependent preferences (Simonson and Tversky, 1992; Kivetz et al., 2004a,b). By highlighting a particular quality level, reference-shifting ads induce consumers to adjust their reference point closer toward the advertised quality level.

Reference point formation has been conceptualized as the centroid of the relevant characteristics of all products considered (Bodner and Prelec, 1994) or a point in the "convex hull of the existing alternatives" (Orhun, 2009). In a similar vein, Bordalo and Shleifer (2017) describe a model of "memory-based reference point," where the reference point is formed as a weighted average of previously observed attribute levels. Adapting theories of memory and recall, Bordalo and Shleifer (2017) posit that larger weights are attached to quality levels that have been observed more frequently.

I invoke such a characterization, and model the consumers' reference quality as

$$\chi \left(a_H^R, a_L^R, \eta_H, \eta_L \right) = \frac{q_H + q_L + a_H^R \eta_H + a_L^R \eta_L}{2 + a_H^R + a_L^R}, \tag{2.2.2}$$

where a_j^R is binary and equals 1 if Firm $j \in \{H, L\}$ chooses reference-shifting ad, and 0 otherwise; and η_j denotes the quality level communicated by Firm j in its reference-shifting ad. A firm can

communicate any quality level in a reference-shifting ad.[4] From a high level perspective, however, I can boil down the reference-shifting ads to two types: one that communicates a high quality level, thereby increasing the reference quality, and another that communicates a low quality level, thereby decreasing the reference quality. For the rest of the paper, I use this simplified formulation. Specifically, I consider two types of reference-shifting ads: $R^\uparrow$, which increases the reference quality by magnitude $\delta^\uparrow$, and $R^\downarrow$, which decreases it by magnitude $\delta^\downarrow$.[5] For further simplicity and without losing any qualitative insights, I assume that $\delta^\uparrow = \delta^\downarrow = \delta$. The assumptions on the ranges of the key model parameters are summarized in Section A.1 of the appendix. For the rest of the paper, I will simply use χ instead of $\chi\left(\cdot, \cdot, \cdot, \cdot\right)$

I clarify that the firm's motivation for communicating a certain quality level through reference-shifting ads is unrelated to the signaling roles studied in games of information asymmetry (e.g., Milgrom and Roberts, 1986; Mayzlin and Shin, 2011; Gardete, 2013). In my paper, Firm j's communication of quality level η_j through reference-shifting ad does not affect consumers' inference of the firm's product quality, as there is no information asymmetry. Instead, Firm j seeks to strategically manipulate the consumer's reference point by communicating quality level η_j. Doing so induces the consumer to put a larger weight on this quality level when forming her reference point. As I will demonstrate in the main analysis, shifting consumer's reference point in this manner will have important implications for competitive strategies.

I can now combine into a single action the firm's choice of whether to advertise or not, and if so,

---

[4]To help develop intuition for how reference-shifting ad affects the reference point in (2.2.2), I provide a few illustrative examples using different (arbitrarily chosen) values of η_j: (a) The default reference quality when neither firm chooses reference-shifting ad (i.e., $a_H^R = a_L^R = 0$) is simply the average of the quality levels in the choice-set: $\chi^0 = \frac{q_H + q_L}{2}$; (b) If only Firm H chooses reference-shifting ad and communicates the highest quality level (i.e., $a_H^R = 1$, $a_L^R = 0$, $\eta_H = q_H$) then the reference point *increases* to $\chi' = \frac{2q_H + q_L}{3}$; (c) If only Firm L chooses reference-shifting ad and communicates the lowest quality level (i.e., $a_H^R = 0$, $a_L^R = 1$, $\eta_L = q_L$), then the reference point *decreases* to $\chi' = \frac{q_H + 2q_L}{3}$; (d) If Firms H and L both choose reference-shifting ad and communicate quality levels q_H and q_L, respectively (i.e., $a_H^R = a_L^R = 1$, $\eta_H = q_H$, $\eta_L = q_L$), then the reference point does not change: $\chi' = \frac{2q_H + 2q_L}{4} = \chi^0$.

[5]The reference formation process (2.2.2) implies that the magnitude of the reference shift induced by a firm's reference-shifting ad depends on its competitor's choice of reference-shifting ad. For example, Firm H's reference-shifting ad has a smaller impact on the reference point if Firm L also chooses reference-shifting ad. This is because the consumer distributes the weights of the reference point across both quality levels communicated by the firms. While the reduced formulation $R^\uparrow$ and $R^\downarrow$ abstracts from this subtlety, it parsimoniously preserves the essence of reference-shifting ads and thus suffices for the purposes of my analysis.

which type of ad to display: that is, Firm $j \in \{H, L\}$ chooses ad strategy $a_j \in \{V^\uparrow, V^\downarrow, R^\uparrow, R^\downarrow, \emptyset\}$, where $\emptyset$ means *no advertising*. I assume that if the firm advertises, it incurs a cost $k > 0$, regardless of the ad type. In Section 2.5 I relax two assumptions about ad-related actions — in Section 2.5.1, I allow a single ad to contain both valuation-shifting and reference-shifting content, and in Section 2.5.2, I allow a firm to implement both types of ad content simultaneously by allocating a fixed budget across different ads. The results that I obtain generalize those in the main model.

Game Timing

The firms play a two-stage game. In the first stage, firms simultaneously decide ad strategy $a_j \in \{V^\uparrow, V^\downarrow, R^\uparrow, R^\downarrow, \emptyset\}$ for $j \in \{H, L\}$. In the second stage, firms observe the decisions from the first stage and simultaneously set prices. I use subgame-perfect Nash equilibrium (SPNE) as the solution concept.

2.3 Benchmark: Monopoly

I begin my analysis with the monopoly case, which will serve as a useful benchmark to help understand more nuanced results under competition. Consider a monopolist that offers two products of distinct quality, q_H and q_L, where $q_H > q_L$.[6] The monopolist first decides the advertising strategy $a_M \in \{V^\uparrow, V^\downarrow, R^\uparrow, R^\downarrow, \emptyset\}$, and then sets prices p_H and p_L for the high- and low-quality products, respectively. If the monopolist advertises, then the consumer's utility is changed according to the advertising effects associated with the content. The consumer then evaluates the two products, as well as the outside option, and purchases whichever yields the highest net utility.

To solve for the SPNE, I first solve the pricing subgame. I provide the details of the analysis in the appendix (see proof of Proposition 1). The optimal pricing strategy is consistent with the solution of the standard price discrimination problem for a monopolist, which has been extensively studied in the market segmentation and product line design literature (e.g., Mussa and Rosen, 1978; Moorthy, 1984, 1988).

---

[6]Another benchmark one could consider is a single-product monopoly. The strategic implications for this case are relatively trivial and, therefore, I omit this analysis.

Next, consider the advertising game stage. Recall the distinct effects of each type of advertising on the consumer's utility: $V^\uparrow$ ($V^\downarrow$) increases (decreases) consumer's taste for quality, and $R^\uparrow$ ($R^\downarrow$) increases (decreases) the reference quality. Given this, what advertising strategy maximizes the monopolist's profit? The following proposition shows that while the monopolist may choose advertising to increase consumer's taste for quality or reduce the reference quality, it will never choose a reference-shifting ad that increases the reference quality.

Proposition 1 (Monopolist Advertising Strategy). *If the monopolist advertises, it either chooses a valuation-shifting ad that increases the likelihood of the consumer having high taste for quality, or a reference-shifting ad that lowers the reference quality (i.e., it only chooses either $V^\uparrow$ or $R^\downarrow$).*

The intuition behind the monopolist's advertising strategy is the following. Increasing consumers' marginal valuation of absolute quality means that, all else equal, consumers would be willing to pay more for the products. The monopolist, therefore, may choose advertising content to increase consumers' valuation of quality and then extract the additional surplus through higher prices. It may also choose reference-shifting ad that lowers the reference quality. The reason is that a lower reference quality induces consumers to derive greater comparative utility from the products: with a lower quality benchmark, consumers perceive the high quality product to be much superior, and the low quality product to be not as bad. Such a shift in the consumer's quality perception also increases consumers' willingness-to-pay (WTP) for the monopolist's products.

Conversely, raising the reference quality has the opposite effect of dampening consumer's WTP for quality. This is because consumers evaluate quality against a higher standard, and thus derive less comparative utility. Figure 2.2 illustrates the decrease in consumer's utility for both products as a result of an upward shift in reference quality. Therefore, the monopolist will never choose to increase the reference quality.

Corollary 1. *The monopolist will never choose reference-shifting ad that increases the reference quality (i.e., it will never choose $R^\uparrow$).*

In the next section, I show that the intuitions and results from the monopoly case do not fully
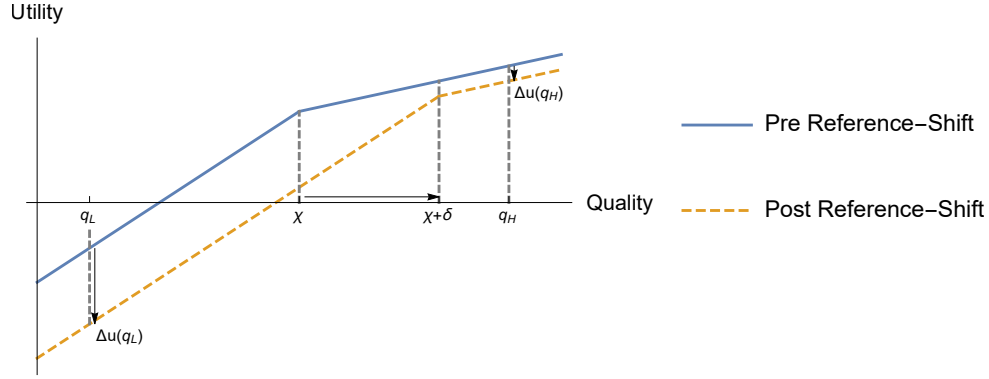
Figure 2.2: Reference-Shift and Consumer Utility for Quality

carry over when competition is introduced. Specifically, under monopoly I see only the ad strategies $V^\uparrow$ and $R^\downarrow$ are used, while in competition I will see that all fmy ad strategies, i.e., $V^\uparrow, V^\downarrow, R^\uparrow$ and $R^\downarrow$ may be used, even though some of these strategies reduce the consumer's total utility for a given quality level.

2.4 Competition

In the duopoly scenario, Firm H offers the product with high quality, q_H, and Firm L offers the product with low quality, q_L, where $q_H > q_L$. Again, I solve by backwards induction, starting from the pricing subgame.

Pricing Subgame

Throughout the discussion of the pricing subgame, I use the *hat* notation to denote post-advertising parameters. For instance, $\hat{\alpha}$ is equal to α increased (decreased) by v if only one firm chose $V^\uparrow$ ($V^\downarrow$) in Stage 1, and $\hat{\chi}$ is equal to the default reference quality $\frac{q_H + q_L}{2}$ increased (decreased) by δ if only one firm chose $R^\uparrow$ ($R^\downarrow$) in Stage 1. To simplify the analysis, I assume that the intrinsic utility V from consuming any product quality is sufficiently large such that the market is covered.

Firm j's demand if it sets price p_j given competitor's price p_k is

$$D_j(p_j, p_k) = \sum_{\theta \in \{1, \theta_l\}} \mathbb{P}\{\theta\} \, \mathbb{P}\{u(q_j; \theta) - p_j > u(q_k; \theta) - p_k \mid \theta\},$$

where the first probability represents the distribution of the consumer's quality taste, and the second probability is with respect to the distribution of the random shock preferences ε_j and ε_k embedded in $u(q_j; \theta_i)$ and $u(q_k; \theta_i)$. For example, by substituting the consumer's utility in equation (2.2.1), I obtain Firm H's demand:

$$D_H(p_H, p_L) = \hat{\alpha} \, \mathbb{P}\{q_H + \beta(q_H - \hat{\chi}) - p_H + \varepsilon_H > q_L - \lambda\beta(\hat{\chi} - q_L) - p_L + \varepsilon_L\}$$

$$+ (1 - \hat{\alpha}) \, \mathbb{P}\{\theta_l q_H + \beta(q_H - \hat{\chi}) - p_H + \varepsilon_H > \theta_l q_L - \lambda\beta(\hat{\chi} - q_L) - p_L + \varepsilon_L\},$$

which simplifies to $\frac{1}{2\sigma} \left((\hat{\alpha} + (1 - \hat{\alpha})\theta_l)(q_H - q_L) + \beta(q_H - \hat{\chi}) + \lambda\beta(\hat{\chi} - q_L) + p_L - p_H + \sigma \right)$. Solving for the best-response prices that maximize $\pi_j(p_j, p_k) = D_j(p_j, p_k) \, p_j$ and computing the fixed point yields the subgame equilibrium prices and profits

$$p_H^* = \sigma + \frac{1}{3} \left((\hat{\alpha} + (1 - \hat{\alpha})\theta_l) \, (q_H - q_L) + \beta(q_H - \hat{\chi}) + \lambda\beta(\hat{\chi} - q_L) \right), \tag{2.4.1}$$

$$p_L^* = \sigma - \frac{1}{3} \left((\hat{\alpha} + (1 - \hat{\alpha})\theta_l) \, (q_H - q_L) + \beta(q_H - \hat{\chi}) + \lambda\beta(\hat{\chi} - q_L) \right), \tag{2.4.2}$$

and

$$\pi_H^* = \frac{1}{18\sigma} \left(3\sigma + (\hat{\alpha} + (1 - \hat{\alpha})\theta_l) \, (q_H - q_L) + \beta(q_H - \hat{\chi}) + \lambda\beta(\hat{\chi} - q_L) \right)^2, \tag{2.4.3}$$

$$\pi_L^* = \frac{1}{18\sigma} \left(3\sigma - (\hat{\alpha} + (1 - \hat{\alpha})\theta_l) \, (q_H - q_L) - \beta(q_H - \hat{\chi}) - \lambda\beta(\hat{\chi} - q_L) \right)^2. \tag{2.4.4}$$

The profit expressions (2.4.3) and (2.4.4) make intuitive sense. Firm H's profit increases with the consumer's likelihood of having higher taste for quality increases (higher $\hat{\alpha}$), because more high-valuation consumers will be drawn to the high-quality product, which in turn allows Firm H to raise its price. Firm L's profit decreases with $\hat{\alpha}$ for the opposite reason. Moreover, Firm L's profit

decreases with the reference quality, $\hat{\chi}$. This intuition is essentially the same as the monopoly case: higher reference quality reduces consumer's WTP for q_L as consumers evaluate quality against a higher standard.

However, the comparative statics for Firm H's profit with respect to the reference quality is less intuitive and different from the monopoly case. Specifically, Firm H's profit *increases* with the reference quality, even though consumer's utility from q_H decreases with the reference quality. The intuition rests on the concavity of the utility function induced by consumer's loss aversion. Loss aversion implies that utility is more sensitive in the loss domain. Therefore, while a higher reference point indeed reduces utility for both quality levels, the utility for the low quality product declines disproportionately more steeply than does the utility for the high quality product. This asymmetric devaluation is visualized in Figure 2.2 where $|\Delta u(q_L)| > |\Delta u(q_H)|$. As a consequence, consumers perceive the low and high quality products to be more differentiated. Enlarged perceived differentiation, in turn, allows Firm H to charge a higher premium above Firm L's price:

$$\frac{\partial}{\partial \chi} \left(p_H^* - p_L^* \right) = \frac{2}{3}\beta(\lambda - 1) > 0, \tag{2.4.5}$$

where the positivity follows from $\lambda > 1$. Hence, Firm H's profit increases with the reference quality. I call this the *premium effect*. The following proposition summarizes the finding.

Proposition 2 (Premium Effect). *The higher-quality firm's profit increases with the reference quality. Furthermore, this premium effect becomes more pronounced as consumers: (i) care more about comparative (vs. absolute) utility (i.e., high β), and (ii) exhibit stronger loss aversion, or have more concave utility (i.e., high λ).*

As I will demonstrate in the analysis of the advertising game, the premium effect described in Proposition 2 will form the basis for Firm H's incentive to increase the reference quality via reference-shifting advertising.[7] I note that the premium effect is driven by diminishing marginal utility from quality. In this formulation, it is driven by loss aversion because loss aversion causes

---

[7]Observe from (2.4.5) that the premium effect disappears if either consumers do not care about relative valuation (i.e., $\beta = 0$), or consumers do not exhibit loss aversion (i.e., $\lambda \leq 1$).

the utility function to show diminishing marginal utility from quality. I find the premium effect, and therefore the subsequent effects that it causes, in other formulations with utility that is concave in quality (for instance, in the formulation in Section A.3 in the appendix).

Advertising Subgame

After solving for the pricing subgame equilibrium, I proceed to examine the firms' equilibrium advertising strategies in Stage 1. Following the discussions above, I can rule out certain dominated strategies. Since Firm H's profit increases in $\hat{\alpha}$, Firm H will never choose advertising to diminish the consumer's taste for quality. Similarly, Firm L will never choose a reference-shifting ad that increases the reference quality, as doing so would only reduce its profit. In sum, Firm H will choose advertising strategy from the set $\{V^{\uparrow}, R^{\uparrow}, \emptyset\}$, whereas Firm L will choose from $\{V^{\downarrow}, R^{\downarrow}, \emptyset\}$. In other words, if Firm H (Firm L) chooses valuation-shifting advertising it only chooses $V^{\uparrow}$ ($V^{\downarrow}$), and if Firm H (Firm L) chooses reference-shifting advertising it only chooses $R^{\uparrow}$ ($R^{\downarrow}$). The tuple (a_H, a_L) denotes the ad strategies of the firms. The following proposition summarizes the equilibrium advertising strategies.

Proposition 3 (Advertising Equilibrium). *Let the thresholds λ_1, λ_2, λ_3, λ_4 and λ_5 and q_1, q_2 and q_3 be as defined in the proof.*

1. *If* $\max\left[\lambda_1, 1 + \frac{(q_H - q_L)v(1 - \theta_l)}{\beta\delta}\right] < \lambda$ *and* $q_1 < q_L$, *then both firms choose reference-shifting advertising, i.e., $(a_H, a_L) = (R^{\uparrow}, R^{\downarrow})$;*

2. *If* $\max\left[\lambda_2, 1 + \frac{(q_H - q_L)v(1 - \theta_l)}{\beta\delta}\right] < \lambda < \lambda_3$ *and* $q_L < q_1$, *then Firm H chooses reference-shifting advertising while Firm L does not advertise, i.e., $(a_H, a_L) = (R^{\uparrow}, \emptyset)$;*

3. *If* $\lambda_4 < \lambda < 1 + \frac{(q_H - q_L)v(1 - \theta_l)}{\beta\delta}$ *and* $q_L < q_2$, *then both firms choose valuation-shifting advertising, i.e., $(a_H, a_L) = (V^{\uparrow}, V^{\downarrow})$;*

4. *If* $\lambda < \min\left[\lambda_5, 1 + \frac{(q_H - q_L)v(1 - \theta_l)}{\beta\delta}\right]$ *and* $q_2 < q_L < q_3$, *then Firm H chooses valuation-shifting advertising while Firm L does not advertise, i.e., $(a_H, a_L) = (V^{\uparrow}, \emptyset)$; and*

19
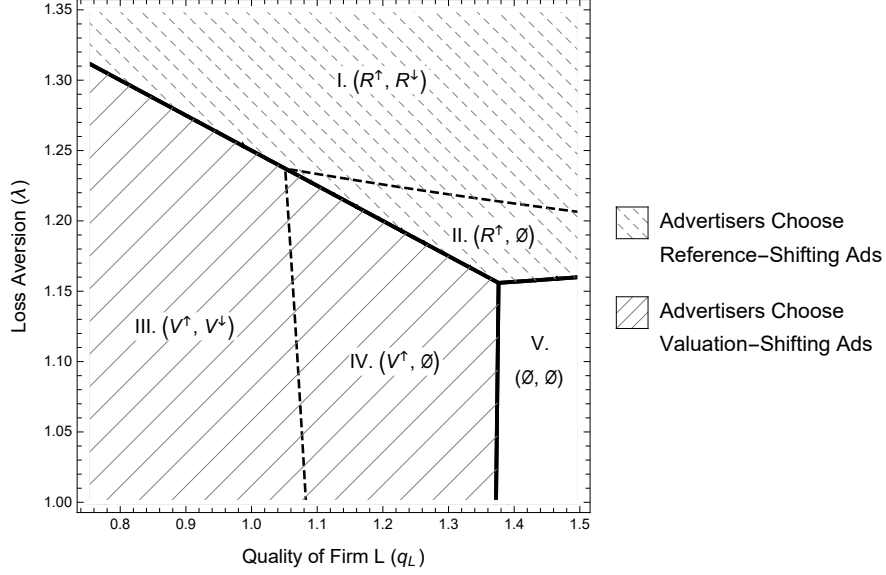
Figure 2.3: Equilibrium Advertising Strategies (tuples denote equilibrium strategies (a_H^*, a_L^*)); $q_H = 2, \alpha = \theta_l = 0.5, \nu = \delta = 0.25, \sigma = 4, \beta = 2, k = 0.03$

5. *Otherwise, neither firm advertises, i.e., $(a_H, a_L) = (\emptyset, \emptyset)$.*

Proposition 3, illustrated in Figure 2.3, sheds light on a number of important insights regarding firms' advertising strategies. Note that in the area above the solid line in Figure 2.3 only reference-shifting advertising is done, either by both firms or by only Firm H. In the area below the solid line, only valuation-shifting advertising is done, either by both firms or by only Firm H. In the bottom right corner, neither firm advertises. In the following, I discuss the different regions in the plot and the insights obtained from these results.

When consumers are highly loss averse (large λ), firms choose reference-shifting advertising (Regions I and II in Figure 2.3). Intuitively, the more loss averse the consumers are, the more concave their utility function, and this strengthens the premium effect (see Proposition 2). Consequently, Firm H chooses reference-shifting advertising to raise the reference point, thereby enhancing the perceived differentiation between the products. Firm L also chooses reference-shifting advertising, but one that lowers the reference point, in order to undo the adverse perception changes that lead consumers to dislike the low-quality product disproportionately more.

Given a large value of λ, Firm L's incentive to adopt such "defensive" strategy is especially

strong when q_L is high (Region I). To see this, suppose q_L is close to q_H. Consumers perceive the *absolute* valuations of q_L and q_H to be practically the same. Thus, consumers will base their purchase decision more heavily on the *comparative* component. As discussed earlier, however, concavity of the utility function implies that an increase in reference quality is disproportionately more detrimental to Firm L's comparative valuation. Therefore, when q_L is high, Firm L has strong incentives to choose reference-shifting advertising to offset Firm H's reference-shifting effect.

On the other hand, when λ is intermediate but q_L is high (Region II), Firm H's reference-shifting ad does not reduce Firm L's profit as much. The reason is that even though a higher reference quality disadvantages Firm L along the comparative dimension, consumers with low valuation for quality will still opt for the cheaper, low-quality product. In this case, offsetting Firm H's reference-shifting advertising is not worth the advertising cost and thus, Firm L foregoes advertising.

When both λ is small, firms opt for valuation-shifting advertising (Regions III, IV, and V). Moreover, both firms advertise at low q_L (Region III), only Firm H advertises at intermediate q_L (Region IV), and neither firm advertises at high q_L (Region V). Why do firms shift from reference-shifting to valuation-shifting advertising when λ is small? Again, the mechanism pertains to the curvature of the utility function. Note that a small λ corresponds to an increasingly linear utility for quality. This reduced curvature dampens the premium effect for Firm H and thus diminishes its returns from reference-shifting advertising. Thus, Firm H turns to valuation-shifting advertising instead, which has relatively higher returns. This in turn motivates Firm L to choose valuation-shifting advertising to prevent consumers from being persuaded by Firm H's ad to value quality more, an outcome that would cast its low quality product in an unfavorable light. In total, in Region III, where λ is small (i.e., the premium effect will be dampended) and q_L is low (i.e., product differentiation is high), both firms seek to pull consumer's preference toward their own product using valuation-shifting ads. This can be viewed as a vertical analogue of "combative advertising" in Chen et al. (2009) wherein firms shift consumer's horizontal ideal points to their favor.

Next, consider the regions when λ is small (i.e., the premium effect will be dampened) and q_L is intermediate or high (i.e., product differentiation is not high) (Regions IV and V). In this case, the returns on valuation-shifting advertising are small for both firms. Intuitively, lower product differentiation invites a stronger competitive response. Therefore, as q_L increases, firms withdraw from valuation-shifting advertising (nor do they choose reference-shifting advertising because consumer utility is not sufficiently concave, as λ is small, and therefore the premium effect is small). Note that as q_L increases, Firm L withdraws from advertising first while Firm H continues to advertise (Region IV). This is because Firm H's higher margin implies a higher return from valuation-shifting advertising than Firm L. For large enough q_L, neither firm advertises (Region V).

Interestingly, Proposition 3 shows that Firm L's advertising strategy may be non-monotonic in its product quality. Specifically, Firm L may advertise at low and high q_L, but not at intermediate q_L (e.g., for $\lambda \approx 1.22$ in Figure 2.3). To understand this, observe that the returns from valuation-shifting advertising decrease with q_L as the absolute valuations of q_H and q_L become similar. Conversely, the returns from reference-shifting advertising increase with q_L. Intuitively, for small quality differentials, comparative valuation with respect to the reference point becomes increasingly important. Thus, Firm L has much to gain from decreasing the reference quality, thereby minimizing consumer's perception of loss from consuming q_L. Taken together, at low (high) q_L, returns from valuation- (reference-) shifting advertising are high, whereas at intermediate ranges of q_L, returns from both advertising types are unprofitably low. Firm L, therefore, foregoes advertising altogether at intermediate q_L.

Firm L's withdrawal from advertising at intermediate ranges of q_L has interesting implications for Firm H's profit. I state the result in the following proposition.

Proposition 4. *If λ is intermediate and q_L low, then Firm H's profit may increase in q_L such that both firms are better off under reduced product differentiation.*

Standard economic theory asserts that as competing firms become less differentiated, price competition intensifies, and as a result, profits fall. In contrast, Proposition 4 suggests that profits of both firms may rise with *less* differentiation, in particular, in the absence of demand expansion

22
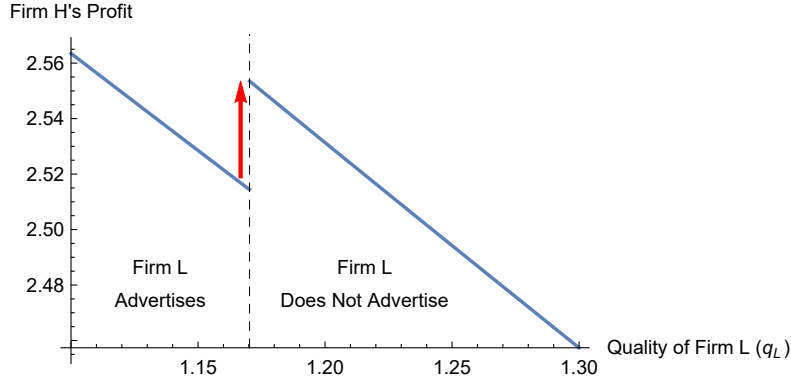
Figure 2.4: Firm H's Equilibrium Profit; $q_H = 2, \lambda = 1.2, \alpha = \theta_l = 0.5, \nu = \delta = 0.25, \sigma = 4, \beta = 2, k = 0.03$

effects. The intuition behind this seemingly counter-intuitive result rests on Firm L's withdrawal from advertising for high q_L. As illustrated in Figure 2.3, for low λ (e.g., $\lambda \approx 1.2$ in the figure), Firm L chooses valuation-shifting advertising for low q_L and then foregoes advertising for high q_L as returns fall. Such withdrawal increases the returns from Firm H's valuation-shifting advertising, which in turn results in a discrete, positive jump in Firm H's profit (see Figure 2.4). In sum, Firm H benefits from an increase in Firm L's quality level due to a favorable shift in its competitor's advertising regime.

Finally, it is interesting to note that both firms may be worse off in the equilibrium wherein both advertise (Regions I and III); they would have been better off had they cooperatively decided not to advertise. This prisoner's dilemma situation arises because even though advertising induces consumers to evaluate the advertising firm's product more favorably, when both firms advertise, their advertisements effects offset one another. Overall, firms end up spending advertising budget for zero net effect. Thus, if products are sufficiently differentiated or consumers exhibit sufficiently high loss aversion, then persuasive advertising could be excessive from a social welfare perspective. Extending on the welfare analysis, I investigate how consumer surplus is affected by the firms' equilibrium behavior.

Consumer Surplus

In my framework, firms' advertisements have a direct influence on consumers' utility for quality; namely, advertisements shift the consumers' taste and reference point for quality. Thus, I expect
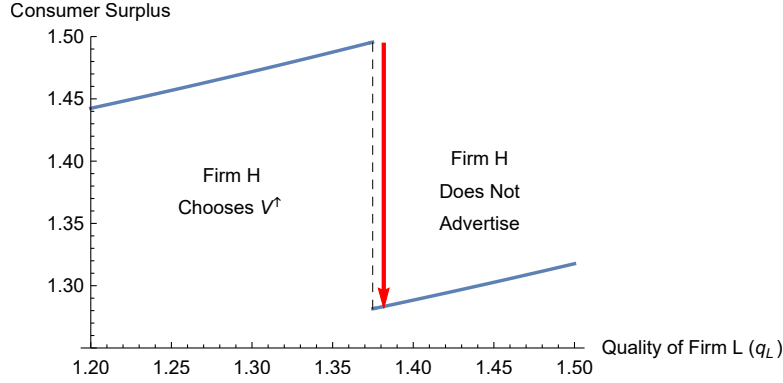
Figure 2.5: Equilibrium Consumer Surplus; $q_H = 2, \lambda = 1.1, \alpha = \theta_l = 0.5, v = \delta = 0.25, \sigma = 4, \beta = 2, k = 0.03, V = 4$

advertisements to have implications for consumer surplus. As discussed in Dixit and Norman (1978), in persuasive advertising the utility function of the consumer changes because of the ads. Therefore, deriving implications for consumer surplus involves making a choice between using pre-ad and post-ad utility functions for calculating post-ad consumer surplus. While there is no clear consensus on this point, I follow Kotowitz and Mathewson (1979) and the arguments in Bagwell (2007) and use post-ad utility functions for calculating post-ad consumer surplus.

Standard economic theory suggests that lower product differentiation implies more intense price competition and, therefore, lower prices and higher consumer surplus. In the context of competitive advertising that I describe, however, I find that this is not necessarily the case. While the standard reasoning rests on the premise that consumer valuation of quality remains constant, in my context, firms alter consumer valuation of quality through ads, which may affect consumer surplus.

Consider an increase in the quality of the low-quality product from an intermediate to high level. As discussed in Proposition 3, for a large enough λ, a reduction in the quality differential motivates firms to shift from valuation-shifting to reference-shifting ads as firms anticipate comparative utilities to factor more importantly in consumer's purchase decision. In particular, Firm H stops using the valuation-shifting ad that induces consumers to derive higher marginal utility from quality and switches to reference-shifting ad that decreases perceived utility from both products. This switch in Firm H's advertising regime results in lower consumer surplus as consumers expe-

rience less utility from the fixed quality levels offered in the market (see Figure 2.5). I summarize this result in the proposition below.

Proposition 5. *The expected consumer surplus may decrease as the average quality in the market increases.*

I conclude the discussion of the main model by highlighting how my equilibrium analysis may inform advertising strategies in practice. At a high level, my findings underscore the importance of adapting advertising content judiciously based on consumer characteristics and to settings of varying product differentiation. Strategic nuances that arise from competition may make one type of advertising content more suitable for certain consumer segments or markets than another. For example, when facing intense product market competition, and when consumers have diminishing marginal utility from quality, firms should seek to shift consumer's reference point by focusing on content that emphasizes a certain quality level.[8]. Furthermore, firms offering different products have the incentive to move the reference point lower or higher. However, in a market where products are sufficiently differentiated along the vertical dimension, a potentially effective advertising strategy for lower end manufacturers is to display content that underscores the importance the quality/price trade-off, thereby persuading consumers to care less about absolute quality.

2.5 Extensions and Robustness

An assumption underlying the main model is that a firm can only choose one type of advertising: either reference-shifting or valuation-shifting. In reality, however, both effects may be operative in a single ad. For example, in Figure 2.1c, Verizon's ad, which highlights its expansive network coverage, may not only shift the consumers' reference quality for network coverage, but also enhance their marginal valuation for network coverage. Therefore, it is important to test the robustness of my main insights when advertisement content is not as clear cut as posited in the

---

[8]A number of studies, such as Klapper et al. (2005) and Nicolau (2012) identify correlation between certain demographic variables such as employment status, age, and household size and the degree of loss aversion, i.e., such observables can be used to target specific segments ith different levels of concavity in utility.

main model. In the first extension, I study this scenario. In the second extension, I consider the possibility of firms running more than one advertising campaign with different campaigns having different ad content. By allowing a firm to allocate a fixed advertising budget across the two types of content, I examine how the results of my main model change when firms can endogenously choose the mixture of advertisement content for a given advertising budget. Finally, I show the robustness of my insights to an alternative model specification.

2.5.1 Multiple and Exogenous Ad Effects

In this section, I assume that a single ad has both reference-shifting and valuation-shifting effects. I use the exogenous parameter $w \in [0, 1]$ to determine the relative strengths of each effect in a particular ad. Specifically, if Firm H (Firm L) advertises, then the reference quality increases (decreases) by $w\delta$, and the likelihood of a consumer having high-valuation increases (decreases) by $(1-w)\nu$. In other words, a large w corresponds to ads that predominantly influence the consumers' reference quality, whereas a small w corresponds to those that predominantly shift valuations for quality. Given these effects, firms make a binary decision whether to advertise or not, and then set prices. The rest of the model specifications remain the same as the main model.

The following proposition characterizes the equilibrium advertising strategies when advertisements can have joint effects.

Proposition 6. *Let $\tilde{\lambda}_1$, $\tilde{\lambda}_2$, $\tilde{\lambda}_3$ and $\tilde{\lambda}_4$ be as defined in the proof.*

1. *If $\max[\tilde{\lambda}_1, \tilde{\lambda}_2] < \lambda < \tilde{\lambda}_3$, then both firms advertise;*

2. *if $\max[\tilde{\lambda}_3, \tilde{\lambda}_4] < \lambda$ or $\tilde{\lambda}_4 < \lambda < \tilde{\lambda}_2$, then only Firm H advertises; and*

3. *otherwise, neither firm advertises.*

Figure 2.6 plots the equilibrium outcomes characterized in Proposition 6. When products are sufficiently differentiated (Figure 2.6a), firms advertise when w is low; i.e., valuation-shifting effect dominates. This reflects the insight from the main model wherein firms choose valuation-shifting ad under high product differentiation in order to adjust consumer's preferences in their

26
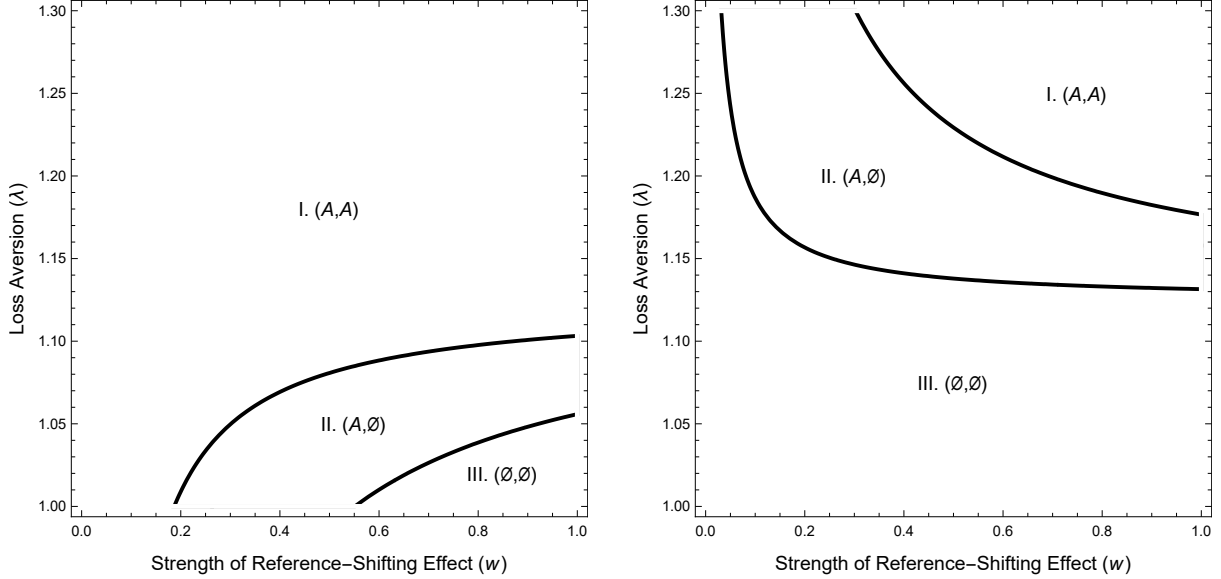
(a) High Product Differentiation; $q_L = 0.75$ (b) Low Product Differentiation; $q_L = 1.4$

Figure 2.6: Equilibrium Region Plot with Exogenous, Multiple Ad Effects; $q_H = 2, \alpha = \theta_l = 0.5, v = \delta = 0.25, \sigma = 4, \beta = 2, k = 0.03$

favor. In addition, when λ and w are both high such that consumer's utility is sufficiently concave and advertising effect is predominantly reference-shifting, then both firms advertise. Intuitively, Firm H advertises to capitalize on the premium effect, while Firm L advertises to offset this effect of being perceived as being disproportionately worse than the high quality product.

As illustrated by the marked reduction in the advertising region going from Figure 2.6a to Figure 2.6b, when product differentiation is low, firms advertise less. This is because price competition lowers the returns from advertising. Nevertheless, when consumer's utility is sufficiently concave (e.g., due to high loss aversion, λ) and reference-shifting effect carries a large weight in ads then, consistent with the results from the main model, both firms advertise to shift the consumer's comparative utility component in their favor.

2.5.2 Multiple and Endogenous Ad Effects

In this section, I allow firms to endogenously choose how much of each effect (valuation- and reference-shifting) it wants to create through its ads. To operationalize this, I allow each firm to allocate a fixed ad budget among the two forms of ads. Suppose firms can invest μ_v proportion of

their budget in valuation-shifting ads and μ_r proportion in reference-shifting ads, where $\mu_v, \mu_r \in [0, 1]$ and $\mu_v + \mu_r \leq 1$. I assume that the effect of investing a certain proportion of the budget to an ad type is linearly related to the associated ad effect discussed in the previous model. That is, an investment of μ_v in valuation-shifting advertising changes the consumer's likelihood of having high valuation for quality by a magnitude of $\mu_v v$, and an investment of μ_r in reference-shifting advertising changes the reference quality by $\mu_r \delta$. I assume a convex marginal cost of ads, $c(\mu) = k\mu^2$.[9]

I solve for the SPNE numerically. Figure 2.7 illustrates the equilibrium ad allocation levels for both firms as a function of loss aversion (top panel) and the low quality level (bottom panel). I see that, qualitatively, the equilibrium outcomes at extreme levels of λ coincide with that of the main model. When λ is small, both firms allocate all of their budget to valuation-shifting advertising. And when λ is large, both firms invest heavily in reference-shifting advertising, mirroring the outcome from the main model. Therefore, broadly speaking, the underlying intuitions from the main model carry over.

Nevertheless, an interesting feature of the equilibrium outcome in this extension model is the partial budget allocations across different types of ads, which occur at intermediate ranges of λ and q_L. Mixing advertising content allows firms to capitalize on the two distinct effects associated with valuation- and reference-shifting advertising. Consider the top panel of Figure 2.7. When λ is in the neighborhood of 1.2, firms distribute their ad budget partially across both types of advertising. In this case, Firm H strikes an optimal balance between the premium effect generated by the reference-shifting content, and the valuation-shifting effect which draws consumers toward higher quality products.

Similarly, the bottom panel of Figure 2.7, which plots the optimal ad weights with respect to q_L, echoes the insights from the main model. As q_L becomes increasingly close to q_H, firms shift their ad budget away from valuation-shifting ads and toward reference-shifting ads. This pattern

---

[9]This specification is consistent with the cost function of the main model with μ's constrained to $\{0, 1\}$. Another consistent specification is a linear cost function such that $c(\mu) = k\mu$. In the linear cost model, however, firms' profits become convex in the weight variables and the problem simplifies to the main model.

(a) Firm H's Ad Strategy: Impact of λ

(b) Firm L's Ad Strategy: Impact of λ

(c) Firm H's Ad Strategy: Impact of q_L

(d) Firm L's Ad Strategy: Impact of q_L
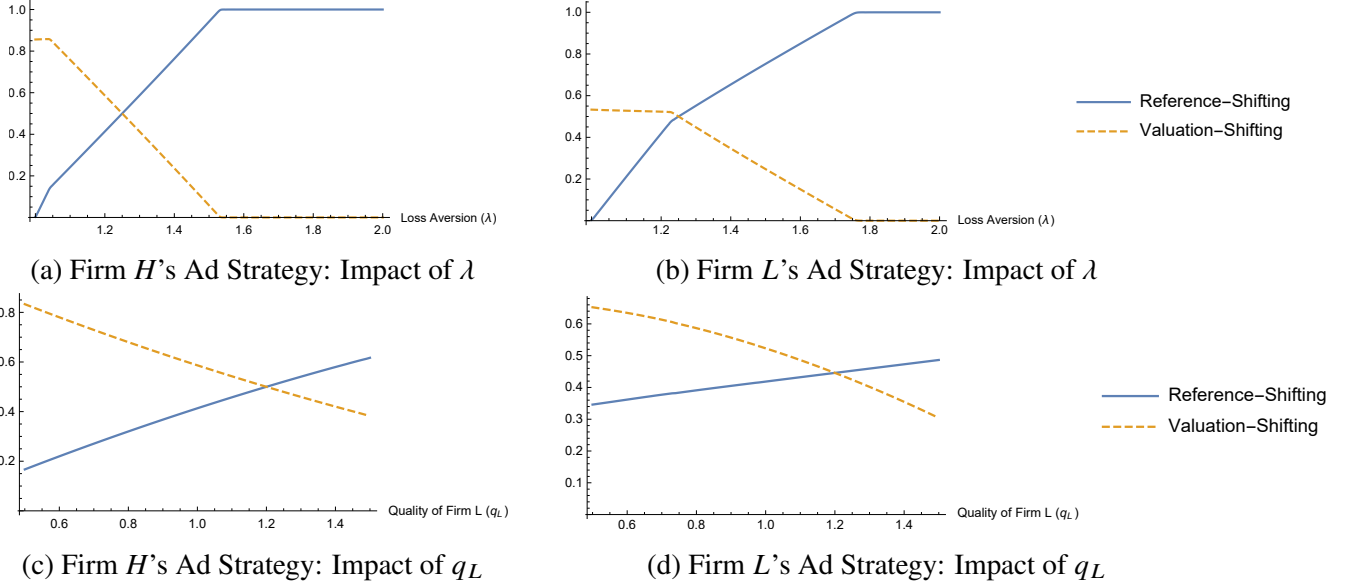
Figure 2.7: Equilibrium Budget Allocation; $q_H = 2, q_L = 1, \lambda = 1.2, \alpha = \theta_l = 0.5, \delta = \nu = 0.25, \sigma = 4, \beta = 2, k = 0.03$

is consistent with the previous finding that as q_L increases, consumers care more about the comparative valuation component, and thus, returns on reference-shifting advertising increase. In sum, while the flexibility afforded by the opportunity to allocate budgets across ad types "smoothens" the discrete advertising outcome from the main model, the main underlying forces and hence qualitative insights remain the same.

2.5.3   Alternative Utility Model

As I have stressed before, a main characteristic of the utility function that drives my results is that there are diminishing marginal returns to quality. In the main model, I have implemented this concavity in consumption utility using a linear loss-aversion model, which also provided a natural way to incorporate the quality reference point. In the appendix, I show the robustness of my main insights by considering an alternative specification of the consumer utility. Specifically, I model the concavity of consumer's utility for quality by relaxing the behavioral assumption that consumers exhibit loss aversion and simply considering a general concave utility function.

Furthermore, another way in which the alternative utility function in the appendix is different is

that I assume it to be entirely deterministic. Note that, in the main model, I had included the random component of utility to make the demand function smooth and simplify the pricing analysis. In the appendix, I show that removing this simplification leads to the same insights, although the pricing analysis is more complicated, essentially because the pricing equilibrium is in mixed strategies and not in pure strategies.

While the details of the analysis are provided in the appendix, I highlight here that my main insights are robust. Specifically, I find that a monopolist offering two products never chooses reference-shifting advertising that increases the quality reference point, as doing so only dampens consumer's WTP. In a duopoly, the incentive of the low-quality firm is similar — it never chooses to increase the reference quality. On the other hand, the high-quality firm may find it profitable to choose reference-shifting advertising that increases the reference point. The underlying intuition is analogous to the main model — due to the concavity of the utility function, an increase in the quality reference point induces a steeper decline in consumer's valuation for the low-quality product than it does for the high-quality product. This increases perceived quality differentiation between the products which, in turn, generates the premium effect wherein the high-quality firm charges a higher premium over its competitor's price.

2.6 Conclusions and Discussion

I study a scenario in which firms offering products of different qualities can use persuasive advertising to influence consumers' perceptions, either by influencing their valuations for quality, or by influencing their reference point for quality. I consider two qualitatively distinct forms of advertising content that has different effects: (i) valuation-shifting effect, that changes the quality valuations of consumers (by inducing a quality focus or a value focus in consumers' minds), and (ii) reference-shifting effect, that changes the reference point with respect to which consumers evaluate quality (by influencing the quality anchor against which consumers evaluate quality). I obtain a number of interesting insights regarding firms' use of these two types of content in their ads.

30

I find that a monopolist, if it advertises, always chooses either valuation-shifting advertising to increase consumer's marginal valuation of quality, or reference-shifting advertising to reduce the reference quality. Importantly, the monopolist never chooses reference-shifting advertising to raise the reference quality, because doing so only reduces the consumer's utility from the product, which in turn reduces the monopolist's profit.

Interestingly, this intuition does not carry over when competition is introduced. My analysis of a duopoly situation reveals that a high-quality firm may choose reference-shifting advertising to increase the consumers' reference quality — even though doing so lowers the utility from the product, due to the property of diminishing marginal utility of quality, it enhances the high-quality product's valuation relative to the low-quality product. This reference-shift-induced differentiation is operative only when consumers have diminishing returns to quality (i.e., the utility function is concave), and are more pronounced the more concave the consumer's utility for quality is (e.g., when consumers exhibit strong loss aversion). I find that when product differentiation is high between the two firms, the low-quality firm chooses valuation-shifting advertising to adjust consumer's quality taste towards its low-quality product. When product differentiation is low, it switches to reference-shifting advertising. Intuitively, as products become less differentiated, consumers rely more heavily on comparative utilities when making their purchase decisions. Therefore, the low-quality firm seeks to lower the reference point via reference-shifting ads, thereby alleviating the perception of loss consumers feel toward the low-quality product. Another counterintuitive insight that I find is that smaller differentiation among firms' qualities, even though it intensifies competition, may lead to lower consumer surplus.

My work offers a number of testable hypotheses regarding firms' choices of advertising content and strategy in quality differentiated markets. First, as consumers exhibit stronger loss aversion, or have steeply diminishing marginal utility of quality (i.e., λ increases), firms uses less valuation-shifting content and more reference-shifting content in its advertising. Second, if the product differentiation between competing firms is high (i.e., q_L is low), then firms change their advertising content from reference-shifting to valuation-shifting content. Third, in markets characterized

31

by intermediate levels of product differentiation and consumer utility concavity, firms prefer to include both valuation-shifting and reference-shifting content in their advertising. To test these hypotheses, one could compare advertising in markets corresponding to different values of the model parameters. For instance, to test the first hypothesis, one could compare advertising strategies in markets corresponding to small and large λ, respectively.

My work is a first step towards understanding persuasive advertising in vertically differentiated markets, and presents many opportunities for further research. For example, it would be interesting to investigate what happens when the reference-updating process is allowed to be dynamic. While my model assumes that reference quality is determined primarily by ad exposure, the literature on reference points suggests that prior consumption experiences may play an important role in reference point formation (Hardie et al., 1993; Kopalle et al., 1996). Therefore, future extensions could explore how the interaction of ad content and consumption experience in the reference updating process affects firms' advertising decisions. Another possibility is to analyze how quality decisions are affected by ad decisions that I study. While my model abstracts from firms' quality choices by imposing them to be exogenously endowed, future work may consider a quality decision stage prior to the stage where firms choose their ad types.

Chapter 3: Learning in Online Advertising

3.1 Introduction

Online advertising, with an annual spending of over $100B, has become the largest category of advertising in the US.[1] Online advertising inventory is sold using two pricing models: performance-based and impression-based. In performance-based (e.g., pay-per-click) pricing, an advertiser pays only if a consumer completes a pre-defined action (e.g., a click). In impression-based pricing, the advertiser pays for its ad being shown to a consumer, regardless of whether the impression leads to an action.

Understanding how an ad performs (e.g., how likely a consumer will take an action after viewing an ad) is crucial for publishers in performance-based pricing, and for advertisers in impression-based pricing. For example, in pay-per-click pricing, it is more profitable for a publisher to accept a payment of $1 per click for an ad with click-through rate (CTR) 10%, for an expected revenue $0.10 per impression, than a payment of $2 per click for an ad with CTR 4%, for an expected revenue $0.08 per impression. Similarly, the probability of action affects an advertiser's willingness to pay (WTP) for an impression in impression-based pricing. The advertiser is willing to pay more per impression if it knows that the impression leads to a desired outcome with a higher probability.

Previous literature on online advertising primarily assumes that the probability of the pre-defined action (e.g., CTR) is known to advertisers and publishers (Edelman et al., 2007; Katona and Sarvary, 2010; Jerath et al., 2011). In practice, however, advertisers and publishers have to *learn* this probability. For example, when a new advertiser joins the market, or when an existing advertiser revamps its ad campaign, the CTRs of its ads are typically unknown to the publisher, other advertisers, and the advertiser itself. They can at best have an expectation of the CTR based

---

[1] https://content-na1.emarketer.com/us-ad-spending

on a few observable characteristics of the advertiser.[2] 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, learning is asymmetric: *participating* in advertising auctions is not sufficient for the advertiser's CTR to be learned; the advertiser has to *win* advertising auctions sufficiently many times before the publisher and the advertisers can learn its CTR.

The learning dynamic can affect advertisers' and publishers' strategies in the market. In particular, winning in an advertising auction has two effects on an advertiser's payoff. First, the advertiser receives an immediate value from showing its ad to a consumer (the direct effect). Second, winning reveals information about the performance of the ad to both the advertiser and the publisher (the indirect effect); this improves the advertiser's and the publisher's estimate of the true CTR of the ad. In performance-based pricing, this ad-performance information is used by the publisher to determine the pricing and allocation of an ad slot, and in impression-based pricing, it is used by the advertiser to determine the advertiser's WTP. While the previous literature has primarily studied the direct effect of winning in an advertising auction, my paper focuses on the indirect effect.

These two effects give rise to interesting trade-offs for advertisers when a new advertiser joins the publisher. I illustrate these trade-offs in the following example.

Example. Suppose an advertiser, A, is the only advertiser bidding on an advertising slot of publisher P. Suppose that the slot is sold in a pay-per-click second price auction, A's bid is \$1 per click, and its CTR is 15%. Assume that B is a new advertiser who wants to advertise on the same slot. B's bid is also \$1 per click, but its CTR is not known to anybody at the time of entry. For the initial auctions, P assigns an average CTR estimate (e.g., based on the performance of advertisers with similar characteristics) of, say, 10%.

However, P can eventually learn the new advertiser's CTR after sufficient impression and click data for the new advertiser become available. Furthermore, B can facilitate this learning process

---

[2]For instance, in pay-per-click pricing, Google assigns an average Quality Score to new advertisers based on the performances of other advertisers using the same keyword. See `https://searchengineland.com/didnt-know-recent-quality-score-changes-259559`.

by bidding aggressively and thereby winning in the early rounds. Doing so allows P to observe more click data for B's ad which would in turn allow P to more accurately estimate B's true CTR.

Importantly, in pay-per-click pricing, P's estimate of B's CTR directly affects the payment and allocation of the advertisers. This is because publishers use *effective bids*, computed as advertisers' submitted bids multiplied by their expected CTRs, to calculate payment and allocation.[3] Given this, would B prefer to have its CTR learned by P quickly or not?

If B privately knew its true CTR, then the answer would be evident. For example, if it knew that its true CTR is 20% (i.e., higher than P's estimate), then B would unambiguously prefer P to quickly update its CTR from the 10% estimate to the true 20%. 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 × 15% and win the auction for a cost-per-click of $0.75; it would have lost the auction to A had its CTR remained at the average of 10% (see Table 3.1). Conversely, B's incentive to facilitate P's learning its CTR would diminish if B knew its true CTR is *lower* than P's prior estimate. In this case, B's long-term payoff would decrease if its low CTR is learned quickly. In sum, B prefers P to update B's CTR estimate more quickly (slowly) if it privately knows that its CTR is higher (lower) than P's prior estimate.

---

[3]In practice, effective bids can also include other factors such as landing page experience and advertiser's reputation; however, for the purpose of this example, I only consider the expected CTR and the submitted bid that are the two most important elements of effective bids.

| | Advertiser B's CTR | | |
	Not known (CTR=10%)	Known (CTR=20%)	Known (CTR=5%)
B	Has to bid (and pay) $1 \times 15\%/10\% = \$1.5$ to win	Wins at cost-per-click $1 \times 15\%/20\% = \$0.75$	Has to bid (and pay) $1 \times 15\%/5\% = \$3$ to win
A	Wins at cost-per-click $1 \times 10\%/15\% = \$0.66$	Has to bid (and pay) $1 \times 20\%/15\% = \$1.33$ to win	Wins at cost-per-click $1 \times 5\%/15\% = \$0.33$

Table 3.1: When the Publisher Knows vs. Does Not Know New Advertiser's CTR

In reality, however, when B first enters the market, 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 P's learning process if B wants to maximize its profit.

Similarly, for the existing advertiser A, P'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 slot; if it turns out to be lower, A can win the auction at a lower cost-per-click than when B's CTR is not known to P (from $0.66 to $0.33 in Table 3.1). Again, given that the existing advertiser A can facilitate (hinder) P'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, I study how the learning incentives affect the advertisers' and the publisher's strategies. I use a game-theoretic model to analyze advertisers' and publisher's strategies in a learning environment. To facilitate exposition, in the main body of the paper, I assume the publisher uses performance-based pricing, which currently accounts for 62% of the online advertising market in the US,[4] and use pay-per-click terminology. In the extensions, I show that my results apply to pay-per-impression pricing model as well. I are interested in answering the following research questions.

---

[4] http://totalaccess.emarketer.com/chart.aspx?r=219092

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

2. Does an existing advertiser (incumbent) benefit from the publisher learning the CTR of the entrant? How does this affect the incumbent's bidding strategy?

3. How does the lack of information about a new advertiser's CTR affect the publisher's revenue? How do learning incentives affect the publisher's optimal strategy?

In answering the first set of questions, I show that a new advertiser's expected payoff when its CTR is learned by the publisher 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 publisher, than in the long run after its CTR becomes known. This finding is in line with what industry experts commonly recommend new advertisers regarding 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.[5]

My result indicates that even for advertisers whose long-run equilibrium cost-per-click is low, the initial cost-per-click (at the time of joining the market) may be above their valuation. In other words, advertisers should be prepared to lose money in the beginning when they start advertising with a publisher for the first time. Moreover, they should not be discouraged from using that publisher even if the initial cost-per-clicks are higher than their WTP.

In answering the second set of questions, I 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 publisher. This is because an incumbent with a high CTR does not want to risk earning a low margin (or worse, losing its ad slot) in the event the entrant's CTR turns out to be high.

---

[5] https://searchengineland.com/4-ways-to-determine-your-your-starting-bids-144616.

This "preemptive" strategy, however, is too expensive for an incumbent with a low CTR. As I show, an incumbent with a low CTR lowers its bid when an entrant joins, so that the entrant's CTR is learned more quickly. Intuitively, competing with an advertiser whose CTR is unknown is too costly for the weak 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 set of questions, interestingly, I find that the publisher 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 publisher's revenue. Under certain conditions, however, the publisher's ignorance could also hurt its revenue. For instance, if the entrant's CTR is high, the publisher 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 publisher from learning the entrant's potentially high CTR.

I find that the publisher can mitigate the loss of not knowing the entrant's CTR by reducing the reserve price of the entrant. By reducing the reserve price, the publisher increases the probability of the entrant winning in the auction, thereby increasing the probability of learning the entrant's CTR. Furthermore, I characterize the optimal mechanism and show that, first, in the presence of learning considerations, a variation of the standard second-price auction with optimal reserve prices is sufficient to achieve the optimal revenue. Second, it is optimal for the publisher to favor the entrant in the beginning, before the entrant's CTR is learned. This manifests in a lower optimal reserve price of the entrant when the publisher does not know the entrant's CTR than when it knows.

In addition, I discuss alternative mechanisms that can help the publisher mitigate its loss of not knowing the entrant's CTR. For example, Google provides $75 ad credit to new advertisers when they spend $25 on AdWords.[6] Facebook also offers ad credit to new accounts that have a

---

[6] https://www.google.com/ads/adwords-coupon.html

sufficiently high audience engagement on their pages.[7] While these programs have traditionally been viewed as promotions to attract new advertisers, my research reveals new strategic incentives beyond new customer acquisition that motivate publishers to offer ad credit.

Theoretical Contribution. While, from a managerial point of view, my work sheds light on advertisers' and publishers' strategies regarding new entries, I also want to highlight two unique aspects of my model from a theoretical point of view. First, in the context of online advertising, I study the *transition* of a game from an incomplete information game to a full information one. While the previous literature on online advertising assumes that the game is either always full information (e.g., Edelman et al., 2007) or always incomplete information (e.g., Edelman and Schwarz, 2010), in practice, the level of information is constantly changing. My paper takes a first step towards bridging this gap by analyzing the transition.[8] I show that the advertisers' and the publishers' strategies regarding the transition are qualitatively distinct from those in full information and incomplete information games.

Second, my analysis demonstrates how some of the standard results from learning theory may be reversed when the subjects of learning are not as "passive" as commonly assumed in the literature (e.g., Gittins and Jones, 1979; Katehakis and Veinott, 1987). For instance, exploration-exploitation trade-off from standard learning theory suggests that knowing less about new advertisers would only hurt the publisher's revenue because the publisher must then learn about new advertisers through costly exploration. In contrast, my model shows that the publisher may be better off knowing less about the new advertiser due to the advertisers' strategic responses during the publisher'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. First, I discuss related literature. In Section 3.2, I present the model. I analyze the model and discuss advertisers' strategies in Section 3.3. The publisher's optimal strategy is discussed in Section 3.4. I explore extensions of the main model in

---

[7]http://www.digitalsitemap.com/free-facebook-ad-coupon/

[8]In fact, since new advertisers constantly join this market, and even existing advertisers frequently revamp 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.

Section 3.5 to establish the robustness of my main results, and conclude in Section 3.6.

Related Literature

My work contributes to the vast literature on display advertising. Empirical works in this area have assessed the effectiveness of display advertising in various contexts. Lambrecht and Tucker (2013) demonstrate that retargeting may not be effective when consumers have not adequately refined their product preferences. Hoban and Bucklin (2015) find that display advertising increases website visitations for a large segment of consumers along the purchase funnel, but not for those who had visited before. Bruce et al. (2017) examine the dynamic effects of display advertising and show that animated (vs. static) ads with price information are the most effective in terms of consumer engagement. On the theoretical front, Sayedi et al. (2018) study advertisers' bidding strategies when publishers allow advertisers to bid for exclusive placement on the website. Sayedi (2018) analyzes the interaction between selling ad slots through real-time bidding and selling through reservation contracts. 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. Berman (2016) explores the effects of advertisers' attribution models on their bidding behavior and their profits. Kuksov et al. (2017) study firms' incentives in hosting the display ads of their competitors on their websites.

Within online advertising, the increasing prevalence of search advertising has motivated a growing body of empirical (e.g., Rutz and Bucklin, 2011; Yao and Mela, 2011; Haruvy and Jap, 2018) and theoretical papers. Katona and Sarvary (2010) and Jerath et al. (2011) 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. 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. 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 my knowledge, my paper is the first on online 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 my 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 publisher by pooling with another type. Despite some similarities, my paper differs in that I do not model information asymmetry. Although I allow players take certain actions to facilitate or hinder the revelation of information, those actions do not signal their types. Furthermore, in signaling games, players mimic other players' strategies in order to hide or reveal information; in contrast, advertisers in my model interfere with the publisher's learning process in order to do so.

There are a few papers in Computer Science and Operations Research literature that address dynamic learning in repeated auctions. 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. 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.

Closest to my paper within this stream is Iyer et al. (2014), which studies bidding strategies of agents who learn their valuations. Under the assumption that the market size is infinitely large, Iyer et al. (2014) adopt a mean-field approximation to solve for equilibrium strategies. They report a similar finding that in a learning environment, an advertiser's bid consists of the present expected value of winning the ad slot and the "marginal future gain from one additional observation regarding [the advertiser's] valuation." The present paper, however, differs along several important dimensions.

First, since I use performance-based pricing, the learning agent in my model is the publisher, not the advertiser. The publisher receives new information about a new advertiser who wins, and incorporates the information to the rules of the subsequent auctions. Thus, a new advertiser bids strategically not to learn its own type *per se*, but to influence the publisher's learning process. Second, my paper sheds light on a novel incentive for existing advertisers to deter the publisher from learning the new advertiser's type. This is distinct from the idea of advertisers adopting (symmetric) bidding strategies to learn their own types. The discrepancies in the incentives across advertisers that are highlighted in my paper do not emerge in a mean-field equilibrium wherein all agents behave in a symmetric manner. Finally, my paper analyzes a small, stylized market with limited number of participants, which allows us to model fully rational behavior of all players. My assumption of a small market is motivated by the fact that, due to the fine-grained targeting available in online advertising, most auctions have a small number of participants; as such, advertisers' one-to-one interactions affect their optimal strategies. Papers that employ mean-field equilibrium (e.g., Iyer et al., 2014; Balseiro et al., 2015) abstract away from advertisers' one-to-one interactions, and characterize an approximate equilibrium wherein agents are assumed to be boundedly rational.

3.2 Model

My model consists of a publisher and two advertisers, the incumbent and the entrant, indexed by P, I and E, respectively. The publisher sells one ad slot in a second-price auction with reserve price R.[9] 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 publisher's webpage in display advertising, or consumer's search query in search advertising.

In my main model, I assume performance-based pricing, which currently accounts for 62% of the online advertising market in the US,[10] and use pay-per-click pricing terminology.[11] In Section 3.5.1, I show that, under some assumptions, my findings apply to impression-based pricing as well. I first assume that both advertisers have the same valuation per click, which I normalize to 1. This assumption is not necessary, but simplifies the discussion of advertisers' strategies in Section 3.3. I relax this assumption in Section 3.4 when analyzing the publisher's strategy. The incumbent (entrant) submits a bid b_{It} (b_{Et}), where t indexes the game stage. The bids indicate how much the advertisers are willing to pay per click.

In performance-based pricing, publishers take advertisers' expected performance into account when determining payment and allocation. In pay-per-click pricing, publishers compute advertisers' *effective bids* as the product of their submitted bids and the estimated CTRs of their ads.[12] Some publishers may also include other parameters such as landing page experience in the effective bids; however, to focus on the role of CTRs, I only take the submitted bids and the CTRs into account, and assume that the two advertisers are the same along other dimensions that a publisher

---

[9]I consider a multiple-slot Generalized Second-Price auction in Section 3.5.

[10]http://totalaccess.emarketer.com/chart.aspx?r=219092

[11]Pay-per-click pricing is the most common form of performance-based pricing; nonetheless, my results can be readily applied to other performance-based pricing models such as pay-per-conversion.

[12]For example, see https://www.facebook.com/business/help/430291176997542 and https://searchengineland.com/guide/ppc/how-the-ppc-ad-auction-works.

may consider.[13] Therefore, the effective bids 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 effective bid wins the auction, provided its effective bid is greater than or equal to the reserve price, R. 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 $\max[c_I b_{It}, R]/c_E$.

I assume that c_E is drawn from a differentiable cumulative distribution function (c.d.f.) F_E. Since the incumbent has been advertising with the publisher for an extended period of time, following Edelman et al. (2007) (and many other papers in the literature), I 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 has not advertised with the publisher in the past. When the entrant joins, the publisher, the incumbent, and the entrant only know the distribution of the entrant's CTR. [14] I assume that c_I and μ_E, the expected value of c_E, are greater than the reserve price, so that the incumbent and the entrant can beat the reserve price in expectation.

Before I proceed, I should elaborate on the meaning of the CTR parameters c_I and c_E. In my 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 the ad's click propensity. Indeed, publishers only take into account advertiser-specific CTRs, controlling for position effects, when computing advertisers' effective bids.[15] Position-specific CTRs will be incorporated in the multi-slot extension in Section 3.5.4.

Next, I describe the timing of the game, which is depicted in Figure 3.1.

Stage 1: The entrant joins the market. The entrant's CTR is initially unknown, and is therefore

---

[13] For a discussion of other parameters in advertisers' effective bids in search advertising see Katona and Zhu (2017).

[14] Note that this model implies two important assumptions on the information structure of the game. First, the assumption that c_I is common knowledge implies that the entrant and the publisher have the same level of information about the incumbent. Second, I are implicitly assuming that the incumbent and the publisher have the same level of information about the entrant. In practice, it is possible that large publishers such as Google and Facebook can estimate advertisers' CTRs more accurately than other advertisers based on their vast troves of data. I relax both of these assumptions in Section 3.5.2 and establish the robustness of my results.

[15] https://support.google.com/google-ads/answer/1659696

set to its expected value μ_E.[16] The incumbent and the entrant simultaneously submit their bids b_{I1} and b_{E1} to the publisher. The incumbent's effective bid is $c_I b_{I1}$ whereas the entrant's is $\mu_E b_{E1}$, since the publisher 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 publisher by the next stage; otherwise, it remains unknown.

To simplify the analysis, I assume that if the entrant wins a single auction (i.e., the auction in Stage 1), then the publisher learns its CTR. In practice, the entrant would have to win sufficiently many times for the publisher to accurately learn its CTR. Stage 1 in my model corresponds to as many auctions as the entrant needs to win for the publisher to learn its CTR. Furthermore, in practice, learning is continuous and gradual such that the publisher's estimate of the entrant's CTR improves incrementally every time the entrant wins. My model can be viewed as a discrete approximation of this learning process: the publisher 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 effective bid is $c_I b_{I2}$. The entrant's effective bid depends on the outcome of the Stage 1 auction. If the entrant had won in Stage 1, then its CTR becomes known to the publisher by Stage 2, and therefore, its effective bid is $c_E b_{E2}$. Otherwise, as in Stage 1, its CTR is not learned and its effective bid is $\mu_E b_{E2}$.[17]

I 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, δ affects how the advertisers trade off short-term revenue (in Stage 1) for long-term revenue (in Stage 2).[18]

---

[16]In Google AdWords, new advertisers received an average Quality Score of 6. See `https://searchengineland.com/minimum-quality-score-can-save-money-adwords-226757`. In Section 3.5.3, I consider an extension in which, instead of using μ_E, the publisher strategically sets the entrant's CTR.

[17]If the entrant wins the auction in Stage 1, the publisher learns c_E; however, I do not make any assumptions on whether the incumbent also learns c_E or not. Specifically, as I show in Lemma 1, the incumbent bids truthfully in Stage 2 regardless of the outcome of Stage 1.

[18]One might argue that the publisher 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, I 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 publisher. 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, δ could be interpreted as the length of time required for the publisher to learn the entrant's CTR if the entrant does not win in Stage 1 (compared to when it wins in Stage 1).
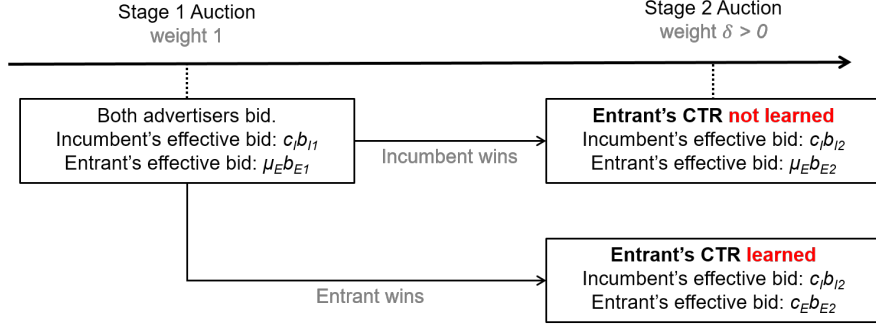
Figure 3.1: Game Timing

The incumbent's expected profit is the sum of its first and second stage payoffs. That is, $\mathbb{E}[\pi_I] = \pi_{I1} + \delta \mathbb{E}[\pi_{I2}]$ where π_{I1} denotes the incumbent's first stage payoff, and π_{I2} its second stage payoff contingent on the realization of c_E, over which expectation is taken. Specifically,

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

where $\tilde{c}_E$ is c_E if c_E is learned (i.e., entrant won in Stage 1 auction), and μ_E otherwise. Similarly, the entrant's expected profit is $\mathbb{E}[\pi_E] = \mathbb{E}[\pi_{E1}] + \delta \mathbb{E}[\pi_{E2}]$, where

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

Finally, the publisher's expected profit is $\mathbb{E}[\pi_P] = \mathbb{E}[\pi_{P1}] + \delta \mathbb{E}[\pi_{P2}]$, where

$$
\pi_{P1} = \begin{cases} \max[\mu_E b_{E1}, R] & \text{if } c_I b_{I1} \geq \max[\mu_E b_{E1}, R], \\ c_E \frac{\max[c_I b_{I1}, R]}{\mu_E} & \text{if } \mu_E b_{E1} > \max[c_I b_{I1}, R], \\ 0 & \text{otherwise}, \end{cases}
\qquad
\pi_{P2} = \begin{cases} \max[\tilde{c}_E b_{E2}, R] & \text{if } c_I b_{I2} \geq \max[\tilde{c}_E b_{E2}, R], \\ c_E \frac{\max[c_I b_{I2}, R]}{\tilde{c}_E} & \tilde{c}_E b_{E2} > \max[c_I b_{I2}, R], \\ 0 & \text{otherwise}. \end{cases}
$$

I use subgame perfect Nash equilibrium as the solution concept and solve by backward induction. Finally, to ensure the existence of a weakly dominant strategy for the incumbent, I assume that $c_I + \delta \left((c_I - \mu_E)^+ - \int_0^1 (c_I - \max[c_E, R])^+ \, dF_E\right) \geq R$, for which a sufficient condition is

$\delta \leq \frac{1}{f_E(R)+F_E(R)}.$[19] This assumption is only needed to facilitate the exposition in Section 3.3, and will be dropped in Section 3.4.

3.3 Advertisers' Strategies

In this section, I analyze the advertisers' bidding strategies and assume that the publisher's mechanism is exogenous. As a benchmark, in Section 3.3.1, I analyze the advertisers' strategies in a full information game. Then, in Section 3.3.2, I study how learning incentives in an incomplete information game affect the advertisers' bidding strategies.

3.3.1 Full Information Setting

As a benchmark, I first consider the case where the entrant's CTR is common knowledge. This corresponds to what most of the previous theoretical papers in online advertising literature assume. Even though the auction is not a standard second-price auction because advertisers' bids are multiplied by their CTRs, 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 7 (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 publisher, respectively, are $\pi_I^F = (1 + \delta)(c_I - \max[c_E, R])^+$, $\pi_E^F = (1 + \delta)(c_E - c_I)^+$, and $\pi_P^F = (1 + \delta) \max[\min[c_I, c_E], R]$, where $x^+ \equiv \max[x, 0]$.*

Proposition 7 shows that when the publisher 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, I show that truthful bidding is no longer an equilibrium strategy when the publisher does not know the entrant's CTR.

---

[19]This is not a restrictive assumption; for example, for $F_E(c) = c$, the condition holds for all $\delta > 0$ and $c_I \geq R$. The sufficient condition derives from the fact that $c_I + \delta \left((c_I - \mu_E)^+ - \int_0^1 (c_I - \max[c_E, R])^+ dF_E \right)$ is equal to R at $c_I = R$, and then imposing that the former increases in c_I.

3.3.2 Incomplete Information Setting

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

I begin my analysis under incomplete information with the second stage bids. I focus on dominant strategy equilibrium where advertisers play weakly dominant strategies. As I show in Lemma 1, Stage 2 auction is straightforward: advertisers bid truthfully. This is because in the last stage there are no strategic considerations of future payoffs; thus, the truthfulness property of standard second-price auctions holds.

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

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

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

$$b_{I1}^* = 1 + \frac{\delta}{c_I} \left((c_I - \mu_E)^+ - \int_0^{c_I} c_I - \max[c_E, R] \, dF_E \right), \tag{3.3.1}$$

$$b_{E1}^* = 1 + \frac{\delta}{\mu_E} \left(\int_{c_I}^1 c_E - c_I \, dF_E - (\mu_E - c_I)^+ \right). \tag{3.3.2}$$

In general, truthful bidding is a weakly dominant strategy in a second-price auction even under incomplete information. Expressions (3.3.1) and (3.3.2), however, show that the advertisers' bids are no longer truthful. What drives the change in advertisers' strategies in my model is the advertisers' incentive (or lack thereof) to help the publisher learn the entrant's CTR. The advertisers' Stage 1 bids are shaped by their preference to play a Stage 2 game in which the entrant's CTR is μ_E vs. c_E, where c_E is randomly drawn from F_E. For example, if the entrant's expected payoff in

Stage 2 is higher when its CTR is c_E (i.e., its CTR is learned), compared to when it is μ_E (i.e., its CTR is not learned), the entrant would raise its Stage 1 bid.

But does the entrant prefer its CTR to be learned by the publisher? I 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 $\mu_E \leq c_I$, a situation in which the entrant would have surely lost in Stage 2 if its CTR were unknown and set to μ_E by the publisher. Second, 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 μ_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 μ_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. Therefore, in expectation, the entrant prefers its CTR to be learned by the publisher.

The following table shows this more formally for the case when $\mu_E > c_I$:

Table 3.2: Entrant's Stage 2 Profit

Publisher does not know c_E	Publisher knows c_E
$\mathbb{E}[\pi_{E2}] = \int_0^1 c_E(1 - c_I/\mu_E)\, dF_E$	$\mathbb{E}[\pi_{E2}] = \int_0^1 c_E(1 - c_I/c_E)^+\, dF_E$
$= (1 - c_I/\mu_E)\int_0^1 c_E\, dF_E = (1 - c_I/\mu_E)\mu_E$	$= \int_{c_I}^1 c_E(1 - c_I/c_E)\, dF_E$
$= \mu_E - c_I = \int_0^1 c_E - c_I\, dF_E$	$= \int_{c_I}^1 c_E - c_I\, dF_E$

From Table 3.2, I see that the entrant's Stage 2 profit when the publisher 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 publisher 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 publisher learns its CTR. Therefore, the entrant bids aggressively in Stage 1 in order to facilitate the publisher's 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 publisher in Stage 2. If c_I is close to μ_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 μ_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. 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 in the event c_E turns out to be low. Thus, a weak incumbent has a strategic incentive to underbid.

On the other hand, if c_I is significantly greater than μ_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 μ_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). However, the reward of a low c_E realization is outweighed by the risk of a high c_E 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 incentive 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 3.3: Incumbent's Stage 2 Profit when $c_I > \mu_E$

Publisher does not know c_E	Publisher knows c_E
$\mathbb{E}[\pi_{I2}] = c_I(1 - \mu_E/c_I)$	$\mathbb{E}[\pi_{I2}] = \int_0^1 c_I(1 - \max[c_E, R]/c_I)^+ \, dF_E$
$= c_I - \mu_E$	$= \int_0^R c_I(1 - R/c_I) \, dF_E + \int_R^{c_I} c_I(1 - c_E/c_I) \, dF_E$
$= \int_0^1 c_I - c_E \, dF_E$	$= \int_0^R c_I - R \, dF_E + \int_R^{c_I} c_I - c_E \, dF_E$

From Table 3.3, I see that the incumbent's Stage 2 profit when the publisher 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 $c_E \in (0, R)$, I have $c_I - R$ integrated; since $R > c_E$, the incumbent is better off when the publisher 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 publisher 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. I summarize these results in the following proposition.

Proposition 8 (Advertisers' Strategies in Stage 1 Under Incomplete 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. See Figure 3.2.*

The advertisers' bidding behavior outlined in Proposition 8 can also be understood from an *asymmetric learning* perspective. Suppose that the publisher always learns the entrant's CTR in Stage 2, regardless of the Stage 1 outcome. In this hypothetical scenario, the advertisers' Stage 2
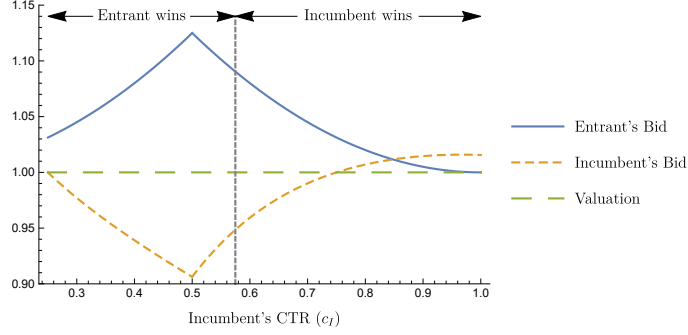
Figure 3.2: Stage 1 Equilibrium Bids; $R = \frac{1}{4}, \delta = \frac{1}{2}, F_E(c) = c$

payoffs would be independent of the Stage 1 outcome. As a result, neither the incumbent nor the entrant would have incentive to deviate from truthful bidding in Stage 1. In my model, however, the fact that the publisher's learning is asymmetric — that is, learning occurs if only if the entrant wins in Stage 1 — creates an important interdependence between the two sequential auctions. This interdependence, which is depicted in Figure 3.1, generates strategic incentives for advertisers to deviate from truthful bidding.

Publisher's Revenue

I turn to the implications of learning incentives on the publisher's revenue. Is the publisher unequivocally better off knowing the entrant's CTR? One may conjecture that being more informed about the bidders can only benefit the publisher as it would allow for more efficient ad slot allocation. Surprisingly, I find that this is not always the case. Under certain conditions, *not* knowing the entrant's CTR increases the publisher's revenue.[20]

The intuition revolves around two effects. First, the publisher'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 is higher if the publisher learns its CTR. This higher bid increases the incumbent's payment if it wins, which results in higher Stage 1 revenue for the publisher.

The second effect is subtler. Consider the publisher's Stage 2 revenue when $c_I > \mu_E$. Recall

---

[20]To be more precise, the common knowledge that the publisher does not know the entrant's CTR may increase its revenue.
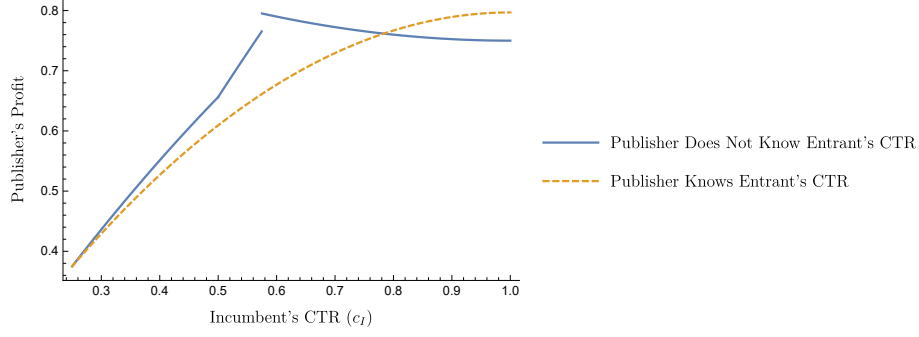
Figure 3.3: Publisher Revenue; $R = \frac{1}{4}, \delta = \frac{1}{2}, F_E(c) = c$

that the advertisers bid truthfully in Stage 2. If c_E is not known, the publisher's expected revenue in Stage 2 is $(\mu_E/c_I)c_I = \mu_E$. Using the definition of μ_E, this can be rewritten as

$$\int_0^1 c_E \, dF_E. \tag{3.3.3}$$

If c_E is known, the publisher's Stage 2 revenue depends on the realization of c_E and can be written as

$$\int_0^R R \, dF_E + \int_R^{c_I} c_E \, dF_E + \int_{c_I}^1 c_I \, dF_E. \tag{3.3.4}$$

Comparing the two integral expressions (3.3.3) and (3.3.4), I see that within the integration range $c_E \in (0, R)$, Expression (3.3.4) is larger; within the range $c_E \in (R, c_I)$, the two expressions are equal, and within the range $c_E \in (c_I, 1)$, Expression (3.3.3) is larger. Thus, if c_I is not too high, then the publisher's revenue when it does not know c_E (i.e., Expression (3.3.3)) is larger than when it does (i.e., Expression (3.3.4)). Intuitively, since the benefit of a high realization of c_E is bounded from above by c_I, i.e., the publisher cannot fully reap the benefits of a high c_E, the publisher's Stage 2 revenue may be higher when c_E is *not* known than when it is. Taken together, the publisher's ignorance of the entrant's CTR can be blissful for moderate values of c_I. This result is formalized in the following proposition.

Proposition 9 (Publisher Revenue: Ignorance is Bliss)**.** *The publisher'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.*

Proposition 9 suggests that publishers 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 publisher's revenue because ignorance induces advertisers to bid aggressively. However, Proposition 9 also reveals conditions under which the publisher's ignorance can be a curse. For instance, if the incumbent is strong (e.g., high c_I in Figure 3.3), then not knowing the entrant's CTR decreases the publisher's revenue. This is because, when c_I is sufficiently high, the entrant, who is the "price setter" in the auction, bids less aggressively. Furthermore, when c_I is high, the publisher does not learn the entrant's CTR in equilibrium due to the incumbent's aggressive bidding. As a result, it suffers from suboptimal allocation of the ad slot (i.e., missing out on a potentially high c_E).

Given that the publisher may incur a revenue loss for not knowing c_E, one may wonder what strategies a publisher can deploy to mitigate this loss. In the next section, I characterize the publisher's optimal strategy in a learning environment. I show that, in the presence of learning incentives, it is optimal for the publisher to favor the entrant in Stage 1 in order to increase the probability of the entrant's winning.

3.4   Publisher's Strategy

In the previous section, I assumed that advertisers have the same, commonly known valuation for the ad slot. Moreover, I focused primarily on the advertisers' strategies, with the publisher passively implementing an exogenously fixed auction mechanism. In this section, I analyze a setting where advertisers have stochastic, private valuations and, more importantly, the publisher optimally chooses the mechanism that maximizes its profit.[21] I show that, in the presence of learning incentives, the publisher can achieve the optimal revenue using a variation of the standard second-price auction with personalized (advertiser-specific) reserve prices. Additionally, the learning incentives induce the publisher to favor the entrant in Stage 1.

---

[21]In order to characterize the optimal mechanism, I have to assume stochastic private valuations for the advertisers; otherwise, the publisher's optimal strategy is to set the reserve price of Stage 2 to 1, leaving no surplus for the advertisers. Stochastic private valuation is a standard assumption in mechanism design literature; e.g., see Myerson (1981) for a general setting, and Edelman and Schwarz (2010) for the context of online advertising.

3.4.1 Optimal Mechanism

Suppose advertiser j's per-click valuation, v_j, is drawn from a c.d.f. G_j with support $[0, \bar{v}_j]$ and is private information, for $j \in \{I, E\}$. Following the literature on auction theory (see Krishna, 2010), I impose the following assumption on G_j.

Assumption 1 (Increasing Hazard Rate). *Let g_j denote the density of G_j. The hazard rate function $\frac{g_j(x)}{1-G_j(x)}$ is increasing in x for $j \in \{I, E\}$.*[22]

Prior to Stage 1, the publisher sets the ad auction rules. In particular, it decides the allocation rule (who wins the ad slot), and the payment rule (how much each bidder pays). The rest of the game proceeds the same as in Section 3.3. The following lemma characterizes the publisher's optimal mechanism.[23]

Lemma 3 (Publisher's Optimal Mechanism). *The publisher's optimal mechanism is as follows.*

1. *Compute the incumbent's and entrant's virtual bids, respectively, as*

$$\psi_{I1}(b_{I1}) = c_I \left(b_{I1} - \frac{1 - H_I(b_{I1})}{h_I(b_{I1})} \right) \text{ and } \psi_{E1}(b_{E1}) = \mu_E \left(b_{E1} - \frac{1 - H_E(b_{E1})}{h_E(b_{E1})} \right) + \delta \Delta_P,$$
(3.4.1)

and set the virtual reserve price to $\delta \Delta_I$.

2. *Compute advertiser j's virtual bid as*

$$\psi_{j2}(b_{j2}) = c_j \left(b_{j2} - \frac{1 - G_j(b_{j2})}{g_j(b_{j2})} \right) \text{ for } j \in \{I, E\},\text{[24]}$$
(3.4.2)

and set the virtual reserve price to 0,

---

[22]Assumption 1 greatly facilitates the derivation of the optimal mechanism. A large class of distributions satisfy this property; e.g., exponential, Weibull, modified extreme value, Gamma (with parameters $\alpha > 1, \lambda > 0$), and truncated normal (with "commonly accepted [parameters]"). See Barlow and Proschan (1965) and Brusset (2009) for details.

[23]The publisher's optimal mechanism is not unique. In this paper, I choose the mechanism that is consistent with the literature, in the sense that the publisher's optimal virtual bid transformation in a learning environment converges to the optimal virtual bid transformation in Myerson (1981) as the publisher's learning incentive goes to zero.

[24]I are slightly abusing notation: "c_E" in Stage 2 is $\tilde{c}_E$, which is c_E if c_E is learned, and μ_E otherwise.

where $H_j(b_{j1}) = G_j\left(b_{j1} - \frac{\delta\Delta_j}{c_j}\right)$, $\Delta_I = \pi_{I2}(\mu_E) - \int_0^1 \pi_{I2}(c_E)\, dF_E$, $\Delta_E = \int_0^1 \pi_{E2}(c_E)\, dF_E - \pi_{E2}(\mu_E)$, $\Delta_P = \int_0^1 \pi_{P2}(c_E)\, dF_E - \pi_{P2}(\mu_E)$, and $\pi_{j2}(c'_E)$ denotes the Stage 2 profit of player $j \in \{I, E, P\}$ under the optimal Stage 2 mechanism when the publisher assigns entrant's CTR as c'_E.

Allocate the ad slot to the advertiser with highest virtual bid, provided it exceeds the virtual reserve price. Payment (per-click) is equal to the minimum bid required for the winning advertiser to win.

The details of the proof are provided in the appendix. I briefly discuss here the intuition behind the optimal mechanism. Variables Δ_j, $j \in \{I, E, P\}$, capture the difference in a full-information Stage 2 vs. an incomplete-information Stage 2 in the players' payoffs; i.e., Δ_E measures the additional Stage 2 payoff the entrant gains from having its CTR learned by the publisher; Δ_P measures the additional Stage 2 payoff the publisher gains from learning the entrant's CTR; and Δ_I represents the additional Stage 2 payoff the incumbent gains if the entrant's CTR is *not* learned. Distributions H_j, $j \in \{I, E\}$, are similar to advertisers' valuation distributions G_j, except that they are shifted to account for the advertisers' incentives to have the entrant's CTR learned or not learned.

The derivation of the optimal mechanism closely follows Myerson (1981). The optimal mechanism in Stage 2, where learning incentives are absent, is a direct application Myerson's lemma. Intuitively, the virtual bid transformation amounts to sorting advertisers based on the marginal revenue they bring to the publisher (Krishna, 2010). Thus, allocating the ad slot to the advertiser with the highest virtual bid maximizes the publisher's profit.

In Stage 1, the presence of learning incentives (for both the advertisers and the publisher) distorts the advertiser's virtual bids compared to the standard format in Myerson (1981). Specifically, I see from (3.4.1) that the publisher additively inflates the entrant's virtual bid by $\delta\Delta_P$. This term represents the additional Stage 2 payoff the publisher gains from learning the entrant's CTR and is proven to be always positive.[25] Intuitively, since the publisher can only learn the entrant's CTR if

---

[25]To see that Δ_P is positive, it suffices to show $\pi_{P2}(c_E) = \iint \max\left[\psi_{I2}(x_{I2}|c_I), \psi_{E2}(x_{E2}|c_E)\right]^+ d\mathbf{G}$ is convex in c_E. The integrand is convex in c_E because it is the maximum of $\psi_{I2}(x_{I2}|c_I)^+$, which is independent of c_E, and $\psi_{E2}(x_{E2}|c_E)$ which is a linear function of c_E. And since any linear combination with positive weights of convex of functions is also convex, I conclude $\pi_{P2}(c_E)$ is convex in c_E.

the entrant wins in Stage 1, the publisher has an incentive to help the entrant win. The publisher accomplishes this by increasing the entrant's virtual bid in Stage 1.[26]

Lemma 3 also sheds light on the nature of the optimal virtual bids. For example, if the advertisers' valuations are uniformly distributed, then it is optimal for the publisher to compute virtual bids by multiplying the advertisers' bids with their expected CTRs (modulo an additive term). This implies that publishers with diffuse priors about advertisers' valuations can achieve near-optimal revenues by ranking advertisers based on CTR × bid. Moreover, the fact that the CTR-multiplier formula is also used in Stage 1 in the presence of learning dynamics attests to the robustness of this particular virtual bid format.

Next, I discuss the advertisers' bidding strategies under the optimal mechanism. Interestingly, I find that the insights from Section 3.3 regarding bid adjustments carry over to the optimal mechanism setting. As shown in Figure 3.4, the entrant overbids in Stage 1. Its motivation closely mirrors that of Section 3.3: its expected payoff in Stage 2 is greater if its CTR is learned by the publisher because the downside risk of a low c_E draw is bounded.

A weak incumbent bids below its valuation and helps the entrant reveal its CTR. In contrast to Section 3.3, however, the heterogeneity in advertisers' valuations necessitates an additional condition for this result to hold. Namely, the valuation distributions G_I and G_E must be such that the weak incumbent's probability of winning in Stage 2 decreases sufficiently slowly in c_E. Roughly, this is equivalent to the incumbent's valuation distribution being more concentrated around higher values than is the entrant's valuation distribution. For then, even if the entrant's CTR turns out to be high in Stage 2, the incumbent, whose valuation is more heavily concentrated on higher values, would still have a considerable chance of winning. This condition ensures the weak incumbent, who effectively helps the entrant win in Stage 1, feels adequately "insured" against the risk of a high c_E draw in Stage 2. The weak incumbent will then forego its Stage 1 profit and help reveal the entrant's CTR, as it creates an opportunity to earn higher profits against an entrant with a low CTR draw.

---

[26]It can be easily verified that the Stage 1 virtual bids in (3.4.1) reduce to the standard format (Myerson, 1981) when the learning dynamics are muted (e.g., $\delta = 0$).
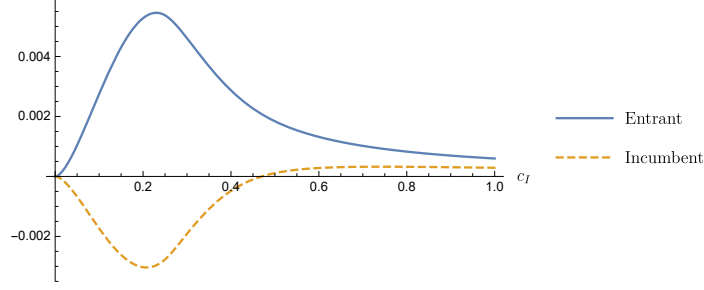
Figure 3.4: Advertisers' Bid Adjustments Relative to Valuation; $G_I(x_I) = \left(\frac{x_I}{2}\right)^4$, $G_E(x_E) = x_E$

Finally, a strong incumbent may overbid under the optimal mechanism (see Figure 3.4). Again, the intuition mirrors that from Section 3.3; however, the added necessary condition is that the incumbent's probability of winning in Stage 2 decrease steeply in the entrant's CTR draw. In this case, the incumbent deems the risk of revealing the entrant's CTR in Stage 2 too high. Therefore, it bids aggressively in Stage 1 and deters the publisher from learning the entrant's CTR. I summarize these findings in the following proposition.

Proposition 10 (Advertisers' Strategies in Stage 1 Under Optimal Mechanism). *Suppose the publisher implements the optimal mechanism characterized in Lemma 3. In Stage 1, the entrant always bids above its valuation. The incumbent bids below its valuation if c_I is low and the valuation distributions G_I and G_E are such that the probability of the incumbent winning in Stage 2 decreases sufficiently slowly in the entrant's CTR draw. The incumbent bids above its valuation if c_I is high and G_I and G_E are such that the incumbent's probability of winning in Stage 2 decreases steeply in the entrant's CTR draw. The exact conditions are provided in the proof in Section B.1.4.*

3.4.2 Optimal Reserve Prices

In this section, I delve deeper into a particular aspect of the optimal mechanism, the (effective) reserve prices. I examine how the learning incentives affect the publisher's optimal reserve prices in Stage 1. Before I proceed, I should clarify two distinct units of reserve prices. *Virtual reserve price* is defined in the virtual bids space, and measures the *minimum virtual bid* required for an advertiser to participate. *Effective reserve price* is defined in the submitted bids space, and refers

to the *minimum submitted bid* required for a particular advertiser to participate (Ostrovsky and Schwarz, 2016). To illustrate, suppose the publisher sets a virtual reserve price $0.15 and an advertiser bids b. Suppose that the advertiser's virtual bid is set to $0.1 \times b$ by the publisher. The advertiser will be considered in the auction if its virtual bid $0.1 \times b$ is greater than or equal to $0.15. Equivalently, the advertiser will be considered if its *submitted* bid b is greater than or equal to the *effective reserve price* $0.15/0.1 = $1.5.

Although the optimal mechanism uses the same virtual reserve price for all advertisers, because it applies different virtual bid transformations, advertisers experience different effective reserve prices. In the following, I consider the effective reserve price as it is a more intuitive concept to discuss advertisers' payments.

My analysis shows that the optimal reserve price depends crucially on two countercurrent forces. On the one hand, the *entrant's overbidding incentive* exerts an upward force on the entrant's reserve price. That is, the higher is the value that the entrant gains from the publisher learning its CTR, the higher it bids in Stage 1. The publisher can, thus, extract more from the entrant by setting a higher reserve price. Therefore, the reserve price increases with the entrant's overbidding incentive. The converse is true for the incumbent: the reserve price for the incumbent increases with the incumbent's value of *deterring* the publisher from learning the entrant's CTR.

On the other hand, the *publisher's learning incentive* pushes the entrant's reserve price downward. The higher is the value for the publisher if it learns the entrant's CTR, the greater its incentive to help the entrant win in Stage 1. This is accomplished by lowering the entrant's reserve price. Therefore, the reserve price decreases with the publisher's learning incentive.

Figure 3.5 illustrates the dynamics of each advertiser's optimal reserve price for particular valuation distributions G_I and G_E. Interestingly, the entrant's reserve price can be non-monotonic in the Stage 2 weight parameter, δ. When δ is small, the publisher learning incentive dominates, and as δ increases, the publisher lowers the reserve price to facilitate learning the entrant's CTR. When δ is large, however, the entrant overbidding incentive dominates. Here, as δ increases, the publisher increases the reserve price to extract more surplus from the entrant.
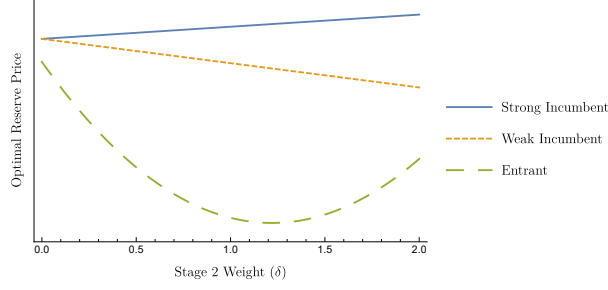
Figure 3.5: Optimal Reserve Prices (not to scale); $G_I(v_I) = \left(\frac{v_I}{2}\right)^3$, $G_E(v_E) = \left(\frac{4v_E}{3}\right)^3$, $F_E(c) = c$

So far, I have discussed the two countercurrent forces that induce the publisher to increase/decrease the entrant's reserve price in a learning environment. But what is the net effect of these forces on the entrant's reserve price? I conclude this section by presenting the conditions under which the publisher sets a lower reserve price for the entrant when it does not know the entrant's CTR than under full information. While a closed-form characterization of the optimal reserve price with respect to δ is intractable, I can analytically delineate the conditions for when the publisher sets a lower reserve price. In Proposition 11, I show that, for any δ, the publisher sets a lower reserve price for the entrant in a learning environment (compared to full information) if and only if the increment in its Stage 2 profit from learning the entrant's CTR is sufficiently high.

Proposition 11. *The publisher sets a lower reserve price for the entrant when the publisher does not know the entrant's CTR than when it does if and only if $\Delta_P > \Delta_I + \frac{\mu_E \rho}{\delta}$ (where Δ_j is as defined in Lemma 3 and $\rho > 0$ is as defined in the proof); i.e., the publisher's gain in Stage 2 from learning the entrant's CTR is sufficiently high.*

The results of Lemma 3 and Proposition 11 show that when the publisher's learning incentives are sufficiently strong, it is optimal for the publisher to *favor* the entrant. Favoring the entrant can be implemented by increasing the entrant's virtual bid as in Lemma 3, or decreasing the entrant's reserve price as in Proposition 11. In Section 3.5.3, I show that other mechanisms that favor the entrant can create a similar effect. For example, giving free advertising credit to new advertisers, or artificially inflating the "estimated" CTR of new advertisers can increase the publisher's revenue.

3.5 Extensions

I present fmy extensions of the main model to assess the robustness of my results. In Section 3.5.1, I show that my results from the main model continue to hold under impression-based pricing. In the main model, I assumed that the advertisers and the publisher have the same level of information about the CTRs; in Section 3.5.2, I relax this assumption to establish the robustness of my results. In Section 3.5.3, I explore other mechanisms that help the publisher increase its revenue in a learning environment. I show that offering free ad credit to new advertisers, or artificially inflating the "estimated" CTR of new advertisers can increase the publisher's revenue. Finally, in Section 3.5.4, I turn to the context of search advertising and discuss how my results change when there are multiple ad slots.

3.5.1 Impression-based Pricing

Suppose the publisher sells a single ad slot through a cost-per-impression (CPM) auction instead of cost-per-click. Consistent with practice, I assume that the publisher receives the advertisers' per-impression bids and assigns the slot to the highest bidder.[27] The winning bidder pays the minimum bid required to win the auction.

To model a CPM auction, it is important to recognize that the relevant performance metric for advertisers is the consumers' "estimated action rates" per ad impression. Consumer actions can range from clicking a link to watching a video longer than a certain amount of time, or completing a certain task on the advertiser's website. For $j \in \{I, E\}$, let $a_j \in [0, 1]$ denote the probability of action of a consumer conditional on viewing advertiser j's ad. I normalize the value of a consumer's action to 1 for both advertisers. Put together, a single impression is worth a_j to advertiser j.

In the spirit of the main model, assume that the action probability associated with the entrant's ad, a_E, is known only up to its c.d.f. $\tilde{F}$ with mean $\tilde{a}_E$, while the incumbent's action probability, a_I, is common knowledge.[28] The true value of a_E can only be learned if the entrant wins the first

---

[27]This contrasts with rankings based on effective bids in CPC auctions, where publishers multiply the advertiser's bid with its Quality Score.

[28]For my analysis, I only need the incumbent's WTP to be common knowledge; this is a standard assumption in

stage auction; only then can the entrant accurately assess the likelihood of a consumer responding to its ad with some pre-defined action. Similar to Section 3.3, I assume that the reserve price is exogenously set at R, and is less than a_I and $\tilde{a}_E$. In addition, as in Section 3.3, I assume $a_I + \delta \left((a_I - \tilde{a}_E)^+ - \int_0^1 (a_I - \max[a_E, R])^+ \, d\tilde{F} \right) \geq R$ to ensure the existence of weakly dominant strategies. Analyzing the advertisers' strategies yields the following proposition, which echoes the results from the main model.

Proposition 12 (Advertisers' Strategies in Stage 1 Under CPM). *In Stage 1, the entrant always bids above its valuation. The incumbent bids below its valuation if a_I, the action probability associated with its ad, is low, and bids above valuation if a_I is high. See Figure 3.6.*

Proposition 12 shows that advertisers' strategies under impression-based pricing are similar to those under performance-based pricing. However, there is an important difference in the advertisers' incentives between the two pricing models. Under performance-based pricing, the entrant does not care about learning the CTR itself; it overbids so that the *publisher* learns the CTR. In contrast, under impression-based pricing, since the CTR determines the entrant's valuation per impression, the entrant overbids so that the entrant itself learns its CTR. Similarly, under performance-based pricing, the incumbent adjusts its bid to affect whether the publisher learns the entrant's CTR or not, whereas under impression-based pricing, the incumbent wants to affect whether the entrant learns its own CTR or not. Despite discrepancies in the advertisers' incentives across these two pricing models, the mathematical expressions that capture the advertisers' payoffs are relatively similar: in performance-based pricing, the publisher includes the advertiser's CTR in the effective bid, whereas in impression-based pricing the advertiser includes the CTR in its submitted bid. As such, I obtain the same strategic behavior under both pricing models.

3.5.2   Information Asymmetry

In this section, I test the robustness of my results when the information symmetry assumption is relaxed. I examine two distinct cases. In Section 3.5.2, I demonstrate that the qualitative insights

---

papers with a full-information setting in online advertising, e.g.,Edelman et al. (2007).
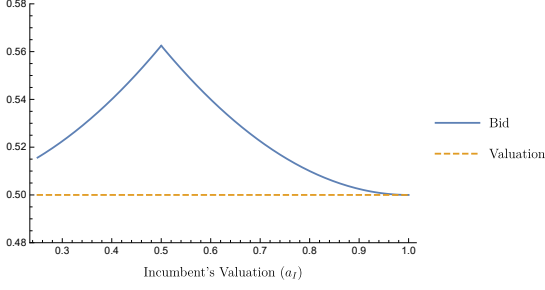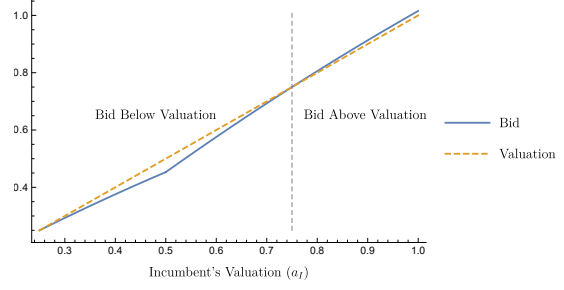
(a) Entrant's Bidding Strategy   (b) Incumbent's Bidding Strategy

Figure 3.6: CPM Bidding Strategies; $R = \frac{1}{4}, \delta = \frac{1}{2}, \tilde{F}(a) = a$

of the main model carry over when the entrant is at an informational disadvantage; i.e., the entrant only knows the incumbent's CTR up to some distribution, while the incumbent and the publisher know the incumbent's true CTR. In Section 3.5.2, I replicate the incumbent's bidding pattern from the main model in a setting where the incumbent only knows the distribution of the entrant's CTR up to a distribution, while the entrant and the publisher know the true distribution of c_E.

Entrant Does Not Know True c_I

In the main model, I assumed that the entrant, incumbent and the publisher possessed the same level of information regarding the incumbent's CTR. In reality, it could be argued that the entrant may not be as knowledgeable about c_I as the incumbent and the publisher. In this section, I analyze a model that captures this information asymmetry more realistically. My objective is to show that the the entrant's overbidding behavior is robust to the setting where the entrant is less informed about c_I than the incumbent and the publisher.

To that end, suppose that the entrant does not know the true value of c_I, but knows that it follows some distribution F_I over the support $[R, 1]$. On the other hand, the publisher and the incumbent both know the true c_I. I find that the entrant's overbidding pattern is robust to this information setting. The following proposition summarizes the finding.

Proposition 13. *Suppose the entrant does not know the true CTR of the incumbent, but only knows its distribution. In Stage 1, the entrant always bids above its valuation.*

I know from Section 3.3 that the entrant overbids for any given value of c_I. Intuitively, when

63

the entrant only knows the distribution of c_I, it integrates over all possible values of c_I to calculate its optimal bid. Since the optimal strategy for any value of c_I is to overbid, the optimal strategy when c_I is not known (i.e., the outcome of the integration) is still to overbid. This is formalized in the proof of Proposition 13.

Incumbent Does Not Know True Distribution of c_E

Another important case to consider is when the incumbent, as opposed to the entrant, is at an informational disadvantage. How would the incumbent having less information about c_E than the entrant and the publisher impact its bidding strategy?[29] To address this question, I allow for the possibility that the incumbent does not know the true distribution of c_E; i.e., the incumbent strategizes based on some prior belief over a range of possible distributions of c_E. The publisher and the entrant, on the other hand, know the true distribution of c_E.

Suppose the incumbent believes that the true distribution of c_E is F_x for some $x \in \mathcal{X}$. Let $P(x)$ denote the c.d.f. of the incumbent's prior over the class of distributions $\{F_x\}_{x \in \mathcal{X}}$, and μ_x the mean of F_x. In addition, as in Section 3.3, assume that the condition for the existence of weakly dominant strategies holds: $c_I + \delta \int_{x \in \mathcal{X}} \left((c_I - \max[\mu_x, R])^+ - \int_0^1 (c_I - \max[c_E, R])^+ \, dF_x \right) dP(x) \geq R$. Then there exists CTR thresholds such that a weak incumbent (i.e., with a low CTR) bids below its valuation, whereas a strong incumbent (i.e., with a high CTR) bids above it. I state this as a proposition.

Proposition 14. *Suppose the incumbent does not know the true distribution of the entrant's CTR. There exists a pair of thresholds $(\underline{c}', \overline{c}')$ such that the incumbent bids below its valuation if $c_I < \underline{c}'$, and bids above valuation if $c_I > \overline{c}'$.*

The intuition for the result of Proposition 14 is similar to that of Proposition 13. For any distribution F_x, I know from Section 3.3 that a weak incumbent underbids and a strong incumbent overbids. When the incumbent does not know the entrant's distribution, it integrates over all pos-

---

[29]Note that the entrant's bidding strategy from the main model is unchanged because it is independent of the incumbent's information structure — it depends primarily on the publisher's information structure.

sible distributions to calculate its optimal bid; however, the same pattern continues to exist. When the incumbent is sufficiently weak it overbids and when it is sufficiently strong it overbids. This is formalized in the proof of Proposition 14.

3.5.3 Other Mechanisms

In this section, I discuss two other mechanisms that can help publishers increase their revenue in a learning environment.

Free Advertising Credit for New Advertisers

Publishers 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.[30] Similarly, Facebook sends promotional codes to new advertisers that a have sufficiently high user engagement on their pages. In this section, I study the implications of offering ad credit on the advertisers' bidding strategies and the publisher's profit.

Suppose the publisher sets the ad credit $\alpha \geq 0$ prior to Stage 1, and then the incumbent's CTR is drawn from c.d.f. F_I with support $[R, 1]$.[31] The advertisers observe α and the rest of the game proceeds identically as in the main model. The effect of the publisher's ad credit is to transfer free ad credit α to the entrant if it wins in Stage 1. Thus, the entrant's Stage 1 payoff when the publisher offers ad credit α is $c_E \left(1 - \frac{\max[c_I b_{I1}, R]}{\mu_E}\right) + \alpha$ if it wins, and 0 otherwise. Next, I present the advertisers' bidding strategies when the publisher offers free ad credit.

Proposition 15 (Advertisers' Strategies with Ad Credit). *If the publisher offers ad credit $\alpha \geq 0$ to the entrant, then compared to the benchmark bids (3.3.1) and (3.3.2), the incumbent's first stage*

---

[30]https://www.google.com/ads/adwords-coupon.html

[31]I assume α is set before c_I is realized because publishers use the same amount of ad credit across many keywords for which incumbents have different CTRs. That is, in practice, α is not a function of c_I. In addition, note that if α is decided after c_I is realized, the truthfulness nature of these second-price auctions will break down. This is because the incumbent would anticipate the publisher to set α high enough to extract all surplus from the incumbent's bid, creating incentives for the incumbent to shade its bid. Finally, the assumption that the ad credit is only available in Stage 1 reflects the fact that these promotions typically expire after a short period of time.

bid remains unchanged, whereas the entrant's bid increases by $\frac{\alpha}{\mu_E}$ to

$$b_{E1}^*(\alpha) = 1 + \frac{\alpha}{\mu_E} + \frac{\delta}{\mu_E}\left(\int_{c_I}^1 c_E - c_I \, dF_E - (\mu_E - c_I)^+\right). \tag{3.5.1}$$

The intuition behind advertisers' bidding strategies is straightforward: the incumbent's bidding strategy does not change because α does not affect its underlying payment mechanism. However, the ad credit increases the entrant's payoff when it wins in Stage 1, and thus incentivizes the entrant to bid more aggressively in the first stage.

I turn to the impact of ad credit on the publisher's revenue. Given the first stage bids of the incumbent and the entrant in (3.3.1) and (3.5.1), respectively, the publisher's expected revenue as a function of ad credit α is $\mathbb{E}[\pi_P(\alpha)] = \int_R^1 \Pi_P(\alpha, c_I) \, dF_I$ where

$$\Pi_P(\alpha, c_I)$$
$$= \begin{cases} \mu_E b_{E1}^*(\alpha) + \delta \min[c_I, \mu_E] & \text{if } c_I b_{I1}^* \geq \mu_E b_{E1}^*(\alpha), \\ \max[c_I b_{I1}^*, R] - \alpha + \delta \left(\int_0^{c_I} \max[c_E, R] \, dF_E + (1 - F_E(c_I))c_I\right) & \text{if } c_I b_{I1}^* < \mu_E b_{E1}^*(\alpha). \end{cases}$$
$$\tag{3.5.2}$$

Expression (3.5.2) reveals the three forces created by the ad credit α. The first is the cost of α that is transferred from the publisher to the entrant when it wins; this has a negative effect on the publisher's revenue, and is represented by $-\alpha$ in the second case of Expression (3.5.2). The two other forces have a positive effect on the publisher's revenue, and are more nuanced; I discuss each of these in turn.

Recall that the entrant's bid increases in proportion to the ad credit α (see Proposition 15). This implies that the incumbent's payment upon winning increases with α. The publisher can thus extract additional surplus from the incumbent by inflating its payment. I call this the *extraction effect*.

The last effect of ad credit concerns the change in the publisher's Stage 2 payoff. To illustrate, suppose the incumbent's CTR is high. In this case, knowing the entrant's CTR leads to a higher publisher revenue than that under ignorance (see Proposition 9). 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 helps the entrant win, thereby facilitating the publisher learning its CTR, it could increase the publisher's Stage 2 revenue. I call this the *learning effect*.

These three effects summarize all the pros and cons of offering ad credit in my model.[32]

Inflating the Bid Multiplier

In this section, I analyze how the publisher's profit would change if, instead of offering ad credit (Section 3.5.3), the publisher artificially inflated the entrant's effective bid by a multiplier $\beta \geq 1$. To begin, suppose that the publisher applies a boosting multiplier β such that for any bid b_E of the entrant, the entrant's Stage 1 effective bid $\mu_E b_E$ increases to $\beta \times \mu_E b_E$. The rest of the game proceeds as in the main model.

I find that the two policies — offering ad credit and multiplicatively boosting the effective bid — have the same qualitative implications for the publisher's profit. The intuition is as follows. In the case of ad credit, the entrant increases its own effective bid by bidding high in anticipation of the ad credit, whereas in the case of boosting multiplier, the publisher increases the effective bid on behalf of the entrant. Thus, the resultant effective bids across the two policies are the same, and the players' payoffs are identical up to a constant. I formalize this finding in the following proposition.

Proposition 16. *The multiplicative boosting policy is isomorphic to the free ad credit policy, in the sense that the publisher can replicate (up to a constant) its profit from one policy using the other.*

In summary, I find that other mechanisms such as offering free ad credit or inflating the entrant's effective bid can increase the publisher's revenue as they facilitate learning by favoring the entrant. These mechanisms, however, are inefficient in the sense that they do always guarantee the publisher the optimal profit. To see this, note that the optimal mechanism sometimes lowers the payment for both the entrant and the incumbent at the same time (e.g., when the incumbent's CTR is low). Such

---

[32]Note that I 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, I focus on the extraction and learning effects of ad credit that are new to the literature. I show that even if the free ad credit does not attract new advertisers, the publisher may still benefit from offering it because of these two positive effects.

an outcome cannot be produced by offering ad credit or inflating the entrant's bid because these instruments unilaterally benefit the entrant at the expense of the incumbent.

Another reason why these alternative mechanisms do not always yield the optimal profit pertains to the virtual bid transformations. Offering ad credit and inflating the entrant's bid do not take into account the advertisers' valuation distributions in the allocation rule. Recall that the optimal mechanism computes the marginal revenue each advertiser generates based on the advertisers' valuation distributions and then allocates the ad slot accordingly. Such efficient allocation cannot always be attained with artificial adjustment of the entrant's effective bid, especially when the valuation distributions are non-uniform.

3.5.4 Multiple Advertising Slots

Another assumption in the main model is that the publisher offers a single ad slot. In search advertising, however, search engines typically sell more than one ad slot, which are allocated via the Generalized Second-Price (GSP) auction. In this section, I test whether the main insights derived from the base model carry over to the multiple-slot GSP setting.

I consider a two-slot, three-player game where two incumbents face an entry from a new advertiser. To simplify the analysis, I assume that the existing advertisers also learn c_E if the entrant wins in Stage 1.[33] As in Section 3.3, all advertisers share a common per-click valuation of 1, and the reserve price R is less than μ_E. I normalize the position-specific CTR of the first ad slot to 1 and denote that of the second slot as $\theta \in (0, 1)$. I index by i and I the incumbent with the lower and higher CTR, respectively (i.e., $c_i < c_I$), and normalize c_I to 1.[34] Since a weakly dominant strategy equilibrium no longer exists, I use the lowest-revenue envy-free (LREF) Nash equilibrium (Edelman et al., 2007) for equilibrium selection.

My analysis shows that the main results are robust to multiple-slot settings under GSP auction.

---

[33]This could be justified by the fact that advertisers can estimate the effective bid of the advertisers below them by observing the amount they are charged.

[34]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.

In particular, the entrant's overbidding strategy carries over, but with the caveat that the reserve price has to be sufficiently high. 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 next lemma summarizes the advertisers' incentives in Stage 1.

Lemma 4 (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.*

Lemma 4 shows that if the second ad slot generates very few clicks (i.e., θ is low), then the incumbent with the higher CTR is better off masking the entrant's CTR, thereby securing the top ad position. The intuition resonates with the insights from the main model. Had the search engine learned the entrant's CTR and it turned out to be high, the strong incumbent would risk being "downgraded" to the low-CTR slot below. Conversely, if the second ad slot generates almost as many clicks as the first slot (i.e., high θ), the strong incumbent benefits from the search engine learning the entrant's CTR. The reason is that the incumbent can capitalize on a potentially low c_E realization, while its loss from possibly being driven down to the second ad slot against a high c_E is mitigated by the high θ.

Next, I examine the advertisers' bidding strategies in Stage 1. Figures 3.7a and 3.8a depict the entrant's bid with respect to the weak incumbent's CTR, c_i. Observe that for high c_i, the entrant's high bid mirrors the pattern from Section 3.3 (see Figure 3.2). Intuitively, 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 an opportunity to receive a positive payoff in Stage 2.

When facing a low c_i, the entrant's bidding strategy may diverge from the main model. In contrast to the single-slot case, the entrant may lower its bid when the incumbent's CTR is low
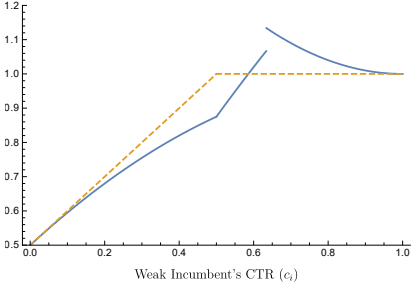
(a) Entrant
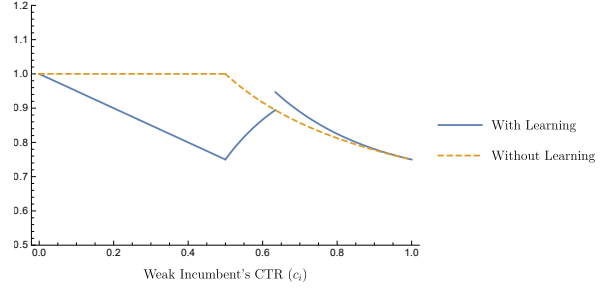
(b) Weak Incumbent

Figure 3.7: Optimal bids in GSP without reserve prices; $\delta = 1$, $\theta = \frac{1}{2}$
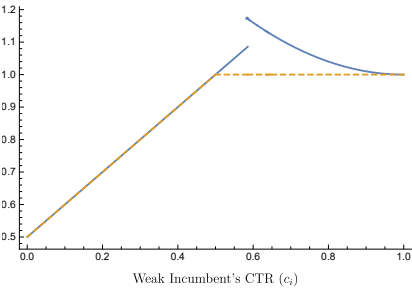
(a) Entrant
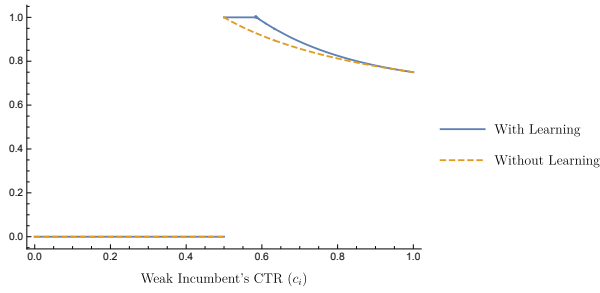
(b) Weak Incumbent

Figure 3.8: Optimal bids in GSP with reserve prices; $\delta = 1$, $\theta = \frac{1}{2}$

(see Figure 3.7a). The intuition revolves around the weak incumbent's incentive to help the search engine learn the entrant's CTR. Figure 3.7b shows that a weak incumbent shades its bid in order to help the entrant secure the second slot in Stage 1. The search engine can then learn the entrant's CTR, which in turn creates an opportunity for the weak incumbent to win in Stage 2. And since in the LREF equilibrium the advertisers' bids change in proportion to their competitors', the entrant also shades its bid for low c_i.

However, when the search engine sets a sufficiently high reserve price, the weak incumbent bids below the reserve price in the LREF equilibrium (see Figure 3.8b). As a result, the entrant's incentive to shade its own bid disappears, and I recover the overbidding pattern from the main model (see Figure 3.8a). I formalize this finding in the next proposition.

Proposition 17. *The entrant always bids (weakly) higher in the setting where the search engine does not know (but can learn) the entrant's CTR compared to the full-information setting if and only if the reserve price is sufficiently high.*

70

3.6 Conclusion

In this paper, I study learning in online advertising. I investigate how a publisher's lack of information about a new advertiser's click-through rate affects the strategies of new and existing advertisers, as well as the publisher. My theoretical analysis offers useful insights for several issues of managerial importance.

Implications for New Advertisers. I show that when a new advertiser starts online advertising with a publisher, 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 publisher 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.

My 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 an existing advertiser. First, if the new advertiser's CTR turns out to be high, the existing advertiser risks losing its ad slot to the new advertiser. Second, since the online advertising slots are sold in auctions, entry of a new advertiser increases the payment of the existing advertiser. I demonstrate that, in response to these entry effects, an existing advertiser with a high CTR — e.g., a trademark owner advertising on its branded keywords, or a manufacturer advertising on its own product pages on an online retailer — 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 Publisher. When a new advertiser enters the market, the publisher 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 publisher's expected revenue. Surprisingly, my result shows that the ignorance may be a boon to the publisher: its ignorance may incentivize the advertisers to bid more aggressively, which in turn may increase the publisher's revenue compared to the full information benchmark.

The publisher'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 publisher's long-term revenue. I show the publisher can mitigate this loss by favoring the new advertiser in the auction. For example, by lowering the reserve price of, offering free ad credit to, or artificially inflating the bid of the new advertiser, the publisher can increase the probability that the new advertiser wins. This allows the publisher to learn the new advertiser's CTR more quickly, which in turn increases the publisher's long-term revenue. In fact, my results show that the optimal selling mechanism favors the new advertiser in the early rounds of the auction.

Future Research. My work is a first step towards understanding how agents strategically respond to a publisher's learning process. Future research could explore other scenarios where agents and publishers interact in a learning environment. For instance, a publisher may want to learn sellers' qualities of products for ranking purposes, or customers' WTP for pricing purposes. In addition, while I allow the transition from an incomplete to a full information game, I make several simplifying assumptions in doing so. For example, the transition is discrete and binary in my model. Analyzing the advertisers' strategies in a model with gradual, continuous learning process could lead to interesting additional insights. Finally, my model assumes that advertisers' CTRs are exogenously given. While this assumption is realistic for a given ad copy, it does not capture advertisers' constant efforts in improving their ad copies (e.g., through experimenting with new ad copies). Modeling advertisers' experiments in improving their CTRs, while the CTRs are being learned by the publisher and the advertisers themselves, is another interesting avenue for future

72

research.

Chapter 4: Customer Purchase Journey, Privacy Choices, and Advertising Strategies

4.1 Introduction

Advances in information technology have led to unprecedented levels of consumer tracking on the Internet (Lerner et al., 2016; Macbeth, 2017; Manjoo, 2019). According to Schelter et al. (2018), 355 third-party domains had installed trackers (e.g., cookies and web beacons) on over 90% of 41 million websites. Moreover, a recent study by Karaj et al. (2019) shows that 82% of the monitored web traffic had third-party scripts owned by Google, making it the largest third-party tracker by reach. These trackers allow firms to monitor not only which sites consumers visit, but also their browsing behavior such as whether the consumers interacted with the firms' ads (Roesner et al., 2012). Firms track consumers' online behavior for many reasons. Tracking helps firms (i) analyze site traffic and browsing patterns in order to deliver personalized content (Goldfarb and Tucker, 2011a; Bleier and Eisenbeiss, 2015), (ii) infer consumers' product preferences to inform pricing decisions (de Cornière and Nijs, 2016; Ichihashi, 2019; Montes et al., 2019; Taylor, 2004), and (iii) target ads to particular consumer segments (Bergemann and Bonatti, 2011; Iyer et al., 2005; Shen and Villas-Boas, 2018).

In particular, tracking helps advertisers to observe consumers' online browsing and purchase activities and infer their purchase journey stages. For instance, (Sahni et al., 2019) use consumer tracking data to infer whether a consumer is a "product viewer" or a "cart creator" and target ads accordingly. Google allows advertisers to use various tags to specify remarketing audiences based on such inferences.[1] Industry experts advocate that advertisers focus on targeting based on consumers' "stages in the decision journey" (Edelman and Schwarz, 2010). Empirical findings that

---

[1] https://support.google.com/google-ads/answer/6335506 (accessed December 2019)

ad effects, as measured by sales (Johnson and Myatt, 2017; Lambrecht and Tucker, 2013; Seiler and Yao, 2017) or website return visits (Hoban and Bucklin, 2015; Sahni et al., 2019), vary widely across consumers' journey stages further highlight the importance of considering the purchase journey in developing advertising strategies (Todri et al., 2019).

While consumer tracking has benefited advertisers (Goldfarb and Tucker, 2011a; Johnson et al., 2019), its rapid expansion has deepened consumers' concerns about their online privacy (McDonald and Cranor, 2010). For instance, 77% of US Internet users indicate that they are "concerned about how tech/social media companies are using [their] online data ... for commercial purposes" (eMarketer, 2019a), 64% of UK Instagram users say "it's creepy how well online ads know me" (eMarketer, 2018a) and 68% of US Internet users report feeling concerned about "social media companies displaying ads based on their data" (eMarketer, 2018b).

In response to the growing outcry from consumers and privacy advocates, advertising organizations and regulators worldwide have sought to curb practices that potentially infringe on privacy, such as online tracking. Notably, in May 2018, the European Union (EU) enacted the General Data Protection Regulation (GDPR). Compared to its predecessors (e.g., Privacy and Electronic Communications Directive), the GDPR is considered the most stringent and comprehensive in terms of geographic and legislative scope.[2] Its hefty violation fines (maximum of $22.5 million and 4% of annual global turnover) are forcing even large firms like Google and Facebook to take compliance seriously.[3] The California Consumer Privacy Act (CCPA), a US analogue of the GDPR, is expected to go into effect in January 2020.[4]

One of the main tenets of the GDPR and the CCPA is the requirement that firms not only inform consumers what data will be collected for what purposes, but also obtain explicit affirma-

---

[2]The regulation applies to all firms processing personal data of European subjects even if the firm operates outside of Europe. Personal data is defined as "any information relating to an identifiable person who can be directly or indirectly identified in particular by reference to an identifier" (`https://eugdpr.org/the-regulation/gdpr-faqs/`; accessed May 2019).

[3]In January 2019, Google was fined $57 million "for not properly disclosing to users how data is collected across its services ... to present personalized advertisements" (Satariano, 2019). Facebook revamped their privacy settings in compliance with the GDPR (`https://marketingland.com/what-marketers-need-to-know-about-facebooks-updated-business-tools-terms-238140`; accessed September 2019).

[4]See Future of Privacy Forum (2018) for a detailed comparison of the two regulations.

tive consent to use their data. In other words, firms are not allowed to collect consumer data by default; consumers themselves must opt-in to their data being collected and processed by firms.[5] If consumers opt-out from tracking, then advertisers cannot monitor consumers' behavior across websites. Consequently, advertisers' targeting capabilities are drastically undermined and ad impressions could be potentially wasted (e.g., repeated exposure to consumers who had already purchased).[6] On the other hand, if consumers opt-in to tracking, advertisers can target ads to specific audiences based on a set of behavioral criteria (e.g., consumers who previously interacted with the ad but did not purchase).

The impact of privacy regulations on the advertising industry is a topic of ongoing debate among practitioners, academics, and policymakers. On one hand, regulations are expected to limit advertisers' tracking capability, thereby reducing ad effectiveness (Aziz and Telang, 2016; Goldfarb and Tucker, 2011b). This may have contributed to the 50% decline in bids coming through sell-side ad platforms, and the 15% reduction in Google ad offerings via its ad exchange, after the GDPR went into effect (Kostov and Schechner, 2018). On the other hand, there is evidence suggesting that despite consumers' *stated* aversion towards tracking, they appear not as reluctant to allow tracking in practice.[7] For example, Johnson et al. (2019) find that less than 0.26% of US and EU consumers opt-out from behavioral targeting in the AdChoices program. Moreover, 67% of US and Canadian consumers report that they would feel "comfortable sharing personal information with a company" if it transparently discloses how their data will be used (Ipsos, 2019). These findings suggest that privacy regulations that endow consumers with the choice to being tracked may not necessarily result in low opt-in rates. In this respect, the net effect of privacy regulations may not be as detrimental to advertisers as they fear.

---

[5]While consumers were able to manually delete cookies even before the regulation, complete tracking prevention was extremely costly, if not impossible. For example, data collectors often used flash cookies technology to re-spawn cookies that were deleted by consumers (Stern, 2018; Angwin, 2010). Moreover, firms were able to purchase personal data from third-party information vendors without consumers' consent — such activities are now subject to GDPR enforcement.

[6]https://www.blog.google/products/marketingplatform/360/privacy-safe-approach-managing-ad-frequency/ (accessed October 2019)

[7]The overstatement of privacy concerns relative to revealed preferences is known as the "privacy paradox." See Norberg et al. (2007) and Athey et al. (2017) for details.

The discussion above on the advancements in tracking technology and shifts in industry regulations raises important questions for marketers and regulators alike. Does consumer tracking, which enables targeting based on a consumer's inferred purchase journey stage, lead to higher or lower levels of advertising intensity? How do advertising intensity and advertising effectiveness influence consumers' privacy choices of whether to allow being tracked? What are the implications of consumers' endogenous privacy choices on the ad network's profit? Which market participants benefit and lose from the regulation?

In this paper, I seek to shed light on these questions by developing a game theory model. I consider a two-period model in which consumers visit content pages, and each consumer creates one opportunity for an ad impression per period. An advertiser buys ad impressions from an ad network that sells ad inventory supplied by the content pages. Motivated from the discussion above, I assume that ad effects depend on the consumers' journey states represented by a "funnel" and that their purchase journey is influenced by advertising (Abhishek et al., 2017, 2018; Kotler and Armstrong, 2012). Based on their preferences for ad exposure and privacy, consumers choose whether to allow advertisers to track their online behavior.

Importantly, I model privacy as a multi-dimensional construct and assume that consumers jointly consider two aspects of privacy: its intrinsic value and its instrumental value (Becker, 1980; Posner, 1981; Farrell, 2012; Wathieu and Friedman, 2009). The intrinsic value of privacy refers to the utility consumers derive from protecting privacy for its own sake. I assume that consumers derive positive utility from protecting their privacy. On the other hand, the instrumental value of privacy stems from the indirect effects of the consumers' privacy choices (e.g., opting-in to tracking may affect the consumers' expected ad experiences). In contrast to the intrinsic aspect of privacy, the instrumental aspect of privacy may be either positive or negative depending on the firms' strategic responses.

Before proceeding, note that in my model, the instrumental value of privacy stems from the advertiser's strategic responses to consumers' purchase journey information (or lack thereof). To the extent that this formulation captures the notion that an individual's utility is dependent on the

actions of the recipient of the individual's information, my conceptualization is largely consistent with the literature. For example, Becker (1980) and Posner (1981) discuss scenarios in which a job applicant's utility is affected by her potential employer's reaction to her criminal records, and Farrell (2012) considers a case where a consumer's utility is dependent on a firm's discriminatory pricing strategy in response to learning the consumer's willingness-to-pay.

My analysis yields a number of interesting insights. First, I find that, under certain conditions, consumers choose to opt-in to being-tracked because they expect to see fewer ads when advertisers can track them and infer their funnel states. In particular, this is the case if ad effectiveness is intermediate. Intuitively, the reason is the following. Consider an opt-out consumer who cannot be tracked. To this consumer, the advertiser shows ads in both periods if ad effectiveness is intermediate: ad effectiveness is high enough such that the first ad is worthwhile, and low enough that the first ad does not render the second ad wasteful. In contrast, for opt-in consumers, the advertiser shows a targeted ad in the second period only to selected consumer segments. Therefore, if ad effectiveness is intermediate, some consumers trade-off their costs from the instrumental and intrinsic aspects of privacy; i.e., they trade-off the benefit of seeing fewer ads by opting-in to tracking (positive instrumental value of privacy) with the disutility they feel from giving up their privacy (intrinsic cost of privacy).[8] Under other conditions, opting-in to tracking (weakly) increases the number of ads seen, in which case there is no such trade-off and both aspects of privacy utility induce consumers to opt-out from being tracked.

Second, I find that the consumers' opt-in decisions have important implications for the ad ecosystem. In particular, due to changes in consumers' opt-in behaviors, the ad network's profit may decrease in ad effectiveness, even though higher effectiveness implies higher purchase conversion probability. Intuitively, high ad effectiveness induces the saturation effect whereby the

---

[8]As firms comply with high standards of transparency enforced by privacy regulations, consumers will become not only aware of the privacy choices they are entitled to, but increasingly knowledgeable about the downstream consequences of their choices. For example, Figure C.1 in Appendix C.1 shows a sample privacy notice from Google shown to consumers in Europe. It describes the potential changes in ad intensity that could result from consumers' privacy choices. In practice, a significant fraction of ad slots indeed can be left unsold; for display ads, ad fill rates (i.e., the ratio of the number of ad slots that are available to get filled to the number that actually get filled) typically range from as low as 67% to 100% (Balseiro et al., 2014; Johnson et al., 2019).

marginal value of successive ads is diminished by previously shown ads. This causes the advertiser to forego showing successive ads to opt-out consumers. In contrast, for opt-in consumers, enhanced targeting efficiency induces the advertiser to show successive ads. Thus, consumers expect to see fewer ads under no tracking, which incentivizes them to opt-out from tracking. As consumers opt-out, targeting efficiency falls, lowering ad valuations. Consequently, the ad network's profit can decrease as ads become more effective.

Third, privacy regulations increase consumer surplus and decrease the ad network's profit compared to a regime in which everyone can be tracked. Interestingly, however, if the advertiser is privately informed about ad valuations, consumers opting out of tracking may be a boon to the ad network. The ad network's inability to track opt-out consumers serves as a commitment mechanism that induces the ad network to sell untargeted ads that reach a larger consumer segment than targeted ads. This supply-side "market thickening" effect sometimes induces the advertiser to bid more aggressively for opt-out consumers than for opt-in consumers. Therefore, the ad network's profit may be higher if some consumers exercise their privacy rights and opt-out from tracking.

I have a number of extensions that relax some of the simplifying assumptions in the main model. I consider cases with (i) information asymmetry between the advertiser and the ad network, (ii) multiple competing advertisers, (iii) imperfect signals of purchase histories for opt-in consumers, and (iv) an infinite time horizon where consumers arrive in overlapping generations. Overall, I find that the main insights are not affected in these extensions, while I obtain certain interesting new insights. For instance, I find from one of the extensions that more consumers may opt-in to being tracked if the signal about purchase behavior is less accurate; as a result, the ad network may be better off having lower signal precision.

In addition to being related to the papers referenced earlier, my paper contributes to two interrelated streams of research: targeted advertising and online privacy. Extant literature on targeted advertising studies various implications of targeting. For example, it examines the impact of targeting on ad supply, ad prices, advertising strategies, ad intensity and adoption of ad avoidance tools (Athey and Gans, 2010; Aziz and Telang, 2017; Bergemann and Bonatti, 2011; Esteban et al.,

2001; Iyer et al., 2005; Johnson et al., 2013; Shen and Villas-Boas, 2018). I extend the existing literature in a novel and important way by modeling the consumer purchase journey, which allows us to study funnel state-dependent ad effects. I show that modeling funnel considerations creates a previously-unstudied link between the effectiveness of cross-period ads, which leads to novel insights pertaining to the impact of tracking on advertising strategies.

I also contribute to the growing literature on online privacy. Research on price-discrimination examines consumers' implicit privacy decisions, whereby consumers strategically time their purchase to control the disclosure of their preferences to the firm, thereby mitigating price-discrimination (Taylor, 2004; Villas-Boas, 2004). Other papers investigate more explicit privacy decisions, whereby consumers take (often costly) actions to control the amount of information disclosed to firms (Acquisti and Varian, 2005; Conitzer et al., 2012; Ichihashi, 2019; Montes et al., 2019). de Cornière and Nijs (2016) investigate the ad network's incentive to disclose consumer information to advertisers, whereas in my paper, the consumers exercise their privacy rights to decide the flow of their personal information. The mechanisms behind my results are orthogonal to market thickness (Bergemann and Bonatti, 2011; Rafieian and Yoganarasimhan, 2018) and market structure (Campbell et al., 2015) as I abstract from advertiser competition in the main model.

D'Annunzio and Russo (2019) study a similar setting to ours where consumers can endogenously decide whether to be tracked or not. However, my paper is different in several important ways. First, I explicitly model consumer tracking along the purchase journey; i.e., advertisers track consumers' progressions through the purchase journey after a series of ad exposures, rather than tracking single- vs. multi-homing consumers across different publishers. Second, the consideration of consumers' transitions down the purchase journey by virtue of previously shown ads gives rise to multi-period dynamics as advertisers consider retargeting consumers along the journey which I explicitly model. On the other hand, D'Annunzio and Russo (2019) consider a reduced-form effect of tracking in a static environment with a focus on publishers' decisions (e.g., ad capacity and outsourcing advertising to ad networks). Finally, while D'Annunzio and Russo (2019) assume that the number of ads a consumer is exposed to is independent of her privacy choice, I explicitly

80

incorporate potential changes in ad intensity as an instrumental aspect of the consumer's privacy choice. In my model, the consumer's instrumental value of privacy is weighed against her intrinsic value of privacy.

At a higher level, my research (i) advances the understanding of the impact of tracking on the advertising ecosystem from a novel purchase journey perspective and (ii) contributes to the ongoing debate on online privacy regulations. My findings suggest that assessing the impact of privacy regulations on the advertising industry is a complex issue. Nevertheless, I identify several robust theoretical insights that can inform various regulatory implications.

The rest of the paper is organized as follows. In Section 4.2, I describe the main model. In Section 4.3, I present the main results including the impact of tracking on advertising intensity, consumers' opt-in behavior, and the implications of endogenous privacy choice on the ad network's profit. In Section 4.4, I assess the robustness of the main insights by analyzing fmy extensions. In Section 4.5, I summarize the key results and conclude. All proofs are relegated to Section C.2 of the appendix.

4.2   Model

The game consists of three players: consumers, an advertiser and an ad network. Consumers sequentially visit content pages where the ad network enables showing ads to them. The advertiser buys ad impressions from the ad network to reach consumers. Before I discuss each player's decisions and payoffs, I first explain a key feature of my model: the consumer purchase journey. I describe the relationship between advertising and consumers' progression down the purchase funnel.

Purchase Journey and Ad Effects

I consider a stylized purchase journey consisting of three distinct states labeled *top*, *middle*, and *bottom* (see Figure 4.1). For ease of exposition, I denote these states by T, M and B, respectively, and the consumers in the respective states by f-consumers for $f \in \{T, M, B\}$. I define funnel state

81
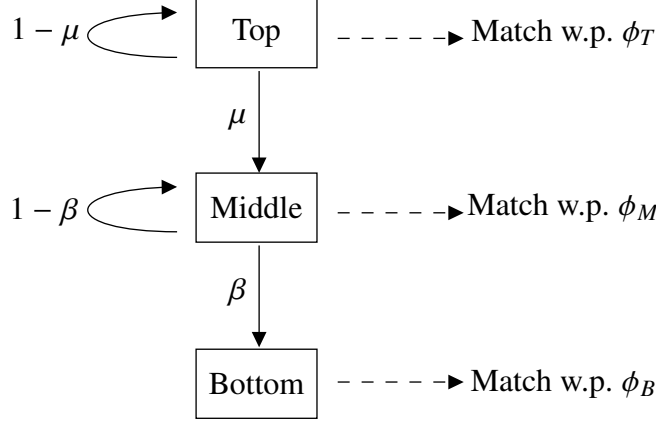
Figure 4.1: Purchase Journey and Ad Effects

$f \in \{T, M, B\}$ with a probability ϕ_f, which measures the likelihood of an f-consumer realizing a product match. In each time period, an f-consumer (who has not purchased yet) realizes a product match with probability (w.p.) ϕ_f, where $0 \leq \phi_T < \phi_M < \phi_B \leq 1$.[9] The three funnel states can be interpreted as follows: the top-funnel corresponds to "awareness" state, wherein the consumer is aware of the product's existence but is not seriously considering purchase; the mid-funnel corresponds to "consideration" or "interest" state, wherein the consumer is potentially considering purchase; and the bottom-funnel corresponds to even higher consideration and purchase interest by the consumer. I normalize ϕ_T and ϕ_B to 0 and 1, respectively.

If a consumer realizes a product match, she derives positive utility v from consuming the product; otherwise, she derives zero utility. In accordance with the empirical literature, I assume that ads affect consumers' likelihood of realizing a match with the advertised product (e.g., Johnson et al., 2016; Lee, 2002; Sahni, 2015; Shapiro et al., 1997; Xu et al., 2014); i.e., ads influence the consumers' progression through the funnel. Ads induce T-consumers to transition to funnel state M w.p. $\mu \in [0, 1]$, and have no effect w.p. $1 - \mu$. Similarly, ads induce M-consumers to transition to funnel state B w.p. $\beta \in [0, 1]$ and have no effect w.p. $1 - \beta$.

Note that my model specifications allow for flexible ad response curvatures using the funnel transition parameters μ, β, and mid-funnel match probability ϕ_M. For instance, Figure 4.2 depicts a convex ad response curve with small ϕ_M, and a concave ad response curve with large ϕ_M. Adver-

---

[9]Taylor (2004) calls this match probability the "intensity of taste for a particular class of goods" (pg. 635).
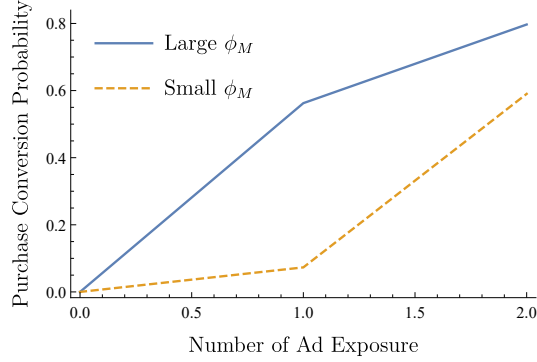
Figure 4.2: Ad Response Shapes

tising strategies, consumer choices and welfare outcomes will depend crucially on the curvature of the consumers' ad responses.

I also note that my results will remain completely unaltered if I scale the effectiveness numbers to vary between ranges different from 0 to 1. For instance, if I consider a sub-population of consumers who potentially respond to ads while the other consumers do not, then the unconditional effectiveness of ads shown to the whole population would be scaled down in proportion to the sub-population of responsive consumers. Therefore, ad effectiveness numbers for the whole population of consumers can be scaled to vary between, say, 0 and 0.05 (i.e., effectiveness rates between 0% and 5%, which are arguably closer to empirical estimates), or any other range, rather than between 0 and 1 (i.e., effectiveness rates between 0% and 100%), without any impact on my results (see Section C.3 of the appendix for more details). However, for model simplicity and expositional clarity, I use the formulation with effectiveness numbers varying between 0 and 1.

Consumers

A unit mass of consumers visit two content pages (both of which are in the ad network), one in each of Period $t \in \{1, 2\}$. Consumers are exposed to at most one ad impression per period from the page they visit. As described above, these ad exposures influence the consumers' progression through the purchase journey. For now, I assume that the initial state of newly arriving consumers is T. In other words, new consumers who visit content pages for the first time are *not* considering purchase. This helps deliver the main insights more cleanly. In Section 4.4.4, I relax this

83

assumption by allowing some fraction of consumers to arrive in funnel state M.

Consumer utility consists of two components, product utility and privacy utility. The product consumption utility of an f-consumer (i.e., consumer in funnel state $f \in \{T, M, B\}$) is

$$u_{\text{prod}} = \tilde{v}_f - p,$$

where $\tilde{v}_f$ represents the stochastic match valuation, which equals v w.p. ϕ_f, and 0 w.p. $1 - \phi_f$, and p denotes the product price. If the consumer does not purchase, she derives the outside option utility 0. Without loss of generality, I normalize the match utility v to 1. Therefore, the consumer purchases if and only if she realizes a match and $p \leq 1$. Note that a consumer makes the purchase decision after realizing her match value. I assume that a consumer purchases at most one unit.

Next, I turn to privacy utility. I assume that consumers dislike being tracked, and that they are heterogeneous in their tracking disutility.[10] This disutility is captured by the privacy cost parameter θ, which has cumulative distribution function F. I assume that consumers can decide whether to opt-in or opt-out of being tracked. If a consumer opts-in, firms can track her identity and online browsing behavior (across content pages and across sessions) for targeting purposes. Later, I describe in more detail how tracking and targeting are implemented. The privacy utility of a consumer with privacy cost θ is

$$u_{\text{priv}}(x) = -\eta \, \tilde{q}(x) - \theta x, \tag{4.2.1}$$

where x denotes the consumer's privacy decision, which equals 1 if she opts-in, and 0 if she opts-out,[11] $\tilde{q}$ the total number of ads she expects to see, η the disutility she incurs per unit of ad impression (Johnson et al., 2013; de Cornière and Taylor, 2014), and θ the disutility she incurs for

---

[10]Heterogeneity may stem from numerous factors such as differences in what consumers believe constitutes personal information (Acquisti et al., 2016) and differences in consumers' perception of their privacy control (Tucker, 2014).

[11]I clarify two implicit assumptions here. First, while consumers may modify their privacy decision at any time in practice, I assume consumers make a one-time privacy decision at the beginning of the game. Second, following the literature on endogenous privacy choices (e.g., Conitzer et al., 2012; Montes et al., 2019), I assume that privacy decision is binary. In practice, consumers may choose varying degrees of information disclosure. These assumptions keep the analysis simple without significantly changing the qualitative insights.

allowing tracking. In sum, consumers' privacy decisions are based on (i) the number of ads they anticipate to see as a result of their privacy decisions, and (ii) the extent to which they value privacy for its own sake. These two components constitute the instrumental and intrinsic aspects of privacy, respectively.

Advertiser

Depending on whether consumers can be tracked or not, the advertiser can buy different types of ads. If tracking is prohibited, then the consumers' identities cannot be matched across content page visits. In this case, the advertiser can only buy untargeted impressions (e.g., ads displayed to all website visitors independent of their browsing histories). In particular, even if an ad is shown in Period 1 and shifts the distribution of consumers along the purchase journey, the advertiser is not able to target ads in Period 2 based on the funnel states.

On the other hand, if consumer tracking is allowed, the advertiser can buy ad impressions at the funnel-stage level. By installing tags on websites and embedding cookies on consumers' browsers, the advertiser can monitor the websites visited by the consumers, their browsing activity within the websites, and their purchase behavior.[12] Based on this information, the advertiser can specify the target audience such that their ads are shown only to consumers who meet some pre-specified criteria.[13] For example, the advertiser can target ads to consumers who are inferred to be in funnel state M and did not purchase. In each period, the advertiser decides which impressions to bid for and the respective bid amounts.

The advertiser also sets product price p_t in Period $t \in \{1, 2\}$. I normalize the marginal cost of the product to zero. Therefore, the advertiser's margin per conversion in Period t is p_t.

---

[12]In Section 4.4.3, I a consider a general setting wherein consumers' purchase histories are observed *imperfectly*, and demonstrate that the main insights are preserved.

[13]For example, Facebook allows advertisers to target consumers "who engaged with any post or ad" or "clicked any call-to-action button" (`https://www.facebook.com/business/help/221146184973131?helpref=page_content`)

Ad Network

The ad network sells ad impressions to advertisers via second-price auctions.[14] It sets reserve price R_t^j in Period $t \in \{1, 2\}$, where j indexes the type of ad impression (e.g., ads targeted to M-consumers or untargeted ads for opt-out consumers). In my paper, this is equivalent to selling through a posted price; however, I choose the auction format to be consistent with how a vast majority of display ads are sold in the market.[15] The ad network maximizes its total profit across two periods, which consists of revenue and cost from ad impression sales. Costs may include operational costs associated with ad inventory management, as well as maintenance costs related to setting up ad auctions and delivering ads.[16] I denote this per-impression cost by $k \geq 0$.

Game Timing

The timing of the game is as follows.

---

[14]Note that, due to the revenue equivalence principle, my results would not change if I considered first-price auctions, a mechanism to which some firms have recently transitioned (e.g., see https://support.google.com/admanager/answer/9298211).

[15]In 2019, over 83.5% of total display ad spend in US were transacted through real-time auctions (eMarketer, 2019b).

[16]This model feature is motivated from my conversations with industry practitioners. In particular, they have indicated that selling ad inventory entails very significant operational costs associated with, for example, (i) storing, retrieving and relaying data to advertisers, and (ii) resolving the auction and announcing the outcome to all the bidders.

Period 0: Consumers decide whether or not to opt-in to being tracked.

Period 1: Ad network sets reserve prices for ads for opt-in consumers and ads for opt-out consumers. Advertiser sets product price and bids for ad impressions.

- If ads are shown, some consumers transition through funnel.

- Consumers make purchase decisions.

Period 2: Ad network sets reserve prices for targeted ads for opt-in consumers, and untargeted ads for opt-out consumers. Advertiser sets product price and bids for ad impressions.

- If ads are shown, some consumers transition through funnel.

- Consumers make purchase decisions.

I solve for the subgame-perfect equilibrium of the above game.

Before I proceed to the analysis, I note that, in Section 4.4, I consider a number of extensions to address several simplifying assumptions that I have made in the main model. I analyze scenarios in which (i) the advertiser has private information about its ad valuations, (ii) there are multiple, competing advertisers, (iii) the purchase histories of opt-in consumers are not perfectly observable, and (iv) consumers arrive in overlapping generations across an infinite time horizon. I also note that the publishers that own the content pages are treated as passive in the model. However, insofar as the ad network and the publishers share the same objective function of maximizing monetization by showing ads and they split these revenues on a commission basis (which is typically the case[17]), this is a reasonable assumption.

4.3 Analysis

First, note that the advertiser's product pricing decision is trivial and the optimal product price is always 1, which is the consumer's product utility on obtaining a match. Thus, the advertiser's

---

[17]For example, see https://support.google.com/adsense/answer/180195 and https://www.adpushup.com/blog/the-best-ad-networks-for-publishers/.

margin per conversion is 1. Intuitively, if $p_t < 1$, then the advertiser leaves money on the table, and if $p_t > 1$, then no products are sold. Since $p_t^* = 1$ for $t \in \{1, 2\}$, the consumer purchases if and only if she realizes a match. This implies that the consumer's utility from product consumption is always zero whether or not she purchases. Therefore, when discussing consumer utility, I hereafter restrict attention to the privacy utility component.

To develop basic insights, I study the case of no tracking in Section 4.3.1 and full tracking in Section 4.3.2. Then I discuss the main analysis with endogenous consumer tracking choice in Section 4.3.3.

4.3.1 No Tracking

In this section, I analyze the case in which consumers cannot be tracked; i.e., advertisers cannot distinguish consumers' funnel states nor their purchase histories. My objective is to establish the baseline forces that determine the equilibrium advertising outcomes in the absence of consumer tracking. To solve for subgame perfect Nash equilibrium, I first analyze the advertiser's bidding problem in Period 2 and then proceed backwards. I assume that the advertiser plays weakly dominant bids, in the sense that the bidding strategies are not affected by "trembles" in various auction parameters, such as reserve prices and number of bidders.

Period 2

In Period 2, there are two possible subgames: one in which ads were shown in Period 1, and another in which they were not. I index the former Period 2 subgame with the subscript "2|ad" and the latter with "2|no ad." Consider the first subgame, in which ads were previously shown. The Period 2 distribution of consumers along the funnel can be characterized by three groups: (i) those who were not impacted by the first ad and remained in T, (ii) those who saw the ad, transitioned to M, and purchased, and (iii) those who transitioned to M but did not purchase. Given that the first ad induces interest w.p. μ, the first group is of size $1 - \mu$. Since M-consumers realize a product match w.p. ϕ_M, the second group is of size $\mu \phi_M$. Finally, M-consumers do not purchase if they do

not realize a product match; therefore, the third group is of size $\mu(1 - \phi_M)$. While the advertiser knows this Period 2 distribution, it cannot identify which consumer belongs to which group in the absence of tracking.

To compute the advertiser's weakly dominant bid for the Period 2 untargeted ad, the advertiser compares its payoff when it wins vs. loses the ad auction. Let R_t denote the reserve price of untargeted ads in Period t.[18] If the advertiser bids b_2 in Period 2, its payoff is

$$
\pi^A_{2|\text{ad}}(b_2) = \begin{cases} (1 - \mu)\mu\phi_M + \mu(1 - \phi_M)(\beta + (1 - \beta)\phi_M) - R_2 & \text{if } b_2 \geq R_2, \\ \mu(1 - \phi_M)\phi_M & \text{if } b_2 < R_2. \end{cases} \tag{4.3.1}
$$

Consider the advertiser's payoff from winning the auction and displaying the ad, shown on the top row of (4.3.1). The first term denotes the conversion of T-consumers induced by Period 2 advertising: of the $1 - \mu$ fraction of consumers who had not been affected by the Period 1 ad, μ fraction transition to funnel state M, of which ϕ_M fraction realize a match and purchase. Similarly, the second term denotes the conversion of M-consumers who had not converted in Period 1.

Note that if the advertiser bids below the reserve price and loses the auction, then its payoff is not 0 but $\mu(1 - \phi_M)\phi_M$, as shown on the bottom row of (4.3.1). This is because even if no additional ads are shown in Period 2, the non-purchasers in funnel state M — who were pushed down from funnel state T after seeing the Period 1 ad — may realize a product match in Period 2 w.p. ϕ_M and purchase.

The payoffs of the second subgame in which ads were not shown in Period 1 can be analyzed in a similar manner. The following lemma states the subgame outcomes in Period 2.

Lemma 5 (Period 2 Bid and Reserve Price Without Tracking)**.**

- *Suppose the advertiser showed ads in Period 1. The advertiser's weakly dominant bid in Period 2 is $b^*_{2|\text{ad}} = (1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta$, and the ad network's optimal reserve price is $R^*_{2|\text{ad}} = \max\left[k, b^*_{2|\text{ad}}\right]$.*

---

[18]For ease of exposition, I suppress the ad type index j for the reserve price as only one type of ads (i.e., untargeted ads) is offered in the absence of tracking.

- *Suppose the advertiser did not show ads in Period 1. The advertiser's weakly dominant bid in Period 2 is $b^*_{2|no\ ad} = \mu\phi_M$, and the ad network's optimal reserve price is $R^*_{2|no\ ad} = \max\left[k, b^*_{2|no\ ad}\right]$.*

The first part of Lemma 5 provides important preliminary insights into the conditions under which the advertiser buys *successive* ads in Period 2, conditional on having shown ads in Period 1. The advertiser buys successive ads if and only if $b^*_{2|ad} \geq R^*_{2|ad}$, which simplifies to

$$(1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta \geq k. \tag{4.3.2}$$

That is, the marginal effectiveness of the successive ad, expressed on the left-hand side of (4.3.2), must be sufficiently large. Analyzing how this object changes with respect to the model primitives reveals two key determinants of a successive ad's marginal effectiveness.

The marginal effectiveness of the successive ad consists of two components: the marginal conversion of T-consumers (denoted by $(1 - \mu)\mu\phi_M$) and the marginal conversion of M-consumers (denoted by $\mu(1 - \phi_M)^2\beta$). It can be shown that $(1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta$ decreases with respect to μ, the probability of an ad inducing interest, if and only if $\mu > \frac{\phi_M + \beta(1-\phi_M)^2}{2\phi_M}$. Moreover, the marginal effectiveness of the successive ad decreases with respect to ϕ_M, the product match probability of consumers in funnel state M, if and only if β, the probability of an ad inducing action, is greater than $\frac{1-\mu}{2(1-\phi_M)}$. The intuition for the first case is as follows. If μ is large, then the first ad exposure causes the Period 2 distribution of consumers to shift toward M. This implies a diminished role of successive ads in pushing T-consumers down to M in Period 2. Therefore, the marginal effectiveness of successive ads decreases in μ for large μ.

Consider the second case. If β is large, then the marginal effectiveness of successive ads is largely determined by their potential to convert M-consumers. Now, increasing ϕ_M has two effects. First, consumers are more likely to purchase after the first ad exposure such that there is a small segment of non-purchasers in funnel state M-consumers in Period 2. Second, if ϕ_M is large, those non-purchasers in funnel state M are likely to convert on their own without a successive ad

exposure. Thus, increasing ϕ_M dampens the value of a successive ad.

Taken together, I see that an effective ad in Period 1 (i.e., large μ and ϕ_M) may diminish the marginal effectiveness of successive advertising in Period 2. I call this the *saturation effect*. It is visualized by the concave ad response curve for large ϕ_M in Figure 4.2.

Period 1

The reserve prices $R^*_{2|\cdot}$ from Lemma 5 imply that the advertiser's Period 2 payoff is $\mu(1 - \phi_M)\phi_M$ if the advertiser shows ads in Period 1, and 0 otherwise. Taking this into account, the advertiser's problem in Period 1 is to determine the bid b_1 that maximizes

$$
\pi^A_1(b_1) = \begin{cases} \mu\phi_M - R_1 + \mu(1 - \phi_M)\phi_M & \text{if } b_1 \geq R_1, \\ 0 & \text{if } b_1 < R_1, \end{cases}
$$

where R_1 is the reserve price for untargeted ads in Period 1. The following lemma states the advertiser's weakly dominant bid and the ad network's optimal reserve price.

Lemma 6 (Period 1 Bid and Reserve Price Without Tracking)**.** *Let* $x^+ \equiv \max[x, 0]$*. The advertiser's weakly dominant bid in Period 1 is* $b^*_1 = \mu(2 - \phi_M)\phi_M$*, and the ad network's optimal reserve price is*

$$
R^*_1 = \max\left[k + (\mu\phi_M - k)^+ - \left(b^*_{2|ad} - k\right)^+, b^*_1\right]. \tag{4.3.3}
$$

I see from (4.3.3) that the ad network sometimes sets the Period 1 reserve price below the marginal cost k, even if that implies the ad network earns a negative payoff in Period 1. This occurs when showing successive untargeted ads in Period 2 is highly valuable for the advertiser; i.e., when $b^*_{2|ad} = (1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta$ is high. Intuitively, by setting a low reserve price in Period 1, the ad network helps the advertiser display ads, thereby creating an opportunity to extract greater surplus from the advertiser in Period 2. This ad pricing strategy can be viewed as the ad network capitalizing on the *convexity* of the ad response curve.

Advertising Strategy Without Tracking

Given the equilibrium reserve prices and bids, I now characterize the conditions under which the advertiser buys ads in (i) both periods, (ii) only Period 1, and (iii) neither period. The following proposition summarizes the equilibrium advertising strategy across two periods.

Proposition 18 (Advertising Without Tracking). *Suppose the advertiser cannot track consumers. For thresholds $\tilde{\beta}, \overline{\beta}, \underline{\beta}, \overline{\mu}$, and $\underline{\mu}$ defined in the proof, the equilibrium advertising strategy is as follows.*

- *Suppose $\phi_M < 1 - \sqrt{(\mu - k)^+/\mu}$. The advertiser buys ads in both periods if $\beta \geq \tilde{\beta}$, and does not buy any ads otherwise.*

- *Suppose $\phi_M \geq 1 - \sqrt{(\mu - k)^+/\mu}$.*

  - *if (i) $\beta \geq \overline{\beta}$ and $\mu \geq \underline{\mu}$, or (ii) $\underline{\beta} \leq \beta < \overline{\beta}$ and $\underline{\mu} \leq \mu < \overline{\mu}$, then advertiser buys ads in both periods;*

  - *if (i) $\beta < \underline{\beta}$ or $\mu \leq \underline{\mu}$, or (ii) $\underline{\beta} \leq \beta < \overline{\beta}$ and $\mu \geq \overline{\mu}$, then advertiser buys ads only in Period 1.*

Consider the case of small ϕ_M, depicted in Figure 4.3a, where the advertiser either buys ads in neither period or in both periods. This "all-or-nothing" pattern emerges when the ad response curve is convex. Specifically, if ϕ_M is small, the first ad exposure does little in terms of increasing the conversion probability. Thus, the advertiser does not find it worthwhile to advertise *only* in Period 1. However, if β is sufficiently large, a successive ad is highly likely to bring M-consumers in Period 2 down to B and induce purchase. This increase in effectiveness of successive ads compensates for the low effectiveness of the ads shown in Period 1. Thus, if ϕ_M is small, the advertiser either buys ads in neither period or in both.

In contrast, if ϕ_M is large, the total ad intensity across Periods 1 and 2 is more nuanced. In particular, the total ad intensity may be non-monotonic in μ (see Figure 4.3b). To understand the changes in ad intensity as μ increases, consider the cross-section of the plot represented by the large
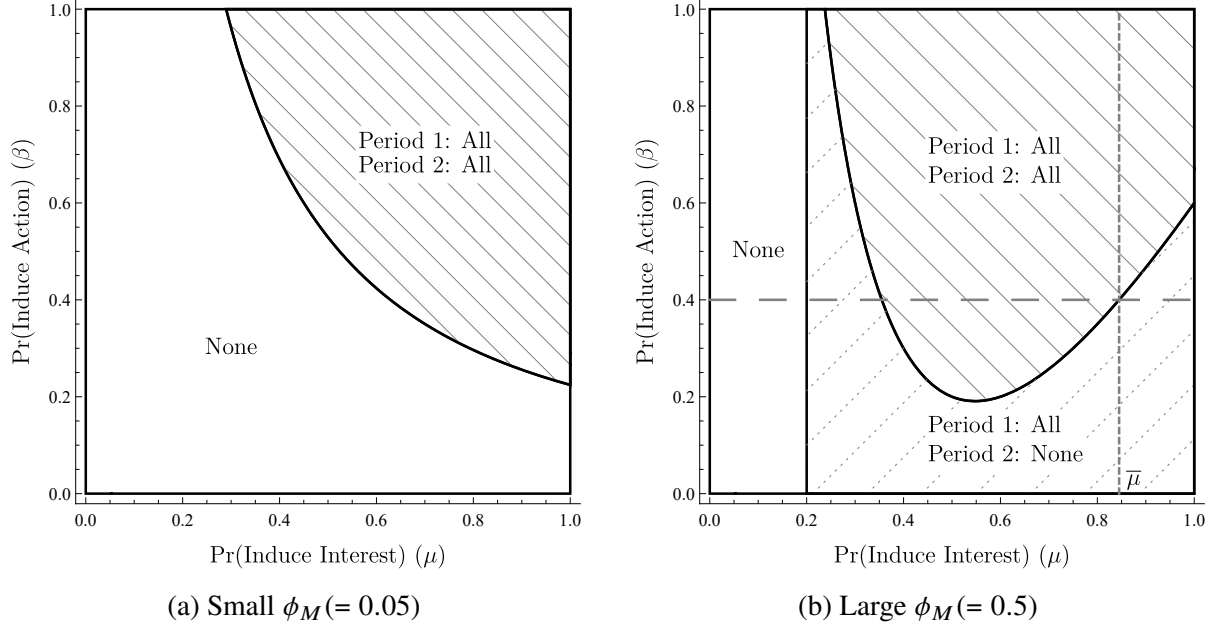
Figure 4.3: Ad Audiences with No Tracking; $k = 0.15$

dashed line in Figure 4.3b for fixed $\beta = 0.4$. Along this line, as μ increases from 0 to $\overline{\mu}$ (denoted by the small dashed line), the ad intensity increases from 0 to 1 to 2 due to increasing effectiveness of ads. Past the $\overline{\mu}$ threshold, however, observe that Period 2 ads are foregone. This is because the combination of large ϕ_M and μ implies a high purchase conversion after Period 1 ads are shown. This diminishes the value of successive ads (i.e., Period 1 ads saturate) and consequently the intensity of ads shown in Period 2 is reduced.

In total, these results highlight the significance of considering the purchase journey in the analysis of advertising strategies, *even when* there is no trackability. In particular, modeling the funnel sheds light on how ads may influence consumer distribution along different funnel states. This distribution determines the marginal effectiveness of successive ads, which in turn affects ad buying decisions across time.

4.3.2   Full Tracking

I now analyze how the ability to track consumers affects the advertiser's strategy and the ad network's profit.

93

Advertising Strategy With Tracking

Without tracking, the advertiser was restricted to buying untargeted ads. With consumer track-ing, however, the advertiser can target ads along two dimensions — consumers' positions in the purchase funnel and their product purchase histories. Specifically, in Period 2, it can target (i) T-consumers and (ii) M-consumers who did not purchase in Period 1 . In the Period 2 subgame for which ads were shown in Period 1, the advertiser's bidding strategy is a pair of bids $\left(b_2^T, b_2^M\right)$ for ad impressions associated with consumer segments (i) and (ii). Given reserve prices R_2^T and R_2^M for segments (i) and (ii), respectively, the advertiser's expected payoff in Period 2 is

$$\pi_{2|\text{ad}}^A \left(b_2^T, b_2^M\right) = \begin{cases} (1-\mu)\left(\mu\phi_M - R_2^T\right) + \mu(1-\phi_M)\left(\beta + (1-\beta)\phi_M - R_2^M\right) & \text{if } b_2^T \geq R_2^T, b_2^M \geq R_2^M, \\ (1-\mu)(\mu\phi_M - R_2^T) + \mu(1-\phi_M)\phi_M & \text{if } b_2^T \geq R_2^T, b_2^M < R_2^M, \\ \mu(1-\phi_M)(\beta + (1-\beta)\phi_M - R_2^M) & \text{if } b_2^T < R_2^T, b_2^M \geq R_2^M, \\ \mu(1-\phi_M)\phi_M & \text{if } b_2^T < R_2^T, b_2^M < R_2^M. \end{cases}$$

The logic of solving for an equilibrium is similar to the case without no tracking, although more complex. Due to space considerations I relegate the full backwards induction analysis to the appendix. The following proposition characterizes the advertiser's strategies in the presence of consumer tracking.

Proposition 19 (Advertising With Tracking). *Suppose the advertiser can track consumers along the purchase funnel. Let $\tilde{\mu} = k \left(\phi_M(2 - \phi_M) + (1 - \phi_M)(\beta(1 - \phi_M) - k)^+\right)^{-1}$. The advertiser's equilibrium advertising strategy is as follows:*

- *if $\mu \leq \tilde{\mu}$, then do not buy any ads in either period;*

- *if $\mu > \tilde{\mu}$, then show ads to all consumers in Period 1. Furthermore,*

    - *if $\mu > \frac{k}{\phi_M}$, then in Period 2, buy ads targeted to T-consumers, and*

    - *if $\beta > \frac{k}{1-\phi_M}$, then in Period 2, buy ads targeted to M-consumers who did not purchase.*
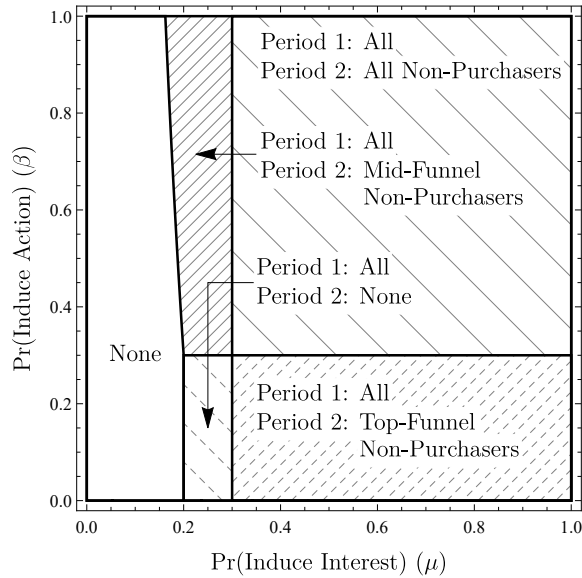
Figure 4.4: Ad Audiences With Tracking; $\phi_M = 0.5$, $k = 0.15$

Figure 4.4 depicts the advertising strategies in the presence of consumer tracking. Before I discuss this figure in detail, I note that a comparison with Figure 4.3 reveals that the advertising strategy with tracking is significantly different from the advertising strategy without tracking. Proposition 19 shows that when the probability of the ad inducing interest is high, the advertiser adopts a *reach strategy*, whereby it targets successive ads to T-consumers who are not considering purchase. When the probability of the ad increasing purchase propensity is high, then the advertiser adopts a *frequency strategy*, whereby successive ads are shown to M-consumers who are already considering purchase.

Overall, Proposition 19 suggests that advertisers should be cognizant of the nuanced ad effects in relation to the consumers' journey down the funnel. As illustrated in Figure 4.4, the combination of consumer trackability and funnel considerations gives rise to various conditions under which one variant of advertising strategy is more profitable than another (e.g., reach vs. frequency).

Comparing No Tracking and Tracking Outcomes

A question of central interest that I can answer using this model is: how does consumer tracking impact overall advertising intensity? My results show that the effect of tracking on ad intensity is
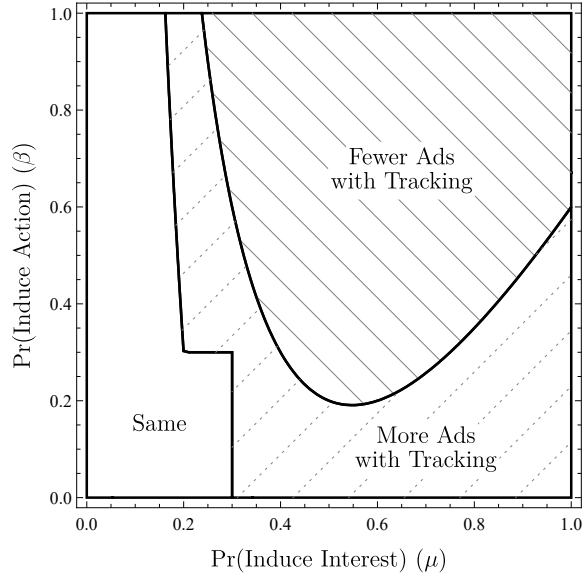
Figure 4.5: Ad Intensity With and Without Tracking; $\phi_M = 0.5$, $k = 0.15$

nuanced (see Figure 4.5). If the funnel transition probabilities μ and β are small, then ad effectiveness is so low that no ads are shown to any consumers regardless of consumer trackability; hence, tracking may not change ad intensity. Otherwise, if either μ or β is sufficiently large, tracking may either increase or decrease the total intensity of ads.

When ad effectiveness is low, the average effectiveness of untargeted ads is low. With trackability, the advertiser can identify and bid for high-valuation impressions such that more ads are shown with tracking. On the other hand, consider the case when ad effectiveness is intermediate. Without tracking, untargeted ads are shown in both periods. However, some of the ads in Period 2 are wasted because they are shown to consumers who already purchased, which does not happen with tracking. Thus, tracking allows the advertiser to reduce spending on wasteful ad impressions, resulting in lower ad intensity with tracking than without.

While the above findings resonate with those from Esteban et al. (2001) and Iyer et al. (2005), my results are different from these papers at high levels of ad effectiveness. Specifically, when ad effectiveness is high, I find that the ad intensity differential *reverses*: more ads are shown under tracking than without. The intuition is that without tracking, high ad effectiveness dampens the value of successive ads, such that only first period ads are shown. On the other hand, when

consumers can be tracked along the purchase journey, ads are targeted to (i) consumers who were not impacted by the first ad and stayed in top-funnel and/or (ii) consumers who moved down to mid-funnel but did not purchase. Therefore, more ads are shown under tracking than without. Put differently, the interdependence between the marginal effectiveness of cross-period ads stemming from funnel considerations reverses the ad intensity differential for high levels of ad effectiveness.

In sum, I establish a non-monotonic relationship between the impact of tracking on ad intensities and ad effectiveness, which is proxied by the funnel transition parameter μ when the match parameter ϕ_M is fixed. I summarize these results in the following proposition.

Proposition 20 (Ad Intensity). *Consumer funnel tracking either increases or decreases the total ad intensity compared to the no tracking case. Specifically, for thresholds $\tilde{\beta}, \overline{\beta}, \underline{\beta}, \overline{\mu}$, and $\underline{\mu}$ defined in the proof,*

- *if (i) $\mu < \tilde{\mu}$ and $\beta < \tilde{\beta}$ or (ii) $\frac{k}{\phi_M(2-\phi_M)} < \mu \leq \frac{k}{\phi_M}$ and $\beta \leq \frac{k}{1-\phi_M}$, then the ad intensities are the same;*

- *if $\phi_M > 1 - \sqrt{(\mu-k)^+/\mu}$ and either (i) $\beta > \overline{\beta}$ and $\mu > \underline{\mu}$, or (ii) $\underline{\beta} \leq \beta < \overline{\beta}$ and $\underline{\mu} \leq \mu < \overline{\mu}$, then tracking reduces ad intensity;*

- *otherwise, tracking increases ad intensity.*

Tracking and Ad Network Profit

How does consumer tracking impact the ad network's profit? I find that tracking weakly increases the ad network's profit. Intuitively, consumer tracking endows the advertiser with more information on which the advertiser can condition its bid. In this case, the ad network can selectively supply ad impressions most highly valued by the advertiser while foregoing unprofitable ones, thereby raising its profit compared to the regime without consumer tracking. I state this result as a proposition.

Proposition 21 (Consumer Tracking and Ad Network Profit). *Consumer funnel tracking weakly increases the ad network's profit.*

It is important to note that the tracking-induced improvement in the ad network's profit hinges on the assumption that the ad network is as knowledgeable about ad valuations as the advertiser. In Section 4.4.1, I show that the result of Proposition 21 does not always carry over to a setting where the advertiser has private information about ad valuations. Surprisingly, information asymmetry between the ad network and the advertiser may result in consumer tracking *lowering* the ad network's profit.

4.3.3 Endogenous Tracking Choice

In the preceding analysis, I examined two distinct cases in which the advertiser was either not able to track any consumers or able to track all consumers. In this section, I investigate the impact of endowing consumers with the choice to be tracked. That is, I analyze how consumers exercise their right to choose whether to allow tracking or not and how this decision affects the ad ecosystem. As discussed in the introduction, the analysis is largely motivated by the recent enactment of data privacy regulations that mandate affirmative consumer consent prior to acquiring and processing consumer data.

Consumers' Opt-In Behaviors

I first characterize the consumers' equilibrium privacy choices given the advertising outcomes under tracking and no tracking. Recall from the consumer privacy utility formulation in (4.2.1) that consumers dislike seeing ads (instrumental aspect) and also dislike being tracked (intrinsic aspect). This implies that consumers will choose to incur the privacy cost from opting-in to being tracked *only if* they expect to see fewer ads from doing so; i.e., only if the positive instrumental value outweighs the intrinsic cost of privacy. The following proposition summarizes the consumer opt-in behavior.

Proposition 22 (Consumer Opt-In Behavior). *Let $q(0)$ and $q(1)$ denote the total number of ads consumers are exposed to when they opt-out and opt-in, respectively. The proportion of consumers who opt-in to being tracked can be non-monotonic in the funnel transition probability μ. In par-*
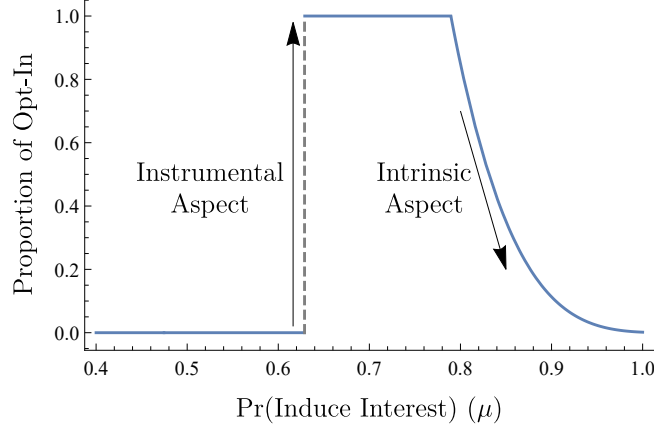
Figure 4.6: Proportion of Opt-In Consumers; $\phi_M = 0.05, \beta = 0.4, F(\theta) = \theta^4, \eta = 4, k = 0.15$

ticular, for thresholds $\overline{\beta}, \underline{\beta}, \overline{\mu}$, and $\underline{\mu}$ defined in the proof, if either (i) $\underline{\beta} < \beta \leq \overline{\beta}$ and $\underline{\mu} < \mu < \overline{\mu}$, or (ii) $\beta > \overline{\beta}$ and $\mu > \underline{\mu}$, then $F(\eta(q(0) - q(1)))$ consumers opt-in; otherwise, all consumers opt-out.

The consumer opt-in pattern is driven by two forces. The first force stems from the changes in number of ads shown to opt-out consumers in Period 2. Recall that in the absence of tracking, for small ϕ_M, as μ increases, the advertiser's strategy sometimes changes from not advertising in either period to any consumer to advertising in both periods to all consumers (see Proposition 18). This pattern emerges from the convexity of the ad response curve: while advertising only once is never profitable, showing successive ads might be. Thus, as consumers expect high ad intensity for large μ when they cannot be tracked, consumers are incentivized to opt-in to being tracked in order to see fewer ads. Figure 4.6 depicts the increase in opt-in rate as μ increases past the threshold $\mu \approx 0.63$, after which the advertiser shows ads to all consumers in both periods under no tracking.

The second force relates to the advertiser's targeting regime for opt-in consumers in Period 2. To illustrate, suppose μ and β are large. In this case, the advertiser adopts the frequency strategy in Period 2 such that successive ads are targeted to M-consumers (see Proposition 19). Now, as the probability of an ad inducing interest increases, consumers are more likely to transition to funnel state M after the first ad exposure, and hence become targets of Period 2 advertising. This dampens consumers' incentives to opting-in to being tracked. Figure 4.6 shows the associated
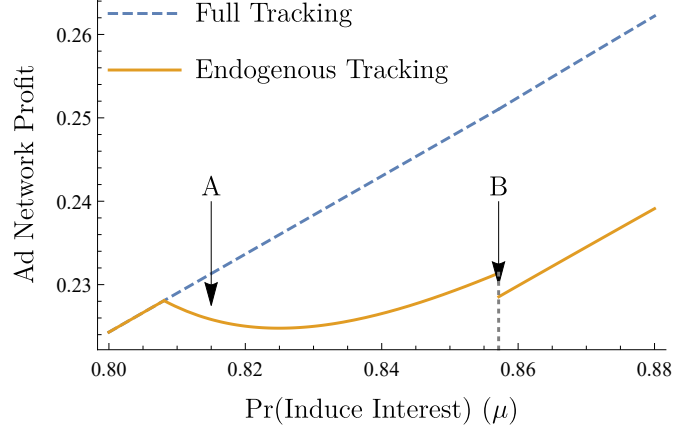
Figure 4.7: Ad Network Profit; $\phi_M = 0.175, \beta = 0.4, \eta = 3, k = 0.15, F(\theta) = \theta^{16}$

decline in opt-in rate as μ increases in the neighborhood of $\mu \approx 0.85$.

Ad Network Profit

Next, I analyze how the consumers' opt-in behaviors characterized above impact the ad network's profit. Interestingly, I find that consumers' opt-in choices lead to non-monotonicities in the ad network's profit with respect to μ. In particular, under certain conditions, the ad network's profit *decreases* in μ, even though larger μ implies higher purchase conversion on average. To understand this, recall from Proposition 21 that the ad network's profit is (weakly) lower under no tracking than under tracking. Since larger μ may result in more consumers opting-out from being tracked (see Proposition 22), it follows that the ad network's profit may decrease in μ.

As described above, the change in the number of opt-in consumers can arise from two distinct forces. First, even if consumers expect to see fewer ads under tracking and thus opt-in to being tracked, if the ad intensity under tracking increases with μ, then less consumers choose to opt-in as μ increases. This decline in opt-in rate induces a continuous decrease in the ad network's profit, as illustrated in the region marked *A* in Figure 4.7. Second, consumers also consider the instrumental aspect of privacy: if consumers expect to see fewer ads without tracking, no consumer chooses to opt-in. This leads to discrete jump in the ad network's profit, as shown in the region marked *B* in Figure 4.7. The following proposition summarizes this finding.

100

Proposition 23 (Equilibrium Ad Network Profit). *Let $q^*(0)$ and $q^*(1)$ denote the equilibrium ad intensity without and with tracking, respectively, and let $\underline{\mu}'$ and $\overline{\mu}'$ be as defined in the proof. Suppose consumers' privacy costs θ are uniformly distributed on $[0, 1]$. Under endogenous tracking, the ad network's profit decreases in μ if and only if either*

- *$q^*(0) = 2$, $q^*(1) = 1 + \mu(1 - \phi_M)$, $\eta(1 - \mu(1 - \phi_M)) < 1$, and either $\mu < \underline{\mu}'$ or $\mu \geq \overline{\mu}'$, or*

- *$\phi_M \geq 1 - \sqrt{(\mu - k)^+/\mu}$ and $\underline{\beta} \leq \beta < \overline{\beta}$.*

What does this mean for the ad network? Conventional wisdom suggests that the ad network would be better off if ads were more effective: ads that yield high purchase conversion are associated with high valuations, which allows the ad network to sell ad slots at higher prices. Proposition 23 provides a countervailing argument. Privacy regulations that allow consumers to choose between being tracked or not may result in more consumers opting-out from being tracked for higher levels of ad effectiveness, in particular if higher ad effectiveness implies more ads being shown to opt-in consumers. In this case, consumers choose to opt-out from tracking, thereby undermining targeting efficiency. This means that ad slots may be sold at lower prices, lowering the ad network's profit.

Consumer Surplus

One of the main objectives of privacy regulations is to protect consumers. Consistent with intuition, I find that giving consumers the choice to being tracked weakly improves consumer surplus, compared to the full tracking benchmark. Intuitively, the regulations allow consumers to make privacy decisions such that their individual surplus is maximized. And since their decision does not impose externalities on other consumers, net consumer surplus weakly increases. I state this as a proposition.

Proposition 24 (Consumer Surplus). *Privacy regulations that allow consumers to choose whether to be tracked or not increase overall consumer surplus compared to the full tracking case.*

In sum, my analysis provides three important takeaways. First, consumers sometimes choose to opt-in to being tracked. Specifically, this happens if the effectiveness of the ads in inducing product interest, μ, is intermediate, where consumers trade-off the benefit of seeing fewer ads (positive instrumental value of privacy) with the disutility of compromising their privacy (intrinsic cost of privacy). Second, the ad network's profit when consumers can choose to be tracked or not may decrease in μ, even if larger μ implies higher purchase conversion. The intuition is that more consumers may choose to opt-out from being tracked as μ increases; this lowers targeting efficiency, which ultimately reduces the ad network's profit. Finally, consumer surplus always increases and the ad network's profit always decreases when consumers have the choice of being tracked or not. As I show in the next section, however, allowing the advertiser to have private information about their ad valuations may reverse the latter result. That is, the ad network's profit may be higher under endogenous tracking than under full tracking.

4.4 Extensions

In this section, I explore fmy extensions that demonstrate the robustness of the qualitative insights obtained from the main model. I also describe some additional insights that emerge from relaxing the assumptions from the main model.

4.4.1 Information Asymmetry about Ad Valuation

In the main model, I assumed that the advertiser's ad valuation is known by the ad network. Consequently, the ad network sets the reserve price such that the advertiser's surplus is extracted fully. In this information asymmetry extension, I allow the advertiser's ad valuation to be private information. In practice, ad valuations may not be fully known to the ad network for several reasons. First, the ad network may not perfectly observe all of the consumers' interactions with the advertiser (e.g., offline interactions in the advertiser's physical store) that may inform the advertiser about consumer valuation. Second, even if the ad network possessed similar same levels of information as the advertiser, it may have less ability to infer consumer's willingness to pay for an
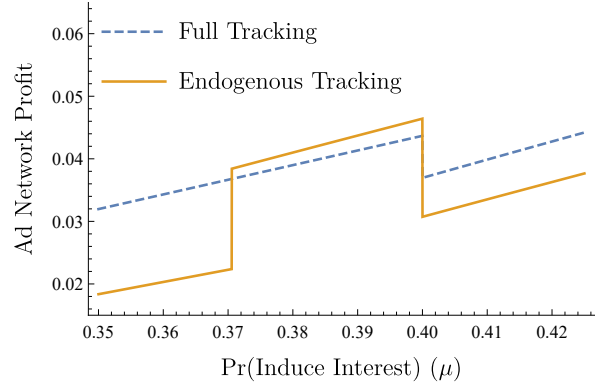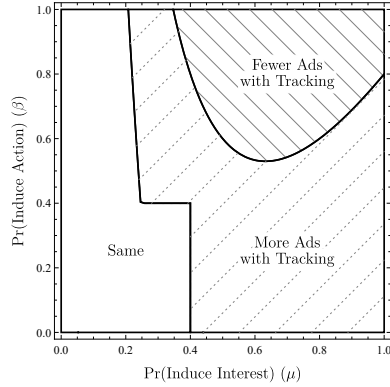
(a) Ad Intensity Differential; $\phi_M = 0.5, k = 0.05$ (b) Ad Network Profit; $\phi_M = \eta = 0.5, \beta = 0.9, k = 0.05, F(\theta) = \theta$

Figure 4.8: Ad Intensity and Ad Network Profit Under Information Asymmetry

advertiser's product (de Cornière and Nijs, 2016).

To that end, I assume that in each period, the advertiser's ad valuation is drawn independently from Uniform[0, 1] and is known privately to the advertiser. In contrast to the main model, in this case the advertiser earns positive surplus. This adds interesting dynamics to the model because in Period 1 the advertiser will anticipate how the Period 1 outcome affects its Period 2 payoff and will bid accordingly.

While the patterns of ad intensity differential between the tracking and no tracking cases are qualitatively unaffected by the advertiser's bidding dynamics (see Figure 4.8a), the result pertaining to the ad network's profit is sometimes reversed. In particular, I find that under certain conditions, the ad network may benefit from regulations that allow consumers to endogenously choose to being tracked (see Figure 4.8b). This occurs when β is sufficiently high and μ intermediate.

The intuition is as follows. With tracking, the ad network has the option to sell select ad impressions targeted to a subgroup of consumers (e.g., M-consumers who did not purchase in Period 1). While such selective ad sales help the ad network to efficiently extract surplus from the advertiser in Period 2, they hurt the advertiser by limiting the size of consumer segments reached. However, as privacy regulations induce some privacy-conscious consumers to opt-out from tracking, the ad network's targetability is reduced. Thus, instead of selling targeted ads, the ad network sells untargeted ads that reach a larger consumer base. I call this the supply-side

"market thickening" effect. Untargeted ads that reach more consumers are more profitable for the advertiser than, say, ads targeted to M-consumers. Thus, in anticipation of higher Period 2 payoffs for opt-out consumer segments, the advertiser bids more aggressively in Period 1 under endogenous tracking than under full tracking. This ultimately leads to higher ad network profit.

The result that the ad network may benefit from consumers opting-out of tracking also informs the policy debate regarding the default privacy choice. Specifically, it suggests that if there is a sufficiently large segment of naive consumers who simply stick with the pre-selected, default privacy settings[19], then the ad network's profit would increase as a result of regulations that mandate an opt-out default. Thus, my model provides a rational explanation, besides the more common "market share" argument[20], for why large ad platforms such as Facebook have been enthusiastically vouching for stringent privacy regulations.[21]

In sum, my analysis sheds light on a novel role of privacy regulations that allow endogenous tracking choices: regulations can serve as a commitment device for the ad network to sell more ad impressions.[22] This in turn better aligns the incentives of the ad network and the advertiser. Under certain conditions, privacy regulations can lead to higher profits for both parties compared to the full tracking benchmark.

4.4.2 Competing Advertisers

I extend the main model by considering two competing advertisers indexed by $i \in \{1, 2\}$. Consumers are heterogeneous in their product preferences: λ proportion of consumers are "loyal" to Advertiser 1, and $1 - \lambda$ proportion to Advertiser 2. A consumer transitions down the purchase funnel according to specifications of the main model (see Section 4.2) only if she sees an ad from the advertiser to which she is loyal. Without loss of generality, I assume that $\lambda > \frac{1}{2}$; i.e.,

---

[19]According to The Harris Poll (2019), 14% of US Internet user respondents indicated that they "have not taken any action" to protect their personal information.

[20]https://www.wsj.com/articles/gdpr-has-been-a-boon-for-google-and-facebook-11560789219

[21]https://www.ft.com/content/0ca8466c-d768-11e8-ab8e-6be0dcf18713

[22]Without such a commitment device, the advertiser would anticipate the ad network to sell only select profitable impressions in Period 2. Thus, the advertiser would expect to gain little surplus in Period 2, such that the ad network would not be able to charge high prices for Period 1 ads.
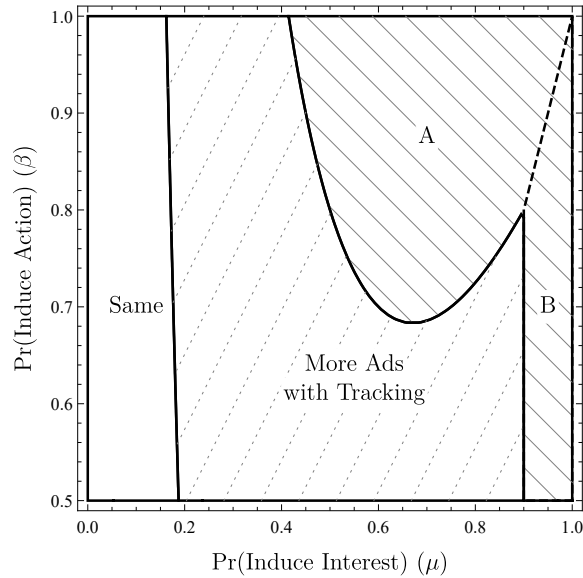
Figure 4.9: Competing Advertisers: Ad Intensity With and Without Tracking; $\phi_M = 0.5, k = 0.15, \lambda = 0.66$

Advertiser 1 is the dominant brand.[23]

I assume that opting-in to tracking reveals not only the consumers' funnel states and purchase histories, but also their *ex ante* product preferences; i.e, whether consumers are loyal to Advertiser 1 or 2. For example, tracking consumers' past visits to and browsing patterns within advertisers' websites may help advertisers infer consumers' product preferences.[24] The rest of the model specifications remain unchanged. Note that the extension model reduces to the main model if $\lambda = 1$.

I find that the qualitative insights from the main model carry over for a large range of parameters (see Figure 4.9). The results diverge if and only if either (a) the effectiveness of ads shown to T-consumers is high (i.e., large μ) or (b) product preference heterogeneity is sufficiently large (i.e., λ close to 0.5). The intuition is the following. Advertiser 2 has high incentive to advertise in Period 2 after Advertiser 1's ads has been shown in Period 1 if Advertiser 2 knows either (a) that its first ad exposure in Period 2 will be highly effective, or (b) that there is a large group of loyal consumers

[23] In the case of symmetry $\left(\lambda = \frac{1}{2}\right)$, the only difference is that the first period outcome is randomly determined between the two advertisers; otherwise, the logic of the analysis is equivalent to the asymmetry case.

[24] https://www.digitaltrends.com/computing/how-do-advertisers-track-you-online-we-found-out/

105

who will respond to Advertiser 2's ad. While in the main model ad slots would have been left unfilled in Period 2 due to the saturation effect, in this case the ad slots are filled by Advertiser 2 (Region B in Figure 4.9). Finally, to avoid seeing ads in both periods, consumers opt-in to tracking for this parameter range. I summarize this finding in the following proposition.

Proposition 25 (Competing Advertisers). *The advertising intensity differential between tracking and no tracking regimes carries over from the main model if either μ is not too large or λ is not too small. Otherwise, opt-out consumers are exposed to ads in both periods: from Advertiser 1 in Period 1, and from Advertiser 2 in Period 2.*

4.4.3 Imperfect Observability of Purchase History

In the main model, I assumed that for an opt-in consumer, both her funnel state $f \in \{T, M, B\}$ and her purchase history were perfectly observable by the ad network. While this assumption helped us deliver the main insights clearly, it may not always reflect the information flow in practice. For example, firms may not be able to perfectly merge a consumer's identity across different website sessions, or they may not be able to match the online identities of consumers who search online but purchase offline. In this section, I relax the assumption that the purchase histories of opt-in consumers are perfectly observable. I make the following assumption which nests the main model as a special case: when an opt-in consumer arrives at a content page, the ad network and the advertiser can infer the consumer's funnel state f perfectly, but her purchase history only imperfectly. In particular, the ad network and the advertiser receive an imperfect signal about the consumer's purchase history.

To that end, let r denote the binary variable which equals 1 if the consumer purchased the advertised product, and 0 otherwise. Let S denote the binary signal that the ad network and the advertiser receive. The signal's accuracy is parametrized by $\rho \in \left[\frac{1}{2}, 1\right]$ where

$$\rho = \mathbb{P}\{S = r | r\}$$
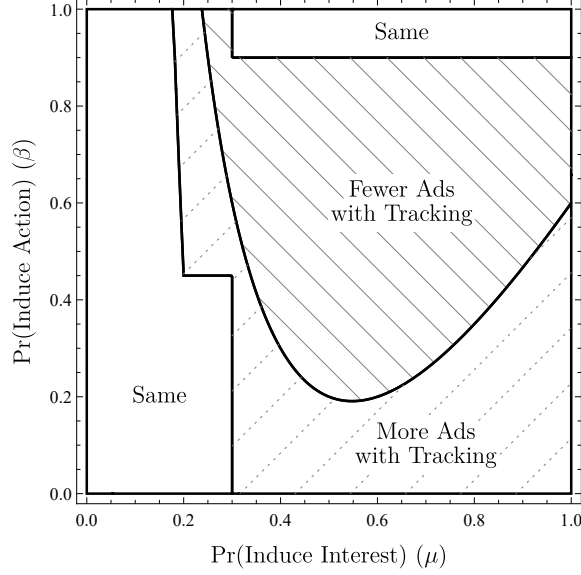
Figure 4.10: Ad Intensity With and Without Tracking Under Imperfect Purchase Observability; $\phi_M = 0.5, k = 0.15, \rho = \frac{2}{3}$

for $r \in \{0, 1\}$. Note that this extension collapses to the base model when $\rho = 1$.

The main departure from the main model is the Period 2 subgame for opt-in consumers when an ad was shown in Period 1. With perfect observability, the ad network forewent selling ad impressions for consumers who had already purchased. In contrast, under imperfect purchase observability, the ad network decides ad supplies based on the signals it receives. In Period 2, the ad network infers that $1 - \mu$ fraction of consumers remain in funnel state T, and μ fraction move down to funnel state M. Of the μ fraction of M-consumers, $\phi_M\rho + (1 - \phi_M)(1 - \rho)$ are associated with the signal that the consumers already purchased the product advertised in Period 1, and $\phi_M(1 - \rho) + (1 - \phi_M)\rho$ are associated with the signal that the consumers did not purchase.

I find that under certain conditions, imperfect purchase observability changes the relative ad intensity between opt-in vs. opt-out consumers. Specifically, when the purchase signal accuracy ρ is low, and ads are highly effective (i.e., μ and β are large), the reduced targetability associated with imperfect observability leads to ads being shown in both periods to all opt-in consumers.

The intuition is that if the ad network is uncertain about a consumer's purchase history, then even if it receives a signal that the consumer had already purchased, the ad network deems the false positive probability to be sufficiently high that it puts up the ad impression for sale. Consequently,

untargeted ads are shown to all opt-in consumers: those who are in state T, those in state M associated with "purchased" signal, and those in state M associated with "not purchased" signal. This implies that the ad intensities across opt-in and opt-out consumers are the same for large μ and β (see top-right corner of Figure 4.10), whereas fewer ads are shown to opt-in consumers with perfect observability (see Figure 4.5).

Next, I conduct comparative statics with respect to the signal accuracy ρ. What is the relationship between consumers' opt-in choices and the accuracy of their purchase history signals? Interestingly, I find that under certain conditions, more consumers choose to opt-in to tracking when purchase signals are less accurate. As the purchase signals become more accurate, more ads may be sold under tracking. Thus, the expectation of more intensive targeted ads motivates consumers with high privacy cost to opt-out from tracking for higher levels of purchase signal accuracy. Furthermore, the lower opt-in rate for high levels of signal accuracy may lower the ad network's profit due to a decline in targeting efficiency. The following proposition summarizes these findings.

Proposition 26. *Suppose $\frac{\mu}{\mu-k}k < \beta$, $\max\left[\frac{k}{\mu}, 1 - \sqrt{\frac{k}{\beta}}\right] < \phi_M < 1 - \frac{k}{\beta}$ and $q^*(0) = 2$ (i.e., ads are shown in both periods to all opt-out consumers). Then, the number of consumers opting-in to tracking decreases with the accuracy of the purchase signal. Consequently, the ad network's profit may decrease in the signal accuracy.*

Proposition 26 sheds light on a strategic force that goes against the lay intuition that accurate information about consumers is beneficial for the ad network. Distinct from the "market thinning" effect (Levin and Milgrom, 2010; Bergemann and Bonatti, 2011), less accurate signals about consumers' behaviors may induce consumers to expect fewer targeted ads should they opt-in to tracking, increasing the appeal of opting-in. And as more consumers opt-in to tracking, the targeting efficiency of the the ad network increases, thereby increasing its profit.

(a) $\phi_M = 0.05$ (b) $\phi_M = 0.5$
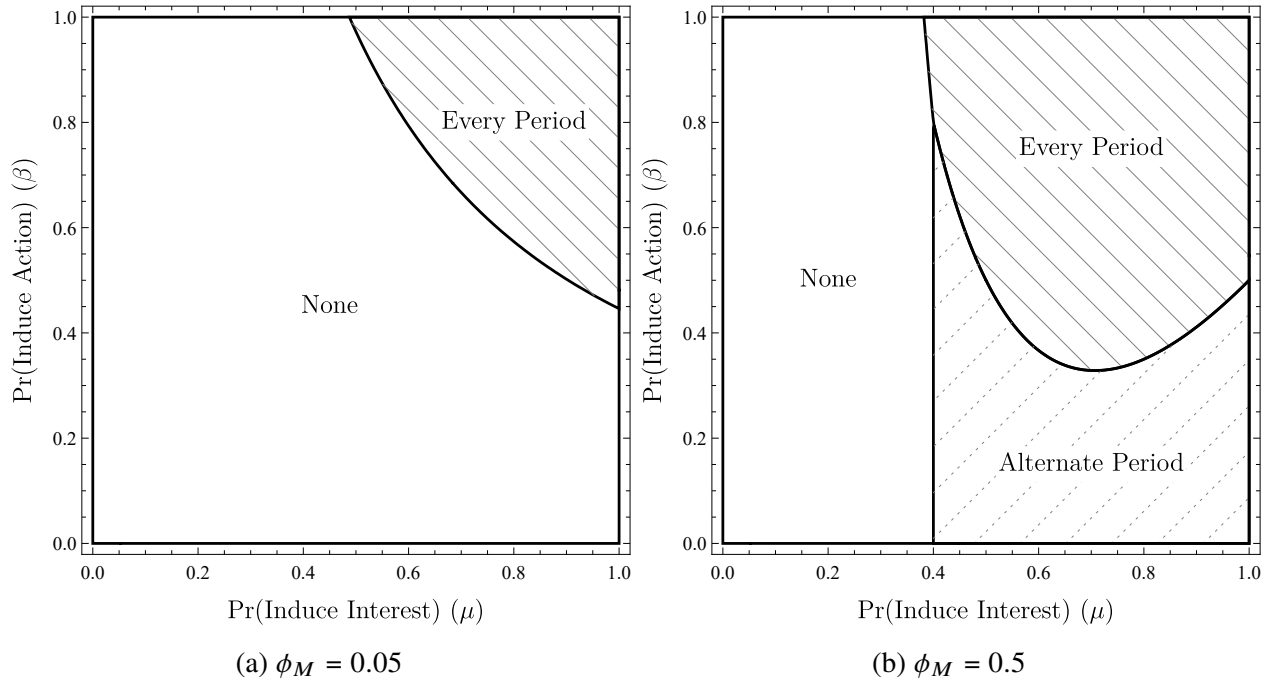
Figure 4.11: Advertising Strategy Without Tracking; $k = 0.25, \sigma = 0, \delta = 1 - 10^{-6}$

4.4.4   Infinite Horizon with Heterogeneous Overlapping Consumer Generations

I extend the game from the main model along two dimensions. First, I relax the assumption that all newly arriving consumers are at funnel state T. In particular, I allow $\sigma \in [0, 1]$ proportion of newly arriving consumers to be in funnel state M, and $1 - \sigma$ in T. Broadly, σ can be interpreted as the advertiser's "brand strength:" the higher the σ, the greater the extent to which the advertiser's product is *a priori* known and considered by consumers. Second, I extend the game horizon from two-period to infinite-period. In each period, a unit mass of consumers — σ mass of M-consumers and $1 - \sigma$ mass of T-consumers — arrive and live for two periods. Thus, in any given period, there are overlapping generations of consumers. The rest of the specifications remain the same as the main model.

I solve for a Markov-perfect equilibrium (MPE) wherein an advertiser's strategy depends only on the payoff-relevant state in that period. The ad network compares the total discounted profit (with discount factor $\delta \in [0, 1)$) obtained from inducing the different advertising outcomes, and then chooses whichever yields the highest profit.

Due to space considerations, I relegate the MPE derivation to Section C.4 of the appendix. Here, I highlight how the qualitative insights obtained here compare to that of the main model. First, Figure 4.11 shows that if consumer heterogeneity is muted (i.e., $\sigma = 0$) and the discount factor is close to 1, the equilibrium outcomes closely mirror that of the two-period model (see Figure 4.3). In particular, the advertiser shows ads to all consumers in both periods if and only if β is sufficiently large and μ is intermediate. The underlying mechanism revolves around the saturation effect, and the qualitative insights remain essentially the same.

Second, I examine how the insights from the main model are moderated by two new parameters: the "brand strength" parameter σ, and the discount factor δ. As illustrated in Figure 4.12, I find that as either σ increases or δ decreases, the parametric region where the advertiser shows ads to all consumers in both periods becomes smaller. The intuition is the following. As σ increases, a larger portion of newly arriving consumers are already mid-way down the funnel in state M. This accentuates the saturation effect: since many newly arriving consumers are already in the consideration phase and will likely convert without additional ad exposures, the value of a successive ad diminishes (see Figure 4.12a). On the other hand, as δ decreases, the advertiser places smaller weight on the value of a successive ad, whose payoff materializes in the future. Therefore, the incentive to buy successive ads decreases, even if convex ad response curves may have otherwise justified showing successive ads (see Figure 4.12b).

4.5 Conclusion

In this paper, I study the impact of tracking consumers' Internet activities on the online advertising ecosystem, and the impact of regulations that, motivated by privacy concerns, endow consumers with the choice to have their online activity be tracked or not (e.g., the GDPR). In particular, I model the consumer "purchase journey" and analyze the impact of consumers' opt-in decisions — co-determined by the intrinsic and instrumental aspects of privacy — on the strategies and profits of advertisers and ad networks.

Among others, I establish the following insights from the analysis. First, when given a choice,
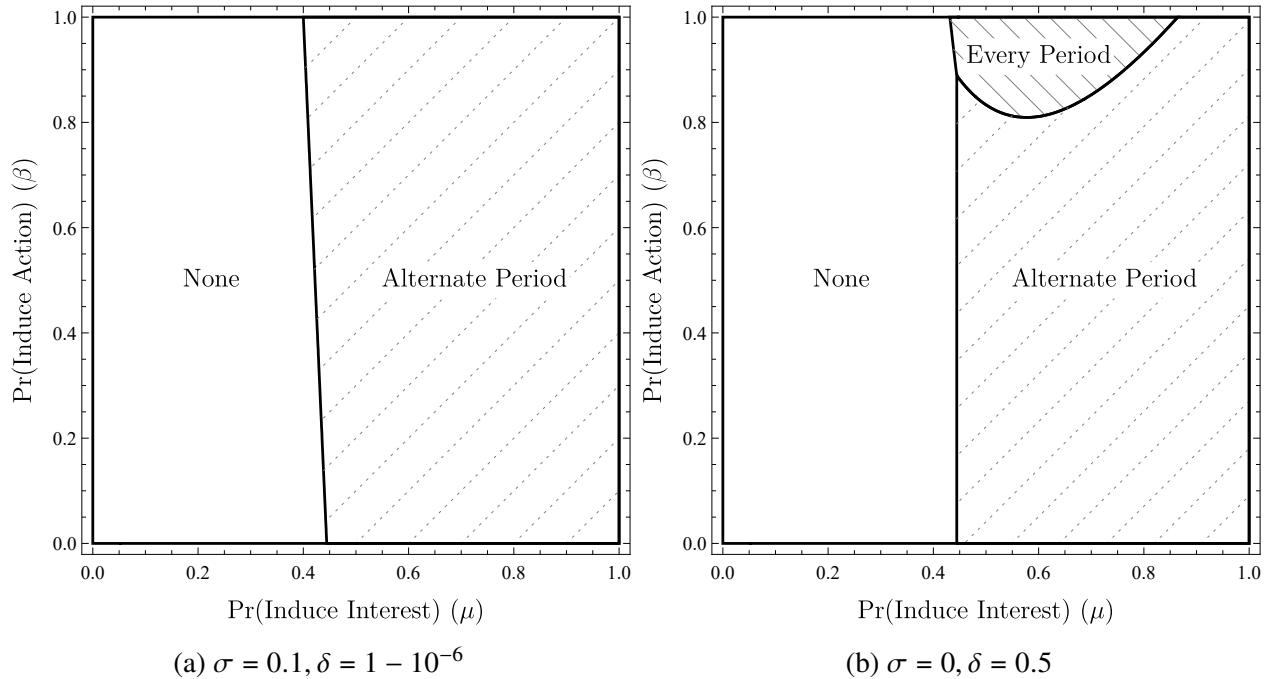
Figure 4.12: Advertising Strategy Without Tracking; $\phi_M = 0.5, k = 0.25$

some consumers will choose to opt-in to tracking because they expect to see fewer ads when advertisers can track them and infer their funnel stages. Specifically when ad effectiveness is intermediate, ad-averse consumers opt-in to tracking, thereby trading-off the benefit of seeing fewer ads (positive instrumental value of privacy) with the disutility they feel from giving up their privacy (intrinsic cost of privacy). Second, consumers' opt-in behaviors have important implications for the ad ecosystem. For example, due to changes in consumers' privacy choices, the ad network's profit may decrease in ad effectiveness, even though higher ad effectiveness implies higher purchase conversion. Finally, I show that privacy regulations improve overall consumer surplus and reduce the ad network's profit. Interestingly, however, if the advertiser has private information about ad valuations, privacy regulations may increase the ad network's profit as well. Intuitively, as privacy-conscious opt-out from tracking, the ad network commits to selling untargeted ads that reach a larger consumer segment than targeted ads. This in turn incentivizes the advertiser to bid more aggressively for opt-out consumers, resulting in higher ad network profit compared to the regime in which all consumers can be tracked.

The results obtained in this paper provide important managerial insights for marketers and reg-

111

ulators alike. My findings suggest that under certain conditions, the ad network and the advertiser could both earn higher profits if the ad network can credibly commit to *not* track consumers. Privacy regulations that allow consumer tracking only under affirmative consent can thus serve as a commitment device that helps the advertiser and the ad network "coordinate" in a mutually profitable manner. Furthermore, my results underscore the need for regulators to consider nuanced approaches to data privacy regulations that are based on various market conditions such as the accuracy of signals pertaining to consumers' online behavior, the degree of information asymmetry, consumer disutility for ads, their value of privacy, and the average effectiveness of ads.

My research generates a number of interesting hypotheses that could be empirically tested. For instance, for ads with intermediate levels of effectiveness, my results suggest that compared to the full-tracking regime prior to the enforcement of privacy regulations, the ad fill rates are likely to increase when consumers have a choice to be tracked or not as more untargeted ads are shown to consumers who choose to opt-out from tracking. In a similar vein, my analysis predicts that ad prices will fall with the advent of privacy regulations due to declines in targeting efficiency associated with consumers opting-out. It would be interesting to investigate these hypotheses across different product categories that are associated with different levels of average ad effectiveness.

I acknowledge several limitations of the paper. First, my model does not account for flexible product pricing decisions by the advertiser because consumers' product utilities assume a "binary" functional form. While this assumption allowed us to focus on the advertising strategies, it would be interesting to consider a finitely elastic product demand that would allow for richer pricing strategies. Second, I implicitly assumed that the advertiser shows the same ad content to all consumers. In practice, advertisers may tailor their messages to consumers in different stages of the journey, insofar as consumers can be tracked (e.g., entice consumers lower down the funnel with price promotions). Thus, another fruitful avenue for future research would be to investigate how personalized ads for opt-in consumers may impact the funnel-transition probabilities. Finally, it would be interesting to examine a more active role of publishers. For example, one could consider publishers acting as information gateways and study the forces that affect the publishers' incentives

to disclose or withhold consumers' information to the ad network.

Conclusion

Advances in information technology in the past few decades have generated a plethora of new opportunities and challenges for the advertising industry. While the vast literature on advertising has examined an extensive array of research questions, still many important questions regarding advertising strategies remain unaddressed. In a collection of three essays, my dissertation seeks to advance our understanding of advertising in three distinct contexts.

Essay 1 shows that when consumers exhibit reference-dependent preferences that yield concave utility functions (e.g., due to loss aversion), then the high-quality firm sometimes displays reference-shifting advertising content that, surprisingly, reduces consumers' absolute valuation of its offering. While this strategy lowers a monopolist's profit, it can increase the high-quality firm's profit in a competitive setting because it increases the *relative* utility of its offerings compared with the low-quality firm's offerings. Essay 2 characterizes the optimal strategies of advertisers and ad sellers in online advertising in a learning environment. In particular, I demonstrate that due to asymmetric learning effects, new advertisers should bid above their valuations in early rounds of the auction. Furthermore, the ad seller should favor new advertisers in the auctions by boosting their bid multipliers in order to learn their ad performances more quickly. Finally, Essay 3 discusses the impact of tracking consumers' online activities on the advertising ecosystem, particularly in the context of recent privacy regulations. Even though consumers incur a privacy cost for opting-in to tracking and disclosing their personal information, they sometimes choose to do so expecting to see fewer ads. A surprising result that follows from consumers' endogenous opt-in decisions is that the ad network's profit may decrease from displaying more effective ads.

References

Abhishek, V., Despotakis, S., and Ravi, R. (2017). Multi-channel attribution: The blind spot of online advertising. *Working paper*. Irvine, CA.

Abhishek, V., Fader, P., and Hosanagar, K. (2018). Media Exposure through the Funnel: A Model of Multi-Stage Attribution. *Working paper*. Irvine, CA.

Acquisti, A., Taylor, C. R., and Wagman, L. (2016). The Economics of Privacy. *Journal of Economic Literature*, 54(2):442–92.

Acquisti, A. and Varian, H. R. (2005). Conditioning prices on purchase history. *Marketing Science*, 24(3):367–381.

Amaldoss, W., Desai, P. S., and Shin, W. (2015a). Keyword Search Advertising and First-Page Bid Estimates: A Strategic Analysis. *Management Science*, 61(3):507–519.

Amaldoss, W. and He, C. (2018). Reference-dependent utility, product variety, and price competition. *Management Science*, 64(9):4302–4316.

Amaldoss, W., Jerath, K., and Sayedi, A. (2015b). Keyword Management Costs and "Broad Match" in Sponsored Search Advertising. *Marketing Science*, 35(2):259–274.

Anderson, S. P. and Renault, R. (2006). Advertising content. *American Economic Review*, 96(1):93–113.

Angwin, J. (2010). The web's new gold mine: Your secrets. *The Wall Street Journal*.

Ariely, D., Loewenstein, G., and Prelec, D. (2003). "coherent arbitrariness": Stable demand curves without stable preferences. *The Quarterly journal of economics*, 118(1):73–106.

Athey, S., Catalini, C., and Tucker, C. (2017). The digital privacy paradox: Small money, small costs, small talk. Technical report, National Bureau of Economic Research.

Athey, S. and Gans, J. S. (2010). The impact of targeting technology on advertising markets and media competition. *American Economic Review*, 100(2):608–13.

Aziz, A. and Telang, R. (2016). What is a digital cookie worth? *Working paper*. Vancouver, BC.

Aziz, A. and Telang, R. (2017). Irrelevant Advertising: How More Consumer Tracking Can Make Ads Less Effective. *Working paper*, pages 1–47. Vancouver, BC.

Bagwell, K. (2007). The economic analysis of advertising. *Handbook of industrial organization*, 3:1701–1844.

Balseiro, S. R., Besbes, O., and Weintraub, G. Y. (2015). Repeated Auctions with Budgets in Ad Exchanges: Approximations and Design. *Management Science*, 61(4):864–884.

Balseiro, S. R., Feldman, J., Mirrokni, V., and Muthukrishnan, S. (2014). Yield optimization of display advertising with ad exchange. *Management Science*, 60(12):2886–2907.

Balseiro, S. R. and Gur, Y. (2017). Learning in Repeated Auctions with Budgets: Regret Minimization and Equilibrium. *Working paper*.

Barlow, R. E. and Proschan, F. (1965). *Mathematical Theory of Reliability*. John Wiley & Sons.

Baucells, M., Weber, M., and Welfens, F. (2011). Reference-point formation and updating. *Management Science*, 57(3):506–519.

Becker, G. S. (1980). Privacy and malfeasance: A comment. *The Journal of Legal Studies*, 9(4):823–826.

Bergemann, D. and Bonatti, A. (2011). Targeting in Advertising Markets : Implications for Offline versus Online Media. *The RAND Journal of Economics*, 42(3):417–443.

Berman, R. (2016). Beyond the Last Touch: Attribution in Online Advertising. *Working paper*.

Berman, R. and Katona, Z. (2013). The Role of Search Engine Optimization in Search Marketing. *Marketing Science*, 32(4):644–651.

Bleier, A. and Eisenbeiss, M. (2015). Personalized online advertising effectiveness: The interplay of what, when, and where. *Marketing Science*, 34(5):669–688.

Bloch, F. and Manceau, D. (1999). Persuasive advertising in hotelling's model of product differentiation. *International Journal of Industrial Organization*, 17(4):557–574.

Bodner, R. and Prelec, D. (1994). The centroid model of context-dependent choice. *Working paper, Sloan School of Management, MIT*.

Bordalo, P., G. N. and Shleifer, A. (2017). Memory, attention, and choice. *Working paper, National Bureau of Economic Research*.

Bruce, N. I., Murthi, B. P. S., and Rao, R. C. (2017). A Dynamic Model for Digital Advertising: The Effects of Creative Format, Message Content, and Targeting on Engagement. *Journal of Marketing Research*, 54(2):202–218.

Brusset, X. (2009). Properties of distributions with increasing failure rate. *Working paper*.

Campbell, J., Goldfarb, A., and Tucker, C. (2015). Privacy regulation and market structure. *Journal of Economics & Management Strategy*, 24(1):47–73.

Cao, X. and Ke, T. T. (2017). Cooperative Search Advertising. *Working paper*.

Chen, Y., Joshi, Y. V., Raju, J. S., and Zhang, Z. J. (2009). A theory of combative advertising. *Marketing Science*, 28(1):1–19.

Chen, Y. and Turut, Ö. (2013). Context-dependent preferences and innovation strategy. *Management Science*, 59(12):2747–2765.

Conitzer, V., Taylor, C. R., and Wagman, L. (2012). Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Science*, 31(2):277–292.

D'Annunzio, A. and Russo, A. (2019). Ad Networks and Consumer Tracking. *Management Science*.

de Cornière, A. and Nijs, R. D. (2016). Online Advertising and Privacy. *The RAND Journal of Economics*, 47(650):48–72.

de Cornière, A. and Taylor, G. (2014). Integration and search engine bias. *The RAND Journal of Economics*, 45(3):576–97.

DellaVigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic literature*, 47(2):315–72.

Desai, P. S., Shin, W., and Staelin, R. (2014). The Company That You Keep: When to Buy a Competitor's Keyword. *Marketing Science*, 33(4):485–508.

Dixit, A. and Norman, V. (1978). Advertising and welfare. *The Bell Journal of Economics*, pages 1–17.

Economist (2006). The ultimate marketing machine. *Economist Special Rep.*, 380(8485):61–64.

Economist (2015). Marketing in the digital age: A brand new game. *Economist*.

Edelman, B., Ostrovsky, M., and Schwarz, M. (2007). Internet Advertising and the Generalized Second-Price Auction : Selling Billions of Dollars Worth of Keywords. *American Economic Review*, 97(1):242–259.

Edelman, B. and Schwarz, M. (2010). Optimal auction design and equilibrium selection in sponsored search auctions. *The American Economic Review*, 100(2):597–602.

eMarketer (2018a). Attitudes toward personalized advertising according to uk instagram users. `https://www.emarketer.com/Chart/Attitudes-Toward-Personalized-Advertising-According-UK-Instagram-Users-June-2018-of-respondents/224451`. (Accessed May 2019).

eMarketer (2018b). Level of concern us internet users have toward social media companies displaying ads based on their data. `https://www.emarketer.com/Chart/Level-of-Concern-US-Internet-Users-Have-Toward-Social-Media-Companies-Displaying-Ads-Based-on-Their-Data-March-2018-of-respondents/218053`. (Accessed May 2019).

eMarketer (2019a). Consumer attitudes on marketing 2019: Privacy concerns mount, and ad blocking isn't going away, aug 2019. `https://content-na1.emarketer.com/consumer-attitudes-on-marketing-2019`. Accessed Dec 2019.

eMarketer (2019b). Programmatic digital display ad spending, nov 2019. `https://content-na1.emarketer.com/programmatic-digital-display-ad-spending`. Accessed Dec 2019.

Epley, N. and Gilovich, T. (2001). Putting adjustment back in the anchoring and adjustment heuristic: Differential processing of self-generated and experimenter-provided anchors. *Psychological science*, 12(5):391–396.

Esteban, L., Gil, A., and Hernandez, J. M. (2001). Informative Advertising and Optimal Targeting in a Monopoly. *The Journal of Industrial Economics*, 49(2):161–80.

Farrell, J. (2012). Can Privacy Be Just Another Good? *Journal on Telecommunications and High Technology Law*, 10(2):251–64.

Frederick, S. and Loewenstein, G. (1999). Hedonic adaptation well-being: The foundations of hedonic psychology (d. kahneman, e. diener, & n. schwarz ed., pp. 302-329). *New York: Russell Sage Foundation*.

Future of Privacy Forum (2018). Comparing privacy laws: Gdpr v. ccpa. `https://fpf.org/wp-content/uploads/2018/11/GDPR_CCPA_Comparison-Guide.pdf`. (Accessed June 2019).

Gardete, P. M. (2013). Cheap-talk advertising and misrepresentation in vertically differentiated markets. *Marketing Science*, 32(4):609–621.

Gittins, J. C. and Jones, D. M. (1979). A Dynamic Allocation Index for the Discounted Multiarmed Bandit Problem. *Biometrika Trust*, 66(3):561–565.

Goldfarb, A. and Tucker, C. (2011a). Online display advertising: Targeting and obtrusiveness. *Marketing Science*, 30(3):389–404.

Goldfarb, A. and Tucker, C. E. (2011b). Privacy regulation and online advertising. *Management science*, 57(1):57–71.

Hardie, B. G., Johnson, E. J., and Fader, P. S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing science*, 12(4):378–394.

Haruvy, E. and Jap, S. (2018). The Long Run Impact of Organic Search on Paid Search Performance. *Working paper*.

Hoban, P. R. and Bucklin, R. E. (2015). Effects of Internet display advertising in the purchase funnel: Model-based insights from a randomized field experiment. *Journal of Marketing Research*, 52(3):375–393.

Hu, Y. J., Shin, J., and Tang, Z. (2016). Incentive Problems in Performance-Based Online Advertising Pricing: Cost per Click vs. Cost per Action. *Management Science*, 62(7):2022–2038.

Hummel, P. and McAfee, R. P. (2016). Machine Learning in an Auction Environment. *Journal of Machine Learning Research*, 17:1–37.

Ichihashi, S. (2019). Online privacy and information disclosure by consumers. *Working paper*. Canada.

Ipsos (2019). Global citizens and data privacy. `https://www.ipsos.com/sites/default/files/ct/news/documents/2019-01/ipsos-wef_-_global_consumer_views_on_data_privacy_-_2019-01-25-final.pptx_lecture_seule_0.pdf`. (Accessed May 2019).

Iyer, G., Soberman, D., and Villas-boas, J. M. (2005). The Targeting of Advertising. *Marketing Science*, 24(3):461–476.

Iyer, K., Johari, R., and Sundararajan, M. (2014). Mean Field Equilibria of Dynamic Auctions with Learning. *Management Science*, 60(12):2949–2970.

Jedidi, K., Mela, C. F., and Gupta, S. (1999). Managing Advertising and Promotion for Long-Run Profitability. *Marketing Science*, 18(1):1–22.

Jerath, K., Ma, L., Park, Y.-H., and Srinivasan, K. (2011). A "Position Paradox" in Sponsored Search Auctions. *Marketing Science*, 30(4):612–627.

Jiang, B., Jerath, K., and Srinivasan, K. (2011). Firm Strategies in the "Mid Tail" of Platform-Based Retailing. *Marketing Science*, 30(5):757–775.

Johnson, E. J., Hassin, R., Baker, T., Bajger, A. T., and Treuer, G. (2013). Can consumers make affordable care affordable? The value of choice architecture. *PloS one*, 8(12):e81521.

Johnson, G., Shriver, S., and Du, S. (2019). Consumer privacy choice in online advertising: Who opts out and at what cost to industry? *Working paper*. Boston, MA.

Johnson, G. A., Lewis, R. A., and Nubbemeyer, E. I. (2016). Ghost Ads : Improving the Economics of Measuring Ad E ff ectiveness. pages 1–56.

Johnson, J. P. and Myatt, D. P. (2017). The Determinants of Product Lines.

Karaj, A., Macbeth, S., Berson, R., and Pujol, J. M. (2019). Who tracks me: Shedding light on the opaque world of online tracking. *arXiv preprint arXiv:1804.08959*.

Katehakis, M. N. and Veinott, A. F. (1987). The Multi-Armed Bandit Problem: Decomposition and Computation. *Mathematics of Operations Research*, 12(2):262–268.

Katona, Z. and Sarvary, M. (2010). The Race for Sponsored Links: Bidding Patterns for Search Advertising. *Marketing Science*, 29(2):199–215.

Katona, Z. and Zhu, Y. (2017). Quality score that makes you invest. *Working paper*.

Kihlstrom, R. E. and Riordan, M. H. (1984). Advertising as a signal. *journal of Political Economy*, 92(3):427–450.

Kivetz, R., Netzer, O., and Srinivasan, V. (2004a). Alternative models for capturing the compromise effect. *Journal of marketing research*, 41(3):237–257.

Kivetz, R., Netzer, O., and Srinivasan, V. (2004b). Extending compromise effect models to complex buying situations and other context effects. *Journal of Marketing Research*, 41(3):262–268.

Klapper, D., Ebling, C., and Temme, J. (2005). Another look at loss aversion in brand choice data: Can we characterize the loss averse consumer? *International Journal of Research in Marketing*, 22(3):239–254.

Kopalle, P. K. and Assunção, J. L. (2000). When (not) to indulge in 'puffery': The role of consumer expectations and brand goodwill in determining advertised and actual product quality. *Managerial and Decision Economics*, 21(6):223–241.

Kopalle, P. K. and Lehmann, D. R. (1995). The effects of advertised and observed quality on expectations about new product quality. *Journal of Marketing Research*, 32(3):280–290.

Kopalle, P. K. and Lehmann, D. R. (2006). Setting quality expectations when entering a market: What should the promise be? *Marketing Science*, 25(1):8–24.

Kopalle, P. K., Rao, A. G., and Assuncao, J. L. (1996). Asymmetric reference price effects and dynamic pricing policies. *Marketing Science*, 15(1):60–85.

Kopalle, P. K. and Winer, R. S. (1996). A dynamic model of reference price and expected quality. *Marketing Letters*, 7(1):41–52.

Kostov, N. and Schechner, S. (2018). Google emerges as early winner from europe's new data privacy law. *Wall Street Journal*.

Kotler, P. and Armstrong, G. (2012). *Principles of Marketing*. Prentice Hall.

Kotowitz, Y. and Mathewson, F. (1979). Advertising, consumer information, and product quality. *The Bell Journal of Economics*, pages 566–588.

Krishna, V. (2010). *Auction Theory (2nd edition)*. Academic Press, San Diego, CA.

Kuksov, D., Prasad, A., and Zia, M. (2017). In-store advertising by competitors. *Marketing Science*, 36(3):402–425.

Lambrecht, A. and Tucker, C. (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research*, 50(5):561–576.

Lee, A. Y. (2002). Effects of implicit memory on memory-based versus stimulus-based brand choice. *Journal of Marketing Research*, 39(4):440–54.

Lerner, A., Simpson, A. K., Kohno, T., and Roesner, F. (2016). Internet jones and the raiders of the lost trackers: An archaeological study of web tracking from 1996 to 2016. In *25th USENIX Security Symposium*.

Levin, J. and Milgrom, P. (2010). Online advertising: Heterogeneity and conflation in market design. *American Economic Review*, 100(2):603–07.

Li, S.-M., Mahdian, M., and McAfee, R. P. (2010). Value of Learning in Sponsored Search Auctions. *International Workshop on Internet and Network Economics*, 6484.

Lu, S., Zhu, Y., and Dukes, A. (2015). Position Auctions with Budget Constraints: Implications for Advertisers and Publishers. *Marketing Science*, 34(6):897–905.

Macbeth, S. (2017). Tracking the trackers: Analysing the global tracking landscape with ghostrank. Technical report, Technical report, Ghostery.

Manjoo, F. (2019). I visited 47 websites. hundreds of trackers followed me. *The New York Times*.

Mayzlin, D. and Shin, J. (2011). Uninformative advertising as an invitation to search. *Marketing Science*, 30(4):666–685.

McDonald, A. M. and Cranor, L. F. (2010). Americans' attitudes about internet behavioral advertising practices. In *Proceedings of the 9th annual ACM workshop on Privacy in the electronic society*, pages 63–72. ACM.

Milgrom, P. and Roberts, J. (1986). Price and advertising signals of product quality. *Journal of political economy*, 94(4):796–821.

Montes, R., Sand-Zantman, W., and Valletti, T. (2019). The Value of Personal Information in Online Markets with Endogenous Privacy. *Management Science*, 65(3):1342–62.

Moorthy, K. S. (1984). Market segmentation, self-selection, and product line design. *Marketing Science*, 3(4):288–307.

Moorthy, K. S. (1988). Product and price competition in a duopoly. *Marketing science*, 7(2):141–168.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317.

Mussweiler, T. and Strack, F. (2001). Considering the impossible: Explaining the effects of implausible anchors. *Social Cognition*, 19(2):145–160.

Myerson, R. B. (1981). Optimal Auction Design. *Mathematics of Operations Research*, 6(1):58–73.

Nelson, P. (1974). Advertising as information. *Journal of political economy*, 82(4):729–754.

Nicolau, J. L. (2012). Asymmetric tourist response to price: Loss aversion segmentation. *Journal of Travel Research*, 51(5):568–676.

Nijs, V. R., Dekimpe, M. G., Steenkamps, J.-B. E., and Hanssens, D. M. (2001). The Category-Demand Effects of Price Promotions. *Marketing Science*, 20(1):1–22.

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, 41(1):100–126.

Orhun, A. Y. (2009). Optimal product line design when consumers exhibit choice set-dependent preferences. *Marketing Science*, 28(5):868–886.

Ostrovsky, M. and Schwarz, M. (2016). Reserve prices in internet advertising auctions: A field experiment. *Working paper*.

Payne, J. W., B.-J. R. and Johnson, E. J. (1993). The adaptive decision maker.

Posner, R. A. (1981). The economics of privacy. *The American economic review*, 71(2):405–409.

Rafieian, O. and Yoganarasimhan, H. (2018). Targeting and Privacy in Mobile Advertising. *Working paper*. Seattle, WA.

Roesner, F., Kohno, T., and Wetherall, D. (2012). Detecting and defending against third-party tracking on the web. In *Proceedings of the 9th USENIX conference on Networked Systems Design and Implementation*, pages 12–12. USENIX Association.

Rutz, O. J. and Bucklin, R. E. (2011). From Generic to Branded: A Model of Spillover in Paid Search Advertising. *Journal of Marketing Research*, 48(1):87–102.

Sahni, N. S. (2015). Effect of temporal spacing between advertising exposures: Evidence from online field experiments. *Quantitative Marketing and Economics*, 13(3):203–247.

Sahni, N. S., Narayanan, S., and Kalyanam, K. (2019). An Experimental Investigation of the Effects of Retargeted Advertising: The Role of Frequency and Timing. *Journal of Marketing Research*, 56(3):401–418.

Satariano, A. (2019). Google is fined $57 million under europe's data privacy law. *The New York Times*.

Sayedi, A. (2018). Real-Time Bidding in Online Display Advertising. *Marketing Science*, 37(4):553–568.

Sayedi, A., Jerath, K., and Baghaie, M. (2018). Exclusive Placement in Online Advertising. *Marketing Science*.

Sayedi, A., Jerath, K., and Srinivasan, K. (2014). Competitive Poaching in Sponsored Search Advertising and Its Strategic Impact on Traditional Advertising. *Marketing Science*, 33(4):586–608.

Schelter, S., Kunegis, J., et al. (2018). On the ubiquity of web tracking: Insights from a billion-page web crawl. *The Journal of Web Science*, 4(4):53–66.

Schwartz, N. and Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45(3):513–523.

Seiler, S. and Yao, S. (2017). The impact of advertising along the conversion funnel. *Quantitative Marketing and Economics*, 15(3):241–278.

Shapiro, S., MacInnis, D. J., and Heckler, S. E. (1997). The Effects of Incidental Ad Exposure on the Formation of Consideration Sets. *Journal of Consumer Research*, 24(1):94–104.

Shen, Q. and Villas-Boas, J. M. (2018). Behavior-Based Advertising. *Management Science*, 64(5):2047–64.

Shin, W. (2015). Keyword Search Advertising and Limited Budgets. *Marketing Science*, 34(6):882–896.

Simonson, I. and Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of marketing research*, 29(3):281–295.

Stern, J. (2018). Facebook really is spying on you, just not through your phone's mic. *Wall Street Journal*.

Subramanian, U., Raju, J. S., and Zhang, Z. J. (2014). The strategic value of high-cost customers. *Management Science*, 60(2):494–507.

Taylor, C. R. (2004). Consumer privacy and the market for customer information. *RAND Journal of Economics*, pages 631–650.

Todri, V., Ghose, A., and Singh, P. V. (2019). Trade-offs in online advertising: Advertising effectiveness and annoyance dynamics across the purchase funnel. *Forthcoming, Information Systems Research*.

Tremblay, V. J. and Martins-Filho, C. (2001). A model of vertical differentiation, brand loyalty, and persuasive advertising. *Advertising and differentiated products*, 10(200):1.

Tremblay, V. J. and Polasky, S. (2002). Advertising with subjective horizontal and vertical product differentiation. *Review of Industrial Organization*, 20(3):253–265.

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of marketing research*, 51(5):546–562.

Tversky, A. and Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Tversky, A. and Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The quarterly journal of economics*, 106(4):1039–1061.

van Heerde, H. J., Gupta, S., and Wittink, D. R. (2003). Is 75% of the Sales Promotion Bump Due to Brand Switching? No, Only 33% Is. *Journal of Marketing Research*, 40(4):481–491.

Villas-Boas, J. M. (2004). Price cycles in markets with customer recognition. *RAND Journal of Economics*, pages 486–501.

Von der Fehr, N.-H. M. and Stevik, K. (1998). Persuasive advertising and product differentiation. *Southern Economic Journal*, pages 113–126.

Wathieu, L. and Friedman, A. (2009). An empirical approach to understanding privacy concerns. *Working Paper*. Washington, DC.

Wilbur, K. C. and Zhu, Y. (2009). Click Fraud. *Marketing Science*, 28(2):293–308.

Xu, L., Duan, J. A., and Whinston, A. (2014). Path to Purchase: A Mutually Exciting Point Process Model for Online Advertising and Conversion. *Management Science*, 60(6):1392–412.

Yao, S. and Mela, C. F. (2011). A Dynamic Model of Sponsored Search Advertising. *Marketing Science*, 30(3):447–468.

Zhu, Y. and Dukes, A. (2017). Prominent attributes under limited attention. *Marketing Science*, 36(5):683–698.

Zhu, Y. and Wilbur, K. C. (2011). Hybrid Advertising Auctions. *Marketing Science*, 30(2):249–273.

Zia, M. and Rao, R. C. (2017). Search Advertising: Budget Allocation Across Search Engines. *Working paper*.

Appendix A: Essay 1

A.1 Assumptions on Parameters

Advertisement Effects (v and δ)

We assume that $v \leq \min[\alpha, 1 - \alpha]$ to ensure that the post-valuation-shift probability of a consumer having high valuation for quality is well-defined. In addition, we assume that $\delta \leq \frac{q_H + q_L}{2}$ to ensure that the resultant (post-advertising) reference point lies in the convex hull of the quality levels of the offerings, consistent with the literature.

Advertising Cost (k)

We assume that

$$k < \frac{v(1 - \theta_l)q_H(6\sigma - q_H((1 - \theta_l)(2\alpha + v) + \beta\lambda + \beta + 2\theta_l))}{18\sigma} \qquad \text{(A.1.1)}$$

to ensure that $V^{\downarrow}$ is not dominated by $\emptyset$ for Firm L when Firm H chooses $V^{\uparrow}$.

Preference Shock Parameter (σ)

We assume

$$\sigma > \frac{1 + \lambda\beta}{3}(q_H - q_L) \qquad \text{(A.1.2)}$$

to ensure interior price solutions.

Intrinsic Utility (V)

We assume that even if the preference shocks are realized most favorably towards the outside option, the intrinsic utility V is sufficiently high that all consumer types prefer to buy either of the products under the competitive equilibrium prices. A sufficient condition can be easily derived as

$V > 2\sigma + \frac{1}{3}\max\left[-q_H(\theta_l - 2\beta\lambda) - 2q_L(\beta\lambda + \theta_l), q_H(\beta\lambda - 3\theta_l + 1) - q_L(\beta\lambda + 1)\right].$

A.2 Proofs

A.2.1 Proof of Proposition 1

The demand for product q_j is

$$D_M^j(p_j, p_k) = \mathbb{P}\{u(q_j; \theta) - p_j > u(q_k; \theta) - p_k, u(q_j; \theta) - p_j > \varepsilon_0\},$$

where $j \neq k \in \{H, L\}$, $u(q_j)$ and $u(q_k)$ share the same brand-specific error ε_M, and $\varepsilon_M - \varepsilon_0 \sim U[-\sigma, \sigma]$. Thus, the monopolist sets prices p_H and p_L that maximizes $\pi_M(p_H, p_L) = D_M^H(p_H, p_L) p_H + D_M^L(p_L, p_H) p_L$. The monopolist cannot observe the consumer's preference (i.e., quality taste θ). We find that the monopolist adopts one of two pricing schemes: it either (i) sets discriminatory prices such that the high-(low-)valuation consumers self-select to choose between q_H (q_L) and the outside option, or it (ii) sets a sufficiently low price for q_H such that both consumer segments choose between q_H and the outside option.[1]

The demand for the high-quality product is

$$D_M(q_H) = \hat{\alpha}\mathbb{P}\{V + q_H + \beta(q_H - \hat{\chi}) - p_H + \varepsilon_M > V + q_L - \beta\lambda(\hat{\chi} - q_L) - p_L + \varepsilon_M,$$

$$V + q_H + \beta(q_H - \hat{\chi}) - p_H + \varepsilon_M > 0 + \varepsilon_0\}$$

$$+ (1 - \hat{\alpha})\mathbb{P}\{\theta_l q_H + \beta(q_H - \hat{\chi}) - p_H + \varepsilon_M > \theta_l q_L - \beta\lambda(\hat{\chi} - q_L) - p_L + \varepsilon_M,$$

$$V + \theta_l q_H + \beta(q_H - \hat{\chi}) - p_H + \varepsilon_M > 0 + \varepsilon_0\},$$

which simplifies to

$$\hat{\alpha}\, \mathbb{I}_{\{q_H - q_L + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L) - p_H + p_L > 0\}} \frac{1}{2\sigma} (V + q_H + \beta(q_H - \hat{\chi}) - p_H + \sigma)$$

$$+ (1 - \hat{\alpha})\mathbb{I}_{\{\theta_l(q_H - q_L) + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L) - p_H + p_L > 0\}} \frac{1}{2\sigma} (V + \theta_l q_H + \beta(q_H - \hat{\chi}) - p_H + \sigma).$$

Similarly, we obtain the demand for the low-quality product, $D_M(q_L)$, as follows:

$$\hat{\alpha}\, \mathbb{I}_{\{q_H - q_L + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L) - p_H + p_L < 0\}} \frac{1}{2\sigma} (V + q_L - \beta\lambda(\hat{\chi} - q_L) - p_L + \sigma)$$

$$+ (1 - \hat{\alpha})\mathbb{I}_{\{\theta_l(q_H - q_L) + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L) - p_H + p_L < 0\}} \frac{1}{2\sigma} (V + \theta_l q_L - \beta\lambda(\hat{\chi} - q_L) - p_L + \sigma).$$

---

[1]Note that due to the firm-specific preference shock ε_M, consumers may sometimes opt for the outside option even when the monopolist price discriminates to induce them to self-select to a particular product.

From the indicator functions, we can see that the monopolist has three pricing strategies at hand: (a) set p_H high and p_L low such that q_H is dominated for all consumer types; (b) set p_H low and p_L high such that q_L is dominated for all consumer types, or (c) set intermediate prices such that q_L is dominated for the high-type consumers, and q_H is dominated for the low-type consumers. Since consumer WTP for q_H is larger than that for q_L, it follows immediately that (b) dominates (a). We thus focus on strategies (b) and (c).

Under (b), the monopolist's profit is $\pi_M^b(p_H) = \frac{1}{2\sigma}(\hat{\alpha}\bar{u}_h + (1-\hat{\alpha})\underline{u}_h - p_H + \sigma)p_H$, where $\bar{u}_j$ $(\underline{u}_j)$ denotes the high(low)-type consumer's deterministic utility for q_j (e.g., $\underline{u}_H = V + \theta_l q_H + \beta(q_H - \hat{\chi})$ and $\bar{u}_L = V + q_L - \beta\lambda(\hat{\chi} - q_L)$). FOC yields the optimal price and profit $p_H^{b*} = \frac{\hat{\alpha}\bar{u}_H + (1-\hat{\alpha})\underline{u}_H + \sigma}{2}$ and $\pi_M^{b*} = \frac{1}{2\sigma}\left(\frac{\hat{\alpha}\bar{u}_H + (1-\hat{\alpha})\underline{u}_H + \sigma}{2}\right)^2$.

Under (c), the monopolist's profit is $\pi_M^c(p_H, p_L) = \frac{\hat{\alpha}}{2\sigma}(\bar{u}_H - p_H + \sigma)p_H + \frac{1-\hat{\alpha}}{2\sigma}(\underline{u}_L - p_L + \sigma)p_L$, subject to the IC constraints $\bar{u}_H - p_H > \bar{u}_L - p_L$ and $\underline{u}_L - p_L > \underline{u}_H - p_H$. Given p_L, the "best-response" p_H is $\max\left[\bar{u}_H - \bar{u}_L + p_L, \min\left[\bar{u}_H - \bar{u}_L + p_L, \frac{\bar{u}_H + \sigma}{2}\right]\right]$, and given p_H, the "best-response" p_L is $\max\left[\bar{u}_L - \bar{u}_H + p_H, \min\left[\underline{u}_L - \underline{u}_H + p_H, \frac{\underline{u}_L + \sigma}{2}\right]\right]$. Algebraic manipulations yield the following equilibrium candidate prices and profits:

1. if $\frac{\bar{u}_H - \underline{u}_L}{2} \in (0, \underline{u}_H - \underline{u}_L)$, then $p_H^* = \frac{\bar{u}_H + \sigma}{2}$ and $p_L^* = \bar{u}_L - \bar{u}_H + p_H^*$ for a profit of

$$\pi_M^{(1)} = \frac{\hat{\alpha}}{2\sigma}\left(\frac{\bar{u}_H + \sigma}{2}\right)^2 + \frac{1-\hat{\alpha}}{2\sigma}\left(\underline{u}_L - \bar{u}_L + \frac{\bar{u}_H + \sigma}{2}\right)\left(\bar{u}_L - \bar{u}_H + \frac{\bar{u}_H + \sigma}{2}\right) \quad (A.2.1)$$

2. if $\frac{\bar{u}_H - \underline{u}_L}{2} \in [\underline{u}_H - \underline{u}_L, \bar{u}_H - \bar{u}_L)$, then

$$\pi_M^{(2)} = \frac{\hat{\alpha}}{2\sigma}\left(\frac{\bar{u}_H + \sigma}{2}\right)^2 + \frac{1-\hat{\alpha}}{2\sigma}\left(\frac{\underline{u}_L + \sigma}{2}\right)^2 \quad (A.2.2)$$

3. if $\frac{\bar{u}_H - \underline{u}_L}{2} \in [\bar{u}_H - \bar{u}_L, \infty)$, then

$$\pi_M^{(3)} = \frac{\hat{\alpha}}{2\sigma}\left(\bar{u}_L - \frac{\underline{u}_L - \sigma}{2}\right)\left(\bar{u}_H - \bar{u}_L + \frac{\underline{u}_L + \sigma}{2}\right) + \frac{1-\hat{\alpha}}{2\sigma}\left(\frac{\underline{u}_L + \sigma}{2}\right)^2 \quad (A.2.3)$$

Therefore, we have

$$
\pi_M^{c*} = \begin{cases} \pi_M^{(1)} & \text{if } \frac{\overline{u}_H - u_L}{2} \in \left(0, \underline{u}_H - \underline{u}_L\right), \\[2mm] \pi_M^{(2)} & \text{if } \frac{\overline{u}_H - u_L}{2} \in \left[\underline{u}_H - \underline{u}_L, \overline{u}_H - \overline{u}_L\right), \\[2mm] \pi_M^{(3)} & \text{if } \frac{\overline{u}_H - u_L}{2} \in \left[\overline{u}_H - \overline{u}_L, \infty\right). \end{cases}
$$

In total, the monopolist's optimal profit is $\pi_M^* = \max\left[\pi_M^{b*}, \pi_M^{c*}\right]$.

Next, we show that π_M^* is increasing in $\hat{\alpha}$ and decreasing in $\hat{\chi}$. To show the first part, it suffices to show that π_M^{b*} and π_M^{c*} are both increasing in $\hat{\alpha}$. But π_M^{b*} increases with $\hat{\alpha}$ because $\overline{u}_H > \underline{u}_H$. Similarly, π_M^{c*} increases with $\hat{\alpha}$ because every possible profit expressions (A.2.1), (A.2.2), and (A.2.3) are increasing in $\hat{\alpha}$. To see this, note that $\frac{\overline{u}_H + \sigma}{2} > \frac{\overline{u}_H + \sigma}{2} - (\overline{u}_L - \underline{u}_L)$ and $\frac{\overline{u}_H + \sigma}{2} > \frac{\overline{u}_H + \sigma}{2} - (\overline{u}_H - \overline{u}_L)$, which shows that $\frac{\partial \pi_M^{(1)}}{\partial \hat{\alpha}} > 0$; $\frac{\overline{u}_H + \sigma}{2} > \frac{u_L + \sigma}{2}$, which shows that $\frac{\partial \pi_M^{(2)}}{\partial \hat{\alpha}} > 0$; and $\overline{u}_L - \frac{u_L - \sigma}{2} > \frac{u_L + \sigma}{2} \iff \overline{u}_L > \underline{u}_L$ and $\overline{u}_H - \overline{u}_L + \frac{u_L + \sigma}{2} > \frac{u_L + \sigma}{2}$, which shows that $\frac{\partial \pi_M^{(3)}}{\partial \hat{\alpha}} > 0$. Thus, $\frac{\partial \pi_M^*}{\partial \hat{\alpha}} > 0$.

To show the second part, we check the derivatives for all possible profit expressions above with respect to $\hat{\chi}$. Since $\overline{u}_H$ and $\underline{u}_H$ are both decreasing in the reference quality, it follows immediately that $\frac{\partial \pi_M^{b*}}{\partial \hat{\chi}} < 0$. Next, consider π_M^{c*}. The prices and expected demands for each consumer segment for each profit expressions (A.2.1), (A.2.2), and (A.2.3) can be shown to be decreasing in $\hat{\chi}$. The only non-trivial case is $\frac{\partial \pi_M^{(3)}}{\partial \hat{\chi}}$, which we analyze more explicitly here. First, note that $\frac{\partial}{\partial \hat{\chi}}\left(\frac{u_L + \sigma}{2}\right) = -\frac{1}{2}\beta\lambda < 0$, which implies that the profit from the low-type consumers decreases with $\hat{\chi}$. Second, consider the derivative of the profit from the high-type consumers:

$$
\frac{\partial}{\partial \hat{\chi}}\left(\overline{u}_L - \frac{u_L - \sigma}{2}\right)\left(\overline{u}_H - \overline{u}_L + \frac{u_L + \sigma}{2}\right) = -\frac{1}{2}\beta\lambda\left(\overline{u}_H - \overline{u}_L + \frac{u_L + \sigma}{2}\right) - \left(\overline{u}_L - \frac{u_L - \sigma}{2}\right)\beta\left(1 - \frac{\lambda}{2}\right).
$$

$$(A.2.4)$$

If $\lambda \leq 2$, then both terms are negative and we obtain that $\frac{\partial \pi_M^{(3)}}{\partial \hat{\chi}} < 0$, as desired. If $\lambda > 2$, then

$$
\begin{aligned}
(A.2.4) &= \beta \left(-\frac{1}{2} \lambda \left(\overline{u}_H - 2\overline{u}_L + \underline{u}_L + \sigma \right) - \left(\overline{u}_L - \frac{\underline{u}_L + \sigma}{2} \right) \right) \\
&< -\beta \left(\left(\overline{u}_H - 2\overline{u}_L + \underline{u}_L + \sigma \right) + \left(\overline{u}_L - \frac{\underline{u}_L + \sigma}{2} \right) \right) \\
&= -\beta \left(\overline{u}_H - \overline{u}_L + \frac{\underline{u}_L + \sigma}{2} \right) \\
&< 0.
\end{aligned}
$$

Therefore, π_M^* is decreasing in $\hat{\chi}$.

A.2.2   Proof of Corollary 1

This follows immediately from the last part of the proof of Proposition 1.

A.2.3   Proof of Proposition 2

We have

$$
\frac{\partial \pi_H^*}{\partial \hat{\chi}} = \frac{1}{9\sigma} (\beta(\lambda - 1)\hat{\chi} + C) \times \beta(\lambda - 1) \tag{A.2.5}
$$

where $C = 3\sigma + (\hat{\alpha} + (1 - \hat{\alpha})\theta_l)(q_H - q_L) + \beta q_H - \lambda \beta q_L$ is positive by Assumption A.1.2. This proves the first part of the proposition. The second part follows from the fact that each of two multiplicative terms in (A.2.5) is increasing in β and λ.

A.2.4   Proof of Proposition 3

From profit expressions (2.4.3) and (2.4.4), we see that the term inside the paratheses are increasing in $\hat{\alpha}$ for Firm H (since $\theta_l < 1$), and decreasing in $\hat{\alpha}$ for Firm L. Similarly, the term is increasing in $\hat{\chi}$ for Firm H and decreasing for Firm L. Combined with the interior-price condition (Assumption A.1.2), this proves that Firm H (L)'s profit increases (decreases) in $\hat{\alpha}$ and increases (decreases) in $\hat{\chi}$. Therefore, we obtain that $V^{\downarrow}$ and $R^{\downarrow}$ are dominated for Firm H, and $V^{\uparrow}$ and $R^{\uparrow}$ are dominated for Firm L.

We are left with $3 \times 3 = 9$ equilibrium candidates. Of these, we show below that $\left(\emptyset, V^\downarrow\right)$, $\left(\emptyset, R^\downarrow\right)$, $\left(R^\uparrow, V^\downarrow\right)$, and $\left(V^\uparrow, R^\downarrow\right)$ cannot be an equilibrium.

$\left(\emptyset, V^\downarrow\right)$ Cannot Be Equilibrium

For $\left(\emptyset, V^\downarrow\right)$ to constitute an equilibrium, it must be that $\pi_L(\emptyset, V^\downarrow) > \pi_L(\emptyset, \emptyset)$ and $\pi_H(\emptyset, V^\downarrow) > \pi_H(V^\uparrow, V^\downarrow)$. We will show that these two conditions cannot be jointly satisfied for any $\lambda > 1$. First, we simplify these inequalities in terms of λ. For the first inequality, we have

$$\frac{\partial}{\partial \lambda}\left(\pi_L(\emptyset, V^\downarrow) - \pi_L(\emptyset, \emptyset)\right) = -\frac{\beta \nu (1 - \theta_l)(q_H - q_L)^2}{18\sigma} < 0. \tag{A.2.6}$$

Therefore, $\pi_L(\emptyset, V^\downarrow) - \pi_L(\emptyset, \emptyset) > 0$ if and only if λ is less than λ_0 which solves $\pi_L(\emptyset, V^\downarrow) = \pi_L(\emptyset, \emptyset)$.

Similarly, for the second inequality, we have $\frac{\partial}{\partial \lambda}\left(\pi_H(V^\uparrow, V^\downarrow) - \pi_H(\emptyset, V^\downarrow)\right) = \frac{\beta \nu (1 - \theta_l)(q_H - q_L)^2}{18\sigma} > 0$, so that $\pi_H(V^\uparrow, V^\downarrow) < \pi_H(\emptyset, V^\downarrow)$ simplifies to λ being less than λ_4 which solves $\pi_H(V^\uparrow, V^\downarrow) = \pi_H(\emptyset, V^\downarrow)$.

Next, we show that $\min[\lambda_0, \lambda_4] < 1$ so that there is no $\lambda \geq 1$ that jointly satisfies the two inequalities above. To that end, note that $\min[\tilde{\lambda}_2, \tilde{\lambda}_3]$ simplifies to

$$\frac{\frac{\min(z, -z)}{\nu(1 - \theta_l)} - \iota(q_H - q_L)^2}{\beta(q_H - q_L)^2}, \tag{A.2.7}$$

where $\iota = (2\hat{\alpha} - \nu)(1 - \theta_l) + \beta + 2\theta_l$ and $z = 6\sigma(3k - \nu(1 - \theta_l)(q_H - q_L)$. Since $\min[z, -z] \leq 0$ for all z, this implies that (A.2.7) is less than $-\frac{\iota}{\beta}$. But $-\frac{\iota}{\beta} < 1$ implies that $\left(\emptyset, V^\downarrow\right)$ cannot be an equilibrium.

$\left(\emptyset, R^\downarrow\right)$ Cannot Be Equilibrium

For $\left(\emptyset, R^\downarrow\right)$ to constitute an equilibrium, it must be that $\pi_H(R^\uparrow, R^\downarrow) < \pi_H(\emptyset, R^\downarrow)$ and $\pi_L(\emptyset, R^\downarrow) > \pi_L(\emptyset, \emptyset)$. Again, we will show that these two conditions cannot be jointly satisfied for any λ. Consider the first inequality. We have $\frac{\partial^2}{\partial \lambda^2}\left(\pi_H(R^\uparrow, R^\downarrow) - \pi_H(\emptyset, R^\downarrow)\right) = \frac{\beta^2 \delta (q_H - q_L - \delta)}{9\sigma} > 0$; there-

fore, the difference is convex in λ. We will show that the difference is increasing at $\lambda = 1$, so that it is increasing for all $\lambda > 1$. To that end, consider $\frac{\partial}{\partial \lambda}\left(\pi_H(R^\uparrow, R^\downarrow) - \pi_H(\emptyset, R^\downarrow)\right) = \frac{\beta\delta(\beta\delta(1-\lambda)+(q_H-q_L)(\alpha+\beta\lambda+(1-\alpha)\theta_l)+3\sigma}{9\sigma}$, which is equal to $\frac{\beta\delta((q_H-q_L)(\alpha+\beta+(1-\alpha)\theta_l)+3\sigma)}{9\sigma}$ at $\lambda = 1$. Since this is positive, the difference is increasing in λ. Therefore, we can simplify the inequality $\pi_H(R^\uparrow, R^\downarrow) < \pi_H(\emptyset, R^\downarrow)$ as $\lambda < \lambda_1$ where λ_1 solves $\pi_H(R^\uparrow, R^\downarrow) = \pi_H(\emptyset, R^\downarrow)$, and is equal to

$$\frac{\sqrt{\rho} - \delta(\beta\delta + \alpha + (1 - \alpha)\theta_l)(q_H - q_L) + 3\sigma)}{\beta\delta(q_H - q_L - \delta)} \tag{A.2.8}$$

where

$$\rho = \delta\left(9\delta\sigma(\sigma - 2k) + 6\sigma(q_H - q_L)(\delta(\alpha + \beta + (1 - \alpha)\theta_l) + 3k) + \delta(q_H - q_L)^2(\alpha + \beta + (1 - \alpha)\theta_l)^2\right).$$

Next, the second inequality can be simplified in a similar manner, recognizing that

$$\frac{\partial}{\partial \lambda}\left(\pi_L(\emptyset, R^\downarrow) - \pi_L(\emptyset, \emptyset)\right) = \frac{\beta\delta(\beta\delta(\lambda - 1) - (q_H - q_L)(\alpha + \beta\lambda + (1 - \alpha)\theta_l) + 3\sigma)}{9\sigma}, \tag{A.2.9}$$

which is positive because $(1+\beta\lambda)(q_H-q_L) > (\alpha+(1-\alpha)\theta_l+\beta\lambda)(q_H-q_L)$ and $3\sigma > (1+\beta\lambda)(q_H-q_L)$ by Assumption A.1.2. Therefore, the inequality $\pi_L(\emptyset, R^\downarrow) > \pi_L(\emptyset, \emptyset)$ simplifies to $\lambda > \lambda_4'$, where λ_4' solves $\pi_L(\emptyset, R^\downarrow) = \pi_L(\emptyset, \emptyset)$ and is equal to $-\frac{\sqrt{\omega}+\delta(\beta\delta+(\alpha(1-\theta_l)+\theta_l)(q_H-q_L)-3\sigma)}{\beta\delta(q_H-q_L-\delta)}$, where

$$\omega = \delta\left(9\delta\sigma(2k + \sigma) - 2q_H\left(3\delta\sigma(\alpha + \beta + (1 - \alpha)\theta_l) + 9k\sigma + \delta q_L(\alpha + \beta + (1 - \alpha)\theta_l)^2\right)\right.$$
$$\left. + 6q_L\sigma(\delta(\alpha + \beta + (1 - \alpha)\theta_l) + 3k) + \delta q_H^2(\alpha + \beta + (1 - \alpha)\theta_l)^2 + \delta q_L^2(\alpha + \beta + (1 - \alpha)\theta_l)^2\right).$$

Next, we show that $\lambda_4' > \lambda_1$, which would imply that the two inequalities above cannot be satisfied for any λ. Consider $\lambda_4' - \lambda_1$, which simplifies to

$$6\beta\delta\sigma - \beta\delta\left(\sqrt{(3\sigma + (q_H - q_L)(\alpha + \beta + (1 - \alpha)\theta_l))^2 + w} + \sqrt{(3\sigma - (q_H - q_L)(\alpha + \beta + (1 - \alpha)\theta_l))^2 - w}\right), \tag{A.2.10}$$

where $w = \frac{18k\sigma(q_H-q_L-\delta)}{\delta}$. Since the function $\sqrt{(x + y)^2 + z} + \sqrt{(x - y)^2 - z}$ is decreasing in z for all $z > -2xy$, we obtain that

$$\sqrt{(3\sigma + (q_H - q_L)(\alpha + \beta + (1 - \alpha)\theta_l))^2 + w} + \sqrt{(3\sigma - (q_H - q_L)(\alpha + \beta + (1 - \alpha)\theta_l))^2 - w}$$

131

attains its maximum at the lowest value of w, but since $w = \frac{18k\sigma(q_H - q_L - \delta)}{\delta}$ is bounded from below by zero, the maximizing w is zero. At $w = 0$, we have

$$\sqrt{(3\sigma + (q_H - q_L)(\alpha + \beta + (1-\alpha)\theta_l))^2 + 0} + \sqrt{(3\sigma - (q_H - q_L)(\alpha + \beta + (1-\alpha)\theta_l))^2 - 0} = 6\sigma.$$

Therefore, for all $w = \frac{18k\sigma(q_H - q_L - \delta)}{\delta} > 0$, the inequality (A.2.10) holds. This proves that $\lambda_4' > \lambda_1$, and we are done.

$\left(R^\uparrow, V^\downarrow\right)$ Cannot Be Equilibrium

For $(R^\uparrow, V^\downarrow)$ to be an equilibrium, it must be that $\pi_H(R^\uparrow, V^\downarrow) > \pi_H(V^\uparrow, V^\downarrow)$ and $\pi_L(R^\uparrow, V^\downarrow) > \pi_L(R^\uparrow, R^\downarrow)$. Consider the first inequality. We have

$$\frac{\partial^2}{\partial v^2}\left(\pi_H(V^\uparrow, V^\downarrow) - \pi_H(R^\uparrow, V^\downarrow)\right) = -\frac{(\theta_l - 1)^2(q_H - q_L)^2}{9\sigma} < 0; \qquad \text{(A.2.11)}$$

i.e., the difference is concave in v. We will show that the difference is increasing for all $v < 1$ by showing that the maximizer v^* lies beyond 1. The maximizer v^* can be easily derived from FOC: $v^* = \frac{2\beta\delta(\lambda - 1) + (q_H - q_L)(2\alpha(1 - \theta_l) + \beta\lambda + \beta + 2\theta_l) + 6\sigma}{2(1 - \theta_l)(q_H - q_L)}$. Furthermore, it can be easily verified that v^* is increasing in α, β and θ_l. Therefore, the smallest value of v^* occurs at $\alpha = \beta = \theta_l = 0$, which yields $v^* = \frac{3\sigma}{q_H - q_L}$. But this is strictly greater than 1 because $3\sigma > (1 + \lambda\beta)(q_H - q_L)$ by Assumption A.1.2. Therefore, the difference is strictly increasing in v for all $v < 1$. We can thus simplify the inequality as $v < \frac{\beta\delta(\lambda - 1)}{(1 - \theta_l)(q_H - q_L)}$ where $\frac{\beta\delta(\lambda - 1)}{(1 - \theta_l)(q_H - q_L)}$ solves $\pi_H(R^\uparrow, V^\downarrow) = \pi_H(V^\uparrow, V^\downarrow)$. Re-arranging in terms of λ yields $\lambda > 1 + \frac{(q_H - q_L)v(1 - \theta_l)}{\beta\delta}$.

Now, consider the second inequality. The difference $\pi_L(R^\uparrow, R^\downarrow) - \pi_L(R^\uparrow, V^\downarrow)$ is concave with respect to λ and has two roots, $r_1 = 1 + \frac{(q_H - q_L)v(1 - \theta_l)}{\beta\delta}$ and $r_2 = \frac{\beta\delta - (q_H - q_L)((2\alpha - v)(1 - \theta_l) + \beta + 2\theta_l) + 6\sigma}{\beta(\delta + q_H - q_L)}$. Assumption A.1.2 imposes an upper-bound on λ; namely, $\lambda < \frac{3\sigma - (q_H - q_L)}{\beta(q_H - q_L)}$. But the second root r_2 cannot be less than this upper-bound: to see this, it suffices to show that $r_2 - \frac{3\sigma - (q_H - q_L)}{\beta(q_H - q_L)}$ is positive. Now, this difference can easily verified to be decreasing in both α and θ_l. Therefore, the minimum value of this difference occurs at $\alpha = \theta_l = 1$, where the difference simplifies to $\frac{(q_H - q_L - \delta)(3\sigma - (\beta + 1)(q_H - q_L))}{(q_H - q_L)(\delta + q_H - q_L)}$. And since $\delta < \frac{q_H - q_L}{2}$ and $3\sigma > (1 + \beta\lambda)(q_H - q_L) > (1 + \beta)(q_H - q_L)$,

we obtain the desired positivity. This means that if r_1 is greater than r_2, both roots lie beyond the upper-bound of λ such that for all parametric regions of our interest, the profit difference $\pi_L(R^\uparrow, R^\downarrow) - \pi_L(R^\uparrow, V^\downarrow)$ is negative. On the other hand, if r_1 is less than r_2, then the difference is negative if and only if $\lambda < r_1$. Under the assumption that $\lambda < \frac{3\sigma - (q_H - q_L)}{\beta(q_H - q_L)}$, these two cases under which the profit difference is negative is equivalent to $\lambda < 1 + \frac{(q_H - q_L)\nu(1-\theta_l)}{\beta\delta}$. This contradicts the necessary condition derived from the first inequality above: $\lambda > 1 + \frac{(q_H - q_L)\nu(1-\theta_l)}{\beta\delta}$. Therefore, we conclude that $(R^\uparrow, V^\downarrow)$ cannot be an equilibrium.

$\left(V^\uparrow, R^\downarrow\right)$ *Cannot Be Equilibrium*

For $(V^\uparrow, R^\downarrow)$ to constitute an equilibrium, it must be that

$$\pi_H(V^\uparrow, R^\downarrow) > \pi_H(R^\uparrow, R^\downarrow) \tag{A.2.12}$$

and

$$\pi_L(V^\uparrow, R^\downarrow) > \pi_L(V^\uparrow, V^\downarrow). \tag{A.2.13}$$

From the analysis above, we have that the first inequality is equivalent to $\lambda < 1 + \frac{(q_H - q_L)\nu(1-\theta_l)}{\beta\delta}$. Consider the second inequality. The difference $\pi_L(V^\uparrow, R^\downarrow) - \pi_L(V^\uparrow, V^\downarrow)$ is concave in λ. And the slope of this difference at the largest value of λ, which is $\lambda = \frac{3\sigma - (q_H - q_L)}{(q_H - q_L)\beta}$ is positive. To see this, the slope at this maximum λ is

$$-\frac{\beta\left(-6\delta^2\sigma - \nu(1-\theta_l)q_H^3 + (1-\theta_l)q_H^2(2\delta(\alpha+\nu-1)+3\nu q_L)+q_H\kappa+q_L^3\nu(1-\theta_l)+2\delta(1-\theta_l)q_L^2(\alpha+\nu-1)-2(\beta+1)\delta^2 q_L\right)}{18\sigma(q_H - q_L)},$$
$$\tag{A.2.14}$$

where $\kappa = \left(2(\beta+1)\delta^2 - 3\nu(1-\theta_l)q_L^2 - 4\delta(1-\theta_l)q_L(\alpha+\nu-1)\right)$. It can be easily verified that (A.2.14) is decreasing in α and increasing in ν. Therefore, (A.2.14) is minimized at $\alpha = 1$ and $\nu = 0$. Then, substituting these values simplifies the slope expression to $\frac{\beta\delta^2(3\sigma - (\beta+1)(q_H - q_L))}{9\sigma(q_H - q_L)}$, which is, again, positive by Assumption A.1.2. Therefore, the difference $\pi_L(V^\uparrow, R^\downarrow) - \pi_L(V^\uparrow, V^\downarrow)$ is increasing for all $\lambda < \frac{3\sigma - (q_H - q_L)}{(q_H - q_L)\beta}$. We thus obtain that the difference is positive if and only if

$\lambda > 1 + \frac{\nu(1-\theta_l)(q_H-q_L)}{\beta\delta}$. This contradicts the first necessary condition.

We are thus left with five equilibrium candidates: $(R^\uparrow, R^\downarrow)$, $(R^\uparrow, \emptyset)$, $(V^\uparrow, V^\downarrow)$, $(V^\uparrow, \emptyset)$, and $(\emptyset, \emptyset)$. We derive necessary and sufficient conditions for each of these in turn.

Equilibrium Condition for $(R^\uparrow, R^\downarrow)$

- Non-deviation conditions for Firm H

    - $\pi_H(R^\uparrow, R^\downarrow) > \pi_H(V^\uparrow, R^\downarrow)$: The difference $\pi_H(V^\uparrow, R^\downarrow) - \pi_H(R^\uparrow, R^\downarrow)$ is concave with respect to λ because its second-derivative is $\frac{\beta^2\delta(\delta-q_H+q_L)}{9\sigma} < 0$. And at $\lambda = 1$, the difference is positive: $(q_H-q_L)(2\alpha(1-\theta_l)+2\beta-\nu\theta_l+\nu+2\theta_l)+6\sigma > 0$. Therefore, the difference is negative if and only if λ is greater than the larger root of $\pi_H(R^\uparrow, R^\downarrow) = \pi_H(V^\uparrow, R^\downarrow)$. Since the two roots are $\frac{\nu(1-\theta_l)(q_H-q_L)}{\beta\delta} + 1$ and $-\frac{\beta\delta+(q_H-q_L)((1-\theta_l)(2\alpha+\nu)+\beta+2\theta_l)+6\sigma}{\beta(q_H-q_L-\delta)}$, where the second root is negative, we obtain that $\pi_H(R^\uparrow, R^\downarrow) > \pi_H(V^\uparrow, R^\downarrow)$ if and only if $\lambda > 1 + \frac{\nu(1-\theta_l)(q_H-q_L)}{\beta\delta}$.

    - $\pi_H(R^\uparrow, R^\downarrow) > \pi_H(\emptyset, R^\downarrow)$: $\lambda > \lambda_1$ from computations above.

- Non-deviation conditions for Firm L

    - $\pi_L(R^\uparrow, R^\downarrow) > \pi_L(R^\uparrow, \emptyset)$: The derivative of the difference $\pi_L(R^\uparrow, R^\downarrow) - \pi_L(R^\uparrow, \emptyset)$ with respect to q_L is
    $$\frac{\beta\delta(\lambda-1)(2\alpha(1-\theta_l)+\beta\lambda+\beta+2\theta_l)}{18\sigma} > 0. \tag{A.2.15}$$

    Therefore, the difference is positive if and only if q_L is greater than the root of the difference, which is equal to
    $$q_1 = \frac{\beta\delta(\lambda-1)-6\sigma}{2\alpha(1-\theta_l)+\beta\lambda+\beta+2\theta_l} + \frac{18k\sigma}{\beta\delta(\lambda-1)(2\alpha(1-\theta_l)+\beta\lambda+\beta+2\theta_l)} + q_H.$$

    - $\pi_L(R^\uparrow, R^\downarrow) > \pi_L(R^\uparrow, V^\downarrow)$: $\lambda > 1 + \frac{(q_H-q_L)\nu(1-\theta_l)}{\beta\delta}$ from computations above.

Equilibrium Condition for $(R^\uparrow, \emptyset)$

- Non-deviation conditions for Firm H

- $\pi_H(R^\uparrow, \emptyset) > \pi_H(V^\uparrow, \emptyset)$: At $\lambda = 1$, the difference $\pi_H(R^\uparrow, \emptyset) - \pi_H(V^\uparrow, \emptyset)$ is equal to $-\frac{(q_H - q_L)v(1-\theta_l)((q_H - q_L)((1-\theta_l)(2\alpha+v)+2\beta+2\theta_l)+6\sigma)}{18\sigma} < 0$. Moreover, the second derivative of the difference with respect to λ is $\frac{\beta^2\delta(\delta + q_H - q_L)}{9\sigma} > 0$. Therefore, the difference is positive if and only if λ is greater than the larger of the two roots of the difference. The two roots are $1 + \frac{v(1-\theta_l)(q_H - q_L)}{\beta\delta}$ and $\frac{\beta\delta - (q_H-q_L)((1-\theta_l)(2\alpha+v)+\beta+2\theta_l)-6\sigma}{\beta(\delta+q_H-q_L)}$, but the second root is negative. This is because Assumption A.1.2 implies $6\sigma > 2(1 + \lambda\beta)(q_H - q_L) > 2\beta(q_H - q_L) > \beta\frac{q_H - q_L}{2} > \beta\delta$. Therefore, the difference is positive if and only if $\lambda > 1 + \frac{v(1-\theta_l)(q_H - q_L)}{\beta\delta}$.

- $\pi_H(R^\uparrow, \emptyset) > \pi_H(\emptyset, \emptyset)$: The difference $\pi_H(R^\uparrow, \emptyset) - \pi_H(\emptyset, \emptyset)$ is convex in λ and negative at $\lambda = 1$; therefore, the difference is positive if and only if λ is greater than larger of the two roots of the difference, which is

$$\lambda_2 = \frac{\beta\delta^2 - 3\delta\sigma + \sqrt{\zeta} - \delta((1-\alpha)\theta_l + \alpha)(q_H - q_L)}{\beta\delta(\delta + q_H - q_L)}. \tag{A.2.16}$$

where

$$\zeta = \delta\left(9\delta\sigma(2k + \sigma) + 6\sigma(q_H - q_L)(\delta((1-\alpha)\theta_l + \alpha + \beta) + 3k)\right.$$
$$\left. + \delta(q_H - q_L)^2((1-\alpha)\theta_l + \alpha + \beta)^2\right).$$

- Non-deviation conditions for Firm L

  - $\pi_L(R^\uparrow, \emptyset) > \pi_L(R^\uparrow, R^\downarrow)$: $q_L < q_1$ from complement of above.

  - $\pi_L(R^\uparrow, \emptyset) > \pi_L(R^\uparrow, V^\downarrow)$: the derivative of the difference with respect to λ is

$$-\frac{\beta v(1 - \theta_l)(q_H - q_L)(2\delta + q_H - q_L)}{18\sigma} < 0. \tag{A.2.17}$$

Therefore, the difference is positive if and only if λ is smaller than the root of the difference, which is $\lambda_3 = \frac{2\beta\delta - \frac{18k\sigma}{v(1-\theta_l)(q_H - q_L)} - (q_H-q_L)((1-\theta_l)(2\alpha-v)+\beta+2\theta_l)+6\sigma}{\beta(2\delta+q_H-q_L)}$.

Equilibrium Condition for $(V^\uparrow, V^\downarrow)$

- Non-deviation conditions for Firm H

- $\pi_H(V^\uparrow, V^\downarrow) > \pi_H(R^\uparrow, V^\downarrow)$: $\lambda < 1 + \frac{(q_H - q_L)\nu(1-\theta_l)}{\beta\delta}$ from computations above.

- $\pi_H(V^\uparrow, V^\downarrow) > \pi_H(\emptyset, V^\downarrow)$: $\lambda > \lambda_4$ from computations above.

- Non-deviation conditions for Firm L

  - $\pi_L(V^\uparrow, V^\downarrow) > \pi_L(V^\uparrow, R^\downarrow)$: $\lambda < 1 + \frac{(q_H - q_L)\nu(1-\theta_l)}{\beta\delta}$ from the computations above.

  - $\pi_L(V^\uparrow, V^\downarrow) > \pi_L(V^\uparrow, \emptyset)$: the the difference $\pi_L(V^\uparrow, V^\downarrow) - \pi_L(V^\uparrow, \emptyset)$ is concave in q_L. For simplicity, we assume that ad cost k is sufficiently low such that Firm L prefers to choose $V^\downarrow$ over not advertising when Firm H chooses $V^\uparrow$ (Assumption A.1.1). Then by concavity, we have that the difference is positive if and only if q_L is less than the larger of the two roots of the difference, which is $q_2 = q_H-$

  $$\frac{3\nu(1-\theta_l)\sigma - 3\sqrt{\sigma}\sqrt{\nu(1-\theta_l)(\nu(1-\theta_l)\sigma - 2k((1-\theta_l)(2\alpha+\nu) + \beta(\lambda+1) + 2\theta_l))}}{\nu(1-\theta_l)((1-\theta_l)(2\alpha+\nu) + \beta(\lambda+1) + 2\theta_l)}.$$
  (A.2.18)

Equilibrium Condition for $(V^\uparrow, \emptyset)$

- Non-deviation conditions for Firm H

  - $\pi_H(V^\uparrow, \emptyset) > \pi_H(R^\uparrow, \emptyset)$: $\lambda < 1 + \frac{\nu(1-\theta_l)(q_H - q_L)}{\beta\delta}$ from the complement of above.

  - $\pi_H(V^\uparrow, \emptyset) > \pi_H(\emptyset, \emptyset)$: the derivative of the difference with respect to q_L is $-\frac{\nu(1-\theta_l)((q_H - q_L)((1-\theta_l)(2\alpha+\nu) + \beta(\lambda+1) + 2\theta_l) + 3\sigma)}{9\sigma} < 0$, and the second derivative with respect to q_L is $\frac{\nu(1-\theta_l)((1-\theta_l)(2\alpha+\nu) + \beta(\lambda+1) + 2\theta_l)}{9\sigma} > 0$. Therefore, the difference is positive if and only if q_L is less than the smaller of the two roots of the difference, which is equal to $q_3 = q_H-$

  $$\frac{6k\sqrt{\sigma}}{\nu(1-\theta_l)\sqrt{\sigma} + \sqrt{\nu}\sqrt{1-\theta_l}\sqrt{\nu(1-\theta_l)\sigma + 2k(2(\alpha+\nu)(1-\theta_l) + \beta(\lambda+1) + 2\theta_l)}}.$$
  (A.2.19)

- Non-deviation conditions for Firm L

136

- $\pi_L(V^\uparrow, \emptyset) > \pi_L(V^\uparrow, V^\downarrow)$: $q_L > q_2$ from the complement of above.

- $\pi_L(V^\uparrow, \emptyset) > \pi_L(V^\uparrow, R^\downarrow)$: the second derivative of the difference $\pi_L(V^\uparrow, R^\downarrow) - \pi_L(V^\uparrow, \emptyset)$ with respect to λ is $\frac{\beta^2 \delta(\delta - q_H + q_L)}{9\sigma} < 0$ and the first derivative with respect to λ is $\frac{\beta\delta(\beta\delta(\lambda-1)+(q_H-q_L)((1-\theta_l)(-(\alpha+v))+\beta\lambda+\theta_l)+3\sigma)}{9\sigma}$, which is positive due to Assumption A.1.2. Therefore, the difference is negative if and only if λ is less than λ_5, the smaller of the two roots, which is equal to $\frac{\beta\delta^2-3\delta\sigma+\sqrt{\psi}+\delta(q_H-q_L)((1-\alpha)\theta_l+\alpha+v(1-\theta_l))}{-\beta\delta(q_H-q_L-\delta)}$, where

$$\psi = \delta\left((q_H - q_L - \delta)\left(\beta^2\delta^2 - 6\beta\delta\sigma - 18k\sigma\right.\right.$$
$$\left. + \beta\delta(q_H - q_L)(2\alpha(1 - \theta_l) + \beta + 2((1 - v)\theta_l + v))\right)$$
$$\left. + \delta(\beta\delta + (q_H - q_L)((\alpha + v)(1 - \theta_l) + \theta_l) - 3\sigma)^2\right).$$

Equilibrium Condition for $(\emptyset, \emptyset)$

Complement of the union of all other equilibrium sets above.

A.2.5 Proof of Proposition 4

Consider the region $\lambda \in \left(\lambda_4, \min\left[\lambda_5, 1 + \frac{(q_H - q_L)v(1 - \theta_l)}{\beta\delta}\right]\right)$ and $q_L < q_3$, wherein Firm H chooses $V^\uparrow$, and Firm L either $V^\downarrow$ or $\emptyset$ in equilibrium. As q_L increases above q_2, Firm L changes its advertising strategy from $V^\downarrow$ to $\emptyset$. Therefore, $\hat{\alpha}$ jumps discretely upward from α to $\alpha + v$, when q_L passes the threshold q_L.

Since Firm H's profit is continuous in q_L, we can find a small $\epsilon > 0$ such that, all else equal, Firm H's profit decreases by no more than ϵ as q_L increases from $q_2 - \epsilon$ to $q_2 + \epsilon$. While the profit change due to increase in q_L is infinitesimally small by construction, the discrete jump in $\hat{\alpha}$ leads to an discrete positive jump in Firm H's profit.

Firm L's equilibrium profit increases with q_L as well because in this region, its equilibrium profit is

$$\max\left[\frac{1}{18\sigma}\left(3\sigma - (\alpha + (1 - \alpha)\theta_l + \beta)(q_H - q_L) - \beta\left(\frac{q_H + q_L}{2}\right)(\lambda - 1)\right)^2 - k,\right.$$
$$\left.\frac{1}{18\sigma}\left(3\sigma - (\alpha + v + (1 - \alpha - v)\theta_l + \beta)(q_H - q_L) - \beta\left(\frac{q_H + q_L}{2}\right)(\lambda - 1)\right)^2\right],$$

from which it is evident that the profit is increasing in q_L.

A.2.6 Proof of Proposition 5

Consider the region $\lambda < \min\left[\lambda_5, 1 + \frac{(q_H - q_L)v(1-\theta_l)}{\beta\delta}\right]$ and $q_2 < q_L$, where Firm L does not advertise and Firm H chooses $V^\uparrow$ for low $q_L < q_3$ and does not advertise for high $q_3 < q_L$. Therefore, $\hat{\alpha}$ jumps discretely downward from $\alpha + v$ to α, when q_L passes the threshold q_L.

Now, consider the consumer surplus:

$$
\begin{aligned}
CS = V + \hat{\alpha} &\left(\frac{(q_L - \beta\lambda(\hat{\chi} - q_L) - p_L)(\sigma - (-p_H + p_L + \Delta_q + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L)))}{2\sigma} \right. \\
&\left. + \frac{(q_H + \beta(q_H - \hat{\chi}) - p_H)(-p_H + p_L + \Delta_q + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L) + \sigma)}{2\sigma} \right) \\
+ (1 - \hat{\alpha}) &\left(\frac{(\theta_l q_L - \beta\lambda(\hat{\chi} - q_L) - p_L)(\sigma - (-p_H + p_L + \theta_l\Delta_q + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L)))}{2\sigma} \right. \\
&\left. + \frac{(\theta_l q_H + \beta(q_H - \hat{\chi}) - p_H)(-p_H + p_L + \theta_l\Delta_q + \beta(q_H - \hat{\chi}) + \beta\lambda(\hat{\chi} - q_L) + \sigma)}{2\sigma} \right)
\end{aligned}
$$

$$\text{(A.2.20)}$$

where $\Delta_q = q_H - q_L$.

If σ is sufficiently large, then the expected consumer surplus increases in $\hat{\alpha}$. To see this, note that the derivative of (A.2.20) with respect to $\hat{\alpha}$ is increasing in $\hat{\chi}$ and β and decreasing in $\hat{\alpha}$, because $\frac{\partial^2 (A.2.20)}{\partial\hat{\alpha}\partial\hat{\chi}} \propto 2\Delta_q\beta(\lambda - 1) > 0$, $\frac{\partial^2 (A.2.20)}{\partial\hat{\alpha}\partial\beta} \propto 2(q_H - \hat{\chi} + \lambda(\hat{\chi} - q_L)) > 0$, and $\frac{\partial^2 (A.2.20)}{\partial\hat{\alpha}\partial\hat{\alpha}} \propto -16\Delta_q^2(1 - \theta_l) < 0$. Therefore, the derivative is minimized at $\hat{\chi} = q_L$, $\beta = 0$, and $\hat{\alpha} = 1$, which upon substitution yields $(9\theta_l - 7)\Delta_q^2 + 9\sigma(q_H + q_L)$. This is positive if and only if

$$\sigma > \frac{\frac{7}{9} - \theta_l}{q_H - q_L}. \tag{A.2.21}$$

Suppose (A.2.21) holds. Since the consumer surplus is continuous in q_L, we can find a small $\epsilon > 0$ such that, all else equal, the consumer surplus decreases by no more than ϵ as q_L increases from $q_3 - \epsilon$ to $q_3 + \epsilon$. By construction, the change in consumer surplus due to increase in q_L is infinitesimally small. However, the change in equilibrium advertising regime as q_L increases

induces a discrete negative jump in α. This in turn results in a discrete negative jump in the consumer surplus.

A.2.7 Proof of Proposition 6

$(\emptyset, A)$ *Cannot Be Equilibrium*

For $(\emptyset, A)$ to constitute an equilibrium, it must be that $\pi_H(\emptyset, A) > \pi_H(\emptyset, A)$ and $\pi_L(\emptyset, A) > \pi_L(\emptyset, \emptyset)$. These two inequalities, respectively, simplify to $k > \frac{1}{18\sigma} A(6\sigma + B)$ and $k < \frac{1}{18\sigma} A(6\sigma - B)$ where $A = v(1 - \theta_l)(1 - w)(q_H - q_L) + \beta\delta(\lambda - 1)w$ and $B = (q_H - q_L)(\beta\lambda + \beta + 2\theta_l + (1 - \theta_l)(2\alpha - v(1 - w))) - \beta\delta(\lambda - 1)w$. Since $A > 0$, the two inequalities cannot hold simultaneously if $B > 0$. But since $\delta \le \frac{q_H - q_L}{2}$, we have $B \ge (q_H - q_L)(\beta\lambda + \beta + 2\theta_l + (1 - \theta_l)(2\alpha - v(1 - w))) - \beta\frac{q_H - q_L}{2}(\lambda - 1)w = (q_H - q_L)\left(\beta\lambda\left(1 - \frac{w}{2}\right) + \frac{\beta w}{2} + \beta + 2\theta_l + (1 - \theta_l)(2\alpha - v(1 - w))\right) > 0$. This completes the proof.

Equilibrium Condition for (A, A)

- Non-deviation condition for Firm H

  $\pi_H(A, A) > \pi_H(\emptyset, A)$: The derivative of the difference $\pi_H(A, A) - \pi_H(\emptyset, A)$ with respect to λ is increasing in α and v because $\frac{\partial^2}{\partial\lambda\partial\alpha} = 2(q_H - q - L)w\delta(1 - \theta_l)$ and $\frac{\partial^2}{\partial\lambda\partial v} = (q_H - q_L)(1 - w)(q_H - q_L - 2w\delta)(1 - \theta_l) > 0$. Therefore, the derivative is minimized at $\alpha = 0$ and $v = 0$, which after substitution yields $2\delta w((q_H - q_L)(\beta\lambda + \theta_l) + 3\sigma - w\beta\delta(\lambda - 1))$. Now, using the fact that $\delta \le \frac{q_H - q_L}{2}$, we obtain $2\delta w((q_H - q_L)(\beta\lambda + \theta_l) + 3\sigma - w\beta\delta(\lambda - 1)) > 2\delta w\left((q_H - q_L)(\beta\lambda + \theta_l) - \frac{w}{2}\beta(\lambda - 1)(q_H - q_L) + 3\sigma\right)$. This in turn is equal to $2\delta w(q_H - q_L)\left(3\sigma + (\theta_l + \beta\lambda\left(1 - \frac{w}{2}\right)) + \frac{\beta w}{2}\right) > 0$. Therefore, the derivative is always positive. Combined with the fact that the difference is convex in λ, we obtain that the difference is positive if and only if λ is greater than the larger of the two roots of the difference, which we denote by $\tilde{\lambda}_1$.

- Non-deviation condition for Firm L

  $\pi_L(A, A) > \pi_L(A, \emptyset)$: the derivative of the difference $\pi_L(A, A) - \pi_L(A, \emptyset)$ with respect to λ is

$-\frac{\beta^2 \delta w (q_H - q_L + \delta w)}{9\sigma} < 0$. Therefore, the difference is positive if and only if $\lambda \in (\tilde{\lambda}_2, \tilde{\lambda}_3)$ where $\tilde{\lambda}_2$ is the smaller and $\tilde{\lambda}_3$ the larger of the two roots of the difference.

Equilibrium Condition for $(A, \emptyset)$

- Non-deviation condition for Firm H

  $\pi_H(A, \emptyset) > \pi_H(\emptyset, \emptyset)$: The derivative of the difference $\pi_H(A, \emptyset) - \pi_H(\emptyset, \emptyset)$ with respect to λ is increasing in α and ν because $\frac{\partial^2}{\partial \lambda \partial \alpha}(\pi_H(A, \emptyset) - \pi_H(\emptyset, \emptyset)) = 2(q_H - q_L)w\delta(1 - \theta_l) > 0$ and $\frac{\partial^2}{\partial \lambda \partial \nu}(\pi_H(A, \emptyset) - \pi_H(\emptyset, \emptyset)) = (q_H - q_L)(1 - w)(q_H - q_L + 2w\delta)(1 - \theta_l) > 0$. Therefore, $\frac{\partial}{\partial \lambda}(\pi_H(A, \emptyset) - \pi_H(\emptyset, \emptyset))$ is minimized at $\alpha = 0$ and $\nu = 0$, which upon substitution yields $2\delta w((q_H - q_L)(\beta\lambda + \theta_l) + 3\sigma + \beta\delta(\lambda - 1)w) > 0$. Therefore, the difference is strictly increasing in λ. Moreover, second derivative of the difference is $\frac{\beta^2 \delta w (q_H - q_L + \delta w)}{9\sigma} > 0$; therefore, the difference is positive if and only if λ is greater than the larger root of the difference, which we denote by $\tilde{\lambda}_4$.

- Non-deviation condition for Firm L

  $\pi_L(A, \emptyset) > \pi_L(A, A)$: $\lambda < \tilde{\lambda}_2$ or $\lambda > \tilde{\lambda}_3$ from the complement of above.

Equilibrium Condition for $(\emptyset, \emptyset)$

Complement of the union of all other equilibrium sets above.

A.3 Alternative Utility Model

A main characteristic of the consumer utility function that drives our results is that there are diminishing marginal returns to quality. In the main model in the paper, we have implemented this concavity in consumption utility using a linear loss-aversion model, which also provided a natural way to incorporate the quality reference point. Here, we show the robustness of our main insights by considering an alternative specification of the consumer utility. Specifically, we model the concavity of consumer's utility for quality by relaxing the behavioral assumption that consumers exhibit loss aversion and simply considering a general concave utility function.

Furthermore, another way in which the alternative utility function assumed here is different from that in the main model is that we assume it to be entirely deterministic. Note that, in the main model, we had included the random component of utility to make the demand function smooth and simplify the pricing analysis. Here, we show that removing this simplification leads to the same insights, although the pricing analysis is more complicated, essentially because the pricing equilibrium is in mixed strategies and not in pure strategies.

Qualitatively, the key insights that we obtain here are the same as those in the main model. Specifically, we find that a monopolist offering two products never chooses reference-shifting advertising that increases the quality reference point, as doing so only dampens consumer's WTP. In a duopoly, the incentive of the low-quality firm is similar — it never chooses to increase the reference quality. On the other hand, the high-quality firm may find it profitable to choose reference-shifting advertising that increases the reference point. The underlying intuition is analogous to the main model — due to the concavity of the utility function, an increase in the quality reference point induces a steeper decline in consumer's valuation for the low-quality product thn it does for the high-quality product. This increases perceived quality differentiation between the products which, in turn, generates the premium effect wherein the high-quality firm charges a higher premium over its competitor's price. We now proceed with the details.

A.3.1 Model

We assume that the consumer utility function is given by

$$u(q_j) = \theta v(q_j - \chi),$$

where $v(\cdot)$ represents the consumer's valuation function with the properties $v'(\cdot) > 0$ and $v''(\cdot) < 0$, that is, the marginal utility for quality diminishes with increasing quality. We assume that the utility is zero for qualities lower than the references point. For setting the reference point, we assume the

following specification, which is similar to that in the main model:

$$\chi\left(a_H^R, a_L^R, \eta_H, \eta_L\right) = \frac{a_0\delta_0 + a_H^R\eta_H + a_L^R\eta_L}{a_0 + a_H^R + a_L^R},$$

where $\delta_0 \leq q_L$ is the reference quality without ads. We assume that values of η_j are such that $\chi \leq q_L$.

The rest of the model specification remains the same as the main model in Section 2.2 of the main paper.

A.3.2 Monopoly

Proposition 27 (Monopolist). *If a monopolist offering two products with distinct quality levels advertises, it never chooses reference-shifting advertising that increases the reference quality.*

Proof. Let the *hat* notations denote the post-advertising variables. The monopolist's profit can be either (i) $\hat{\alpha}v(q_H - \hat{\chi})$, (ii), $\hat{\alpha}(v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \theta_l v(q_L - \hat{\chi})) + (1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$, or (iii) $\theta_l v(q_L - \hat{\chi})$. Of the three possibilities, if either $v(q_L - \hat{\chi})$ or $\hat{\alpha}v(q_H - \hat{\chi})$ is the highest, then the monopolist's profit decreases with $\hat{\chi}$. Therefore, it will never choose $R^\uparrow$, let alone at a cost k.

Suppose $\hat{\alpha}(\theta_h(v(q_H - \hat{\chi}) - v(q_L - \hat{\chi})) + v(q_L - \hat{\chi})) + (1 - \hat{\alpha})v(q_L - \hat{\chi})$ is the highest profit (i.e., the monopolist offers both products and sets prices such that the h-type's incentive compatibility constraint binds, and l-type's rationality constraint binds). We will show that the fact that $\hat{\alpha}(v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \theta_l v(q_L - \hat{\chi})) + (1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$ is higher than the other two profits will imply $R^\uparrow$ is dominated.

First, note that

$$\hat{\alpha}(v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \theta_l v(q_L - \hat{\chi})) + (1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi}) > \hat{\alpha}v(q_H - \hat{\chi}) \iff \hat{\alpha} < \theta_l \quad \text{(A.3.1)}$$

and

$$\hat{\alpha}(v(q_H-\hat{\chi})-v(q_L-\hat{\chi})+\theta_l v(q_L-\hat{\chi}))+(1-\hat{\alpha})\theta_l v(q_L-\hat{\chi}) > \theta_l v(q_L-\hat{\chi}) \iff v(q_H-\hat{\chi}) > v(q_L-\hat{\chi}).$$

$$(A.3.2)$$

Therefore, $\frac{\partial}{\partial \hat{\chi}}\hat{\alpha}(v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \theta_l v(q_L - \hat{\chi})) + (1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$ is equal to

$$\hat{\alpha}\underbrace{\frac{\partial v(q_H - \hat{\chi})}{\partial \hat{\chi}}}_{-} + \underbrace{(\theta_l - \hat{\alpha})}_{+ \text{ by } (A.3.1)}\underbrace{\frac{\partial v(q_L - \chi)}{\partial \chi}}_{-} < 0$$

which implies that $R^{\uparrow}$ only reduces profit. This completes the proof. $\qquad\square$

A.3.3 Duopoly

When there are no error terms in the consumer utility, there is no pure strategy equilibrium in the pricing subgame. Instead, firms play mixed strategies as specified in the following lemma.

Lemma 7 (Mixed Strategy Price Equilibrium). *There does not exist a pure strategy equilibrium. There exists a unique mixed strategy equilibrium characterized by the cumulative distributions functions:*

$$F_H^*(p) = \begin{cases} 0 & \text{if } p < \underline{p}_H, \\ \frac{1}{\hat{\alpha}} - \frac{(1-\hat{\alpha})\theta_l v(q_L-\hat{\chi})}{\hat{\alpha}(p-(v(q_H-\hat{\chi})-v(q_L-\hat{\chi})))} & \text{if } \underline{p}_H \leq p < \overline{p}_H, \\ 1 & \text{if } \overline{p}_H \leq p, \end{cases}$$

$$\text{and } F_L^*(p) = \begin{cases} 0 & \text{if } p < \underline{p}_L, \\ \frac{p-(1-\hat{\alpha})\theta_l v(q_L-\hat{\chi})}{p+(v(q_H-\hat{\chi})-v(q_L-\hat{\chi}))} & \text{if } \underline{p}_L \leq p < \overline{p}_L, \\ 1 & \text{if } \overline{p}_L \leq p, \end{cases}$$

where $\underline{p}_H = v(q_H-\hat{\chi})-v(q_L-\hat{\chi})+(1-\hat{\alpha})\theta_l v(q_L-\hat{\chi}), \overline{p}_H = v(q_H-\hat{\chi})-v(q_L-\hat{\chi})+\theta_l v(q_L-\hat{\chi}), \underline{p}_L = (1-\hat{\alpha})\theta_l v(q_L - \hat{\chi})$ and $\overline{p}_L = \theta_l v(q_L - \hat{\chi})$.

Proof. We first show that there does not exist a pure strategy equilibrium. Suppose, towards a contradiction, there exists one, and denote it by (p_H^*, p_L^*). Since all p_L beyond $v(q_L - \hat{\chi})$ lead to zero demand, they must be dominated. Thus, it must be that $p_L^* \in [0, v(q_L - \hat{\chi}))$. Secondly, due to the assumption in Section A.3.4, Firm H will not lower its price from $v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + p_L$ to $\min[\theta_l(v(q_H - \hat{\chi}) - v(q_L - \hat{\chi})) + p_L, \theta_l v(q_H - \hat{\chi})]$ to serve the whole market. Therefore, Firm H's best response to p_L is

$$p_H(p_L) = v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + p_L, \tag{A.3.3}$$

which is the highest price Firm H can charge the h-segment without inducing them to switch to Firm L. Given Firm H's best response (A.3.3), however, Firm L has an incentive to undercut its price and attract the $\hat{\alpha}$-size h-segment. Firm L has incentive to undercut Firm H so long as $p_L^* > 0$. Since this implies that the only equilibrium candidate is $p_L^* = 0$ and $p_H^* = p_H(0) = v(q_H - \hat{\chi}) - v(q_L - \hat{\chi})$, this must constitute the hypothesized pure strategy equilibrium. In this case, however, Firm L can again increase its profit by unilaterally deviating to $p_L = \theta_l v(q_L - \hat{\chi})$. This positive deviation contradicts that (p_H^*, p_L^*) constitutes an equilibrium.

Next, we construct the unique mixed strategy equilibrium. To that end, we first establish the equilibrium support by eliminating dominated strategies. As explained above, Firm L's price in the interval $(v(q_L - \hat{\chi}), \infty)$ yields zero demand and is thus dominated. Now suppose there exists a non-empty set $\mathcal{P}_L^0$ of non-dominated prices on the interval $(\theta_l v(q_L - \hat{\chi}), v(q_L - \hat{\chi})]$. By non-emptiness, we can pick the largest non-dominated price for Firm L, which we denote by $\overline{p}_L = \max\{p_L : p_L \in \mathcal{P}_L^0\}$. This implies the largest non-dominated price for Firm H can be at most $\overline{p}_H = v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \overline{p}_L$: this is the price at which the h-type consumer is indifferent between buying from Firm L and Firm H; any higher price will yield zero demand. However, if $p_H \leq \overline{p}_H$, then $p_L = \overline{p}_L$ yields zero profit for Firm L, while $p_L = \theta_l v(q_L - \hat{\chi})$ yields a positive payoff of either $\theta_l v(q_L - \hat{\chi})$ (if $v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \theta_l v(q_L - \hat{\chi}) < p_h \leq v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \overline{p}_L$ such that all consumers buy from Firm L) or $(1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$ (if $p_H \leq v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \theta_l v(q_L - \hat{\chi})$ such that l-type consumers buy from Firm L). In other words, $p_L = \theta_l v(q_L - \hat{\chi})$ dominates $p_L = \overline{p}_L$. This is a contradiction because $\overline{p}_L$ was a non-dominated price. Therefore, it must be that the non-

emptiness assumption of $\mathcal{P}_L^0$ is false; i.e., there do not exist non-dominated prices on the interval $(\theta_l v(q_L - \hat{\chi}), \infty)$.

The remaining set of non-dominated prices is $[0, \theta_l v(q_L - \hat{\chi})]$. But we can reduce Firm L's support further by arguing as follows: if $p_L \in [0, \theta_l v(q_L - \hat{\chi})]$, then Firm H sets price on the interval $[v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}), v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + \theta_l v(q_L - \hat{\chi})]$. This, in turn, implies that any price p_L below $(1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$ is dominated by $p_L = \theta_l v(q_L - \hat{\chi})$. To see this, note that when Firm L sets price $p_L = \theta_l v(q_L - \hat{\chi})$, it exclusively attracts the l-type consumers for a profit of $(1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$, regardless of Firm H's price. Thus, any prices that yield payoff less than $(1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$ will be dominated. On the other hand, the maximum profit that Firm L can receive for $p_L \in [0, (1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})]$, even if Firm L manages to attract the whole market, is $(1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})$; hence the result.

The equilibrium supports of Firms H and L simplify, respectively, to

$$\mathcal{P}_H = [(1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi}) + v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}), \theta_l v(q_L - \hat{\chi}) + v(q_H - \hat{\chi}) - v(q_L - \hat{\chi})],$$

$$\mathcal{P}_L = [(1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi}), \theta_l v(q_L - \hat{\chi})].$$

The standard mixed equilibrium indifference conditions over these equilibrium supports yield the stated price distributions. □

Given the equilibrium price distributions above, the subgame equilibrium profits of Firms H and L are, respectively,

$$\pi_H^* = \hat{\alpha}\left(v(q_H - \hat{\chi}) - v(q_L - \hat{\chi}) + (1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi})\right) - \hat{k}, \tag{A.3.4}$$

$$\pi_L^* = (1 - \hat{\alpha})\theta_l v(q_L - \hat{\chi}) - \hat{k}, \tag{A.3.5}$$

where $\hat{k}$ is k if the firm advertises and 0 otherwise.

From these profit expressions, we can characterize the firms' incentives to choose particular advertising strategies.

Proposition 28 (Premium Effect). *While Firm L never chooses reference-shifting advertising that increases the reference quality, Firm H may find it profitable to do so if the premium effect, $v'(q_L - \hat\chi) - v'(q_H - \hat\chi)$, is sufficiently large.*

Proof. From (A.3.5), we can see that π_L decreases as the reference quality increases; therefore, Firm L will never choose $R^\uparrow$. On the other hand, Firm H's profit may increase in the reference quality if $\frac{\partial \pi_H}{\partial \hat\chi} = \hat\alpha(v'(q_L - \hat\chi) - v'(q_H - \hat\chi) - (1 - \hat\alpha)\theta_l v'(q_L - \hat\chi))$ is sufficiently large. In particular, this happens when $v'(q_L - \hat\chi) - v'(q_H - \hat\chi)$ is sufficiently large. This difference in the rates at which the valuations for products q_H and q_L fall as a result of a reference-shift captures the premium effect discussed in the main model. Note that this effect depends crucially on the concavity of the utility function. The effect disappears if $v(\cdot)$ is linear. □

A.3.4 Parameter Space for Duopoly Analysis

To simplify analysis, we restrict attention to the parameter space for which Firm H is not always better off serving the whole market at a low price than serving h-type consumers at a high price. To derive a sufficient condition for this, we impose that for all non-dominated price range of Firm L and for all first stage ad decisions, Firm H's maximum attainable profit when it serves the h-type is greater than that when it serves the whole market. Given Firm L's price p_L, Firm H's profit maximizing prices for each of these cases are, respectively, $p_{h\text{-only}} = v(q_H - \chi) - v(q_L - \chi) + p_L$ and $p_{\text{whole}} = \min[\theta_l(v(q_H - \chi) - v(q_L - \chi)) + p_L, \theta_l v(q_H - \chi)]$. Thus, we want to impose that

$$\min_{p_L \in \mathcal{P}_L} \pi_H(p_{h\text{-only}}) - \pi_H(p_{\text{whole}}) \geq 0,$$

where $\mathcal{P}_L$ denotes the set of Firm L's non-dominated prices. Writing out the expression for $\pi_H(p_{h\text{-only}}) - \pi_H(p_{\text{whole}})$, we obtain that the difference is equal to

$$
\begin{cases}
\alpha(v(q_H - \chi) - v(q_L - \chi) + p_L) - \theta_l(v(q_H - \chi) - v(q_L - \chi)) - p_L & \text{if } p_L \le \theta_l v(q_L - \chi), \\
\alpha(v(q_H - \chi) - v(q_L - \chi) + p_L) - \theta_l v(q_H - \chi) & \text{if } p_L > \theta_l v(q_L - \chi).
\end{cases}
$$
(A.3.6)

And since this difference is (i) continuous in p_L, (ii) decreasing in p_L for $p_L \le \theta_l v(q_L - \chi)$, and (iii) increasing in p_L for $p_L > \theta_l v(q_L - \chi)$, we obtain that the minimum is attained at $p_L = \theta_l v(q_L - \chi)$, which upon substitution yields

$$
\min_{p_L \in \mathcal{P}_L} \pi_H(p_{h\text{-only}}) - \pi_H(p_{\text{whole}}) = (\alpha - \theta_l)v(q_H - \chi) - \alpha(1 - \theta_l)v(q_L - \chi) \quad \text{(A.3.7)}
$$

But (A.3.7) is increasing in χ because the derivative with respect to χ is greater than $v'(q_L - \chi)\theta_l(1 - \alpha) > 0$. Thus, (A.3.7) is minimized at the lowest value of χ which is q_L. At $\chi = q_L$, (A.3.7) is $(\alpha - \theta_l)v(q_H - q_L)$, which is positive if and only if $\alpha < \theta_l$.

147

Appendix B: Essay 2

B.1 Proofs

B.1.1 Proof of Proposition 7

Proof. The incumbent's Stage 2 payoff is $c_I \left(1 - \frac{\max[c_E b_{E2}, R]}{c_I}\right)$ if $c_I b_{I2} \geq \max[c_E b_{E2}, R]$, and zero otherwise. 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.

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.

The incumbent's Stage 1 payoff 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 $\pi_{I2}(b_{I2} = 1|b_{E2} = 1)$ denotes the Stage 2 payoff. Since the incumbent's Stage 2 payoff is the same regardless of the outcome of the Stage 1 auction, it is immaterial when the incumbent determines its Stage 1 bid. Therefore, by the same reasoning as above, we can show that a weakly

dominant strategy in Stage 1 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, consider the publisher's revenue. In any stage, the publisher receives $c_I \, (\max[c_E, R]/c_I)$ if the incumbent wins, and $c_E \, (c_I/c_E)$ if the entrant wins. The publisher receives nothing if both advertisers' effective bids are below the reserve price. The result follows. $\qquad\square$

B.1.2 Proof of Proposition 8

We state two intermediary results which will be used for the proof (see Online Appendix for proofs).

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 $x \in (\tilde{x}_2, b]$.*

Claim 2. *If F_E is continuous, then $\frac{\partial}{\partial c_I} \int_{c_I}^{1} 1 - \frac{c_E}{c_I} dF_E = \frac{1}{c_I^2} \int_{c_I}^{1} c_E \, dF_E$.*

Proof of Proposition 8. Whether the incumbent's bid is below or above valuation depends on the sign of $g(c_I) \equiv \left(1 - \frac{\mu_E}{c_I}\right)^+ - F_E(R)\left(1 - \frac{R}{c_I}\right)^+ - \int_R^1 \left(1 - \frac{c_E}{c_I}\right)^+ dF_E$. 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

$$g(c_I) = \left(1 - \frac{\mu_E}{c_I}\right) - F_E(R)\left(1 - \frac{R}{c_I}\right) - \int_R^{c_I} 1 - \frac{c_E}{c_I} dF_E. \tag{B.1.1}$$

We will show that (B.1.1) 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.

Differentiability:

$$g'(c_I) = \frac{\partial}{\partial c_I} \left(1 - \frac{\mu_E}{c_I} - \int_0^R 1 - \frac{R}{c_I} \, dF_E - \int_R^{c_I} 1 - \frac{c_E}{c_I} \, dF_E \right)$$

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

$$= -\frac{1}{c_I^2} \int_0^1 \max[R - c_E, 0] \, dF_E + \frac{1}{c_I^2} \int_{c_I}^1 c_E \, dF_E, \qquad (B.1.2)$$

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.

Single-peakedness: From (B.1.2), it follows that the sign of $g'(c_I)$ is equal to the sign of $h(c_I) \equiv \int_{c_I}^1 c_E \, dF_E - \int_0^1 \max[R - c_E, 0] \, dF_E$. At $c_I = 0^{+1}$, h is positive because $\int_{c_I}^1 c_E \, dF_E - \int_0^1 \max[R - c_E, 0] \, dF_E = \int_{0^+}^1 c_E \, dF_E - \int_0^1 \max[R - c_E, 0] \, dF_E = \int_0^1 c_E - \max[R - c_E, 0] \, dF_E \geq \int_0^1 c_E - R \, dF_E = \mu_E - R > 0$. At $c_I = 1$, h is negative because $\int_{c_I}^1 c_E \, dF_E - \int_0^1 \max[R - c_E, 0] \, dF_E = \int_1^1 c_E \, dF_E - \int_0^1 \max[R - c_E, 0] \, dF_E = -\int_0^1 \max[R - c_E, 0] \, dF_E < 0$. Finally, $h(c_I)$ is non-increasing because for any $\delta > 0$, $h(c_I + \delta) - h(c_I) = \int_{c_I + \delta}^1 c_E \, dF_E - \int_{c_I}^1 c_E \, dF_E = -\int_{c_I}^{c_I + \delta} c_E \, dF_E \leq 0$.

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 $c_I \geq \xi$.

Endpoint values: We have $g(\mu_E) = -F_E(R) \left(1 - \frac{R}{\mu_E} \right) - \int_R^{\mu_E} \left(1 - \frac{c_E}{\mu_E} \right) dF_E < 0$ and $g(1) = 1 - \mu_E - F_E(R)(1 - R) - \int_R^1 1 - c_E \, dF_E = \int_0^R R - c_E \, dF_E > 0$.

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

Second, whether the entrant bids below or above valuation depends on the sign of $k(c_I) \equiv \int_{c_I}^1 (c_E - c_I) \, dF_E - (\mu_E - c_I)^+$. If $\mu_E \leq c_I$, then $k(c_I) = \int_{c_I}^1 c_E - c_I \, dF_E \geq 0$. If $\mu_E > c_I$, then $k(c_I) = \int_{c_I}^1 (c_E - c_I) \, dF_E - (\mu_E - c_I)^+ = -\int_0^{c_I} c_E - c_I \, dF_E > 0$. $\qquad \square$

---

[1] We evaluate at the right-limit 0^+ because $g'(c_I)$ is undefined at $c_I = 0$.

B.1.3 Proof of Proposition 9

Proof. Consider the difference in effective bids $D_A(c_I) \equiv c_I b_{I1}^*(c_I) - \mu_E b_{E1}^*(c_I)$. If $c_I \leq R$, then

$$
\begin{aligned}
D_A(c_I) &= c_I - \mu_E + \delta \left(\mu_E - R - \int_R^1 c_E - R \, dF_E \right) \\
&= c_I - \mu_E + \delta \left(\int_0^1 c_E - R \, dF_E - \int_R^1 c_E - R \, dF_E \right) \\
&= -(\mu_E - c_I) - \delta \int_0^R R - c_E \, dF_E < 0.
\end{aligned}
$$

Therefore, the entrant wins the Stage 1 auction for all $c_I \leq R$. Note that the entrant also beats the reserve price because $b_{E1}^*(c_I) \geq 1$ (cf. Lemma 2) and $\mu_E > R$.

If $R < c_I < \mu_E$, then $D_A(c_I) =$

$$
c_I - \mu_E - \delta \int_{c_I}^1 c_E - c_I \, dF_E + \delta \left(c_I - \mu_E - \int_0^1 (c_I - \max[c_E, R])^+ \, dF_E \right) < 0. \tag{B.1.3}
$$

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

Finally, if $\mu_E \leq c_I \leq 1$, then

$$
D_A(c_I) = (1 + \delta)(c_I - \mu_E) - \delta \int_{c_I}^1 c_E - c_I \, dF_E - \delta \int_0^1 (c_I - \max[c_E, R])^+ \, dF_E. \tag{B.1.4}
$$

Thus,

$$
D_A'(c_I) = 1 + \delta(1 - F_E(c_I)) + \delta \left(1 - \frac{\partial}{\partial c_I} \int_0^1 (c_I - \max[c_E, R])^+ \, dF_E \right), \tag{B.1.5}
$$

which simplifies to $1 + \delta(1 - F_E(c_I)) + \delta(1 - 1)$ if $c_I \geq \max[c_E, R]$, and $1 + \delta(1 - F_E(c_I)) + \delta(1 - 0)$, otherwise. In either case, the derivative is positive. Therefore, $D_A(c_I)$ is strictly increasing in the interval $[\mu_E, 1]$. Combined with the fact that

$$
D_A(\mu_E) = -\delta \int_{\mu_E}^1 c_E - \mu_E \, dF_E - \delta \int (\mu_E - \max[c_E, R])^+ \, dF_E < 0 \tag{B.1.6}
$$

151

and $D_A(1) = 1 - \mu_E + \delta \left(1 - \mu_E - \int 1 - \max[c_E, R]\, dF_E\right) = 1 - \mu_E + \delta \int \max[c_E, R] - c_E\, dF_E > 0$,

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 Stage 1 bid for all $c_I < \hat{c}$ and the incumbent wins for all $c_I \geq \hat{c}$.

Next, we characterize the publisher's expected payoff when it knows the entrant's CTR. If the entrant's CTR is known, then advertisers bid truthfully. Therefore, the publisher's total expected revenue is $\mathbb{E}[\pi_P^F] = (1 + \delta)\left(\int_{c_I}^1 c_I\, dF_E + \int_0^{c_I} \max[c_E, R]\, dF_E\right)$. On the other hand, if c_E is *a priori* unknown, then the publisher's expected revenue is

$$\mathbb{E}[\pi_P] = \begin{cases} \mu_E b_{E1}^* + \delta \max[\min[c_I, \mu_E], R] & \text{if } c_I b_{I1}^* \geq \mu_E b_{E1}^*, \\ \max[c_I b_{I1}^*, R] + \delta \left(\int_m^1 m\, dF_E + \int_0^m \max[c_E, R]\mathbb{I}_{\{c_I \geq \max[c_E, R]\}}\, dF_E\right) & \text{if } c_I b_{I1}^* < \mu_E b_{E1}^*. \end{cases}$$
$$(B.1.7)$$

Next, define the difference $D_\pi(c_I) \equiv \mathbb{E}[\pi_P] - \mathbb{E}[\pi_P^F]$. If $R < c_I \leq \mu_E$, then

$$D_\pi(c_I) = \max\left[c_I - \delta \int_0^{c_I} c_I - \max[c_E, R]\, dF_E, R\right] - \left(\int_0^{c_I} \max[c_E, R]\, dF_E + \int_{c_I}^1 c_I\, dF_E\right),$$

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

If $\mu_E < c_I < \hat{c}$, then $D_\pi(c_I) =$

$$\max\left[c_I + \delta\left(c_I - \mu_E - \int_0^{c_I} c_I - \max[c_E, R]\, dF_E\right), R\right] - \left(\int_0^{c_I} \max[c_E, R]\, dF_E + \int_{c_I}^1 c_I\, dF_E\right),$$

which is positive if and only if

$$c_I - \int_R^{c_I} c_E\, dF_E - \int_{c_I}^1 c_I\, dF_E - \int_0^R R\, dF_E + \delta\left(c_I - \mu_E - \int_R^{c_I} c_I - c_E\, dF_E - \int_0^R c_I - R\, dF_E\right) > 0.$$

Now, 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: $D_\pi(R) =$

$\delta(R-\mu_E) < 0$ and $D_\pi(1) = 1 - \int_R^1 c_E \, dF_E - \int_0^R R \, dF_E + \delta\left(1 - \mu_E - \int_R^1 1 - c_E \, dF_E - \int_0^R 1 - R \, dF_E\right) = 1 - \int_0^1 \max[c_E, R] \, dF_E + \delta\int_0^1 \max[c_E, R] - c_E \, dF_E > 0$. 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 $\underline{c} \equiv \max[\mu_E, \min[\hat{c}, \tilde{c}_1]]$.

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

$$D_\pi(c_I) = (1 + \delta)\left(\mu_E - \left(\int_0^R R \, dF_E + \int_R^{c_I} c_E \, dF_E + \int_{c_I}^1 c_I \, dF_E\right)\right) + \delta\int_{c_I}^1 c_E - c_I \, dF_E.$$

Similarly as above, we invoke the IVT to prove the unique existence of a root. We have $D'_\pi(c_I) = -(1 + 2\delta)(1 - F_E(c_I)) < 0$, $D_\pi(\mu_E) = (1 + \delta)\left(\int_R^{\mu_E} \mu_E - c_E \, dF_E + \int_0^R \mu_E - R \, dF_E\right) + \delta\int_{\mu_E}^1 c_E - \mu_E \, dF_E > 0$, and $D_\pi(1) = (1+\delta)\left(\mu_E - \int_R^1 c_E \, dF_E - \int_0^R R \, dF_E\right) = (1+\delta)\int_0^R c_E - R \, dF_E < 0$. Therefore, by the IVT, there exists a unique $\tilde{c}_2 \in (\mu_E, 1)$ such that $D_\pi(c_I) > 0$ for all $c_I \in (\mu_E, \tilde{c}_2)$ and $D_\pi(c_I) < 0$ 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 $\overline{c} \equiv \max[\hat{c}, \tilde{c}_2]$. Putting together all the sets for which $D_\pi > 0$ yields the result. $\qquad\square$

B.1.4 Proof of Proposition 10

Proof. The entrant overbids iff $\Delta_E > 0$; i.e., the entrant earns a higher Stage 2 profit if the publisher learns its CTR. It suffices to show that the entrant's Stage 2 profit is convex in its true CTR, for then Jensen's inequality would imply the desired result. To that end, consider the entrant's Stage 2

profit when the publisher assigns it its true CTR:

$$\pi_{E2}(c_E) = \int_0^{\bar{v}_E} U_E(x_E|c_E)\,dG_E$$

$$= \int_0^{\bar{v}_E} c_E \int_0^{x_E} q_{E2}^*(t_E)g_E(t_E)\,dt_E\,dG_E$$

$$= \int_0^{\bar{v}_E} c_E q_{E2}^*(t_E)(1 - G_E(t_E))\,dG_E(t_E)$$

$$= \int_{R_{E2}^*}^{\bar{v}_E} \int_0^{\bar{v}_I} c_E \mathbb{I}_{\{c_E > \chi(x_I, x_E|c_I)\}}(1 - G_E(x_E))\,dG_I\,dG_E$$

where $R_{E2}^* = \inf\left\{b \geq 0 : b - \frac{1 - G_E(b)}{g_E(b)} \geq 0\right\}$, $\mathbb{I}_{\{\mathcal{E}\}}$ is the indicator function which is equal to 1 if $\mathcal{E}$ is true, and 0 otherwise, and

$$\chi(x_I, x_E|c_I) = c_I\left(x_I - \frac{1 - G_I(x_I)}{g_I(x_I)}\right)\Big/\left(x_E - \frac{1 - G_E(x_E)}{g_E(x_{E1})}\right). \tag{B.1.8}$$

Since χ is independent of c_E, we obtain that for any given x_I and x_E, $c_E\mathbb{I}_{\{c_E > \xi(x_I, x_E|c_I)\}}$ is convex in c_E. And since any linear combination with positive weights of convex functions is also convex, we conclude that $\pi_{E2}(c_E)$ is convex in c_E.

Next, we turn to the incumbent. We will work with the following subgradient argument:

Claim 3. *Let* $\mathbb{E}[X] = \mu$. *If a function* $f(x)$ *that is differentiable at* $x = \mu$ *satisfies* $f(x) \geq f'(\mu)(x - \mu) + f(\mu) \geq f(x)$ *for all* x, *then* $\mathbb{E}[f(X)] \geq f(\mathbb{E}[X])$.

This follows immediately from $\mathbb{E}[f(X)] \geq \mathbb{E}[f'(\mu)(X - \mu) + f(\mu)] = f'(\mu)\mathbb{E}[X - \mu] + f(\mu) = f(\mu)$.

The incumbent's Stage 2 profit when the publisher assigns c_E is

$$\pi_{I2}(c_E) = \int_0^{\bar{v}_I} U_I(x_I|c_E, c_I)\,dG_I$$

$$= \int_0^{\bar{v}_I} c_I \int_0^{x_I} q_{I2}^*(t_I)g_I(t_I)\,dt_I\,dG_I$$

$$= \int_0^{\bar{v}_I} c_I q_{I2}^*(x_I)(1 - G_I(x_I))\,dG_I,$$

154

which simplifies to

$$
\begin{cases}
\int_{R_{I2}^*}^{\bar{v}_I} c_I(1 - G_I(x_I))z\left(\frac{c_I\eta_I(x_I)}{c_E}\right) dG_I & \text{if } \frac{c_I\bar{v}_I}{\bar{v}_E} < c_E, \\[2ex]
\int_{R_{I2}^*}^{h_I^{-1}\left(\frac{c_E\bar{v}_E}{c_I}\right)} c_I(1 - G_I(x_{I1}))z\left(\frac{c_I\eta_I(x_I)}{c_E}\right) dG_I + \int_{h_I^{-1}\left(\frac{c_E\bar{v}_E}{c_I}\right)}^{\bar{v}_I} c_I(1 - G_I(x_I)) dG_I & \text{otherwise,}
\end{cases}
$$

where $z(y) = G_E\left(\eta_E^{-1}(y)\right)$, and $\eta_j(x) = x - \frac{1-G_j(x)}{g_j(x)}$ for $j \in \{I, E\}$.

Suppose $c_I < \frac{\bar{v}_E\mu_E}{\bar{v}_I}$ such that at $\pi_{I2}(\mu_E) = \int_{R_{I2}^*}^{\bar{v}_I} c_I(1 - G_I(x_I))z\left(\frac{c_I\eta_I(x_I)}{\mu_E}\right) dG_I$ and $\pi'_{I2}(\mu_E) = \int_{R_{I2}^*}^{\bar{v}_I} c_I(1 - G_I(x_I))z'\left(\frac{c_I\eta_I(x_I)}{\mu_E}\right)\left(-\frac{c_I\eta_I(x_I)}{\mu_E^2}\right) dG_I$. By Claim 3, it suffices to show that

$$
\pi_{I2}(c_E) \geq \pi'_{I2}(\mu_E)(c_E - \mu_E) + \pi_{I2}(\mu_E) \text{ for all } c_E. \tag{B.1.9}
$$

Note that $z(\cdot) \leq 1$, which implies that

$$
\int_{h_I^{-1}\left(\frac{c_E\bar{v}_E}{c_I}\right)}^{\bar{v}_I} c_I(1 - G_I(x_I)) dG_I \geq \int_{h_I^{-1}\left(\frac{c_E\bar{v}_E}{c_I}\right)}^{\bar{v}_I} c_I(1 - G_I(x_I))z\left(\frac{c_I\eta_I(x_I)}{c_E}\right) dG_I
$$

for all c_E. Therefore, a sufficient condition for (B.1.9) is that for all c_E,

$$
\int_{R_{I2}^*}^{\bar{v}_I} (1 - G_I(x_I))z\left(\frac{c_I\eta_I(x_I)}{c_E}\right) dG_I \geq \int_{R_{I2}^*}^{\bar{v}_I} (1 - G_I(x_I))\left(z'\left(\frac{c_I\eta_I(x_I)}{\mu_E}\right)\left(-\frac{c_I\eta_I(x_I)}{\mu_E^2}\right)(c_E - \mu_E)\right.
$$
$$
\left. + z\left(\frac{c_I\eta_I(x_I)}{\mu_E}\right)\right) dG_I. \tag{B.1.10}
$$

Finally, a sufficient condition for (B.1.10) is that $\int_{R_{I2}^*}^{\bar{v}_I} (1 - G_I(x_I))z\left(\frac{c_I\eta_I(x_I)}{c_E}\right) dG_I$ be convex in c_E for all c_E. The convexity condition simplifies to

$$
\int_{R_{I2}^*}^{\bar{v}_I} (1 - G_I(x_I))z''\left(\frac{c_I\eta_I(x_I)}{c_E}\right) dG_I \geq 0 \text{ for all } c_E. \tag{B.1.11}
$$

Note that $z\left(\frac{c_I\eta_I(x_I)}{c_E}\right) = \mathbb{P}\{c_E\eta_E(x_E) \leq c_I\eta_I(x_I)\}$, which is the probability that the entrant's valuation draw is such that the incumbent wins in Stage 2. This probability can be easily verified to be decreasing in c_E. Now, condition (B.1.11) can be interpreted as this probability being "sufficiently

155

convex" for all x_I. This is equivalent to the condition that the rate of decline of the incumbent's winning probability in c_E be sufficiently low.

Finally, the sufficient condition for overbidding around the neighborhood of $c_I = 1$ follows immediately from Claim 3 and the continuity of $\pi_{I2}(c_E)$ with respect to c_I:

$$\pi_{I2}(c_E) \leq \pi'_{I2}(\mu_E)(c_E - \mu_E) + \pi_{I2}(\mu_E) \tag{B.1.12}$$

This means that the incumbent's Stage 2 profit when c_E turns out to be high is considerably low; i.e., the risk of revealing the entrant's CTR is high.[2]   □

B.1.5 Proof of Proposition 11

Proof. Let R_E^F be the entrant's reserve price under full information. R_E^F satisfies

$$R_E^F - \frac{1 - G_E(R_E^F)}{g_E(R_E^F)} = 0. \tag{B.1.13}$$

When the publisher does not know the entrant's CTR, the optimal reserve price R_E satisfies $R_E - \left(1 - G_E\left(R_E - \frac{\delta\Delta_E}{\mu_E}\right)\right)/g_E(R_E - \frac{\delta\Delta_E}{\mu_E}) = \frac{\delta}{\mu_E}(\Delta_I - \Delta_P)$. Now Assumption 1 implies that for all R, $R - \left(1 - G_E\left(R - \frac{\delta\Delta_E}{\mu_E}\right)\right)/g_E(R - \frac{\delta\Delta_E}{\mu_E}) < R - (1 - G_E(R))/g_E(R)$. Therefore, the condition $R_E < R_E^F$ is equivalent to

$$\frac{\delta}{\mu_E}(\Delta_I - \Delta_P) < R_E^F - \frac{1 - G_E\left(R_E^F - \frac{\delta\Delta_E}{\mu_E}\right)}{g_E(R_E^F - \frac{\delta\Delta_E}{\mu_E})}. \tag{B.1.14}$$

Using (B.1.13), the right-hand side simplifies to

$$\left(1 - G_E(R_E^F)\right)/g_E(R_E^F) - \left(1 - G_E\left(R_E^F - \frac{\delta\Delta_E}{\mu_E}\right)\right)/g_E(R_E^F - \frac{\delta\Delta_E}{\mu_E}), \tag{B.1.15}$$

---

[2]Numerical analyses suggest that conditions (B.1.11) and (B.1.12) are satisfied for a large class of valuation distributions G_j. For instance, if $G_E(x_E) = x_E$, condition (B.1.11) holds for any G_I, and condition (B.1.12) holds for G_I that is relatively skewed to the left; i.e., the incumbent is "strong" in the sense that it is likely to have high valuation (see Figure 3.4). This includes the "power distributions" $G_I(x_I) = \left(\frac{x_I}{\overline{v}_I}\right)^3$ for all $\overline{v}_I \geq 3/2$.

156

which is negative by Assumption 1. Labeling this negative object as $-\rho$ simplifies (B.1.14) to

$$\Delta_P > \Delta_I + \frac{\mu_E \rho}{\delta}. \qquad \qquad \Box$$

B.1.6 Proof of Proposition 12

The entrant's payoff is $\tilde{a}_E - \max[b_I, R] + \delta \int_{a_I}^{1} a_E - a_I \, d\tilde{F}$ if $b_E > \max[b_I, R]$, and $\delta(\tilde{a}_E - a_I)^+$ otherwise. The entrant's weakly dominant bid is $b_E^* = \tilde{a}_E + \delta \left(\int_0^1 (a_E - a_I)^+ \, d\tilde{F} - (\tilde{a}_E - a_I)^+ \right)$. Since $(a_E - a_I)^+$ is convex in a_E, Jensen's inequality implies that $\int_0^1 (a_E - a_I)^+ \, d\tilde{F} - (\tilde{a}_E - a_I)^+ \geq 0$; therefore, the entrant bids above its average Stage 1 per-impression valuation, $\tilde{a}_E$.

Finally, consider the incumbent's payoff:

$$\pi_I = \begin{cases} a_I - b_E + \delta(a_I - \tilde{a}_E)^+ & \text{if } b_I \geq \max[b_E, R], \\ 0 + \delta \int_0^1 (a_I - \max[a_E, R])^+ \, d\tilde{F} & \text{if } b_I < \max[b_E, R], b_E \geq R, \\ 0 + \delta(a_I - \tilde{a}_E)^+ & \text{if } b_I < \max[b_E, R], b_E < R. \end{cases} \qquad (\text{B.1.16})$$

Following the reasoning from the proof of Lemma 2 in Section **??**, we obtain that the incumbent's weakly dominant bid is $b_I^* = a_I + \delta \left((a_I - \tilde{a}_E)^+ - \int_0^1 (a_I - \max[a_E, R])^+ \, d\tilde{F} \right)$. First, note that if $a_I \leq \tilde{a}_E$, then $b_I^* = a_I - \delta \int_0^1 (a_I - \max[a_E, R])^+ \, d\tilde{F} \leq a_I$; i.e., the incumbent underbids for low a_I. Second, consider $a_I > \tilde{a}_E$. We have that b_I^* is strictly increasing in a_I in this region because $\frac{\partial b_I^*}{\partial a_I} = \frac{\partial}{\partial a_I} \left(a_I + \delta \left(a_I - \tilde{a}_E - \int_0^1 (a_I - \max[a_E, R])^+ \, d\tilde{F} \right) \right) = 1 + \delta(1 - \tilde{F}(a_I)) > 0$.

Finally, $b_I^* > a_I$ at $a_I = 1$, because $\frac{1}{\delta}(b_I^* - a_I)$ is equal to $(1 - \tilde{a}_E) - \int_0^1 (1 - \max[a_E, R])^+ \, d\tilde{F} = 1 - \tilde{a}_E - \int_0^R 1 - R \, d\tilde{F} - \int_R^1 1 - a_E \, d\tilde{F} = 1 - \left(\int_0^R 1 + a_E - R \, d\tilde{F} + (1 - \tilde{F}(R)) \right) = \tilde{F}(R) - \int_0^R 1 + a_E - R \, d\tilde{F} \geq \tilde{F}(R) - (1 + R - R)\tilde{F}(R) = 0$. In sum, b_I^* is less than a_I at $a_I = \tilde{a}_E$, greater than a_I at $a_I = 1$, and strictly increasing in a_I. Therefore, by the IVT, there exists a unique root $\tilde{a} \in (\tilde{a}_E, 1)$ such that $b_I^* < a_I$ for all $a_I < \tilde{a}$, and $b_I^* > a_I$ for all $a_I > \tilde{a}$.

B.1.7 Proof of Proposition 13

The entrant's Stage 1 payoff is $\mu_E \left(1 - \frac{\max[\tilde{c}_I b_I, R]}{\mu_E} \right) + \delta \int_R^1 \left(\int_{c_I}^1 c_E - c_I \, dF_E \right) dF_I$ if $\mu_E b_E > \tilde{c}_I b_I$, and $\delta \int_R^1 (\mu_E - c_I)^+ \, dF_I$ otherwise, where $\tilde{c}_I \sim F_I$. Even if the entrant does not know the

realization of $\tilde{c}_I$, its weakly dominant bid is $b_E^* = 1 + \frac{\delta}{\mu_E} \int_R^1 \left(\int_0^1 (c_E - c_I)^+ \, dF_E \right) - (\mu_E - c_I)^+ \, dF_I$.

And since $(c_E - c_I)^+$ is convex in c_E for all realizations of c_I, Jensen's inequality implies that $\int_0^1 (c_E - c_I)^+ \, dF_E - (\mu_E - c_I)^+ \geq 0$ for all c_I. Hence, $b_E^* \geq 1$; i.e., the entrant overbids.

B.1.8 Proof of Proposition 14

Proof. Following the argument in the main model, whether the incumbent bids below or above valuation depends on the sign of $\int_{x \in \mathcal{X}} \left((c_I - \max[\mu_x, R])^+ - \int_0^1 (c_I - \max[c_E, R])^+ \, dF_x \right) dP(x)$. But we have shown in the main model that for any distribution F_x of c_E, there exists a pair of thresholds $(\tilde{c}_1(x), \tilde{c}_2(x))$ such that the incumbent underbids for all $c_I < \tilde{c}_1(x)$ and overbids for all $c_I > \tilde{c}_2(x)$ (see proof of Proposition 8 in Section B.1.2). It follows that the integral above is negative for all $c_I < \underline{c}' \equiv \inf_{x \in \mathcal{X}} \tilde{c}_1(x)$ and positive for all $c_I > \overline{c}' \equiv \sup_{x \in \mathcal{X}} \tilde{c}_2(x)$. This completes the proof. □

B.1.9 Proof of Proposition 15

Proof. From (??), the entrant's payoff is $\mu_E \left(1 - \frac{\max[c_I b_{I1}, R]}{\mu_E} \right) + \delta \int_{c_I}^1 c_E \left(1 - \frac{c_I}{c_E} \right) dF_E + \alpha$ if $b_{E1} > \frac{\max[c_I b_{I1}, R]}{\mu_E}$, and $\delta \mu_E \left(1 - \frac{c_I}{\mu_E} \right)^+$ otherwise. Thus, a weakly dominant Stage 1 bid is $b_{E1}^*(\alpha) = 1 + \frac{\delta}{\mu_E} \left(\int_{c_I}^1 c_E - c_I \, dF_E - (\mu_E - c_I)^+ \right) + \frac{\alpha}{\mu_E}$. □

B.1.10 Proof of Proposition 16

Proof. First, we establish that from the publisher's profit perspective, offering ad credit α is equivalent to artificially increasing the entrant's effective bid by α.

Consider the latter mechanism. The Stage 2 outcomes are identical for both cases. In Stage 1, the advertisers' weakly dominants can be easily verified to be the same as the main model; i.e., the entrant bids b_{E1}^* as in (3.3.2) and the incumbent bids b_{I1}^* as in (3.3.1). Note that this implies that the effective bids are the same as the former mechanism where the publisher gives free ad credit α if the entrant wins. To see this, under the former mechanism, the entrant's effective bid is $\mu_E b_{E1}^*(\alpha)$ as in (3.5.1). But from (3.5.1), we have $\mu_E b_{E1}^*(\alpha) = \mu_E b_{E1}^* + \alpha$, which is equivalent to entrant's effective bid under the artificial additive boosting mechanism. The incumbent's effective

bids are trivially the same. Therefore, the mechanisms have the same allocation rule.

Moreover, the payoffs of the two mechanisms are identical. If the incumbent wins in Stage 1, the publisher's Stage 1 profit is $\mu_E b_{E1}^* + \alpha$ in both cases, and if the entrant wins, it is $\max[c_I b_{I1}^*] - \alpha$ in both cases. In sum, the two mechanism yield the same profit for the publisher.

Now, consider the boosting multiplier β. For any given additive term α, if the publisher sets $\beta(\alpha) = 1 + \frac{\alpha}{\mu_E b_{E1}^*}$, then the advertisers' effective bids are the same as in the mechanism wherein the publisher adds α to the entrant's effective bid. Therefore, the two mechanisms have the same allocation rule.

Furthermore, note that the advertisers' bids are the same for both mechanisms: the incumbent (entrant) bids b_{I1}^* (b_{E1}^*). This means that both mechanisms can be cast as a "direct mechanism" where advertisers bid their true "type." Therefore, by the Revenue Equivalence principle (Myerson, 1981), the two mechanisms yield the same expected profit up to a constant. $\square$

B.1.11 Proof of Proposition 17

Proof. From the derivation of the LREF Stage 1 bids (see Section **??** of the online appendix), we see that an important condition that shapes the outcome of the auction is

$$\delta \left(\mathbb{E}[\pi_E^l] - \left(\mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l] \right) \right) \le \theta(c_i - \mu_E). \tag{B.1.17}$$

We can write this condition in terms of the reserve price R. First, note that $\mathbb{E}[\pi_E^l] = \int_{c_i}^{c_I} \theta(c_E - c_i) \, dF_E + \int_{c_i}^{c_I} (c_E - c_I) + \theta(c_I - c_i) \, dF_E$ is independent of R, and

$$\mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l] = c_I - \mu_E - \int_0^1 (c_i - \max[c_E, R])^+ \, dF_E = \int_0^R R - c_E \, dF_E - \int_{c_i}^1 c_E - c_i \, dF_E. \tag{B.1.18}$$

Second, since (B.1.18) is strictly increasing in R, we obtain that (B.1.17) is equivalent to $R \ge \hat{R}$, where $\hat{R}$ solves $\delta \left(\mathbb{E}[\pi_E^l] - \left(\mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l] \right) \right) = \theta(c_i - \mu_E)$.

Suppose $\mu_E \le c_i < c_I$. Following the derivation of LREF bids in Section **??**, if $R > \hat{R}$, then the LREF equilibrium is $[I, i, e]$ and the entrant bids $b_E^*(\delta) = 1 + \frac{\delta}{\theta \mu_E} \mathbb{E}[\pi_E^l]$. Therefore, $b_E^*(\delta) \ge b_E^*(0)$

for all $\delta > 0$. For the weak incumbent, we obtain $b_i^*(\delta) = 1 - \theta + \frac{\theta\mu_E}{c_i}b_E^*(\delta)$; therefore, $b_i^*(\delta) > b_i^*(0)$.

If $R < \hat{R}$, then the LREF equilibrium is $[I, E, i]$ and $b_E^*(\delta) = 1 - \theta + \frac{\theta}{\mu_E}\max[c_i b_i^*(\delta), R]$ and $b_i^*(\delta) = \left(1 + \frac{\delta}{\theta c_i}\left(\mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l]\right)\right)^+$. To determine whether the advertisers bid below or above their learning-free benchmarks, we need to determine the sign of $\mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l]$. Since (i) $\frac{\partial (B.1.18)}{\partial R} > 0$, (ii) $(B.1.18) = -\int_{c_i}^1 c_E - c_i \, dF_E < 0$ at $R = 0$, and (iii) $(B.1.18) = c_i - \mu_E > 0$ at $R = c_i$, there exists a unique $\tilde{R} \in (0, c_i)$ such that $\mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l] < 0$ for all $R < \tilde{R}$ and $\mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l] > 0$ for all $R > \tilde{R}$. Thus, $R < \tilde{R} \Rightarrow b_E^*(\delta) \le b_E^*(0)$ and $b_i^*(\delta) \le b_i^*(0)$ for all $\delta > 0$; and $R > \tilde{R} \Rightarrow \mathbb{E}[\pi_i^0] - \mathbb{E}[\pi_i^l] > 0$, so that $b_E^*(\delta) \ge b_E^*(0)$ and $b_i^*(\delta) \ge b_i^*(0)$ for all $\delta > 0$.

Suppose $c_i < \mu_E \le c_I$. The LREF equilibrium is $[I, E, i]$. The EF conditions for the strong incumbent and the entrant are $c_I(1 - \max[\mu_E b_E, R]/c_I) + \delta\mathbb{E}[\pi_i^l] \ge \theta c_I(1 - \max[c_i b_i, R]/c_I) + \delta\mathbb{E}[\pi_i^l] \iff \mu_E b_E \le (1 - \theta)c_I + \theta\max[c_i b_i, R]$ and $\mu_E(1 - \max[\mu_E b_E, R]/\mu_E) + \delta\mathbb{E}[\pi_E^l] \le \theta\mu_E(1 - \max[c_i b_i, R]/\mu_E) + \delta\mathbb{E}[\pi_E^l] \iff \mu_E b_E \ge (1 - \theta)\mu_E + \theta\max[c_i b_i, R]$. Therefore, the entrant's LREF bid is $b_E^*(\delta) = 1 - \theta + \frac{\theta}{\mu_E}\max[c_i b_i^*, R]$.

The EF conditions for the the weak incumbent and the entrant are $\theta c_i(1 - \max[c_i b_i, R]/c_i) + \delta\mathbb{E}[\pi_i^0] \le \delta\mathbb{E}[\pi_i^l] \iff \theta c_i - \delta\mathbb{E}[\pi_i^l] \le \theta\max[c_i b_i, R]$ and $\theta\mu_E(1 - \max[c_i b_i, R]/\mu_E) + \delta\mathbb{E}[\pi_E^l] \ge \delta\mathbb{E}[\pi_E^0] \iff \theta\max[c_i b_i, R] \le \theta\mu_E + \delta(\mathbb{E}[\pi_E^l] - \mathbb{E}[\pi_E^0])$. Since the LREF bid binds at the lower-bound, we have $\max[c_i b_i^*, R] = c_i - \frac{\delta}{\theta}\mathbb{E}[\pi_i^l]$. Thus, $R < c_i - \frac{\delta}{\theta}\mathbb{E}[\pi_i^l] \Rightarrow b_i^*(\delta) = 1 - \frac{\delta}{\theta c_i}\mathbb{E}[\pi_i^l]$ and $b_E^*(\delta) = 1 - \theta + \frac{\theta}{\mu_E}\left(c_i - \frac{\delta}{\theta}\mathbb{E}[\pi_i^l]\right)$, such that $b_E^*(\delta) \le b_E^*(0)$ and $b_i^*(\delta) \le b_i^*(0)$ for all $\delta > 0$.

On the other hand, if $R \ge c_i - \frac{\delta}{\theta}\mathbb{E}[\pi_i^l]$, then the lowest bid b_i that satisfies the EF condition, $c_i - \frac{\delta}{\theta}\mathbb{E}[\pi_i^l] \le \max[c_i b_i, R] \le \mu_E + \frac{\delta}{\theta}(\mathbb{E}[\pi_E^l] - \mathbb{E}[\pi_E^0])$, is $b_i^* = 0$. In this case, the LREF bids are $b_E^*(\delta) = 1 - \theta + \frac{\theta}{\mu_E}R$ and $b_I^*(\delta) = 0$, such that the bids are the same across environments with and without learning. For this to hold for all $\delta > 0$, we must have $R \ge c_i$.

Taken together, we can define the threshold for the reserve price R above which the stated result holds: $\overline{R}$ equals $\min\left[\hat{R}, \tilde{R}\right]^+$ if $\mu_E \le c_i$, and equals c_i otherwise. $\qquad\square$

Appendix C: Essay 3
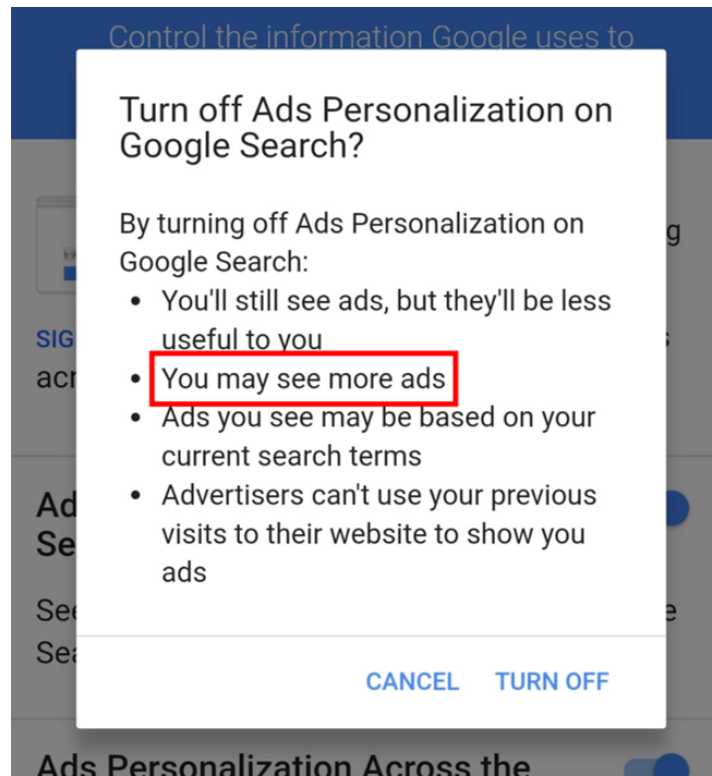
C.1 Sample Privacy Notice

Figure C.1: Google's Privacy Notice in Europe (July 2019)

C.2 Proofs

C.2.1 Statement and Proof of Claim 4

Claim 4. *Suppose a player's payoff from bidding b in an auction parametrized by tuple* (x, y, z, p) *is*

$$
\pi(b) = \begin{cases} x - y\,p & \text{if } b \geq p, \\ z & \text{if } b < p, \end{cases}
$$

161

where $y > 0$ and $p > 0$. Then the player's weakly dominant bid (i.e., robust to any p) is $b^* = (x - z)^+/y$.

Proof. First, if $z \geq x$, then winning leads to strictly lower profit than losing. Therefore, the optimal bid is to lose for any p; hence $b^* = 0$.

Second, suppose $x > z$. We show that there is no strictly dominant deviation strategy for $b^* = (x - z)^+/y$. To that end, consider a deviation b' that is strictly less than $\frac{x-z}{y}$. Then for $p \in \left(b', \frac{x-z}{y}\right)$, we have $\pi(b') = z = x - y\left(\frac{x-z}{y}\right) < x - yp = \pi(b^*)$. For all other ranges of p, the two strategies yield the same payoff. Therefore, b^* weakly dominates b'.

Next, consider another deviation b'' that is strictly greater than $\frac{x-z}{y}$. Then for $p \in \left(\frac{x-z}{y}, b''\right)$, we have $\pi(b') = x - yp < x - y\left(\frac{x-z}{y}\right) = z = \pi(b^*)$. Again, the two strategies yield the same payoff for all other ranges of p. This completes the proof. $\square$

C.2.2 Statement and Proof of Claim 5

Claim 5. *Let $f(x) = \max[x, 0]$ for all $x \in \mathbb{R}$. Then $f(x) + f(y) \geq f(x + y)$ for all $x, y \in \mathbb{R}$.*

Proof.

$$\frac{1}{2}\left(f(x) + f(y)\right) = f\left(\frac{x}{2}\right) + f\left(\frac{y}{2}\right) \geq f\left(\frac{x}{2} + \frac{y}{2}\right) = \frac{1}{2}f(x + y)$$

where the equalities are due to linearity and inequality due to convexity. $\square$

Proof of Lemma 5

Proof. Consider the first subgame wherein the advertiser had shown ads in Period 1. Following Claim 4, the advertiser's weakly dominant bid (against any reserve price R_2) in Period 2 is

$$b^*_{2|\text{ad}} = (1 - \mu)\mu\phi_M + \mu(1 - \phi_M)(\beta + (1 - \beta)\phi_M) - \mu(1 - \phi_M)\phi_M$$

$$= (1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta.$$

Similarly, the advertiser's weakly dominant Period 2 bid in the second subgame, wherein it did not advertise in Period 1, is $b^*_{2|\text{no ad}} = \mu\phi_M$. For each of the Period 2 subgames described above, the ad network sets R_2 as high as b^*_2, provided it is larger than k. Thus, we obtain the optimal Period 2 reserve prices $R^*_{2|\text{ad}} = \max\left[k, b^*_{2|\text{ad}}\right]$ and $R^*_{2|\text{no ad}} = \max\left[k, b^*_{2|\text{no ad}}\right]$. $\qquad\square$

C.2.3 Proof of Lemma 6

Proof. The advertiser's weakly dominant bid b^*_1 in Period 1 follows directly from Claim 4. For the ad network's optimal reserve price, consider its Period 1 payoff:

$$\pi^N(R_1) = \begin{cases} R_1 - k + (b^*_{2|\text{ad}} - k)^+ & \text{if } R_1 \leq b^*_1, \\ 0 + (b^*_{2|\text{no ad}} - k)^+ & \text{otherwise.} \end{cases}$$

It follows that $R^*_1 = b^*_1$ if $b^*_1 - k + (b^*_{2|\text{ad}} - k)^+ \geq (b^*_{2|\text{no ad}} - k)^+$, and $R^*_1 \in (b^*_1, \infty)$ otherwise. The reserve price stated in the lemma satisfies this property. $\qquad\square$

C.2.4 Proof of Proposition 18

Proof. Given the reserve prices derived above, the ad network's profit in Period 2 if ads were shown in Period 1 is $(1-\mu)\mu\phi_M + \mu(1-\phi_M)(\beta + (1-\beta)\phi_M) - \mu(1-\phi_M)\phi_M - k = (1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k$, if the ad network sells Period 2 ads, and 0 otherwise. Therefore, the ad network's Period 2 profit given ads were shown in Period 1 is $\left((1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k\right)^+$. Similarly, if ads were not shown in Period 1, then the ad network's Period 2 profit is $(\mu\phi_M - k)^+$.

Thus, the ad network's total profit from setting reserve price R_1 in Period 1 is

$$\pi^N(R_1) = \begin{cases} R_1 - k + \left((1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k\right)^+ & \text{if } R_1 \leq b^*_1, \\ 0 + (\mu\phi_M - k)^+ & \text{if } R_1 > b^*_1, \end{cases}$$

from which we obtain

$$R_1^* = \begin{cases} b_1^* & \text{if } b_1^* - k + \left((1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k\right)^+ \ge (\mu\phi_M - k)^+, \\[2mm] (b_1^*, \infty) & \text{otherwise.} \end{cases}$$

Since R_1^* can be any number greater than b_1^* when

$$b_1^* < k + (\mu\phi_M - k)^+ - \left((1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k\right)^+, \tag{C.2.1}$$

we can write

$$R_1^* = \max\left[k + (\mu\phi_M - k)^+ - \left((1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k\right)^+, \; \mu\phi_M + \mu(1-\phi_M)\phi_M\right].$$

Next, we derive the conditions under which the advertiser's weakly dominant bids exceed the optimal reserve prices set by the ad network.

Ads Shown Only in Period 2

We first show that showing ads only in Period 2 is never an equilibrium outcome. Towards a contradiction, suppose the conditions for such an equilibrium hold; i.e., $b_1^* < R_1^*$ and $\mu\phi_M - k \ge 0$. But $\mu\phi_M - k \ge 0$ implies that

$$R_1^* = \max\left[k + (\mu\phi_M - k) - \left((1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k\right)^+, \; \mu\phi_M + \mu(1-\phi_M)\phi_M\right], \tag{C.2.2}$$

which simplifies to

$$\max\left[\mu\phi_M - \left((1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta - k\right)^+, \; \mu\phi_M + \mu(1-\phi_M)\phi_M\right]. \tag{C.2.3}$$

This is strictly greater than $b_1^* = \mu\phi_M + \mu(1 - \phi_M)\phi_M$ if and only if $\mu\phi_M + \mu(1 - \phi_M)\phi_M <$ $\mu\phi_M - \left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta - k\right)^+$, which is equivalent to

$$\mu(1 - \phi_M)\phi_M < -\left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta - k\right)^+. \tag{C.2.4}$$

Since the left-hand side is strictly positive while the right-hand side is non-positive, this inequality never holds. A contradiction.

Ads Shown in Periods 1 and 2

The advertiser buys untargeted ads in both periods if and only if $b_1^* \geq R_1^*$ and $b_{2|\text{ad}}^* \geq R_{2|\text{ad}}^*$, which are equivalent to

$$\mu\phi_M + \mu(1 - \phi_M)\phi_M \geq k + (\mu\phi_M - k)^+ - \left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta - k\right)^+ \tag{C.2.5}$$

and

$$(1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta \geq k, \tag{C.2.6}$$

respectively. Note that (C.2.6) implies that (C.2.5) simplifies to

$$(1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta - k \geq k + (\mu\phi_M - k)^+ - (\mu(2 - \phi_M)\phi_M), \tag{C.2.7}$$

which can be re-arranged in terms of β as

$$\beta \geq \tilde{\beta} \equiv \frac{2k + (\mu\phi_M - k)^+ - \mu(3 - \mu - \phi_M)\phi_M}{\mu(1 - \phi_M)^2}.$$

The intersection of conditions (C.2.6) and (C.2.7) simplifies to (C.2.6) if $\mu > \frac{k}{\phi_M(2 - \phi_M)}$, and to (C.2.7) if $\mu < \frac{k}{\phi_M(2 - \phi_M)}$. These branching conditions in turn can be re-written as $\phi_M > 1 - \frac{\sqrt{(\mu - k)^+}}{\sqrt{\mu}}$ and $\phi_M < 1 - \frac{\sqrt{(\mu - k)^+}}{\sqrt{\mu}}$, respectively.

Ads Shown Only in Period 1

The advertiser buys only Period 1 ads if and only if $b_1^* \geq R_1^*$ and $b_{2|ad}^* < R_{2|ad}^*$. But if the second condition holds, the first simplifies to

$$\mu\phi_M - k + \mu(1 - \phi_M)\phi_M \geq (\mu\phi_M - k)^+, \tag{C.2.8}$$

which holds if $\mu\phi_M \geq k$. If $\mu\phi_M < k$, then (C.2.8) simplifies to $k \leq \mu\phi_M(2 - \phi_M)$. This last inequality can be re-arranged in terms of ϕ_M as $\phi_M \geq 1 - \frac{\sqrt{(\mu-k)^+}}{\sqrt{\mu}}$. In total, the intersection of the two conditions simplifies to $\phi_M \geq 1 - \frac{\sqrt{\mu(\mu-k)^+}}{\mu}$ and $(1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta < k$. The latter condition can be simplified using its concavity. To that end, let

$$\underline{\mu} = \frac{\phi_M + \beta(1 - \phi_M)^2 - \sqrt{\left(\beta(1 - \phi_M)^2 + \phi_M\right)^2 - 4k\phi_M}}{2\phi_M},$$

$$\overline{\mu} = \frac{\phi_M + \beta(1 - \phi_M)^2 + \sqrt{\left(\beta(1 - \phi_M)^2\right)^2 - 4k\phi_M} + \phi_M}{2\phi_M}$$

be the two roots of $(1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta = k$. The larger root $\overline{\mu}$ is greater than 1 for all β greater than $\overline{\beta} = \frac{k}{(1-\phi_M)^2}$, and the roots do not exist for all β smaller than $\underline{\beta} = \frac{2\sqrt{k\phi_M}-\phi_M}{(1-\phi_M)^2}$. Algebraic manipulations yield the conditions stated in the proposition.

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

C.2.5   Proof of Proposition 19

Proof. We derive the equilibrium strategies for two subgames: one in which the advertiser showed its ad in Period 1, and the other in which it did not.

First, consider the advertiser's Period 2 bidding problem when it has shown ads in Period 1. Let R_2^i be the reserve prices for impression type $i \in \{T, M, TM\}$. Impression type T (M) denotes the impression for which the consumer is in funnel state T (M), and TM denotes the impression for which the consumer is in either funnel state T or M.

Suppose the advertiser submits bid b_2^i for impression type i. The advertiser's payoff is $\pi_2 =$

$$
\begin{cases}
(1-\mu)\left(\mu\phi_M - R_2^T\right) + \mu(1-\phi_M)\left(\beta + (1-\beta)\phi_M - R_2^M\right) & \text{if } b_2^T \geq R_2^T, b_2^M \geq R_2^M, \\[2mm]
(1-\mu)(\mu\phi_M - R_2^T) + \mu(1-\phi_M)\phi_M & \text{if } b_2^T \geq R_2^T, b_2^M < R_2^M, \\[2mm]
\mu(1-\phi_M)(\beta + (1-\beta)\phi_M - R_2^M) & \text{if } b_2^T < R_2^T, b_2^M \geq R_2^M, \\[2mm]
\mu(1-\phi_M)\phi_M & \text{if } b_2^T < R_2^T, b_2^M < R_2^M.
\end{cases}
$$

Whether $b_2^M \geq R_2^M$ or $b_2^M < R_2^M$, the weakly dominant bid for the T-impression is $b_2^{T*} = \mu\phi_M$. And regardless of b_2^T, the weakly dominant bid for the M-impression is $b_2^{M*} = \beta(1-\phi_M)$.

Next, consider the advertiser's payoff from bidding for TM: $\pi_{2|\text{ad}}^A\left(b_2^{TM}\right) =$

$$
\begin{cases}
(1-\mu)\mu\phi_M + \mu(1-\phi_M)\left(\beta + (1-\beta)\phi_M\right) - ((1-\mu) + \mu(1-\phi_M))\,R_2^{TM} & \text{if } b_2^{TM} \geq R_2^{TM}, \\[2mm]
\mu(1-\phi_M)\phi_M & \text{if } b_2^{TM} < R_2^{TM}.
\end{cases}
$$

It follows that $b_2^{TM*} = \frac{(1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta}{(1-\mu) + \mu(1-\phi_M)}$.

The ad network anticipates b_2^{i*} for $i \in \{T, M, TM\}$ and sets R_2^i that maximizes its Period 2 profit. There are four candidates that the ad network considers:

$$
\left(R_2^T, R_2^M, R_2^{TM}\right) =
\begin{cases}
(\max[k, \mu\phi_M], \infty, \infty) & \text{induces } T\text{-ad sales,} \\[2mm]
(\infty, \max[k, \beta(1-\phi_M)], \infty) & \text{induces } M\text{-ad sales,} \\[2mm]
(\max[k, \mu\phi_M], \max[k, \beta(1-\phi_M)], \infty) & \text{induces } T\text{- and } M\text{-ad sales,} \\[2mm]
\left(\infty, \infty, \max\left[k, \frac{(1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta}{(1-\mu)+\mu(1-\phi_M)}\right]\right) & \text{induces } TM\text{-ad sales.}
\end{cases}
$$

If the ad network chooses the first candidate, then only Period 2 impressions for consumers in funnel state T are potentially sold. Since the size of T-consumers in Period 2 is $1 - \mu$, this strategy yields ad network profit $(1-\mu)(\mu\phi_M - k)^+$. Similarly, the second candidate yields profit $\mu(1-\phi_M)(\beta(1-\phi_M) - k)^+$, the third $(1-\mu)(\mu\phi_M - k)^+ + \mu(1-\phi_M)(\beta(1-\phi_M) - k)^+$, and

the fourth $((1 - \mu)(\mu\phi_M - k) + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k))^+$. From Claim 5, it follows that the third candidate $(R_2^T, R_2^M, R_2^{TM}) = (\max[k, \mu\phi_M], \max[k, \beta(1 - \phi_M)], \infty)$ yields the highest payoff. Therefore, provided ads are shown in Period 1, ads are shown to T-consumers in Period 2 if and only if $\mu\phi_M \geq k$, and ads are shown to M-consumers if and only if $\beta(1 - \phi_M) \geq k$.

Next, consider the second subgame wherein the advertiser did not show ads in Period 1. Then in Period 2, the advertiser's payoff from bidding b_2, given reserve price R_2 is

$$\pi^A_{2|\text{no ad}}(b_2) = \begin{cases} \mu\phi_M - R_2 & \text{if } b_2 \geq R_2 \\ 0 & \text{if } b_2 < R_2. \end{cases}$$

By similar reasoning as above, it follows that $b_2^* = \mu\phi_M$ and $R_2^* = \max[k, \mu\phi_M]$. The ad network's Period 2 payoff in this subgame is $(\mu\phi_M - k)^+$.

With the subgame results at hand, we can solve for the Period 1 game. The advertiser's total payoff from bidding b_1 in Period 1, given reserve price R_1, is

$$\pi^A(b_1) = \begin{cases} \mu\phi_M - R_1 + \mu(1 - \phi_M)\phi_M & \text{if } b_1 \geq R_1, \\ 0 & \text{if } b_1 < R_1, \end{cases}$$

where the term $\mu(1 - \phi_M)\phi_M$ represents the advertiser's Period 2 payoff when it shows ads in Period 1. Claim 4 implies that the advertiser's weakly dominant bid is $b_1^* = \mu\phi_M + \mu(1 - \phi_M)\phi_M$. The ad network anticipates this and sets the reserve price as high as b_1^*, provided $R_1 - k + (1 - \mu)(\mu\phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+ \geq (\mu\phi_M - k)^+$; i.e.,

$$R_1^* = \max\left[k - \left((1 - \mu)(\mu\phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+\right) + (\mu\phi_M - k)^+, b_1^*\right].$$

Therefore, Period 1 ads are shown if and only if $b_1^* \geq R_1^*$, which is equivalent to

$$\mu\phi_M + \mu(1 - \phi_M)\phi_M \geq k - \left((1 - \mu)(\mu\phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+\right) + (\mu\phi_M - k)^+.$$

$$\text{(C.2.9)}$$

Suppose $\mu\phi_M \geq k$. Then (C.2.9) simplifies to

$$\mu(1 - \phi_M)\phi_M \geq -\left((1 - \mu)(\mu\phi_M - k) + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+\right), \qquad \text{(C.2.10)}$$

which is true. Suppose $\mu\phi_M < k$. Then (C.2.9) simplifies to

$$\mu \geq \tilde{\mu} \equiv k\left(\phi_M(2 - \phi_M) + (1 - \phi_M)(\beta(1 - \phi_M) - k)^+\right)^{-1}. \qquad \text{(C.2.11)}$$

Thus, Period 1 ads are shown if and only if either $\mu \geq \frac{k}{\phi_M}$ or $\tilde{\mu} \leq \mu < \frac{k}{\phi_M}$. Since $\frac{k}{\phi_M} \geq \tilde{\mu} \Leftrightarrow$ $\phi_M \leq \phi_M(2 - \phi_M) + (1 - \phi_M)(\beta(1 - \phi_M) - k)^+ \Leftarrow \phi_M \leq \phi_M(2 - \phi_M) \Leftrightarrow 1 \leq 2 - \phi_M$, which is true for all $\phi_M \in [0, 1]$, we obtain that Period 1 ads are shown if and only if $\mu \geq \tilde{\mu}$. $\qquad \square$

C.2.6 Proof of Proposition 20

Proof. Let $q^*(0)$ and $q^*(1)$ denote the equilibrium ad intensities without and with tracking, respectively. We being the proof with three observations. First, note that $q^*(1) < 2$ because with tracking, ads are not shown to consumers who had already purchased. Second, if $q^*(0) > 0$, then $q^*(1) > 0$. To see this, suppose that $q^*(0) = 1$. Then under tracking, the ad network can replicate this no-tracking payoff by showing ads only in Period 1. Similarly, if $q^*(0) = 2$, then under tracking, the ad network can generate a weakly higher profit by showing ads to all consumers except those who already purchased. In either case, the ad network's profit under tracking when it shows ads is higher than not showing any ads, because $q^*(0) > 0$ implies showing ads generates positive surplus. Therefore, $q^*(0) > 0$ implies $q^*(1) > 0$.

Put together, we obtain that $q^*(0) > q^*(1)$ if and only if $q^*(0) = 2$. The condition for $q^*(0) = 2$ is given in Proposition 18. Moreover, $q^*(0) = q^*(1)$ if and only if either $q^*(0) = q^*(1) = 0$ or

$q^*(0) = q^*(1) = 1$. The ad intensities are both zero if and only if $\mu < \tilde{\mu}$ (such that $q^*(1) = 0$) and $\beta < \tilde{\beta}$ and $\phi_M < 1 - \frac{\sqrt{(\mu-k)^+}}{\sqrt{\mu}}$ (such that $q^*(0) = 0$). But $\mu < \tilde{\mu}$ implies $\phi_M < 1 - \frac{\sqrt{(\mu-k)^+}}{\sqrt{\mu}}$, so the condition for $q^*(0) = q^*(1) = 0$ simplifies to $\mu < \tilde{\mu}$ and $\beta < \tilde{\beta}$.

Next, we derive the conditions under which the ad intensities are 1 in either tracking scenario. First, note that if $q^*(1) = 1$, then $q^*(0) < 2$. This is because $q^*(1) = 1$ implies that not showing ads in Period 2 under tracking is better than showing. And since showing ads in Period 2 with tracking yields weakly higher profit than showing ads in Period 2 without tracking, we obtain by transitivity that without tracking, not showing ads in Period 2 is more profitable than showing ads. Therefore, the condition $q^*(1) = 1$ and $q^*(0) = 1$ are jointly satisfied if and only if $\tilde{\mu} < \mu \leq \frac{k}{\phi_M}$ (such that $q^*(1) = 1$) and $\mu > \frac{k}{\phi_M(2-\phi_M)}$ (such that $q^*(0)$ is either 1 or 2). In total, $q^*(0) = q^*(1) = 1$ if and only if $\frac{k}{\phi_M(2-\phi_M)} < \mu \leq \frac{k}{\phi_M}$ and $\beta \leq \frac{k}{1-\phi_M}$. $\qquad\square$

C.2.7 Proof of Proposition 21

Proof. If ads are not shown in Period 1, then the ad network's Period 2 payoffs with and without tracking are the same at $(\mu\phi_M - k)^+$. On the other hand, if ads are shown in Period 1, then the Period 2 subgame under tracking yields the following ad network payoff $\pi^N_{2|\text{ad}} = (1 - \mu)(\mu\phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+$. The ad network's payoff under no tracking is $\pi^N_{2|\text{no ad}} = ((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)\beta(1 - \phi_M) - k)^+$. But we have

$$
\begin{aligned}
\pi^N_{2|\text{no ad}} &= ((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)\beta(1 - \phi_M) - k)^+ \\
&\leq ((1 - \mu)(\mu\phi_M - k) + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k))^+ \\
&\leq (1 - \mu)(\mu\phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+ \\
&= \pi^N_{2|\text{ad}}.
\end{aligned}
$$

Finally, in Period 1, the ad network faces the same problem with and without tracking, except that it anticipates a higher Period 2 payoff with tracking if ads are shown in Period 1. Therefore, the total profit is weakly greater with tracking than without. $\qquad\square$

C.2.8 Proof of Proposition 22

Proof. Consumers opt-in to tracking only if $q^*(0) > q^*(1)$. But recall from from Proposition 20 that $q^*(0) > q^*(1)$ if and only if $q^*(0) = 2$. Therefore, the necessary condition for opting-in is $q^*(0) = 2$. The sufficient condition is that the consumer's privacy cost is low enough that the benefit of seeing fewer ads outweighs the privacy cost of opting-in. The marginal consumer is the consumer with cost $\min[1, \tilde{\theta}]$ such that $-\eta q^*(1) - \tilde{\theta} = -\eta q^*(0)$. $\qquad\square$

C.2.9 Proof of Proposition 23

Proof. Ad network's profit can decrease in μ due to two and only two reasons: (a) higher μ implies lower opt-in rate such that ad network profit decreases towards the opt-out profit, which is lower than opt-in profit, and (b) large μ implies higher ad intensity under tracking such that consumers opt-out.

The first part occurs if and only if $q^*(0) = 2$ and $q^*(1) = 1 + \mu(1 - \phi_M)$; i.e., under tracking, ads are only shown to M-consumers. If ads were shown to T-consumers as well, $q^*(1)$ would decrease in μ such that opt-in rate increases with μ. The opt-in rate is $F(\eta(2 - (1 + \mu(1 - \phi_M)))) = F(\eta(1 - \mu(1 - \phi_M)))$, which decreases in μ if and only if $\eta(1 - \mu(1 - \phi_M)) \in (0, 1)$. For the uniform distribution $F(\theta) = \theta$, the ad network's profit is

$$\pi_N = \mu\phi_M + \mu(1 - \phi_M)(\beta + (1 - \beta)\phi_M) - (1 + \mu(1 - \phi_M))k \tag{C.2.12}$$
$$+ (1 - \eta(1 - \mu(1 - \phi_M)))((1 - \mu)\mu\phi_M - (1 - (1 - \mu)\phi_M)k).$$

We want to find the conditions under which (C.2.12) decreases in μ. Note that

$$\frac{\partial \pi_N}{\partial \mu} = \beta + 2\eta k(\mu - 1) - \phi_M \left(2\beta + \eta\left(4k\mu - 2k + 3\mu^2 - 4\mu + 1\right) + 2\mu - 3\right) + \phi_M^2(\beta + \eta\mu(2k + 3\mu - 2) - 1).$$

Since the second derivative of the above is $-6(1 - \phi_M)\phi_M < 0$, we have that the derivative is concave in μ. Therefore, (C.2.12) is decreasing in μ for $\mu < \underline{\mu}'$ and $\mu > \overline{\mu}'$ where the thresholds are respectively given by the two roots of (C.2.12) in increasing order.

The second part follows from Proposition 22: if $\phi_M > 1 - \sqrt{(\mu - k)^+/\mu}$ and $\underline{\beta} \leq \beta < \overline{\beta}$, then for $\mu = \overline{\mu}^-$, consumers opt-in, and for $\mu = \overline{\mu}^+$, consumers opt-out. The efficiency loss

associated with the increase in opt-out rate creates downward jump in the ad network's profit (cf. Proposition 21). □

C.2.10 Proof of Proposition 24

Proof. Denote by $q(1)$ and $q(0)$ the total expected ad intensity with and without tracking, respectively. Furthermore, denote by $CS(1)$ and $CS(e)$ the total consumer surplus with full and endogenous tracking, respectively. Let $\tilde{\theta} = \max[0, \min[1, \eta(q(0) - q(1))]]$. Then the result follows from

$$CS(e) = \int_0^{\tilde{\theta}} -\eta q^*(1) - \theta \, dF + \int_{\tilde{\theta}}^1 -\eta q^*(0) \, dF \geq \int_0^{\tilde{\theta}} -\eta q^*(1) - \theta \, dF + \int_{\tilde{\theta}}^1 -\eta q^*(1) - \theta \, dF = CS(1).$$

□

C.2.11 Proof of Proposition 25

Proof. We first show that opting-out of tracking does not signal the consumer's types. Let ρ_i and ρ_j denote advertiser i and advertiser j's beliefs, respectively, that the consumer behind the opt-out impression is type i. By Bayes' rule, the beliefs must satisfy

$$\rho_i = \frac{\lambda S_i(\rho_i, \rho_j)}{\lambda S_i(\rho_i, \rho_j) + (1 - \lambda) S_j(\rho_i, \rho_j)},$$

where $S_i(\rho_i, \rho_j)$ denotes the mass of type i consumers who choose to opt-out given advertisers' beliefs ρ. But a type i consumer will opt-out if and only if

$$-\theta - \eta \, q_i(1) < -\eta \, q_i(0; \rho_i, \rho_j),$$

where $q_i(1)$ and $q_i(0; \rho_i, \rho_j)$ is the total number of ads a type i consumer expects to see if she opts-in and -out, respectively. But $q_i(1)$ is independent of consumer's type i because if a consumer opts-in to tracking, the number of ads she expects to see depends only on the parameters μ, β and

172

ϕ_M. Similarly, $q_i(0; \rho_i, \rho_j)$ is independent of consumer's type i because by definition, advertisers cannot base their strategies on consumers' types if they opt-out. Therefore, we obtain

$$S_i(\rho_i, \rho_j) = |\{\theta : -\theta - \eta \, q_i(1) < -\eta \, q_i(0; \rho_i, \rho_j)\}| \equiv S(\rho_i, \rho_j),$$

which implies

$$\rho_i^* = \frac{\lambda S(\rho_i^*, \rho_j^*)}{\lambda S(\rho_i^*, \rho_j^*) + (1 - \lambda)S(\rho_i^*, \rho_j^*)} = \lambda.$$

Next, we derive the conditions under which the advertising outcomes diverge from the single-advertiser main model. Since advertiser i has more loyal consumers, the only new outcome that is possible is the following: in the opt-out market, advertiser i advertises in Period 1 and then advertiser j advertises in Period 2. This occurs if and only if the following three conditions hold:

1. advertiser j's Period 2 bid, conditional on advertiser i's ad begin shown in Period 1, (a) exceeds that of advertiser i and (b) is greater than or equal to the reserve price,

2. advertiser i's bid in Period 1 exceeds the reserve price, and

3. the ad network's profit is higher selling Period 1 ads that not selling them.

Condition 1(a) is equivalent to $(1 - \lambda)\mu\phi_M > \lambda \left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta\right)$. But the difference $(1 - \lambda)\mu\phi_M - \lambda \left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta\right)$ is convex with with respect to μ with two roots 0 and $-\frac{1}{\lambda} + \beta\left(\phi_M + \frac{1}{\phi_M} - 2\right) + 2$. And since $\lambda > \frac{1}{2}$ implies $-\frac{1}{\lambda} + \beta\left(\phi_M + \frac{1}{\phi_M} - 2\right) + 2 > \beta\left(\phi_M + \frac{1}{\phi_M} - 2\right) + 2 - 2 = \beta\left(\phi_M + \frac{1}{\phi_M} - 2\right) > 0$, we obtain that Condition 1(a) simplifies to $\mu > -\frac{1}{\lambda} + \beta\left(\phi_M + \frac{1}{\phi_M} - 2\right) + 2$.

Condition 1(b) is equivalent to $(1 - \lambda)\mu\phi_M \geq k$, which, combined with $\lambda > \frac{1}{2}$, implies $\lambda\mu\phi_M \geq k$; this in turn implies Condition 2. Finally, Condition 3, provided Conditions 1 and 2, is equivalent to $\lambda\mu\phi_M + \lambda\mu(1 - \phi_M)\phi_M - k + (1 - \lambda)\mu\phi_M - k > \lambda\mu\phi_M - k$. This simplifies to $(1 - \lambda)\mu\phi_M - k + \lambda\mu(1 - \phi_M)\phi_M > 0$, which is implied by Condition 1(b): $(1 - \lambda)\mu\phi_M \geq k \iff \lambda < 1 - \frac{k}{\mu\phi_M}$.

In sum, the conjunction of Conditions 1 through 3 simplify to $\mu > -\frac{1}{\lambda} + \beta\left(\phi_M + \frac{1}{\phi_M} - 2\right) + 2 \equiv \tilde{\mu}$ and $\lambda < 1 - \frac{k}{\mu\phi_M} \equiv \tilde{\lambda}$. $\qquad\qquad\square$

C.2.12 Proof of Proposition 26

Proof. Next, it suffices to characterize the conditions under which (i) ads are shown to all opt-out consumers in both periods, and (ii) the ad intensity for opt-in consumers increases with signal accuracy ρ. If both conditions hold, then fewer ads are shown under tracking, and more consumers opt-out from tracking as ρ increases. The first condition is derived from Proposition 18. For the second condition, we begin by characterizing the ad network's ad supply decisions for opt-in consumers with imperfect purchase observability. For expositional ease, denote by N- and P-impressions the impressions associated with "not purchased" and "purchased" signals, respectively.

Given the advertiser's weakly dominant bids for T-, N-, and P-impressions, the ad network's profits from selling each type of impressions are $\mu\phi_M - k$, $\frac{(1-\phi_M)(1-\rho)}{\phi_M\rho+(1-\phi_M)(1-\rho)}\beta(1 - \phi_M) - k$, and $\frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}\beta(1 - \phi_M) - k$, respectively. The ad network sells whichever ad impressions yield positive profit.

Note that it is never profitable for the ad network to sell P-impressions but not N-impressions. The reason is that the fact that signals are at least partially informative imply N-impressions are valued more by the advertiser than P-impressions are.

First, note that $\frac{(1-\phi_M)(1-\rho)}{\phi_M\rho+(1-\phi_M)(1-\rho)} \leq \frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}$ for all $\frac{1}{2} \leq \rho \leq 1$, because $\frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho} - \frac{(1-\phi_M)(1-\rho)}{\phi_M\rho+(1-\phi_M)(1-\rho)} = \frac{(2\rho-1)(\phi_M-1)\phi_M}{((2\rho-1)\phi_M-\rho)(-\rho+(2\rho-1)\phi_M+1)}$ and the latter term's sign is equivalent to that of $\frac{1-\phi_M}{\rho+(1-2\rho)\phi_M}$. Now, the denominator $\rho + (1 - 2\rho)\phi_M$ is always positive because it is a linear function of ρ and is positive at each endpoint $\rho = \frac{1}{2}$ and $\rho = 1$.

Second, note that

$$\frac{\partial}{\partial\rho}\frac{(1 - \phi_M)(1 - \rho)}{\phi_M\rho + (1 - \phi_M)(1 - \rho)} = -\frac{(1 - \phi_M)\phi_M}{(-\rho + (2\rho - 1)\phi_M + 1)^2} < 0 \tag{C.2.13}$$

and

$$\frac{\partial}{\partial\rho}\frac{(1 - \phi_M)\rho}{\phi_M(1 - \rho) + (1 - \phi_M)\rho} = \frac{(1 - \phi_M)\phi_M}{(\rho - 2\rho\phi_M + \phi_M)^2} > 0. \tag{C.2.14}$$

Third, since the bounds $\frac{(1-\phi_M)(1-\rho)}{\phi_M\rho+(1-\phi_M)(1-\rho)}\beta(1-\phi_M)$ and $\frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}\beta(1-\phi_M)$ coincide at $\rho = \frac{1}{2}$, at which point the bounds equal $\beta(1-\phi_M)^2$, we obtain the following:

1. If $\mu\phi_M \geq k$, and $k > \beta(1-\phi_M)^2$, then as ρ increases from $\frac{1}{2}$ to 1, the regime changes from

   $k > \frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}\beta(1-\phi_M)$ to $\frac{(1-\phi_M)(1-\rho)}{\phi_M\rho+(1-\phi_M)(1-\rho)}\beta(1-\phi_M) < k < \frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}\beta(1-\phi_M)$;

   i.e., only T-impressions are shown for low ρ, and then N-impressions are also shown for high

   ρ.

2. If $\mu\phi_M < k$, and $k > \beta(1-\phi_M)^2$, then as ρ increases from $\frac{1}{2}$ to 1, the regime changes from

   $k > \frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}\beta(1-\phi_M)$ to $\frac{(1-\phi_M)(1-\rho)}{\phi_M\rho+(1-\phi_M)(1-\rho)}\beta(1-\phi_M) < k < \frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}\beta(1-\phi_M)$;

   i.e., no impressions are shown for low ρ, and then N-impressions are shown for high ρ.

This constitutes the second condition (i.e., the ad intensity for opt-in consumers increases with signal accuracy ρ).

However, the conditions $\mu\phi_M < k$ and $k > \beta(1-\phi_M)^2$ cannot hold jointly with $q^*(0) = 2$, which requires $(1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta > k$. To see this, it suffices to show that $(1-\mu)\mu\phi_M + \mu(1-\phi_M)^2\beta \leq \max[\beta(1-\phi_M)^2, \mu\phi_M]$, which holds because the left-hand side of the inequality is a linear combination of $\mu\phi_M$ and $\beta(1-\phi_M)^2$, so it must be smaller than the larger of $\mu\phi_M$ and $\beta(1-\phi_M)^2$. Therefore, for the conditions (i) and (ii) above to hold simultaneously, it must be that $\mu\phi_M \geq k$, and $k > \beta(1-\phi_M)^2$.

Finally, to ensure that the threshold of ρ past which the advertising regime changes from fewer to more advertising is between $\frac{1}{2}$ and 1, we must bound k by the largest value attained by the upper bound $\frac{(1-\phi_M)\rho}{\phi_M(1-\rho)+(1-\phi_M)\rho}\beta(1-\phi_M)$, which occurs at $\rho = 1$ and equals $\beta(1-\phi_M)$.

Re-arranging the condition $\beta(1-\phi_M)^2 < k < \min[\mu\phi_M, \beta(1-\phi_M)]$ with respect to ϕ_M and β yields the conditions in the proposition. $\qquad\square$

C.3 Parameter Scaling

We demonstrate the robustness of our main insights to smaller values of advertising effectiveness. To that end, suppose there exist two consumer segments: a potentially responsive segment

and a non-responsive segment, whose sizes are given by α and $1 - \alpha$, respectively, for some small $\alpha \in (0, 1)$. We assume that the potentially responsive consumers respond to ads in the manner described in the main model, while the non-responsive consumers always ignore ads; i.e., they never respond to ads.

Without consumer tracking, the advertiser cannot distinguish between these segments, while with tracking, it can. Therefore, the ad intensity under no tracking is

- 2 if $\alpha \left(\mu \phi_M + \mu(1 - \phi_M)\phi_M \right) - k + \alpha \left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta \right) - k \geq (\alpha\mu\phi_M - k)^+$ and $\alpha \left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta \right) - k \geq 0$,

- 1 if $\alpha \left(\mu\phi_M + \mu(1 - \phi_M)\phi_M \right) - k \geq (\alpha\mu\phi_M - k)^+$ and $\alpha \left((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta \right) - k < 0$, and

- 0 otherwise.

Similarly, the ad intensity under tracking is

- $1 + \alpha(1 - \mu)\mathbb{I}_{\{\mu\phi_M - k \geq 0\}} + \alpha\mu(1 - \phi_M)\mathbb{I}_{\{\beta(1 - \phi_M) - k \geq 0\}}$ if
  $\alpha \left(\mu\phi_M + \mu(1 - \phi_M)\phi_M \right) - k + \alpha \left((1 - \mu)(\mu\phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+ \right) \geq 0$,
  and

- 0 otherwise.

Note that if we let $k' = k/\alpha$, then the ad intensity under no tracking is equivalent to the main model with ad cost k', and the ad intensity under tracking is either 0, 1 or between 1 and 2 under the same conditions as the main model. Thus, the conditions for the ad intensity differential are preserved from the main model.

As illustrated in Figure C.2, we can replicate the advertising intensity differential patterns of the main model (Figure 4.5) for small values of α, ϕ_M, and k. Since the main insights rest on the ad intensity differential pattern, this suffices to show that the insights are robust to parameter scaling.
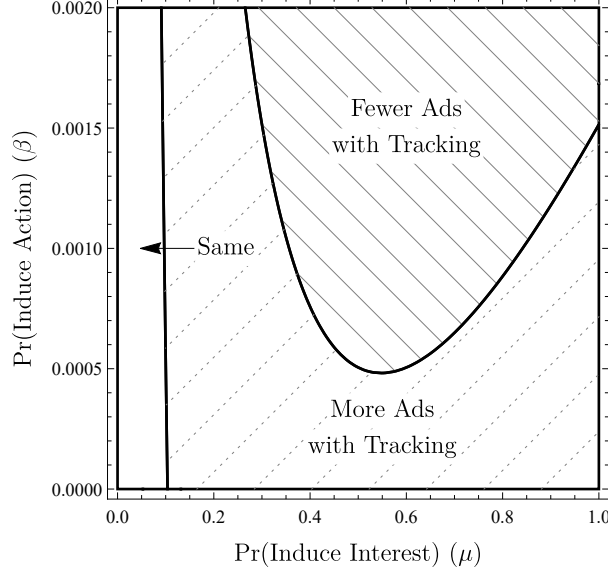
Figure C.2: Ad Intensity with Parameter Scaling; $\alpha = 0.001$, $\phi_M = 0.005$, $k = 1.5 \times 10^{-6}$

C.4 Markov-Perfect Equilibrium

For any given Period t, define "old generation" as the mass of consumers who arrived in Period $t - 1$, and "new generation" as those who arrive in Period t. In our setting, the payoff-relevant states can be fully characterized by the distribution of old generation non-converters in funnel states T and M. Consider the no tracking case where the advertiser cannot target ads based on the consumers' funnel states, nor their purchase history. Let λ_f^{old} denote the proportion of old-generation non-converters in funnel state $f \in \{T, M\}$. There are two possible states in each period: one in which the advertiser showed ads in the previous period, and another in which it did not show ads in the previous period.

To elaborate, suppose the advertiser showed ads in Period $t-1$. The old generations in Period $t-1$ (i.e., those who arrived in Period $t - 2$) leave by Period t because consumers only live for two periods. Therefore, these consumers are irrelevant in the analysis of determining the successive distribution of old generation non-converters in Period t. Of the $1 - \sigma$ T-consumers who arrived in Period $t - 1$, $1 - \mu$ fraction are not influenced by the ad and stay in T, μ fraction transition to M, of which ϕ_M convert and $1 - \phi_M$ do not. Moreover, of the σ M-consumers who joined in Period $t - 1$, $1 - \beta$ stay in M, and still $1 - \phi_M$ of those $\sigma(1 - \beta)$ M-consumers do not convert. Therefore, the

177

distribution of old generation non-converters in Period t would be

$$\left(\lambda_T^{\text{old}}, \lambda_M^{\text{old}}\right) = \left((1 - \sigma)(1 - \mu), ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)\right).$$

We label this state as λ_1, where the subscript 1 indicates the advertiser showed ads in the previous period.

On the other hand, suppose the advertiser did not show ads in Period $t - 1$. Without any ad exposures, the $1 - \sigma$ T-consumers who arrived in Period $t - 1$ would all remain in T by Period t. However, ϕ_M fraction of σ M-consumers convert and $1 - \phi_M$ fraction remain in M in Period t. Therefore, the distribution of old-generation non-converters in this case is

$$\left(\lambda_T^{\text{old}}, \lambda_M^{\text{old}}\right) = \left(1 - \sigma, \sigma(1 - \phi_M)\right).$$

We label this state as λ_0, where the subscript 0 indicates the advertiser did not show ads in the previous period.

Given two states and two possible advertising strategies at each state (i.e., advertise or not advertise), there are four Markov-perfect equilibrium (MPE) candidates: (i) always advertise regardless of the state; (ii) advertise only when the state is λ_0, which is equivalent to "pulse advertising" (i.e., alternate advertising with a single-period break in between; (iii) advertise only when the state is λ_1, which is effectively equivalent to (i); and (iv) never advertise. We compare the ad network's profits for the respective strategies.

I. Always advertise

For always advertising to be MPE, the advertiser's payoff from buying untargeted ads in Period t, given the state is either $\lambda_1 \equiv ((1 - \sigma)(1 - \mu), ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M))$ or $\lambda_0 \equiv (1 - \sigma, \sigma(1 - \phi_M))$, should be greater than that from not buying:

$$(1 - \sigma)(1 - \mu)\mu\phi_M + ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)(\beta + (1 - \beta)\phi_M) + (1 - \sigma)\mu\phi_M - 2R_1 + \delta V_1$$
$$\geq ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)\phi_M + \delta V_0,$$

and

$$2(1 - \sigma)\mu\phi_M + (\sigma(1 - \phi_M) + \sigma)(\beta + (1 - \beta)\phi_M) - 2R_0 + \delta V_1 \geq \sigma(1 - \phi_M)\phi_M + \sigma\phi_M + \delta V_0,$$

where V_1 is the continuation value from having shown ads in the previous stage, and V_0 is the continuation value from not having shown any ads in the previous stage. In equilibrium, the ad network will set reserve prices R_1 and R_0 such that these conditions bind; otherwise, it leaves money on the table. Therefore, from the second condition, we obtain

$$2R_{0}* = 2(1 - \sigma)\mu\phi_M + (\sigma(1 - \phi_M) + \sigma)\beta(1 - \phi_M) + \delta(V_1 - V_0).$$

But if the second condition holds, it must be that the continuation value from not showing ads is the continuation value from showing ads, such that

$$V_0 = 2(1 - \sigma)\mu\phi_M + (\sigma(1 - \phi_M) + \sigma)(\beta + (1 - \beta)\phi_M) - 2R_0^* + \delta V_1.$$

Then, substituting R_0^* yields $V_0 = \sigma(1 - \phi_M)\phi_M + \sigma\phi_M + \delta V_0$, which in turn implies

$$V_0 = \frac{\sigma(2 - \phi_M)\phi_M}{1 - \delta}.$$

Similarly, after substituting $V_0 = \frac{\sigma(2-\phi_M)\phi_M}{1-\delta}$ into the first condition and letting it bind, we obtain

$$2R_1^* = (1 - \sigma)(1 - \mu)\mu\phi_M + ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)(\beta + (1 - \beta)\phi_M)$$
$$+ (1 - \sigma)\mu\phi_M + \delta V_1 - (((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)\phi_M + \delta V_0),$$

which simplifies to

$$2R_1^* = (1 - \sigma)(2 - \mu)\mu\phi_M + ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)^2\beta + \delta V_1 - \delta\frac{\sigma(2 - \phi_M)\phi_M}{1 - \delta}.$$

Since the continuation value of having shown ads is the continuation value from showing ads, we obtain

$$V_1 = (1 - \sigma)(1 - \mu)\mu\phi_M + ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)(\beta + (1 - \beta)\phi_M) + (1 - \sigma)\mu\phi_M - 2R_1^* + \delta V_1$$

which, upon substitution of R_1^* yields $V_1 = ((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)\phi_M + \delta\frac{\sigma(2-\phi_M)\phi_M}{1-\delta}$.

Therefore,

$$2R_1^* = (1-\sigma)(2-\mu)\mu\phi_M + ((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)^2\beta$$
$$+ \delta\left(((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)\phi_M + \delta\frac{\sigma(2-\phi_M)\phi_M}{1-\delta}\right) - \delta\frac{\sigma(2-\phi_M)\phi_M}{1-\delta}.$$

Since this strategy induces the state to be perpetually λ_1, the ad network's total profit is

$$\pi_N^I = \frac{1}{1-\delta}\left((1-\sigma)(2-\mu)\mu\phi_M + ((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)^2\beta\right.$$
$$\left. + \delta\left(((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)\phi_M + \delta\frac{\sigma(2-\phi_M)\phi_M}{1-\delta}\right) - \delta\frac{\sigma(2-\phi_M)\phi_M}{1-\delta} - 2k\right).$$

II. Advertise Only When State is $\lambda_0 = (1-\sigma, \sigma(1-\phi_M))$

For this pulsing strategy to be MPE, we need advertiser's payoff to be higher buying ads given $(1-\sigma, \sigma(1-\phi_M))$, and not buying ads given $((1-\sigma)(1-\mu), ((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M))$, which respectively translate to:

$$2(1-\sigma)\mu\phi_M + (\sigma(1-\phi_M)+\sigma)(\beta + (1-\beta)\phi_M) - 2R_0 + \delta V_1 \geq (\sigma(1-\phi_M)+\sigma)\phi_M + \delta V_0$$

and

$$(((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)+\sigma)\phi_M + \delta V_0 \geq 2(1-\sigma)\mu\phi_M + (\sigma(1-\phi_M)+\sigma)(\beta+(1-\beta)\phi_M) - 2R_1 + \delta V_1.$$

The ad network sets $R_1 = \infty$ (such that no ads are bought at state λ_1) and $2R_0^* = 2(1-\sigma)\mu\phi_M + (\sigma(1-\phi_M)+\sigma)(1-\phi_M)\beta + \delta(V_1 - V_0)$. This implies $V_0 = \sigma(2-\phi_M)\phi_M + \delta V_0$, which means $V_0 = \frac{\sigma(2-\phi_M)\phi_M}{1-\delta}$. Similarly, since $V_1 = (((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)+\sigma)\phi_M + \delta V_0$, we have

$$2R_0^* = 2(1-\sigma)\mu\phi_M + (\sigma(1-\phi_M)+\sigma)(1-\phi_M)\beta$$
$$+ \delta\left(\left((((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)+\sigma)\phi_M + \delta\frac{\sigma(2-\phi_M)\phi_M}{1-\delta}\right) - \frac{\sigma(2-\phi_M)\phi_M}{1-\delta}\right).$$

Since this strategy yields alternating states, the ad network's profit under this strategy is

$$\pi_N^{II} = \frac{1}{1-\delta^2}\left(2(1-\sigma)\mu\phi_M + (\sigma(1-\phi_M)+\sigma)(1-\phi_M)\beta\right.$$
$$\left. + \delta\left(\left((((1-\sigma)\mu + \sigma(1-\beta))(1-\phi_M)+\sigma)\phi_M + \delta\frac{\sigma(2-\phi_M)\phi_M}{1-\delta}\right) - \frac{\sigma(2-\phi_M)\phi_M}{1-\delta}\right) - 2k\right).$$

III. Advertise Only When State is $\lambda_1 = ((1 - \sigma)(1 - \mu), ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M))$

Same as strategy I: set $R_0^* = \infty$ and the rest follows.

IV. Never advertise

This strategy yields 0 payoff.

Finally, comparing the payoffs π_N^I, π_N^{II} and 0 yield the presented equilibrium regions.