Customer Purchase Journey, Privacy, and Advertising Strategies

W. Jason Choi  Kinshuk Jerath  Miklos Sarvary *

Job Market Paper

October 2019

For latest version, click here

*W. Jason Choi (jason.choi@columbia.edu) is a Ph.D. candidate, Kinshuk Jerath (jerath@columbia.edu) is Associate Professor of Marketing, Miklos Sarvary (miklos.sarvary@columbia.edu) is the Carson Family Professor of Business, all at the Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027.
Customer Purchase Journey, Privacy, and Advertising Strategies

Abstract

We investigate the impact on the online advertising ecosystem of tracking consumers’ activities on the Internet. We also study 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 General Data Protection Regulation passed by the European Union in 2018). The consumers’ strategic decisions to (dis)allow advertisers from tracking their activities depend on two aspects of privacy: its intrinsic value (protect privacy for its own sake) and its instrumental value (compromise privacy if doing so indirectly leads to some utility-enhancing outcome). This opt-in decision impacts the precision of inferences by advertisers about how far down a consumer is in the “purchase funnel” for a product by virtue of ads shown previously. The structure of the purchase funnel creates an interdependence between the effectiveness of the sequence of ads shown, which in turn affects advertising strategies. For instance, we find that the intensity of advertising by advertisers is non-monotonic in the effectiveness of ads. Consequently, consumers may opt-in to be tracked when ad effectiveness is intermediate. While privacy regulations generally increase consumer surplus, the implications for the ad network are mixed. Interestingly, the ad network’s profit may (i) be higher under endogenous tracking than under full tracking, and (ii) decrease as ads become more effective. We discuss managerial implications for advertisers as well as policy implications for regulators.

Keywords: customer purchase journey, online advertising, consumer tracking, data privacy, consumer privacy choice
In the physical world, users wouldn’t expect hundreds of vendors to follow them from store to store, spying on the products they look at or purchase. Users have the same expectations of privacy on the web, and yet, in reality, they are tracked wherever they go.

— Nguyen (2018)

1 Introduction

Advances in information technology have led to unprecedented levels of consumer tracking on websites (Lerner et al., 2016). In 2010, the 50 most popular US websites had 2,224 tracking files (e.g., “cookies” and “web beacons”) installed by 131 firms (Angwin, 2010). In 2011, 89% of Alexa top 500 websites had embedded in them at least one cross-site trackers. 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 (Hauser et al., 2009; Urban et al., 2013), (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 infer consumers’ purchase journey stages. Advertisers can observe consumers’ online browsing and shopping activities, infer whether they are, say, “product viewers” or “cart creators,” and target ads accordingly (Sahni et al., 2019). Indeed, industry experts recommend that advertisers focus more on targeting based on consumers’ “stages in the decision journey” than on media allocation (Edelman, 2010). Empirical findings that ad effects — as measured by sales (Johnson et al., 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, 64% of UK Instagram users say “it’s creepy how well online ads know me” (eMarketer, 2018a) and “86% of young [US] adults ... don’t want tailored advertising if it is the result of following their behavior on websites other than [the] one they are visiting” (Turow et al., 2009). Fur-

1 For example, various Google tags can be used to specify remarketing audiences (https://support.google.com/google-ads/answer/6335506).
thermore, 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. 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. The California Consumer Privacy Act (CCPA), a US analogue of the GDPR, is expected to go into effect in January 2020.

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 affirmative consent to use their data. In other words, firms interested in processing consumer data must abide by a transparent opt-in policy. Firms are not allowed to collect consumer data by default; consumers themselves must opt-in to their data being collected and processed by firms. If consumers opt-out from tracking, then advertisers cannot monitor consumers’ behavior across websites. Consequently, the advertisers’ targeting capabilities would be severely undermined. As a result, ad impressions could be potentially wasted (e.g., repeated exposure to consumers who had already purchased). On the other hand, if consumers opt-in to tracking, advertisers may target ads to a narrower audience based on a pre-specified set of behavioral criteria (e.g., consumers who interacted with the ad from previous browsing sessions 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’

---

2 The regulation applies to all firms processing personal data of European subjects (even if the firm operates outside of Europe) and uses a broad definition of personal data: “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). See Future of Privacy Forum (2018) for a detailed comparison of the two regulations.

4 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 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.

5 See https://www.blog.google/products/marketingplatform/360/privacy-safe-approach-managing-ad-frequency/
tracking capacity, thereby reducing ad effectiveness \cite{Aziz, Telang, 2016, Goldfarb, and, Tucker, 2011}. 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 \cite{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.\footnote{The overstatement of privacy concerns relative to revealed preferences is known as the “privacy paradox.” See Norberg \textit{et al.} \cite{2007} and Athey \textit{et al.} \cite{2017} for details.} For example, Johnson \textit{et al.} \cite{2019} find that less than 0.26% of US and EU consumers opt-out from behavioral targeting. Moreover, 67% of US and Canadian consumers report that they would feel “comfortable sharing personal information with a company” if it transparently discloses how the data will be used \cite{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.

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, we seek to shed light on these questions by developing a game theory model. We consider a two-period model in which consumers visit content pages, creating opportunities for ad impressions. An advertiser buys ad impressions from an ad network, who auctions off ad inventory supplied by the content pages. Motivated from the discussion above, we assume that ad effects depend on the consumers’ journey states and that their purchase journey is influenced by advertising \cite{Abhishek, 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, we assume that consumers jointly consider two aspects of privacy: its intrinsic value (protect privacy for its own sake) and its instrumental value (compromise privacy if it indirectly leads to some utility-enhancing outcome) \cite{Farrell, 2012, Wathieu, and, Friedman, 2009}.

Our analysis yields three main insights. First, consumers choose to opt-in to being-tracked if ad effectiveness is intermediate. Intuitively, for opt-out consumers who cannot be tracked, the advertiser
shows mass-advertising 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 targeted ads only to the most responsive consumer segments consisting of (i) those who were not impacted by the first ad and stayed at the top of the funnel and/or (ii) those who moved down to the middle of the funnel but did not purchase. Therefore, if ad effectiveness is intermediate, some consumers trade-off their privacy cost with the benefit of seeing fewer ads by opting-in to tracking.

Second, under the endogenous tracking regime, the ad network’s profit may decrease in ad effectiveness, even though higher effectiveness implies higher average conversion probability. The reason is that high ad effectiveness may induce the saturation effect, whereby the marginal value of successive ads is diminished by previously shown ads. This causes the advertiser to forego successive advertising for opt-out consumers; in contrast, enhanced targeting efficiency induces the advertiser to show successive ads for opt-in consumers. 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, ad slots are sold at lower prices and the ad network’s profit decreases.

Third, privacy regulations increase consumer surplus and decrease the ad network’s profit compared to the full tracking regime. Interestingly, however, if the advertiser is privately informed about ad valuations, consumers opting out of tracking may be a boon to not just the consumers but the ad network as well. Intuitively, the ad network’s inability to track opt-out consumers serves as a commitment mechanism that induces the ad network to auction off untargeted ads that reach a larger consumer segment than targeted ads. This supply-side “market thickening” effect induces the advertiser to bid more aggressively for opt-out consumers than for opt-in consumers, leading to higher ad network profit with endogenous tracking.

We assess the robustness of our insights by analyzing three extensions. First, we allow the advertiser to have private information about ad valuations. As discussed above, we find that information asymmetry leads to bidding dynamics that sometimes results in the ad network’s profit being higher under endogenous tracking. Second, we consider a multiple advertiser scenario with consumer product preference heterogeneity and examine the associated forces that moderate the main results. Finally, we extend the time horizon from two to infinity with heterogeneous, overlapping generations of consumers, and characterize a Markov perfect equilibrium.

---

8In practice, ad slots indeed can be left unsold; in fact, 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 50% to 90%.
In addition to being related to the papers referenced earlier, our 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 2013; Shen and Villas-Boas 2018). We 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. We 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.

We 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 explicit privacy decisions, whereby consumers take (often costly) action to control the amount of information disclosed to firms (Acquisti and Varian 2005; Conitzer et al. 2012; Ichihashi 2019; Montes et al. 2019). D’Annunzio and Russo (2017) study consumer opt-in policy in a setting where firms can track cross-website behavior, and de Cornière and Nijs (2016) investigate the ad network’s incentive to disclose consumer information to advertisers. Our paper is different from these works in two ways: first, that we model the purchase journey, and second, that we allow consumers to jointly consider the instrumental and intrinsic aspects of privacy when deciding whether to allow tracking or not. These features of the model generate unique implications for consumer opt-in behavior and the ensuing advertising outcomes. The mechanisms behind our results are orthogonal to market thickness (Bergemann and Bonatti 2011; Rafieian and Yoganarasimhan 2018) and market structure (Campbell et al. 2015) as we abstract from advertiser competition in the main model.

At a higher level, our 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. Our findings suggest that assessing the impact of privacy regulations on the advertising industry is a complex issue. Nevertheless, we identify several robust theoretical insights that inform various regulatory implications.

The rest of the paper is organized as follows. In Section 2 we describe the main model. In Section 3 we present the main equilibrium 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, we assess the robustness of the main insights by analyzing three extensions. In Section 5, we summarize the key results and conclude. All proofs are relegated to Section D of the appendix.

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. The ad network sells the impressions on the content pages via second price auctions. Before we discuss each player’s decisions and payoffs, we first explain a key feature of our model: the consumer purchase journey. We describe the relationship between advertising and consumers’ progression down the purchase funnel.

Purchase Journey and Ad Effects

We consider a stylized purchase journey consisting of three distinct states labeled top, middle, and bottom (see Figure 1). For ease of exposition, we denote these states by T, M and B, respectively, and the consumers in the respective states by f-consumers for $f \in \{T, M, B\}$. We define funnel state $f \in \{T, M, B\}$ with a probability $\phi_f$, which measures the likelihood of an $f$-consumer realizing a product match. At each time period, an $f$-consumer (who has not purchased yet) realizes a product match with probability (w.p.) $\phi_f$, where $0 \leq \phi_T < \phi_M < \phi_B \leq 1$.

Taylor (2004) calls this match probability the “intensity of taste for a particular class of goods” (pg. 635).
interpreted as follows: the top-funnel corresponds to “awareness” state, wherein the consumer is aware of the product’s existence but not 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. We normalize $\phi_T$ and $\phi_B$ to 0 and 1, respectively.

If a consumer realizes a product match, she derives positive utility $v$ from consuming the product; otherwise, she derives utility 0. In accordance with the empirical literature, we 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. $1 - \mu$. 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 our model specifications allow for flexible ad response curvatures using the funnel transition parameters $\mu$, $\beta$, and mid-funnel match probability $\phi_M$. For instance, Figure 2 depicts a convex ad response curve for low $\phi_M$, and a concave ad response curve for high $\phi_M$. We will show how advertising strategies, consumer choices and welfare outcomes depend crucially on the curvature of consumer’s ad response.

We also note that our results will remain completely unaltered if we scale the effectiveness numbers to vary between ranges different from 0 to 1. For instance, if we consider a sub-population of consumers who potentially respond to ads, 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 our results (see Section B of the appendix for more details). However, for model simplicity and expositional clarity, we 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, we 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 4.3, we relax this 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. $\phi_f$, 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. We 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. We assume that a consumer purchases at most one unit.

Next, we turn to privacy utility. We assume that consumers dislike being tracked, and that they are heterogeneous in their tracking disutility.\footnote{Heterogeneity may stem from numerous factors such as differences in what consumers believe constitutes personal information \cite{Acquisti2016} and differences in consumers’ perception of their privacy control \cite{Tucker2014}} This disutility is captured by the privacy cost parameter $\theta$, which has cumulative distribution function $F$. We assume that consumers can decide whether to opt-in or opt-out of tracking. If a consumer opts-in, firms can track her identity and online browsing behavior (across content pages and across sessions) for targeting purposes. Later, we describe in more detail how tracking and targeting are implemented. The privacy utility of a consumer with privacy cost $\theta$ is

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

(1)
where $x$ denotes the consumer’s privacy decision, which equals 1 if she opts-in, and 0 if she opts-out\textsuperscript{11}. $\tilde{q}$ the number of ads she expects to see, $\eta$ the disutility she incurs per unit of ad impression (Johnson \textit{et al.}, 2013; de Cornière and Taylor, 2014), and $\theta$ the disutility she incurs for 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 decision, 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 bid for different types of ads. If tracking is prohibited, then consumer 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). 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, in the presence of consumer tracking, the advertiser can buy ad impressions at the funnel-stage level. By installing tags on its website and embedding cookies on consumer browsers, the advertiser can monitor the websites visited by the consumers, their browsing activity within the websites, and their purchase behavior. Based on this information, the advertiser can specify various targeting criteria such that their ads are shown only to consumers who meet some pre-specified criteria. For example, the advertiser can target ads to consumers who are 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 $p_t$ at each Period $t \in \{1, 2\}$. We normalize the marginal cost of the product to zero. Therefore, the advertiser’s margin per conversion at Period $t$ is $p_t$.

**Ad Network**

The ad network sells ad impressions to advertisers via second price auctions. It sets reserve price $R_{jt}$ for each 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). The ad network maximizes its total profit across two periods, which consists of revenue from ad impression sales and cost associated with selling impressions. Costs may include operational costs associated with ad inventory management, as well as maintenance costs.

\textsuperscript{11}We clarify two implicit assumptions here. First, while consumers may modify their privacy decision at any time in practice, we 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 \textit{et al.}, 2012; Montes \textit{et al.}, 2019), we 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.
related to setting up ad auctions and delivering ads. We denote this per-impression cost by $k \geq 0$.

**Game Timing**

The timing of the game is as follows.

**Period 0:** Consumers decide whether or not to opt-in for tracking.

**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-ins 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.

## 3 Analysis

We solve for the subgame perfect equilibrium. First, we note that the 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 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. Thus, the net effect of an ad on a consumer’s conversion probability is driven by (a) the probability of the ad changing the consumer’s funnel state, and (b) the associated change in the consumer’s match probability. Note that the consumer’s product utility is always zero whether or not she purchases. Therefore, when discussing consumer utility, we hereafter restrict attention to the privacy utility component.

To develop basic insights, we study the case of no tracking in Section 3.1 and full tracking in Section 3.2. Then we discuss the main analysis with endogenous consumer tracking choice in Section 3.3.

---

12 This model feature is motivated from our conversations with industry practitioners. In particular, they have indicated that selling ad inventory entails various forms of 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.

13 One could argue that leaving positive consumer surplus by setting $p_t < 1$ may induce favorable opt-in outcomes which compensate the advertiser for its lower margin. However, because privacy choices are made prior to observing product prices, to the extent that prices can be flexibly changed upon the consumer’s website visit, such low prices cannot be sustained in equilibrium.
3.1 No Tracking

In this section, we analyze the case in which consumers cannot be tracked; i.e., advertisers cannot distinguish consumers’ funnel states nor their purchase histories. Our 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, we first analyze the advertiser’s bidding problem in Period 2 and then proceed backwards. We assume that the advertiser plays weakly dominant bids, in the sense that the bids are robust to any reserve prices.

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. We 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. \( \mu \), the first group is of size \( 1 - \mu \). Since \( M \)-consumers realize a product match w.p. \( \phi_M \), 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 \( t \).

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

Consider the advertiser’s payoff from winning the auction and displaying the ad, shown on the top row of (2). 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, \( \mu \) fraction transition to funnel state \( M \), of which \( \phi_M \) fraction realize a match and convert. Similarly, the second term denotes the conversion of \( M \)-consumers who had not converted in Period 1.

\footnote{For ease of exposition, we 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.}
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 (2). This is because even if no additional ads are shown in Period 2, the non-converters in $M$ — who were pushed down from $T$ after seeing the Period 1 ad but did not convert — may realize a product match w.p. $\phi_M$ 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 1** (Period 2 Bids and Reserve Prices Without Tracking).

- Suppose the advertiser showed ads in Period 1. The advertiser’s weakly dominant bid in Period 2 is $b^*_2|_{ad} = (1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2 \beta$, and the ad network’s optimal reserve price is $R^*_2|_{ad} = \max[k, b^*_2|_{ad}]$.
- 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 $R^*_2|_{no\ ad} = \max[k, b^*_2|_{no\ ad}]$.

The first part of Lemma 1 provides important preliminary insights into the conditions under which the advertiser will buy 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

$$
(1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2 \beta \geq k. \tag{3}
$$

That is, the marginal effectiveness of the successive ad, expressed on the left-hand side of (3), 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.

To begin, note that 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 $\mu$, 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 $\phi_M$, the product match probability of consumers in funnel state $M$, if and only if $\beta$, 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 $\mu$ is high, 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 $\mu$ for high $\mu$. 

12
Consider the second case. If $\beta$ is high, then the marginal effectiveness of successive ads is largely determined by their potential to convert $M$-consumers. Now, increasing $\phi_M$ has two effects. First, consumers are more likely to convert after the first ad exposure such that there is a small segment of non-converted $M$-consumers in Period 2. Second, if $\phi_M$ is high, those non-converted $M$-consumers are likely to convert on their own without a successive ad exposure. Thus, increasing $\phi_M$ dampens the value of a successive ad.

Taken together, we see that under certain conditions, a first ad that is increasingly effective in converting consumers (i.e., high $\mu$ and $\phi_M$) diminishes the marginal effectiveness of successive advertising in Period 2. We call this the *saturation effect*. It is visualized by the concave ad response curve for high $\phi_M$ in Figure 2.

**Period 1**

The reserve prices $R^*_2$ from Lemma 1 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 Period 2 payoff into account, the advertiser’s problem in Period 1 is to determine the bid $b_1$ that maximizes

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

where $R_1$ 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 2** (Period 1 Bid and Reserve Price Without Tracking). *The advertiser’s weakly dominant bid in Period 1 is $b^*_1 = \mu\phi_M + \mu(1 - \phi_M)\phi_M$, and the ad network’s optimal reserve price is*

$$
R^*_1 = \max \left[ k + (\mu\phi_M - k)^+ - (b^*_2|_{ad} - k)^+, b^*_1 \right],
$$

*where $x^+ \equiv \max\{x, 0\}$.*

We see from (4) that the ad network sometimes sets the Period 1 reserve price below the marginal cost $k$, even if that implies the ad network may earn 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 win, 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
Advertising Strategy Without Tracking

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

**Proposition 1 (Advertising Without Tracking).** Suppose the advertiser cannot track consumers. For thresholds 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 $\beta \geq \bar{\beta}$, and does not buy any otherwise.

- Suppose $\phi_M \geq 1 - \sqrt{(\mu - k)^+/\mu}$.
  - if (i) $\beta \geq \bar{\beta}$ and $\mu \geq \underline{\mu}$, or (ii) $\underline{\beta} \leq \beta < \bar{\beta}$ and $\underline{\mu} \leq \mu < \bar{\mu}$, then advertiser buys ads in both periods;
  - if (i) $\beta < \underline{\beta}$ or $\mu \leq \underline{\mu}$, or (ii) $\underline{\beta} \leq \beta < \bar{\beta}$ and $\mu \geq \bar{\mu}$, then advertiser buys ad only in Period 1.

Consider the case of low $\phi_M$, depicted in Figure 3a, where the advertiser either buys ads in neither periods or buys in both periods. This “all-or-nothing” pattern emerges when the ad response curve is convex (see dashed curve in Figure 2). Specifically, if $\phi_M$ is low, 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 $\beta$ is sufficiently high, a successive ad is highly likely to bring $M$-consumers in Period 2 down to $B$ and induce purchase. This sharp increase in the effectiveness of a successive ad may compensate for the low effectiveness of the first ad. Thus, if $\phi_M$ is low, the advertiser either buys ads in neither periods or in both.

In contrast, if $\phi_M$ is high and $\mu$ is either medium or high, the advertiser may advertise only in Period 1 and not successively in Period 2 (see Figure 3b). If $\mu$ is medium, a successive ad is not effective enough in inducing interest (i.e., bringing $T$-consumers down to $M$) and the size of $M$-consumers in Period 2 is not large enough. Therefore, for medium $\mu$, the value of showing a successive ad is low such that only Period 1 ads are shown.

The advertiser also foregoes Period 2 advertising if $\mu$ is high. In this case, there would be many $M$-consumers, who have transitioned from $T$ after the first ad exposure, and only a few $T$-consumers who are not considering purchase. And if $\beta$ is not high, the successive ad has little effect in increasing the $M$-consumers’ match probabilities. In other words, the saturation effect—in the form of reducing the size of $T$-consumers in Period 2—dampens the marginal impact of the second ad. Therefore, only Period 1 ads are shown for high $\mu$.

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

### 3.2 Full Tracking

We now analyze how the ability to track consumers affects the advertiser’s strategy and the ad network’s profit. In practice, advertisers have the technology to monitor an extensive array of consumer online behavior, ranging from clicks on website links to the view duration of a video ad. As described above, advertisers can install event tracking tags on their websites that help them not only count the number of times an event occurs, but also draw meaningful probabilistic inferences about consumer’s journey through the funnel (e.g., whether the consumer is aware of the product, whether she is close to purchase, etc.). Based on this information, the advertiser bids in the auction run by the ad network.

---

15 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)
Advertising Strategy With Tracking

Without tracking, the advertiser was restricted to buying untargeted ads. In the presence of consumer tracking, however, the advertiser can target successive ads along two dimensions—consumers’ product purchase histories, and their positions in the funnel. Specifically, in Period 2, it can target (i) $T$-consumers who are not considering purchase, (ii) $M$-consumers who did not purchase in Period 1 and are considering purchase, or (iii) both consumer segments in (i) and (ii). Note that since consumers purchase at most one unit, the advertiser will never buy impressions for consumers who had already purchased. The following proposition characterizes the advertiser’s strategies in the presence of consumer tracking.

**Proposition 2** (Advertising With Tracking). Suppose the advertiser can track consumers along the purchase funnel. Let $\tilde{\mu} = k(\phi_M(2 - \phi_M) + (1 - \phi_M)(\beta(1 - \phi_M) - k)^+)^{-1}$. 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 convert.

Figure 4 depicts the advertising strategies in the presence of consumer tracking. Before we discuss this figure in detail, we note that a comparison with Figure 3 shows that the advertising strategy with
tracking is significantly different from the advertising strategy without tracking. Proposition 2 shows that when the probability of the ad inducing interest is high, the advertiser adopts a reach strategy, whereby it targets successive ads to 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 consumers who are already considering purchase.

Overall, Proposition 2 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, 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**

Another question we are interested in is: how does consumer tracking impact the overall advertising intensity? Our results show that the effect of tracking on ad intensity is mixed (see Figure 5). If the funnel transition probabilities $\mu$ and $\beta$ are low, then ad effectiveness is so low that no ads are shown regardless of the tracking capability; hence, tracking may not change ad intensity. Otherwise, if either $\mu$ or $\beta$ is sufficiently high, tracking may either increase or decrease the total volume of ads.

Tracking enables ads to be delivered more efficiently as they can be targeted based on funnel states and purchase history. When ad effectiveness is low, an increase in per-impression valuation due to tracking raises the demand for ad slots, such that more ads are shown with tracking than without. On the other hand, when ad effectiveness is intermediate, even Period 2 untargeted ads (which are potentially wasted to consumers who already purchased) may be shown. In this case, tracking allows the advertiser to reduce spending on wasteful ad impressions, resulting in lower ad intensity with tracking than without.

While these findings resonate with those from Esteban et al. (2001) and Iyer et al. (2005), our results diverge from these papers at high levels of ad effectiveness. Specifically, when ad effectiveness is high, 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 funnel, 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, we establish a non-monotonic relationship between the impact of tracking on ad intensities and ad effectiveness, which is proxied by the funnel transition parameter $\mu$ when the match parameter $\phi_M$ is fixed. We summarize these results in the following proposition.

**Proposition 3 (Ad Intensity).** Consumer funnel tracking either increases or decreases the total ad intensity compared to the no tracking case. Specifically,

- if (i) $\mu < \tilde{\mu}$ and $\beta < \tilde{\beta}$ or (ii) $k \frac{\phi_M(2-\phi_M)}{2-\phi_M} < \mu \leq k \frac{1}{1-\phi_M}$ and $\beta \leq k \frac{1}{1-\phi_M}$, then the ad intensities are the same;
- if $\phi_M > 1 - \sqrt{(\mu - k)^+} / \mu$ and either (i) $\beta > \tilde{\beta}$ and $\mu > \tilde{\mu}$, or (ii) $\beta \leq \beta < \tilde{\beta}$ and $\mu \leq \mu < \tilde{\mu}$, 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? We find that tracking weakly increases the ad network’s profit. Intuitively, consumer tracking endows the advertiser with more information, which increases advertising ROI. The advertiser values targeted ads more highly than untargeted ads, which implies that the ad network sometimes sells more impressions in the presence of tracking, and also extracts larger surplus per-impression. Both of these factors raise the ad network’s profit. We state this result as a proposition.
Proposition 4 (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 4.1, we show that the result of Proposition 4 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.

3.3 Endogenous Tracking Choice

In the preceding analysis, we examined two distinct cases in which the advertiser was either able to track consumers or not. In this section, we investigate the impact of endowing consumers with choice to be tracked. That is, we analyze how consumers exercise their right to choose whether to allow tracking or not and how the privacy decisions affect the ad network’s profit.

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. Interestingly, we find that under endogenous tracking choice, the ad network’s profit can be non-monotonic in ad effectiveness. Furthermore, the ad network’s profit can be higher under privacy regulations compared to the full tracking benchmark, when we allow the advertiser to possess private information about ad valuations (see Section 4.1).

Consumer Opt-In Behavior

We first characterize the consumers’ equilibrium privacy decisions given the advertising outcomes under tracking and no tracking. Recall from the consumer privacy utility (1) that consumers dislike seeing ads and also dislike being tracked. 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. The following proposition summarizes the consumer opt-in behavior.

Proposition 5 (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 \( \mu \). In particular, if either (i) \( \beta < \beta \leq \overline{\beta} \) and \( \mu < \mu < \overline{\mu} \), or (ii) \( \beta > \overline{\beta} \) and \( \mu > \overline{\mu} \), 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 pertains to the shift in Period 2
advertising regime for opt-out consumers. Recall that in the absence of tracking, for low $\phi_M$, the advertiser’s strategy sometimes changes from not advertising in either period to mass-advertising in both periods as $\mu$ increases (see Proposition 1). 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 high $\mu$ when they cannot be tracked, consumers are incentivized to opt-in to being tracked in order to see fewer ads, at the expense of their privacy cost. Figure 6 depicts the precipitous increase in opt-in rate as $\mu$ increases past the threshold $\mu \approx 0.63$, after which the advertiser shows mass-advertising under no tracking.

The second force relates to the advertiser’s targeting regime in Period 2 for opt-in consumers. To illustrate, suppose $\mu$ and $\beta$ are high. In this case, the advertiser adopts the frequency strategy in Period 2 for opt-in consumers such that successive ads are targeted to $M$-consumers (see Proposition 2). 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 be targets of Period 2 advertising. This dampens consumers’ incentives to opting-in to being tracked. Figure 6 illustrates the associated decline in opt-in rate as $\mu$ increases in the neighborhood of $\mu \approx 0.85$.

**Ad Network Profit**

Next, we analyze how the consumer opt-in behavior characterized above impacts the ad network’s profit. We find that consumers’ opt-in choices lead to non-monotonicities in the ad network’s profit with respect to $\mu$. In particular, under certain conditions, the ad network’s profit decreases in $\mu$, even though higher $\mu$ implies higher consumer conversion on average.

Recall from Proposition 4 that the ad network’s profit is lower under no tracking than under tracking:
the advertiser has lower valuations for untargeted ads, which implies that under no tracking, ad slots are sold at lower prices, and sometimes fewer slots are filled. Since higher $\mu$ may result in more consumers opting-out from being tracked (see Proposition 5), we obtain that the ad network’s profit may decrease in $\mu$. As described above, the change in opt-in rate 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 $\mu$, then less consumers choose to opt-in. 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 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 7. The following proposition summarizes this finding.

**Proposition 6** (Equilibrium Ad Network Profit). Let $q^*(0)$ and $q^*(1)$ denote the equilibrium ad intensity without and with tracking, respectively, and let $\mu'$ and $\bar{\mu}'$ be as defined in the proof. Suppose consumers’ privacy costs $\theta$ are uniformly distributed on $[0,1]$. Under endogenous tracking, the ad network’s profit decreases in $\mu$ if and only if either

- $q^*(0) = 2, q^*(1) = 1 + \mu(1 - \phi_M), \eta(1 - \mu(1 - \phi_M)) < 1$, and either $\mu < \mu'$ or $\mu \geq \bar{\mu}'$, or
- $\phi_M \geq 1 - \sqrt{(\mu - k)/\mu}$ and $\beta \leq \beta < \bar{\beta}$.

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 consumer conversion increase advertiser’s valuation, which allows the ad network to sell more ad slots at higher prices. Proposition 6 provides a countervailing argument. Privacy regulations that allow consumers to choose 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, undermining targeting efficiency. This means that fewer ad slots may be sold and at lower prices. As a result, the ad network’s profit decreases.

**Consumer Surplus**

One of the main objectives of privacy regulations is to protect consumers. Consistent with intuition, we 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.

**Proposition 7** *(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, our analysis provides three important takeaways. First, consumers choose to opt-in to being-tracked if the effectiveness of the ads in inducing product interest, \( \mu \), is intermediate. This is driven by consumers’ incentives to avoid seeing mass-advertising shown to opt-out consumers. Second, the ad network’s profit under endogenous privacy choices may decrease in \( \mu \), even if higher \( \mu \) implies higher average conversion. The intuition is that more consumers may choose to opt-out from being tracked as \( \mu \) 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 we show in the next section, however, allowing the advertiser to have private information about 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 Extensions

In this section, we explore three extensions of the main model. First, we relax the assumption that ad valuation is common knowledge. Specifically, we allow the advertiser to have private information about its ad valuation, while the ad network only knows the valuation up to a distribution. Therefore, unlike the main model, the ad network leaves positive surplus to the advertiser. We show that the advertiser’s positive surplus introduces intricate dynamics to the model that, while maintaining the key results of the focal model, also reverses some results. Specifically, we find that the ad network’s profit
may decrease due to consumer tracking.

Second, we test the robustness of our results in a setting with competing advertisers. We consider a second advertiser who competes with the first advertiser for the advertising slot. Our analysis indicates that while competition reduces the parametric region for which ad slots are left unfilled across two time periods, the underlying insights from the main model are unchanged.

Finally, we consider an infinite-horizon game with heterogeneous, overlapping generations of consumers. This extension not only captures the reality of consumer flux more accurately, but also confirms that the obtained insights do not depend qualitatively on end-period effects inherent in finite-horizon models.

4.1 Information Asymmetry about Ad Valuation

In the main model, we assumed that the advertiser’s ad valuation is known by the ad network. Consequently, the ad network set the reserve price such that the advertiser’s surplus was entirely extracted. In this information asymmetry extension, we allow the advertiser’s ad valuation to be private information. In practice, ad valuations may not be fully known to the ad network. 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 capacity to infer consumer’s willingness to pay for a specific product offered by the advertiser (de Cornière and Nijssen, 2016).

To that end, we assume that at each period, the advertiser’s ad valuation is drawn independently from Uniform[0, 1] and is known privately by the advertiser. Therefore, in contrast to the main model, 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 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 8a), the result pertaining to the ad network’s profit is sometimes reversed. In particular, we find that under certain conditions, the ad network may benefit from regulations that allow consumers to endogenously choose to being tracked (see Figure 8b). This occurs when $\beta$ is sufficiently high and $\mu$ 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 helps the ad network efficiently extract surplus from the advertiser in Period 2,
it hurts 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 falls. This undermines the ad network’s ability to extract surplus efficiently from the advertiser. Thus, instead of selling targeted ads, the ad network sells untargeted ads that reach a larger consumer base. We 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 in 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.

In sum, our analysis sheds light on a novel role of privacy regulations that allow endogenous tracking choices: regulations can serve as a commitment device that 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.

### 4.2 Two Advertisers

In this section, we extend the main model by considering two advertisers indexed by $j \in \{1, 2\}$. Consumers are heterogeneous in their product preferences: $\lambda$ proportion of consumers are “loyal” to ad-
Figure 9: Two Advertisers: Ad Intensity With and Without Tracking; \( \phi_M = 0.5, k = 0.15, \lambda = 0.66 \)

Adverter 1, and \( 1 - \lambda \) proportion to advertiser 2. A consumer transitions down the purchase journey according to specifications of the main model (see Section 2) only if she sees an ad from the advertiser to which she is loyal. Without loss of generality, we assume that \( \lambda > \frac{1}{2} \); i.e., Advertiser 1 is the dominant brand.

We assume that opting-in to tracking reveals not only the consumers’ funnel states and purchase history, but also their ex ante product preference; 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.\(^{16}\) The rest of the model specifications remain unchanged. Note that the extension model reduces to the main model if \( \lambda = 1 \).

We find that the qualitative insights from the main model carry over for a large range of parameters (see Figure 9). The results diverge if and only if either (a) the effectiveness of top-funnel advertising is very high (i.e., high \( \mu \)) or (b) product preference heterogeneity is sufficiently large (i.e., \( \lambda \) close to 0.5). The intuition is that Advertiser 2 has high incentive to advertise in Period 2 after Advertiser 1 has shown ads 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 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, the ad slots are filled by Advertiser 2 in the extension (Region B in Figure 9). Finally, to avoid seeing

\(^{16}\)https://www.digitaltrends.com/computing/how-do-advertisers-track-you-online-we-found-out/
mass-advertising, consumers opt-in to tracking for this parameter range. We summarize this finding in the following proposition.

**Proposition 8** (Two Advertisers). *The advertising intensity differential between tracking and no tracking regimes carries over from the main model if either the effectiveness of top-funnel ads is not too high or product preference heterogeneity is not too large. Otherwise, opt-out consumers are exposed to ads in both periods: from Advertiser 1 in Period 1, and from Advertiser 2 in Period 2.*

### 4.3 Infinite Horizon with Heterogeneous Overlapping Consumer Generations

We extend the game from the main model along two dimensions. First, we relax the assumption that all newly arriving consumers are at funnel state $T$. In particular, we allow $\sigma \in [0,1]$ proportion of newly arriving consumers to be in funnel state $M$, and $1-\sigma$ in $T$. Broadly, $\sigma$ can be interpreted as the advertiser’s “brand strength”: the higher the $\sigma$, the greater the extent to which the advertiser’s product is *a priori* known and considered by consumers. Second, we extend the game horizon from two-period to infinite-period. At each period, a unit mass of consumers — $\sigma$ 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.

We solve for a Markov-perfect equilibrium (MPE) wherein an advertiser’s strategy depends only on the payoff-relevant state in that period (Villas-Boas, 1999). Following the analysis of the main model, we assume that the advertiser submits weakly dominant bids, and that the ad network sets reserve prices anticipating these bids. The ad network compares the total discounted profit (with discount factor $\delta \in [0,1]$) obtained from inducing the four advertising outcomes above, and then chooses whichever yields the highest profit.

Due to space considerations, we relegate the MPE derivation to Section [C] of the appendix. Here, we highlight how the qualitative insights obtained here compare to that of the main model. First, Figure [10] shows that if consumer heterogeneity is muted (i.e., $\sigma = 0$) and the discount factor is close to 1, the equilibrium outcomes essentially mirror that of the two-period main model (see Figure [3]). In particular, the advertiser buys mass-advertising in every period if and only if $\beta$ is sufficiently large and $\mu$ is intermediate. The underlying mechanism revolves around the saturation effect, and the qualitative insights remain essentially the same.

Second, we examine how the insights from the main model are moderated by two new parameters: the “brand strength” parameter $\sigma$, and the discount factor $\delta$. As illustrated in Figure [11], we find that as either $\sigma$ increases or $\delta$ decreases, the parametric region where the advertiser shows mass-advertising in
every period becomes smaller.

The intuition is the following. As $\sigma$ increases, a larger portion of newly arriving consumers are already mid-way down the funnel at state $M$. This has a similar effect as the saturation effect: since many newly arriving consumers are already in the consideration phase and will likely convert without ad exposures, the value of a successive ad diminishes (see Figure 11a). On the other hand, as $\delta$ 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 11b).

5 Conclusion

In this paper, we study the impact on the online advertising ecosystem of tracking consumers’ activities on the Internet, 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., GDPR). In particular, we model the consumer “purchase journey” and analyze the impact of consumers’ strategic opt-in behavior on the strategies and profits of advertisers and ad networks.

Among others, we establish the following insights from the analysis. First, consumers choose to opt-in...
to tracking when ad effectiveness is intermediate. For opt-out consumers, if ad effectiveness is high enough such that the advertiser finds advertising in the first period worthwhile, yet low enough such that successive advertising is not rendered wasteful by the first ad, then the advertiser shows mass advertising. For opt-in consumers, the advertiser shows targeted ads. Therefore, ad-averse consumers opt-in to tracking, trading-off their privacy costs with the benefit of seeing fewer ads under tracking.

Second, the ad network’s profit may decrease in ad effectiveness, even though higher ad effectiveness implies higher consumer conversion on average. Without tracking, if ad effectiveness is high, the first ad reduces the marginal effectiveness of the successive ad. With tracking, on the other hand, second period ads are shown to: (i) consumers who did not respond to the first ad and stayed in top of the funnel and/or (ii) consumers who transitioned to middle of the funnel but did not purchase. Overall, more ads are shown to consumers who can be tracked, which incentivizes consumers to opt-out from tracking. The resultant decline in consumer trackability reduces advertiser’s valuation for ads, which in turn results in ad slots being sold at lower prices. Therefore, the ad network’s profit may decline in ad effectiveness.

Finally, we 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 such as the GDPR may increase the ad network’s profit as well — even if some consumers
opt-out from tracking, thereby undermining targetability. The intuition is that lack of trackability forces the ad network to sell 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 in early rounds, resulting in higher ad network profit compared to the full tracking case. Thus, privacy regulations can serve as a commitment mechanism that better aligns the incentives of the ad network and the advertiser.

The results obtained in this paper provide important managerial insights for marketers and regulators alike. In particular, our 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, our results underscore the need for regulators to consider nuanced approaches to data privacy regulations that are based on various market conditions such as the degree of information asymmetry, consumer disutility for ads, their value of privacy, and the average effectiveness of ads.

We acknowledge several limitations of the paper. First, our model does not account for flexible pricing decisions 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 (e.g., monopolist’s price-discrimination and product price competition between multiple advertisers). Second, we implicitly assumed that the advertiser shows single 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 (a) how the advertiser’s content personalization for opt-in consumers may impact the funnel-transition probabilities, and (b) how the difference in transition probabilities across opt-in vs. opt-out consumers affects the equilibrium outcomes.

References


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


Appendix

A Micro-Model of Consumer Behavior

We describe a micro-model in which consumers’ ad responses and purchase behavior map to the consumer behavior described in the main model. Based on an information processing framework, we show that the micro-model below implies that in equilibrium, (i) $\mu$ fraction of consumers in state $T$ processes the ad, thereby moving to state $M$, while $1-\mu$ ignore the ad and stay in $T$, (ii) whenever a consumer is in state $M$, she purchases w.p. $\phi_M$ without ad exposure, whereas with ad exposure, her match realization probability increases to 1 w.p. $\beta$, and w.p. $1-\beta$ she buys w.p. $\phi_M$.

To that end, consider consumers as information processing agents who process ad content at some cost. Specifically, if consumers process the ad, they acquire relevant product information but pay cost $c \geq 0$. Consumers then probabilistically realize a match value. If consumers do not process the ad, then they do not consider purchasing the product. We assume that the information processing cost $c$ is realized independently at each Period $t \in \{1, 2\}$. In particular, it is equal to 0 w.p. $\mu$ and $\bar{c}$ w.p. $1-\mu$, for some large $\bar{c}$. The cost captures both the time-varying opportunity cost (e.g., consumers can sometimes afford to pay more attention to ads and process presented information) as well as the cognitive cost involved in browsing the advertiser’s website and acquiring more product information.

To replicate the outcomes of the main model as an endogenous, rational process, it suffices to establish the following: (i) if ads are shown, then $\mu$ consumers, who realize an information processing cost of 0, choose to process the ad. Additionally, these consumers do not benefit from strategically foregoing or delaying the processing of ads; and (ii) consumers who realize a match choose to purchase, whereas consumers who do not realize a match do not; in particular, a matched consumer does not strategically delay product purchase, nor does an unmatched consumer have strategic incentive to purchase the product.

First, does a $T$-consumer with zero information processing cost ever benefit from not processing the presented ad, thereby potentially avoiding ads targeted to $M$-consumers? No, because if zero-cost consumers do not process the ads, then everybody will remain in funnel state $T$ in Period 2 and the advertiser will target ads to $T$-consumers if the ad delivery cost $k$ is sufficiently small (i.e., $k \leq \mu\phi_M$). Therefore, $T$-consumers with zero information processing cost have no incentive to not process the ad in Period 1.
Second, consumers are not strategic in their purchasing decision. Delaying purchase even if a consumer has realized a match is not beneficial for the consumer because her product utility remains zero, and only opens the possibility of being exposed to additional ads in Period 2 since she has not purchased; she would have precluded this possibility had she purchased after realizing a match. Furthermore, purchasing even when a consumer does not realize a match is not beneficial if we assume that the maximal potential gain from seeing fewer ads $\eta$ is smaller than the product disutility of an unmatched consumer: $|0 - p^*| = 1$.

### B 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 $\alpha$ 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 $\alpha (\mu \phi_M + \mu (1 - \phi_M) \phi_M) - k + \alpha ((1 - \mu) \mu \phi_M + \mu (1 - \phi_M)^2 \beta) - k \geq (\alpha \mu \phi_M - k)^+$ and $\alpha ((1 - \mu) \mu \phi_M + \mu (1 - \phi_M)^2 \beta) - k \geq 0$,
- $1$ if $\alpha (\mu \phi_M + \mu (1 - \phi_M) \phi_M) - k \geq (\alpha \mu \phi_M - k)^+$ and $\alpha ((1 - \mu) \mu \phi_M + \mu (1 - \phi_M)^2 \beta) - k < 0$, and
- $0$ otherwise.

Similarly, the ad intensity under tracking is

- $1 + \alpha (1 - \mu) I_{\mu \phi_M - k \geq 0} + \alpha \mu ((1 - \phi_M) - k \geq 0) I_{(1 - \phi_M) - k \geq 0}$ if $\alpha (\mu \phi_M + \mu (1 - \phi_M) \phi_M) - k + \alpha ((1 - \mu) (\mu \phi_M - k)^+ + \mu (1 - \phi_M) (\beta (1 - \phi_M) - k)^+ \geq 0$, 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 12, we can replicate the advertising intensity differential patterns of the main model (Figure 5) for small values of $\alpha$, $\phi_M$, 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.

C 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 $\lambda_f^{\text{old}}$ 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 $T$, $\mu$ fraction transition to $M$, of which $\phi_M$ convert and $1 - \phi_M$ do not. Moreover, of the $\sigma$ $M$-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 distribution of old generation
non-converters in Period \( t \) would be
\[
\left( \lambda_T^{\text{old}}, \lambda_M^{\text{old}} \right) = ( (1 - \sigma)(1 - \mu), (1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M) .
\]
We label this state as \( \lambda_1 \), 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, \( \phi_M \) fraction of \( \sigma \) \( M \)-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
\[
\left( \lambda_T^{\text{old}}, \lambda_M^{\text{old}} \right) = ( 1 - \sigma, \sigma(1 - \phi_M) ) .
\]
We label this state as \( \lambda_0 \), 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 \( \lambda_0 \), which is equivalent to “pulse advertising” (i.e., alternate advertising with a single-period break in between; (iii) advertise only when the state is \( \lambda_1 \), which is effectively equivalent to (i); and (iv) never advertise. We compare the ad network’s profits for the respective strategies.

\textbf{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:
\[
(1 - \sigma)(1 - \mu)\phi_M + ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)\beta + (1 - \beta)\phi_M) + (1 - \sigma)\mu\phi_M - 2R_1 + \delta V_1 \geq ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)\phi_M + \delta V_0 ,
\]
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,

$$2R_1^* = (1 - \sigma)(2 - \mu)\mu\phi_M + ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)^2 \beta + \delta \left(\frac{((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)\phi_M + \delta \frac{\sigma(2 - \phi_M)\phi_M}{1 - \delta}}{1 - \delta} - \frac{\sigma(2 - \phi_M)\phi_M}{1 - \delta}\right).$$

Since this strategy induces the state to be perpetually $\lambda_1$, the ad network’s total profit is

$$\pi^I_N = \frac{1}{1 - \delta} \left((1 - \sigma)(2 - \mu)\mu\phi_M + ((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)^2 \beta + \delta \left(\frac{((1 - \sigma)\mu + \sigma(1 - \beta))(1 - \phi_M)\phi_M + \delta \frac{\sigma(2 - \phi_M)\phi_M}{1 - \delta}}{1 - \delta} - \frac{\sigma(2 - \phi_M)\phi_M}{1 - \delta}\right) - 2k\right).$$

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 $\lambda_1$) 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

$$2R_0^* = 2(1-\sigma)\mu\phi_M+(\sigma(1-\phi_M)+\sigma)(1-\phi_M)\beta+\delta \left( ((1-\sigma)\mu+\sigma(1-\beta))(1-\phi_M)+\sigma)\phi_M+\delta \frac{\sigma(2-\phi_M)\phi_M}{1-\delta} \right) - \frac{\sigma(2-\phi_M)\phi_M}{1-\delta}.$$  

Since this strategy yields alternating states, the ad network’s profit under this strategy is

$$\pi_{II}^N = \frac{1}{1-\delta^2} \left( 2(1-\sigma)\mu\phi_M+(\sigma(1-\phi_M)+\sigma)(1-\phi_M)\beta+\delta \left( ((1-\sigma)\mu+\sigma(1-\beta))(1-\phi_M)+\sigma)\phi_M+\delta \frac{\sigma(2-\phi_M)\phi_M}{1-\delta} \right) - \frac{\sigma(2-\phi_M)\phi_M}{1-\delta} \right) - 2k.$$  

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 $\pi_{I}^N$, $\pi_{II}^N$ and 0 yield the presented equilibrium regions.

D Proofs

Statement and Proof of Claim 1

Claim 1. Suppose a player’s payoff from bidding $b$ in an auction parametrized by tuple $(x, y, z, p)$ is

$$\pi(b) = \begin{cases} 
  x - yp & \text{if } b \geq p, \\
  z & \text{if } b < p,
\end{cases}$$

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 $b^* = (x-z)^+/y$.  

39
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.

Statement and Proof of Claim 2

Claim 2. Let $\phi(x) = \max[0, x]$ for all $x \in \mathbb{R}$. Then $f(x) + f(y) \geq f(x + y)$ for all $x, y \in \mathbb{R}$.

Proof.

$$\frac{1}{2} (f(x) + f(y)) = f \left( \frac{x}{2} \right) + f \left( \frac{y}{2} \right) \geq f \left( \frac{x}{2} + \frac{y}{2} \right) = \frac{1}{2} f(x + y)$$

where the equalities are due to linearity and inequality due to convexity.

Proof of Lemma 1

Proof. Consider the first subgame wherein the advertiser had shown ads in Period 1. Following Claim 1, the advertiser’s weakly dominant bid (against any reserve price $R_2^*$) in Period 2 is

$$b_{2|\text{ad}}^* = (1 - \mu)\mu \phi_M + \mu (1 - \phi_M) (\beta + (1 - \beta)\phi_M) - \mu (1 - \phi_M)^2 \beta$$

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 $b_{2|\text{ad}}^*$, 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 $R_{2|\text{no ad}}^* = \max \left[ k, b_{2|\text{no ad}}^* \right]$.

Proof of Lemma 2

Proof. The advertiser’s weakly dominant bid $b_1^*$ in Period 1 follows directly from Claim 1. 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 $b_1^* - k + (b_{|ad}^* - k)^+ \geq (b_{|no\ ad}^* - k)^+$, and $R_1^* \in (b_1^*, \infty)$ otherwise. The reserve price stated in the lemma satisfies this property.

Proof of Proposition 1

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 $((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+$. 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

$$\pi^N(R_1) = \begin{cases} 
R_1 - k + ((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+ & \text{if } R_1 \leq b_1^*, \\
0 + (\mu \phi_M - k)^+ & \text{if } R_1 > b_1^*,
\end{cases}$$

from which we obtain

$$R_1^* = \begin{cases} 
b_1^* & \text{if } b_1^* - k + ((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+ \geq (\mu \phi_M - k)^+, \\
(b_1^*, \infty) & \text{otherwise.}
\end{cases}$$

Since $R_1^*$ can be any number greater than $b_1^*$ when $b_1^* < k + (\mu \phi_M - k)^+ - ((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+$, we can write

$$R_1^* = \max \left[ k + (\mu \phi_M - k)^+ - ((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+, \mu \phi_M + \mu(1 - \phi_M)\phi_M \right].$$

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 $\mu \phi_M - k \geq 0$. But $\mu \phi_M - k \geq 0$ implies that $R_1^* = \max \left[ k + (\mu \phi_M - k)^+ - ((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+, \mu \phi_M + \mu(1 - \phi_M)\phi_M \right]$, which simplifies to max $\left[ \mu \phi_M - ((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+, \mu \phi_M + \mu(1 - \phi_M)\phi_M \right]$. This is strictly greater than $b_1^* = \mu \phi_M + \mu(1 - \phi_M)\phi_M$ if and only if $\mu \phi_M + \mu(1 - \phi_M)\phi_M < \mu \phi_M - ((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+$, which is equivalent to $\mu(1 - \phi_M)\phi_M < -((1 - \mu)\mu \phi_M + \mu(1 - \phi_M)^2\beta - k)^+$. 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 \( b_{2|ad}^* \geq R_{2|ad}^* \), which are equivalent to

\[
\mu \phi_M + \mu (1 - \phi_M) \phi_M \geq k + (\mu \phi_M - k)^+ - ((1 - \mu) \mu \phi_M + \mu (1 - \phi_M)^2 \beta - k)^+ \tag{5}
\]

and

\[
(1 - \mu) \mu \phi_M + \mu (1 - \phi_M)^2 \beta \geq k, \tag{6}
\]

respectively. Note that (6) implies that (5) simplifies to

\[
(1 - \mu) \mu \phi_M + \mu (1 - \phi_M)^2 \beta - k \geq k + (\mu \phi_M - k)^+ - (\mu (2 - \phi_M) \phi_M), \tag{7}
\]

which can be re-arranged in terms of \( \beta \) 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 (6) and (7) simplifies to (6) if \( \mu > \frac{k}{\phi_M (2 - \phi_M)} \), and to (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

\[
\mu \phi_M - k + \mu (1 - \phi_M) \phi_M \geq (\mu \phi_M - k)^+, \tag{8}
\]

which holds if \( \mu \phi_M \geq k \). If \( \mu \phi_M < k \), then (8) simplifies to \( k \leq \mu \phi_M (2 - \phi_M) \). This last inequality can be re-arranged in terms of \( \phi_M \) as \( \phi_M \geq 1 - \frac{\sqrt{(\mu - k)^+}}{\mu} \). 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

\[
\mu = \frac{\phi_M + \beta (1 - \phi_M)^2 - \sqrt{(\beta (1 - \phi_M)^2 + \phi_M)^2 - 4k \phi_M}}{2 \phi_M},
\]

\[
\bar{\mu} = \frac{\phi_M + \beta (1 - \phi_M)^2 + \sqrt{(\beta (1 - \phi_M)^2 + \phi_M)^2 - 4k \phi_M + \phi_M}}{2 \phi_M}
\]

be the two roots of \( (1 - \mu) \mu \phi_M + \mu (1 - \phi_M)^2 \beta = k \). The larger root \( \bar{\mu} \) is greater than 1 for all \( \beta \) greater than \( \bar{\beta} = \frac{k}{(1 - \phi_M)^2} \), and the roots do not exist for all \( \beta \) smaller than \( \beta = \frac{2 \sqrt{\phi_M - \phi_M}}{(1 - \phi_M)^2} \). Algebraic manipulations yield the conditions stated in the proposition.
Proof of Proposition 2

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

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

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 $TM$:

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

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:

$$(R_2^T, R_2^M, R_2^{TM}) = \begin{cases} 
(\max[k, \mu\phi_M], \infty, \infty) & \text{induces } T\text{-ad sales}, \\
(\infty, \max[k, \beta(1 - \phi_M)], \infty) & \text{induces } M\text{-ad sales}, \\
(\max[k, \mu\phi_M], \max[k, \beta(1 - \phi_M)], \infty) & \text{induces } T\text{- and } M\text{-ad sales}, \\
(\infty, \infty, \max[k, (1 - \mu)\mu\phi_M + \mu(1 - \phi_M)^2\beta \frac{1}{(1 - \mu) + \mu(1 - \phi_M)}]) & \text{induces } TM\text{-ad sales}.
\end{cases}$$

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 2, 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 1 implies that the advertiser’s weakly dominant bid is \(b_1^* = \mu \phi_M + (1 - \phi_M) \phi_M\). 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.,

\[
R_1^* = \max[k - ((1 - \mu)(\mu \phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+) + (\mu \phi_M - k)^+, b_1^*].
\]

Therefore, Period 1 ads are shown if and only if \(b_1^* \geq R_1^*\), which is equivalent to

\[
\mu \phi_M + (1 - \phi_M) \phi_M \geq k - ((1 - \mu)(\mu \phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+) + (\mu \phi_M - k)^+. \tag{9}
\]

Suppose \(\mu \phi_M \geq k\). Then (9) simplifies to \(\mu(1 - \phi_M) \phi_M \geq -((1 - \mu)(\mu \phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+)\), which is true. Suppose \(\mu \phi_M < k\). Then (9) simplifies to \(\mu \geq \mu \equiv \tilde{\mu} \equiv \frac{k}{\phi_M} \equiv \phi_M(2 - \phi_M) + (1 - \phi_M)(\beta(1 - \phi_M) - k)^+ \iff \phi_M \leq \phi_M(2 - \phi_M) \iff 1 \leq 2 - \phi_M\), which is true for all \(\phi_M \in [0, 1]\), we obtain that Period 1 ads are shown if and only if \(\mu \geq \tilde{\mu}\). ■

44
Proof of Proposition 3

Proof. Let $q^*(0)$ and $q^*(1)$ denote the equilibrium ad intensities without and with tracking, respectively. We begin 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 1. 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 $\mu < \tilde{\mu}$ and $\phi_M < 1 - \frac{\sqrt{(\mu-k)^+}}{\sqrt{\mu}}$ (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 $\beta \leq \frac{k}{1-\phi_M}$.

Proof of Proposition 4

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 $\pi_{2|ad}^N = (1-\mu)(\mu\phi_M - k)^+ + \mu(1-\phi_M)(\beta(1-\phi_M) - k)^+$. The ad network’s payoff under no tracking is $\pi_{2|no\ ad}^N = (1-\mu)\mu\phi_M + \mu(1-\phi_M)\beta(1-\phi_M) - k)^+$. 45
But we have

\[
\pi_{2|\text{no ad}}^N = ((1 - \mu)\mu\phi_M + \mu(1 - \phi_M)\beta(1 - \phi_M) - k)^+ \\
\leq ((1 - \mu)(\mu\phi_M - k) + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k))^+ \\
\leq (1 - \mu)(\mu\phi_M - k)^+ + \mu(1 - \phi_M)(\beta(1 - \phi_M) - k)^+ \\
= \pi_{2|\text{ad}}^N.
\]

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. □

**Proof of Proposition 5**

Proof. Consumers opt-in to tracking only if \(q^*(0) > q^*(1)\). But recall from from Proposition 3 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 \(-\eta q^*(1) - \tilde{\theta} = -\eta q^*(0)\). □

**Proof of Proposition 6**

Proof. Ad network’s profit can decrease in \(\mu\) due to two and only two reasons: (a) higher \(\mu\) 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) high \(\mu\) 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 \(\mu\) such that opt-in rate increases with \(\mu\). The opt-in rate is \(F(\eta (2 - (1 + \mu(1 - \phi_M)))) = F(\eta (1 - \mu(1 - \phi_M)))\), which decreases in \(\mu\) 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

\[
\pi_N = \mu\phi_M + \mu(1 - \phi_M)(\beta + (1 - \beta)\phi_M) - (1 + \mu(1 - \phi_M))k \\
+ (1 - \eta(1 - \mu(1 - \phi_M)))((1 - \mu)\mu\phi_M - (1 - (1 - \mu)\phi_M)k). \tag{10}
\]

We want to find the conditions under which (10) decreases in \(\mu\). 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 \(\mu\). Therefore, \(10\) is decreasing in \(\mu\) for \(\mu < \mu'\) and \(\mu > \mu'\) where the thresholds are respectively given by the two roots of \(10\) in increasing order.

The second part follows from Proposition 5: if \(\phi_M > 1 - \sqrt{(\mu - k)/\mu}\) and \(\beta < \beta\), then for \(\mu = \mu^-\), consumers opt-in, and for \(\mu = \mu^+\), 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 4).

Proof of Proposition 7

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

\[
CS(e) = \int_0^{\tilde{\theta}} -\eta q^*(1) - \theta \, dF + \int_{\tilde{\theta}}^1 -\eta q^*(0) \, dF \geq \int_0^{\tilde{\theta}} -\eta q^*(1) - \theta \, dF + \int_1^{1 - \eta q^*(1) - \theta} dF = CS(1).
\]

Proof of Proposition 8

Proof. We first show that opting-out of tracking does not signal the consumer’s types. Let \(\rho_i\) and \(\rho_j\) 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 \(\rho\). 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 \(\mu\), \(\beta\) and \(\phi_M\). 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 \((1 - \lambda)\mu\phi_M > \lambda ((1 - \mu)\mu\phi_M + \mu (1 - \phi_M)^2 \beta)\). But the difference \((1 - \lambda)\mu\phi_M - \lambda ((1 - \mu)\mu\phi_M + \mu (1 - \phi_M)^2 \beta)\) is convex with respect to \( \mu \) 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 \( \lambda > \frac{1}{2} \), 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 \( \lambda < 1 - \frac{k}{\mu \phi_M} \equiv \tilde{\lambda} \).