Privacy and Polarization: An Inference-Based Framework

Tommaso Bondi
Cornell Tech

Omid Rafieian
Cornell Tech

Yunfei (Jesse) Yao
CUHK

February 20, 2024

Abstract

Advances in behavioral targeting allow news publishers to monetize based on advertising. However, behavioral targeting requires consumer tracking, which has heightened privacy concerns among consumers and regulators. In this paper, we examine how stricter privacy regulations that ban consumer tracking affect news publishers' content strategies. We develop a theoretical framework that captures a change in privacy policies as a shift in publishers' inference about consumer types. We consider a model where news publishers choose the content and advertising, and ideologically heterogeneous consumers select their preferred content based on their ideology and idiosyncratic shocks. We compare two salient informational environments: (1) behavioral targeting, where perfect inference about consumers is allowed, and (2) contextual targeting, where consumer tracking is banned due to privacy regulations, and publishers can only infer consumer types based on their content choice. We show that privacy regulations that ban behavioral targeting incentivize publishers to shift towards more extreme and polarizing content in both monopoly and duopoly settings, even though the shift to more extreme content can hurt both demand and consumer welfare. In summary, our research uncovers a previously unexplored relationship between privacy and polarization, shedding light on the potential unintended consequences of privacy regulations in media markets.

JEL Codes: M37, L82, L13, D83.
Keywords: advertising, targeting, privacy, polarization.

1 Introduction

Digital publishers increasingly use advertising as a monetization strategy. At the core of ad-based monetization is behavioral ad targeting that creates a sustainable revenue stream for publishers and keeps the online content mostly free. However, behavioral ad targeting naturally requires the collection and use of consumer-level data, thereby leading to privacy concerns among consumers. According to a recent survey by Pew Research, over 80% of US adults are concerned about how companies use the data they collect from them, making the need for privacy regulation an issue with bipartisan support (McClain et al., 2023). In response to consumers' privacy concerns in the US and globally, regulatory bodies and even some private firms have started taking actions to restrict consumer tracking online and protect consumer privacy. A few prominent examples include the California Consumer Privacy Act (CCPA) and the European Union's General Data Protection Regulation (GDPR) in the public sector, and Apple's App Tracking Transparency (ATT) and Google's Privacy Sandbox in the private sector.

Although the main intent of privacy regulations is to safeguard consumer privacy, a consistent finding from past empirical research in this domain is that privacy regulation hurts digital publishers (Goldfarb and Tucker, 2011; Alcobendas et al., 2021; Johnson et al., 2023). More specifically, prior research suggests that the revenue loss due to privacy regulations is more pronounced for general interest (vs. specialized) publishers who would have greater uncertainty about their consumer types in the absence of consumer tracking (Goldfarb and Tucker, 2011). For example, in the absence of consumer tracking, a mainstream news website like the New York Times has a harder time inferring consumer types and interests to show them relevant ads, compared to a niche, ideologically extreme website like Infowars, which has more precise information about their consumers. Thus, the negative impact of privacy regulation on news publishers largely depends on the content they create.

The content-dependent impact of privacy regulation on news publishers gives rise to an important question: does privacy regulation affect content strategies employed by news publishers? To counter the loss imposed by privacy regulations, a news website can shift from general news coverage to more specific and niche content, thereby drawing more accurate inferences about users who consume that content and showing them more relevant ads. For example, a news publisher knows more about a consumer who clicked on an ideological opinion piece than a consumer who clicked on a daily news. Such a shift in news content has important implications for consumer welfare in the news landscape, especially given the rise in political polarization and media bias. Nevertheless, the extant research on privacy has focused on measuring the impact of regulations on publishers' market *outcomes* and largely ignored their strategic response in terms of *content design*.

In this paper, we bridge this gap and endogenize content decisions by news publishers in markets with and without privacy regulations that ban consumer tracking. We explicitly model the choice of mainstream and niche content using a simple Hotelling model of product design and consumer demand. Specifically, we seek to answer the following questions:

1. How does privacy regulation affect content strategies employed by a monopolist news publisher who monetizes based on advertising? How does the equilibrium change in a duopoly news market?

2. How do the aforementioned strategies affect consumer polarization? When do news publishers have greater incentive to create polarizing content?

3. What are the implications in terms of both demand, consumer welfare, and publisher profits?

To answer these questions, we develop a model in which profit-maximizing news publishers choose their content design and advertising, and ideologically heterogeneous readers consume it based on both its slant and their own idiosyncratic taste shocks. In our model, publishers monetize by selling ad impressions to advertisers. As such, their utility is the product of two separate components: (1) an extensive margin that captures the *quantity of impressions*, and (2) an intensive margin that captures the *quality of impressions*.

To study the impact of privacy regulations that ban consumer tracking (e.g., GDPR), we compare two different information environments: *behavioral targeting* and *contextual targeting*. The difference between the two comes from the possibility of tracking. Under *behavioral targeting*, publishers possess perfect knowledge of each consumer's type as they are allowed to track consumers. On the other hand, the contextual targeting scenario mimics the situation in the presence of privacy regulations that ban consumer tracking, where publishers are only allowed to use contextual information to target their ads to consumers. The comparison between contextual and behavioral targeting allows us to examine the downstream impact of privacy regulations on publishers' equilibrium choices and overall market outcomes.

Under behavioral targeting, since publishers can perfectly match ads to consumers on a one-to-one basis, they can achieve maximum quality of impressions regardless of what content each consumer chooses to consume. As a result, the publisher's profit-maximization problem is greatly simplified: because perfect ads-to-readers matching is possible, the profit-maximizing content choice is the one that maximizes the quantity of impressions.

Conversely, under contextual targeting, publishers no longer have access to perfect information about consumer types, so they need to rely solely on a single piece of information to infer consumer types: consumer's self-selection into content. As such, publishers may have an incentive to deviate from the demand-maximizing content strategy (equilibrium under behavioral targeting) and create content that helps increase the quality of impressions. In other words, publishers may benefit from creating content with lower demand but sharper signals about consumers, thereby achieving higher profits by balancing the quantity and quality of impressions.

We consider a monopolist news publisher and examine its equilibrium strategies under both behavioral and contextual targeting regimes. In our analysis, we show that deviating from the demand-maximizing content strategy can be an equilibrium under contextual targeting when consumers are sufficiently ideologically differentiated, and their sensitivity to imperfect ad targeting (mismatches between ad and consumer type) is sufficiently high. Interestingly, we find that the deviation is more toward extreme (vs. moderate) content as publishers can sharpen the signal about consumer types by moving towards the extreme ends of consumers' ideological preferences. Notably, this incentive persists even when such a shift towards extreme content reduces both consumer welfare and total welfare. This is a particularly important finding as it does not depend on an explicit consumer (or collective) distaste for increased polarization.

We then extend our model to a duopoly setting where two news publishers compete. In line with our main finding in the monopoly case, we show that the move from behavioral to contextual targeting due to privacy regulations results in a shift to more extreme and polarizing content.

Under behavioral targeting, our setting is consistent with the canonical Hotelling model: when enough consumers are centrists, both publishers choose a moderate (centrist) position, split demand equally, and make equally precise inference about their (identically distributed) readers. The picture changes with contextual targeting: in this case, even with a relatively uniformly distributed set of consumers, one or both publishers have incentives to switch to extreme content, sacrificing demand for inferential accuracy. Depending on the consumers' ideological distribution and the importance of targeting, we identify both extreme-moderate and extreme-extreme equilibria, and show that these are always more likely to exist under contextual targeting. Overall, we find that market forces will become stronger for publishers to produce more extreme and polarizing content.

In summary, our paper makes several contributions to the literature. We study the link between privacy and polarization and highlight the possibility of an unintended consequence of privacy regulations in increasing polarization. Our finding is important as it goes against extensive media speculations and policy memos that cite personalization as a key contributor to the increased polarization over the past few decades (Pariser, 2011). In our model, we reverse this and emphasize that it is exactly the *inability to personalize* that leads firms to move to more extreme and polarizing content to sharpen their inference about consumers. A key innovation of our framework is in explicitly modeling privacy as an inference problem. From a game theoretical standpoint, this modeling framework allows us to examine players' equilibrium responses to the privacy shifts. For empirical studies, this provides a framework to quantify the magnitude of information gain/loss in various data environments. Finally, in the context of media bias, we introduce inference motives as an important determinant of actions chosen by strategic players, which is largely ignored in the prior literature on media markets.

2 Literature Review

Our work relates to the literature on media markets. Early theoretical work in this domain considers the competition between broadcasters in the presence of the advertising market and shares insights into equilibrium outcomes in this market in terms of content provision and advertising strategies (Dukes and Gal-Or, 2003; Gal-Or and Dukes, 2003; Anderson and Coate, 2005; Godes et al., 2009).

A separate stream of work in this literature has examined content strategies as they relate to media bias and polarization (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2006; Xiang and Sarvary, 2007). With the growth of digital news consumption, a series of recent studies have focused on the specific aspects of the digital context and examined pricing and content strategies by media markets (Ambrus et al., 2016; Athey et al., 2018; Lin, 2020; Amaldoss et al., 2021; Jain and Qian, 2021; Amaldoss et al., 2023; Amaldoss and Du, 2023). Our work adds to this stream of work by studying a key aspect of the digital context: behavioral targeting and the possibility of privacy regulations. In particular, we study the impact of privacy regulations on media markets and examine the equilibrium outcomes in terms of content strategies and their implications for media bias and polarization.

Our work relates to the literature on behavioral targeting and personalization. With the advancements in targeting technologies in advertising markets, a series of papers have studied the impact of targeting accuracy on equilibrium market outcomes (Chen et al., 2001; Iyer et al., 2005; Levin and Milgrom, 2010; Bergemann and Bonatti, 2011; Sayedi, 2018; Rafieian and Yoganarasimhan, 2021; Shin and Yu, 2021; Shin and Shin, 2023). In the news media context, many have speculated that greater personalization results in more polarization, citing the positive correlation between the rise of political polarization in the US and the surge in personalized content delivery through online platforms (Pariser, 2011). Despite the widespread lay belief that causally connects personalization and polarization, empirical findings in this domain do not present a consistent viewpoint. Surprisingly, the demographic groups in the US that are least likely to use the Internet experienced the greatest increase in polarization (Boxell et al., 2017). Moreover, studies analyzing users' browsing histories reveal that despite social media and the Internet being associated with greater ideological divergence among users, they also increase exposure to opposing views (Flaxman et al., 2016). Investigations focused on specific platforms and their personalized features offer conflicting results concerning the link between personalization and polarization. Notably, studies investigating Facebook's news feed algorithm, Google's search personalization, and YouTube have found limited evidence suggesting that personalization contributes to content bias (Bakshy et al., 2015; Ribeiro et al., 2020; Hosseinmardi et al., 2021). In our paper, we build on the theoretical literature on ad targeting and develop a model to study the impact of personalization on the supply of polarizing content. Our work extends this literature by providing an inference-based theoretical account that presents a more nuanced view of the link between personalization and polarization.

Our paper relates to the literature on privacy. A vast body of theoretical work has examined different issues related to consumer identification, privacy, and information markets (Villas-Boas, 1999, 2004; Taylor, 2004; Acquisti and Varian, 2005; Bergemann and Bonatti, 2015; Bergemann et al., 2018; Yang, 2022; Ke and Sudhir, 2023; Yao, 2023). Empirical papers in this domain have studied the impact of privacy regulations on market outcomes in different settings (Goldfarb and Tucker, 2011; Johnson, 2022; Johnson et al., 2023). In particular, Goldfarb and Tucker (2011) study a change in tracking and targeting regulations and document lower response rates to ads and, therefore, lower ad revenues for publishers. Notably, they demonstrate a heterogeneous effect of privacy regulation on digital publishers, with general interest publishers such as the New York Times experiencing higher revenue loss than specialized publishers such as Car and Driver Magazine. We extend this literature by endogenizing the content design decision to allow publishers to respond optimally to the change in privacy policies. We present a generic theoretical framework that characterizes privacy as an inference problem, whereby stricter privacy policies have a more negative impact on the accuracy of inference about consumers. Importantly, we identify the possibility of increased polarization as an unintended consequence of privacy regulations in digital markets.

3 Model

We theoretically characterize a market where media firms (e.g., news publishers) create content for consumers and monetize by placing ads. To reflect consumers' ideological preferences, we assume three discrete consumer types $\{0, 1/2, 1\}$, where 0 and 1 refer to the opposing ideological ends (e.g., left vs. right) and 1/2 refers to the centrist position.[1] Consistent with consumer preferences, the set $\{0, 1/2, 1\}$ defines the action set for the media firm's content choice, which allows us to

[1] Our monopoly case can be entirely characterized by two ideological positions. However, this is not the case for duopoly, as whenever firms choose to provide partisan content, they naturally do so by choosing opposite positions. In order to have a unified framework and to match well-known phenomena such as political polarization, we use three positions throughout our model section.

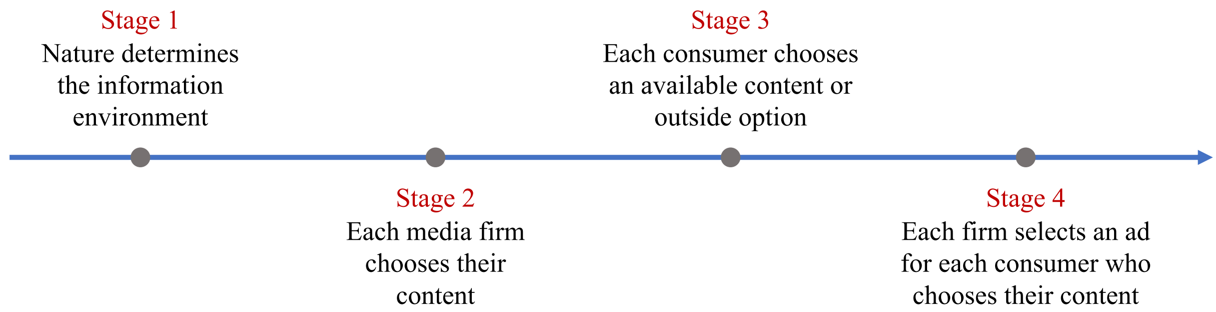


Figure 1: Timeline of the game

characterize the match between the consumer type and content choice by the media firm. While politics is a natural application for our model, our framework encompasses broader settings with ideological preferences about issues (e.g., animal rights vs. hunting websites). We unify our concept by referring to the content in the middle ($= 1/2$) as *mainstream* and the content at the two extremes ($\in \{0, 1\}$) as *niche*.[2]

We denote the consumer type or private signal by θ and assume a symmetric distribution for it that depends on a single parameter λ, with $1 - 2\lambda$ proportion of users having a centrist or mainstream position ($\theta = 1/2$), and λ proportion of consumers being at each extreme or niche end of consumer types ($\theta \in \{0, 1\}$). As such, we can view $\lambda \in [0, 1/2]$ as a measure of ideological partisanship among consumers, which is a key parameter in our analysis. We denote the media firm's content choice by x which comes from $\{0, 1/2, 1\}$. If the consumer chooses the content, an impression will be generated and the media firm can place an ad $a \in [0, 1]$ in that impression catering to different consumer types. As such, although the content x is the same for all consumers, ad a can be targeted at the impression level depending on the information available about the consumer.

Figure 1 presents the timeline of our game in four stages. We present the details of each stage as follows:

1. In the first stage, nature determines the information environment. We characterize the information environment with two components: (1) the distribution of consumer types, which is

[2]We stress that our use of "niche" and "mainstream" is in line with that in Johnson and Myatt (2006), and needs not have a market share interpretation. In other words, niche content is always more polarizing than mainstream one, but it might have lower demand.

characterized by λ, and (2) the observability of each consumer's type or private signal, which depends on whether behavioral tracking and targeting are allowed. The former component is common knowledge in all regimes, but privacy regulations can restrict the latter aspect of the information environment. When *behavioral targeting* is allowed, we assume that the media firm knows the consumer type for each consumer on its platform, as it has rich prior information about the consumer through tracking. On the other hand, when behavioral targeting is banned due to privacy regulations, the media firm needs to engage in *contextual targeting* and infer consumer type based on the context, e.g., the content consumed by the consumer.

2. In the second stage, each media firm chooses content x to maximize its expected profits. This content is broadcast to all consumers. As such, one could view the content choice as the firm's positioning, where the firm can choose one mainstream position at $x = 1/2$ and two niche positions at $x = 0$ and $x = 1$. In the political context, the niche positions can be interpreted as left- and right-leaning ideological positions.

3. In the third stage, each consumer consumes at most one content unit.[3] We define a type θ consumer's utility of consuming content x as follows:

$$U(x; \theta) = 1/2 - |x - \theta| + \epsilon, \tag{1}$$

where the first term is the base utility from media consumption that is set to $1/2$, the second term is the consumer's distaste for content mismatch as measured by the distance between the content from consumer type, and the third term is an idiosyncratic error term with mean zero that comes from the distribution $f(\cdot)$ such that $f(\epsilon) > 0,\ \forall \epsilon \in [-1/2, 1/2]$. We can generalize the base utility to any v, but fixing it at $v = 1/2$ conveys the main insights while keeping the model simple. We normalize the consumer's outside option to 0.

The utility framework in Equation (1) implies that the consumer located at one extreme ($\theta = 0$ or 1) will never consume the content located at the other extreme ($x = 1$ or 0), but

[3]The assumption that only one content unit is consumed is without loss of generality. As long as individual consumption does not depend on θ, all of our findings remain unchanged.

may consume the mainstream content ($x = 1/2$) if her idiosyncratic term for the mainstream content is sufficiently high. Mainstream consumers ($\theta = 1/2$) may consume the mainstream content or either niche content, depending on the relative values of the idiosyncratic terms.

4. In the final stage, the media firm chooses ad a if the consumer chooses to consume content x. The expected value generated from showing ad a to a consumer of type θ is given by the match value function $\mathcal{M}$ as follows:

$$\mathcal{M}(a;\theta) = 1 - \gamma(a-\theta)^2, \tag{2}$$

where the match value convexly decreases in the mismatch between ad and consumer $|a - \theta|$, and the parameter γ captures the value of targeting and is an instrumental parameter for our analysis. We assume that the match value function $\mathcal{M}(\cdot;\cdot)$ is a continuous and smooth function in both a and θ. As is clear from Equation (2), the maximum value of the match value function is one. We do not restrict ourselves to the case of $\gamma \leq 1$, so a large ad-consumer mismatch $(a-\theta)^2 > 1/\gamma)$ can result in negative ad profits. These could be interpreted, for instance, as the ad being repulsive to consumers (e.g., a hunting rifle ad for a vegetarian consumer), alienating them in the future.

We now define the media firm's profit maximization problem. Let $D(x)$ denote the total consumer demand given content choice x. This is the *quantity* of impressions the media firm is able to generate. For each impression, the firm can place an ad and obtain ad revenue that depends on the *quality* of the impression characterized by the match value function. We assume that there is a perfectly competitive ad marketplace, so the revenue-per-impression is the same as the expected match value of the ad given content. Let $\mathcal{I}$ denote the information available for each impression. We define the firm's optimal profits π given information $\mathcal{I}$ as follows:

$$\pi(x, a; \mathcal{I}) = \max_{x,a} D(x) \cdot \mathbf{E}\left[\mathcal{M}(a;\theta) \mid \mathcal{I}\right], \tag{3}$$

where the firm jointly maximizes content x and ad a given the information $\mathcal{I}$ available. In any event, the firm has contextual information about an impression, so we have $x \in \mathcal{I}$. Under behavioral targeting, we assume that the firm also has information about the exact consumer type θ. Since we are interested in cases under *behavioral* and *contextual* targeting, we characterize the information available under each as follows:

- The information under *behavioral* targeting is defined as $\mathcal{I}_b$ and includes the contextual information x as well as behavioral information θ, i.e., $\mathcal{I}_b = \{x, \theta\}$.

- The information under *contextual* targeting is defined as $\mathcal{I}_b$ and includes the contextual information x, i.e., $\mathcal{I}_c = \{x\}$.

We specifically model the information under behavioral and contextual targeting as extreme cases to better understand the incentives under these two information environments. However, our framework is flexible, and one could easily consider more middle-ground cases where there is more contextual information or less behavioral information.

4 Consumer Privacy and Firm's Inference Problem

A natural way to think about consumer privacy in our setting is to examine how uncertain the firm is about the consumer's type given the information $\mathcal{I}$, i.e., how private is the consumer's private signal. We present our definition of consumer privacy as follows:

Definition 1 (Consumer Privacy)**.** *The consumer's privacy given the information $\mathcal{I}$ is defined as* $\text{Privacy}(\mathcal{I})$ *and is equal to the conditional variance of consumer type given information $\mathcal{I}$, that is,* $\text{Privacy}(\mathcal{I}) = \mathbf{Var}(\theta|\mathcal{I})$.

We can easily verify that the consumer's privacy is equal to zero under behavioral targeting as $\mathcal{I}_b$ fully reveals the consumer's type. On the other hand, the consumer's privacy under contextual targeting depends on the content choice. Some choices can reveal more about the consumer type, thereby reducing the posterior variance of type and consumer privacy. This is an important consideration for the media firm when making a decision about the content. We characterize the relationship between consumer privacy and content choice in the following definition:

Definition 2 (Privacy Enhancing/Reducing Choice)**.** *Content choice x is more privacy-enhancing (privacy-reducing) than content choice x' if $Privacy(x) > (<) Privacy(x')$.*

We now revisit the firm's profit maximization problem as presented in Equation (3). For ad choice a, the firm needs to infer the consumer type given information $\mathcal{I}$. Since the ad choice is made conditional on content choice, we start with the firm's optimal ad choice and show the following lemma:

Lemma 1. *The firm's optimal ad choice a^* is the posterior mean of consumer type given information $\mathcal{I}$, i.e., $a^* = \mathbf{E}[\theta \mid \mathcal{I}]$.*

This lemma follows from the property of the variance and the quadratic specification of $\mathcal{M}$.[4] It indicates that the optimal ad a^* is the same as consumer type θ under behavioral targeting: $a^*(\theta) = \theta$, while it is equal to the conditional mean of θ under contextual targeting: $a^*(x) = \mathbf{E}(\theta|x)$.

Corollary 1. *The firm chooses the content x that maximizes $D(x) \cdot [1 - \gamma \mathrm{Privacy}(\mathcal{I})]$.*

The corollary implies that the firm's profits are negatively correlated with consumer privacy. Under behavioral targeting, the firm fully observes consumer type, so the firm chooses the content that maximizes its total demand:

$$\pi^b(x) = \max_x D(x),$$

where $\pi^b(x)$ is shorthand for the profits under behavioral targeting, and the firm is only interested in maximizing the *quantity* of impressions. Under contextual targeting, the firm needs to infer consumer type given their self-selection into content x, so the profit maximization problem will simplify to:

$$\pi^c(x) = \max_x D(x)[1 - \gamma Privacy(x)],$$

where $\pi^c(x)$ is shorthand for the profits under contextual targeting, and the firm has to trade-off the *quantity* of impressions with the *quality* of impressions. As a consequence, the demand-maximizing content choice is not necessarily the optimal content choice, as the firm may trade off more popular

[4]In other words, with a different functional form for $\mathcal{M}(a, \theta)$, the firm might find it optimal to deviate from $a^* = \mathbf{E}[\theta \mid \mathcal{I}]$ in the contextual targeting case.

content with a more accurate inference about the consumer. In our equilibrium analysis, we examine this inference incentive as it relates to polarization.

5 Equilibrium

5.1 Equilibrium Concept

Due to the setup of a multi-stage game with incomplete information, we consider the Perfect Bayesian Equilibrium. Specifically, when the firm chooses the advertising location in the last period, it updates its belief about consumer type based on the equilibrium strategy of the consumer, the content choice, and the consumption decision using Bayes' rule.

5.2 Monopoly

5.2.1 Demand

When there is only one firm, the consumer compares the utility of consuming the monopoly's content with the utility of choosing the outside option. Thus, she will consume the content if and only if the utility from content consumption is positive. The following lemma calculates the consumer demand given different content choices.

Lemma 2. *The total demand for a monopoly's mainstream content, $x = 1/2$, is $1-2F(0)\lambda$. Among them, $1 - 2\lambda$ are mainstream consumers, $[1 - F(0)]\lambda$ are type 0 consumers, and $[1 - F(0)]\lambda$ are type 1 consumers.*

The total demand for a monopoly's niche content, $x = 0$ or 1, is $1-F(0)+[2F(0)-1]\lambda$. Among them, λ are type $\theta = x$ consumers and $[1 - F(0)](1 - 2\lambda)$ are mainstream consumers.

This lemma highlights an important, general and natural property. The average distance between consumers to niche content is longer than between consumers to mainstream content. If the firm chooses mainstream content, $x = 1/2$, both types of niche consumers, $\theta = 0, 1$, may consume it because they are located not too far way from the content. In contrast, if the firm chooses niche content, say $x = 0$, mainstream consumers may consume it, while type $\theta = 1$ consumers will not, independent on their idiosyncratic taste shock, because they are too ideologically distant from it. This endogenous selection leads to the symmetry in the firm's inference ability. By choosing niche content $x = 0$, the firm knows for sure that it does not attract any type $\theta = 1$ consumers, and can better infer the consumer's type by Bayesian updating.

5.2.2 Content Choice

The monopoly's content choice is straightforward under behavioral advertising. It chooses the content that maximizes the total demand because there is no inference problem, and profit simply equals total demand.

Proposition 1. *Under behavioral advertising, the monopoly chooses mainstream content if* $[4F(0)-1]\lambda < F(0)$ *and niche content if* $[4F(0)-1]\lambda > F(0)$.

As we have discussed, the firm can perfectly target consumers in the advertising market and does not consider inference strategically under behavioral advertising. In contrast, it relies on Bayesian updating to infer consumer type and consider inference strategically under contextual advertising. Two conditions are necessary for the inference problem to be non-trivial. First, the total demand for the mainstream content is higher than that for the niche content, $[4F(0) - 1]\lambda < F(0)$. The firm will choose niche content under both behavioral and contextual advertising if this condition is violated. Second, the demand for mainstream content from mainstream consumers is lower than the demand for niche content from niche consumers of the same type, $\lambda > 1/3$. The firm will choose mainstream content under both behavioral and contextual advertising if the first condition holds but the second condition does not hold. In the subsequent analyses, we will focus on the interesting case where both conditions hold: $[4F(0) - 1]\lambda < F(0)$ and $\lambda > 1/3$.

Proposition 2. *Suppose* $[4F(0) - 1]\lambda < F(0)$ *and* $\lambda > 1/3$*, mainstream content is more privacy-enhancing than niche content. Under contextual advertising, there exists* $\gamma^m > 0$ *such that the monopoly chooses niche content if* $\gamma > \gamma^m$ *and mainstream content if* $\gamma < \gamma^m$*. In addition,* γ^m *decreases in* λ *if* $F(0) = 1/2$.

Condition	Equilibrium under Behavioral Advertising	Equilibrium under Contextual advertising
$[4F(0)-1]\lambda > F(0)$	niche	niche
$[4F(0)-1]\lambda < F(0)$ & $\lambda < 1/3$	mainstream	mainstream
$[4F(0)-1]\lambda < F(0),\ \lambda > 1/3,\ \&\ \gamma < \gamma^m$	mainstream	mainstream
$[\mathbf{4F(0)-1}]\lambda < \mathbf{F(0)},\ \lambda > \mathbf{1/3},\ \&\ \gamma > \gamma^{\mathbf{m}}$	**mainstream**	**niche**

Table 1: Monopoly Equilibrium Under Behavioral and Contextual Advertising

Table 1 compares the equilibrium under behavioral and contextual advertising. The last row in bold text highlights the case where the monopoly chooses mainstream content under behavioral advertising and niche content under contextual advertising. By choosing niche content rather than mainstream content, it faces a lower uncertainty about consumer type. Therefore, it can infer the consumer type and target consumers more accurately. Though the total demand is lower, the advertising revenue per consumer is higher. So, the firm faces a quantity and quality tradeoff: more eyeballs is equivalent (from an advertising perspective) to *worse* eyeballs. The optimal content then depends on how valuable better targeting is (the size of γ).

Intuitively, one would expect a more partisan consumer base (higher λ) to increase the likelihood the firm switches to niche content, since this is more in demand to begin with. Consequently, a higher λ leads to a lower γ^m, thus expanding the set of γ's that lead the firm to produce niche content ($\gamma > \gamma^m$).

5.2.3 Consumer Welfare

We have shown that the monopoly may trade off the total demand for a more accurate inference when behavioral tracking is banned. How does the switch from mainstream content to niche content affect consumer welfare?

Proposition 3. *Niche content leads to lower consumer welfare if and only if* $\lambda < \hat{\lambda} := \frac{1-2[1-F(0)]\mathbf{E}[\epsilon|\epsilon>0]}{3-8[1-F(0)]\mathbf{E}[\epsilon|\epsilon>0]}$. *In addition,* $\hat{\lambda} > 1/3$ *for any distribution of* ϵ.

Corollary 2. *Consumer welfare is lower under contextual advertising than under behavioral advertising if* $1/3 < \lambda < \hat{\lambda},\ [4F(0)-1]\lambda < F(0),$ *and* $\gamma > \gamma^m$.

Intuitively, niche (mainstream) consumers receive a higher surplus from consuming niche (mainstream) content. So, niche content leads to lower consumer welfare if most consumers are mainstream ones (low λ), while it leads to higher consumer welfare if most consumers are niche ones (high λ).

The cutoff threshold $\hat{\lambda}$ depends critically on the distribution of the idiosyncratic term. In particular, it is immediate to verify that $\hat{\lambda}$ is increasing in $\mathbf{E}(\epsilon|\epsilon > 0)$. A distribution of ϵ that is concentrated around 0 results in low values of $\mathbf{E}(\epsilon|\epsilon > 0)$, while the opposite is true for a bimodal distribution that concentrates mass near the extremes $-1/2$ and $1/2$.

The reason why this is a key determinant of consumer welfare is that $\mathbf{E}(\epsilon|\epsilon > 0)$ captures the average idiosyncratic match between readers and their chosen content, *conditional on them choosing mismatched content*.[5] Thus, the higher this quantity, the (absolutely and relatively) higher the consumer welfare associated with mainstream content, which attracts the highest degree of mismatches.

Proposition 3 also shows that there always exists a parameter range such that the optimal content choice by the monopoly (niche) decreases consumer welfare. Therefore, banning behavioral tracking may drive the monopolist publisher to switch from mainstream to niche content, even when doing so reduces both total demand and consumer welfare. This highlights an important unintended consequence of privacy regulation.

Figure 2 illustrates different content choices and consumer welfare changes under contextual advertising in non-trivial cases where $[4F(0) - 1]\lambda < F(0)$ and $\lambda > 1/3$. As we can see, all combinations of the content choice and welfare effect can occur in equilibrium.

[5] Given our utility specification, consumers always prefer ideologically aligned content to the outside option. Thus, in this case there is no positive selection in their idiosyncratic taste ϵ.

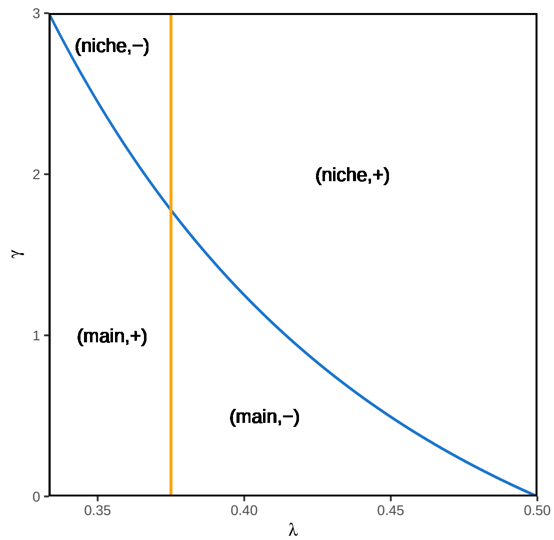


Figure 2: Monopoly's content choice and consumer welfare under contextual advertising when $\epsilon \sim U[-1/2, 1/2]$.
In each label (niche/main,+/-), the first element denotes the content choice, and the second element denotes whether the content choice increases or decreases consumer welfare compared to the opposite choice.

5.3 Duopoly

Now suppose that the first chooses x_1 content, while the second chooses x_2 content. Denote their content choices by (x_1, x_2). WLOG, $x_1 \leq x_2$. By symmetry, (0,0) is equivalent to (1,1), and (0,1/2) is equivalent to (1/2,1). So, we only consider (0,0), (0,1/2), (0,1), and (1/2,1/2). One can see that (0,1) dominates (0,0). So, (0,0) will never be an equilibrium.

5.3.1 Demand

Differently from the monopoly case, consumers not only compare the utility of consuming a firm's content with the outside option, but also with the utility of consuming the other firm's content. We characterize the demand given different content choices in the appendix.

It is important to note that, even without inference motivation (behavioral advertising), competing firms have an incentive to choose more polarized content to increase differentiation and soften the competition, which increases the total demand. Therefore, it is, in some sense, harder for them to choose more polarized content when the information environment shifts from behavioral targeting to contextual targeting. Nevertheless, the next section shows that the inference motivation becomes stronger with competition.

5.3.2 Content Choice

Due to tractability reasons, we make an additional assumption that $\epsilon \sim U[-1/2, 1/2]$ to simplify the demand function. The following result summarizes the equilibria under behavioral and contextual advertising.

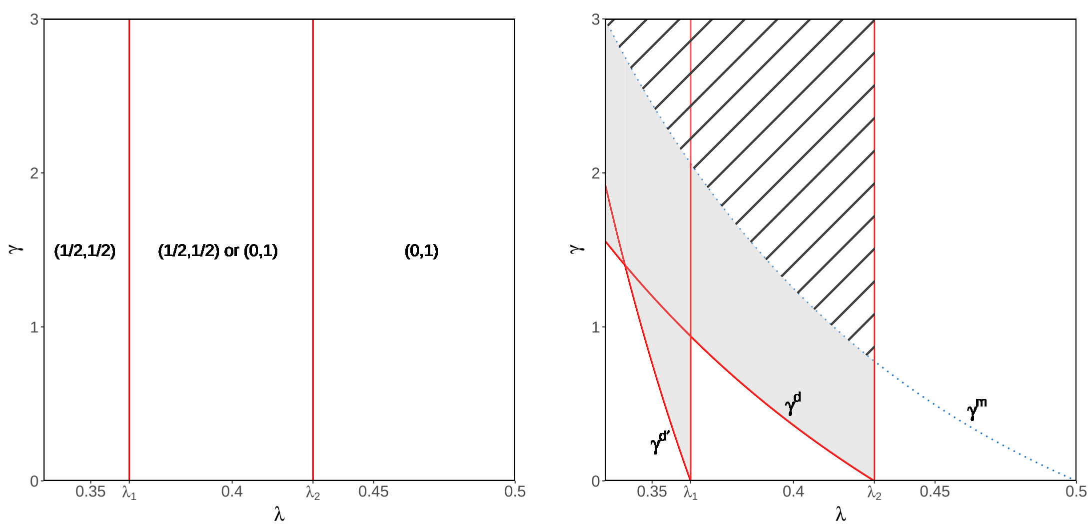
Proposition 4. *Suppose $\epsilon \sim U[-1/2, 1/2]$. The equilibria under behavioral and contextual advertising are summarized in the following table, where $1/3 < \lambda_0 < \lambda_1 < \lambda_2 = 3/7$, $\gamma^d > 0$ for $\lambda < \lambda_2$, and $\gamma^{d'} > 0$ for $\lambda < \lambda_1$.*

Condition	*Equilibrium under Behavioral Advertising*	*Equilibrium under Contextual advertising*
$1/3 < \lambda < \lambda_1$ & $\gamma < \min\{\gamma^d, \gamma^{d'}\}$	*(1/2,1/2)*	*(1/2,1/2)*
$\mathbf{1/3} < \lambda < \lambda_\mathbf{0}$ & $\gamma^\mathbf{d} < \gamma < \gamma^{\mathbf{d}'}$	***(1/2,1/2)***	***(0,1/2)***
$\lambda_\mathbf{0} < \lambda < \lambda_\mathbf{1}$ & $\gamma^{\mathbf{d}'} < \gamma < \gamma^\mathbf{d}$	***(1/2,1/2)***	***(1/2,1/2) or (0,1)***
$\mathbf{1/3} < \lambda < \lambda_\mathbf{1}$ & $\gamma > \max\{\gamma^\mathbf{d}, \gamma^{\mathbf{d}'}\}$	***(1/2,1/2)***	***(0,1)***
$\lambda_0 < \lambda_1 < \lambda < \lambda_2$ & $\gamma < \gamma^d$	*(1/2,1/2) or (0,1)*	*(1/2,1/2) or (0,1)*
$\lambda_\mathbf{1} < \lambda < \lambda_\mathbf{2}$ & $\gamma > \gamma^\mathbf{d}$	***(1/2,1/2) or (0,1)***	***(0,1)***
$\lambda > \lambda_2$	*(0,1)*	*(0,1)*

The rows in bold text highlight cases where the duopoly chooses more polarized content under contextual advertising than under behavioral advertising. Figure 3 also illustrates the duopoly's content choices under behavioral (left figure) and contextual (right figure) advertising. The content choices are the same under both behavioral and contextual advertising in the white region, while the extent of media polarization is larger under contextual advertising than under behavioral advertising in the solid and striped shaded region.[6]

[6] The extent of media polarization increases from (1/2,1/2) to (0,1/2) to (0,1).

When $\lambda > \lambda_2$, there are enough niche consumers so that the duopoly will choose the polarized content strategy (0,1) even under behavioral advertising. So, there is no room for the duopoly to be more polarized when switching to contextual advertising. Now consider the interesting case in which $\lambda < \lambda_2$. In both monopoly and duopoly cases, contextual advertising leads to more polarization when $\gamma > \gamma^m$ (the striped shaded region in Figure 3b). In addition, it leads to more polarization when $\min\{\gamma^d, \gamma^{d'}\} < \gamma < \gamma^m$ and $\lambda < \lambda_1$ or when $\gamma^d < \gamma < \gamma^m$ and $\lambda_1 < \lambda < \lambda_2$ (the solid shaded region in Figure 3b) under duopoly but not under monopoly. The intuition is that, on the one hand, duopolistic firms directly cannibalize each other's demand if they choose mainstream content. In contrast, they can cover different niche consumers by choosing the opposite content locations and soften the competition. As a result, the benefit of a higher total demand from mainstream content in the duopoly case is lower than in the monopoly case. On the other hand, the benefit of a more accurate inference by choosing niche content is not affected by competition. So, firms are leaning more toward quality in the quantity-quality trade-off under duopoly than under monopoly.



(a) Behavioral advertising (b) Contextual advertising

Figure 3: Duopoly's content choices when $\epsilon \sim U[-1/2, 1/2]$.
White region: the same equilibrium as the behavioral advertising case; Striped shaded region: more polarized equilibrium than the behavioral advertising case, under both monopoly and duopoly; Solid shaded region: more polarized equilibrium than the behavioral advertising case, under duopoly but not under monopoly.

It is interesting to notice that, unless both firms pick the mainstream strategy, an *inferential complementarity* arises: the presence of each firm aids the other firm's inference problem. Competition erodes demand, but improves the selection of consumers: on average, consumers of each firm are more ideologically aligned with it compared to the monopoly case. For instance, in the $(0, 1)$ case, both firms lose relatively more demand from consumers of opposite ideology. Clearly, this complementarity motive is entirely absent when both firms choose mainstream strategies, $(1/2, 1/2)$. In this case, each firm's demand and profit is simply cut in half. In other words, firms' inferential ability is exactly the same as in monopoly, as no additional consumer ideological self-selection arises.

The fact that an increase in privacy results in an increase in equilibrium content polarization is interesting in light of the well-discussed link between content personalization (that is, lack of privacy) and polarization. It is important to stress that, according to this view, personalized content does not *polarize* consumers; rather, it simply *matches* their prior ideological positions. One could say that in this case content positioning is the consequence, rather than cause, of polarization.

The drivers of Proposition 4 are quite different: the desire to make precise inference to increase advertising profits pushes two competing firms away from the standard Hotelling equilibrium, $(1/2, 1/2)$, and toward content $(0, 1)$ that is more ideologically polarized than the underlying distribution of the consumers' ideological preferences, for any value of $\lambda < 1/2$. More privacy leads to more polarization. Polarization in content is driven by firms' desire to monetize ads, not (just) consumer ideology; in fact, the content distribution often becomes more polarized than consumers.

However, with two firms, polarization becomes desirable for consumers. The next proposition shows that consumer welfare always increases in the extent of polarization under duopoly.

Proposition 5. *Suppose* $\lambda > 1/3$, $[4F(0)-1]\lambda < F(0)$, *and* $\epsilon \sim U[-1/2, 1/2]$. *(0,1) content choices always lead to higher consumer welfare than (0,1/2) content choices. (0,1/2) content choices always lead to higher consumer welfare than (1/2,1/2) content choices.*

6 General Base Utility

We fix the consumer's base utility from media consumption at 1/2 in the main model to keep the model simple. That setup can convey all the main insights about the inference role of content choice and the quantity-quality tradeoff. In this extension, we generalize the base utility to v, and show the robustness of our findings. In addition, we demonstrate how the size of the base utility strengthens/weakens the firm's inference incentive.

Proposition 6. *Suppose $\epsilon \sim U[-1/2, 1/2]$. There exists $\underline{v} < 1/2, \overline{v} > 1/2, \gamma^m > 0, \gamma^{d''}, \lambda_1 < \lambda_2$ and $\overline{\lambda}(v) \in (1/3, 1/2]$ such that the following holds if $\underline{v} < v < \overline{v}$:*

1.(Monopoly) The firm chooses mainstream content under behavioral advertising and niche content under contextual advertising if $1/3 < \lambda < \overline{\lambda}(v)$ and $\gamma > \gamma^m$. The threshold γ^m decreases in λ and increases in v. In addition, there exists $\hat{\lambda}(v) > 1/3$ such that niche content leads to lower consumer welfare if $\lambda < \hat{\lambda}(v)$.

2.(Duopoly) The equilibrium under behavioral advertising is (1/2,1/2) if $\lambda < \lambda_1$, (1/2,1/2) or (0,1) if $\lambda_1 < \lambda < \lambda_2$, and (0,1) if $\lambda > \lambda_2$. The equilibrium under contextual advertising is more polarized than that under behavioral advertising if $\gamma > \gamma^{d''}$ and $1/3 < \lambda < \lambda_2$. Moreover, $\gamma^{d''} < \gamma^m$ for all $\lambda \in (1/3, \lambda_2)$.

As we can see, for an interval range of the base utility that contains 1/2, we recover all the main insights from Proposition 1 to 4 in the main model. We still have that content choice has an inference role for the firm under contextual advertising, a monopoly may choose more polarized content to increase profit per consumer despite the lower total demand and potentially lower consumer welfare, and competition strengthens firms' inference incentive.

The general base utility setup provides an additional insight on how the base utility affect the firm's inference incentive. The firm has a smaller incentive to aid its inference by choosing more polarized content as the base utility increases because γ^m increases in v. The base utility can be viewed as vertical quality of the media firm, whereas the match between consumer type and media/ad location is horizontal. When the vertical quality is high, both consumers whose tastes perfectly match the content location and those whose tastes differ from the content location will consume the media. Consequently, the horizontal match quality becomes less important and it becomes harder for the firm to distinguish different consumer types based on their self-selection. The firm is more incentivized to favor quantity in the quantity-quality tradeoff by choosing mainstream content with larger total demand. In contrast, when the base utility is low, consumers are more selective in their consumption decision. Mainstream consumers will not consume niche content that has a low vertical quality and poor horizontal match unless the random shock is very high (a large idiosyncratic term). So, by choosing the more polarized niche content, the firm can screen consumer types very accurately, and thus obtain a high profit per consumer by offering well-matched ads to most consumers. In this case, it has a stronger incentive to favor quality in the quantity-quality tradeoff by choosing niche content.

7 Conclusion

Over the last few years, increasingly strict privacy regulations have altered the ability of online publishers to monetize on advertising.

How does this affect publishers' content strategies? We examine this question theoretically by building a simple model of media product design and consumer demand. In our baseline model, a monopolist media publisher designs content and consumers select their preferred content based on both their ideology and idiosyncratic shocks. Advertisers pay the publisher according to the expected efficacy of their ads for different pieces of content.

We focus on the comparison of two salient informational environments: under *behavioral targeting*, publishers can perfectly track readers, and thus personalize ads on a one-to-one basis. Under *contextual targeting*, publishers can not track readers, and are therefore limited to tailoring ads to content, relying on the information contained in readers' self-selection. Thus, we show, firm's profits are inversely proportional to the amount of its consumers' privacy.

We then show that banning behavioral targeting incentivizes publishers to shift towards niche (or polarizing) content to aid inference about readers (thus decreasing their privacy), thereby increasing the publisher's ad revenue per reader.

In monopoly, this holds true even when a switch to niche content decreases total demand and hurts consumer welfare. In duopoly, contextual targeting increases polarization in equilibrium content provision, often pushing both publishers to produce (ideologically opposite) niche content. This holds true even when duopoly strategies hurt demand; however, in this case consumer welfare is higher than it would be were both firms to choose the mainstream strategy.

The findings of this study shed light on the intricate connection between privacy and polarization. In recent years, a growing body of research has underscored the pivotal role played by increasingly precise personalization, driven by behavioral targeting, in the surge of political polarization within the United States. This perspective posits that reduced privacy leads to heightened personalization, ultimately intensifying polarization (Pariser, 2011). However, our research reverses this insight by offering an alternative account. In our model, the inability to personalize content compels firms to employ an alternative strategy—consumer self-selection. This strategic shift creates a unique incentive for the production of more partisan content, particularly in duopoly scenarios, where polarization intensifies (as noted in the references).

Our research underscores an important shift from viewing content polarization as a mere reflection of consumers' existing ideological polarization. Instead, we demonstrate the active role played by content polarization in several scenarios. The equilibrium content consumed becomes more extreme and polarized than the initial ideological distribution. In essence, the content itself actively contributes to the escalation of polarization, a significant finding that challenges prior assumptions about this relationship.

These findings contribute to the ongoing discourse on privacy, personalization, and political polarization. They highlight the complex dynamics at play and encourage further exploration into the underlying mechanisms of this relationship. As privacy regulations continue to evolve and shape the digital landscape, understanding the impact of these changes on content strategies and societal polarization remains a pertinent area for both research and policy.

In summary, our research uncovers a previously unexplored relationship between privacy and polarization, shedding light on the potential unintended consequences of privacy regulations in media markets.

References

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

Miguel Alcobendas, Shunto Kobayashi, and Matthew Shum. The impact of privacy measures on online advertising markets. *Available at SSRN 3782889*, 2021.

W Amaldoss and J Du. How can publishers collaborate and compete with news aggregators? *Journal of Marketing Research*, 2023.

Wilfred Amaldoss, Jinzhao Du, and Woochoel Shin. Media platforms' content provision strategies and sources of profits. *Marketing Science*, 40(3):527–547, 2021.

Wilfred Amaldoss, Jinzhao Du, and Woochoel Shin. Pricing strategy of competing media platforms. *Marketing Science*, 2023.

Attila Ambrus, Emilio Calvano, and Markus Reisinger. Either or both competition: A "two-sided" theory of advertising with overlapping viewerships. *American Economic Journal: Microeconomics*, 8(3):189–222, 2016.

Simon P Anderson and Stephen Coate. Market provision of broadcasting: A welfare analysis. *The review of Economic studies*, 72(4):947–972, 2005.

Susan Athey, Emilio Calvano, and Joshua S Gans. The impact of consumer multi-homing on advertising markets and media competition. *Management science*, 64(4):1574–1590, 2018.

Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.

Dirk Bergemann and Alessandro Bonatti. Targeting in advertising markets: implications for offline versus online media. *The RAND Journal of Economics*, 42(3):417–443, 2011.

Dirk Bergemann and Alessandro Bonatti. Selling cookies. *American Economic Journal: Microeconomics*, 7(3):259–294, 2015.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The design and price of information. *American economic review*, 108(1):1–48, 2018.

Levi Boxell, Matthew Gentzkow, and Jesse M Shapiro. Greater internet use is not associated with faster growth in political polarization among us demographic groups. *Proceedings of the National Academy of Sciences*, 114(40):10612–10617, 2017.

Yuxin Chen, Chakravarthi Narasimhan, and Z John Zhang. Individual marketing with imperfect targetability. *Marketing Science*, 20(1):23–41, 2001.

Anthony Dukes and Esther Gal-Or. Negotiations and exclusivity contracts for advertising. *Marketing Science*, 22(2):222–245, 2003.

Seth Flaxman, Sharad Goel, and Justin M Rao. Filter bubbles, echo chambers, and online news consumption. *Public opinion quarterly*, 80(S1):298–320, 2016.

Esther Gal-Or and Anthony Dukes. Minimum differentiation in commercial media markets. *Journal of Economics & Management Strategy*, 12(3):291–325, 2003.

Matthew Gentzkow and Jesse M Shapiro. Media bias and reputation. *Journal of political Economy*, 114(2):280–316, 2006.

David Godes, Elie Ofek, and Miklos Sarvary. Content vs. advertising: The impact of competition on media firm strategy. *Marketing Science*, 28(1):20–35, 2009.

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

Homa Hosseinmardi, Amir Ghasemian, Aaron Clauset, Markus Mobius, David M Rothschild, and Duncan J Watts. Examining the consumption of radical content on youtube. *Proceedings of the National Academy of Sciences*, 118(32), 2021.

Ganesh Iyer, David Soberman, and J Miguel Villas-Boas. The targeting of advertising. *Marketing Science*, 24(3):461–476, 2005.

Sanjay Jain and Kun Qian. Compensating online content producers: A theoretical analysis. *Management Science*, 67(11):7075–7090, 2021.

Garrett Johnson. Economic research on privacy regulation: Lessons from the gdpr and beyond. 2022.

Garrett A Johnson, Scott K Shriver, and Samuel G Goldberg. Privacy and market concentration: intended and unintended consequences of the gdpr. *Management Science*, 2023.

Justin P Johnson and David P Myatt. On the simple economics of advertising, marketing, and product design. *American Economic Review*, 96(3):756–784, 2006.

T Tony Ke and K Sudhir. Privacy rights and data security: Gdpr and personal data markets. *Management Science*, 69(8):4389–4412, 2023.

Jonathan Levin and Paul Milgrom. Online advertising: Heterogeneity and conflation in market design. *American Economic Review*, 100(2):603–607, 2010.

Song Lin. Two-sided price discrimination by media platforms. *Marketing Science*, 39(2):317–338, 2020.

Colleen McClain, Michelle Faverio, Monica Anderson, and Eugenie Park. How Americans View Data Privacy, 2023. URL `https://www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy/`.

Sendhil Mullainathan and Andrei Shleifer. The market for news. *American economic review*, 95 (4):1031–1053, 2005.

Eli Pariser. *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin, 2011.

Omid Rafieian and Hema Yoganarasimhan. Targeting and privacy in mobile advertising. *Marketing Science*, 40(2):193–218, 2021.

Manoel Horta Ribeiro, Raphael Ottoni, Robert West, Virgílio AF Almeida, and Wagner Meira Jr. Auditing radicalization pathways on youtube. In *Proceedings of the 2020 conference on fairness, accountability, and transparency*, pages 131–141, 2020.

Amin Sayedi. Real-time bidding in online display advertising. *Marketing Science*, 37(4):553–568, 2018.

Jiwoong Shin and Woochoel Shin. A theory of irrelevant advertising: An agency-induced targeting inefficiency. *Management Science*, 69(8):4481–4497, 2023.

Jiwoong Shin and Jungju Yu. Targeted advertising and consumer inference. *Marketing Science*, 40 (5):900–922, 2021.

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

J Miguel Villas-Boas. Dynamic competition with customer recognition. *The Rand Journal of Economics*, pages 604–631, 1999.

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

Yi Xiang and Miklos Sarvary. News consumption and media bias. *Marketing Science*, 26(5): 611–628, 2007.

Kai Hao Yang. Selling consumer data for profit: Optimal market-segmentation design and its consequences. *American Economic Review*, 112(4):1364–1393, 2022.

Yunfei Jesse Yao. Failure of reputation for privacy. 2023.

Appendix

Proof of Lemma 1. Let p_j denote the posterior probability of $\theta = j$ conditional on the information $\mathcal{I}$, that is, $p_j = P(\theta = j \mid \mathcal{I})$. Therefore, we can write:

$$\begin{aligned}\mathbf{E}[\theta \mid \mathcal{I}] &= p_0 \times 0 + p_{1/2} \times \frac{1}{2} + p_1 \times 1 \\ &= \frac{p_{1/2} + 2p_1}{2} \end{aligned} \tag{4}$$

Now, we write the optimal ad choice as follows:

$$\begin{aligned}\max_a \mathbf{E}\left[\mathcal{M}(a;\theta) \mid \mathcal{I}\right] &= \max_a \left[1 - \gamma \mathbf{E}\left[(a-\theta)^2 \mid \mathcal{I}\right]\right] \\ &= \max_a \left[1 - \gamma\left(p_0 a^2 + p_{1/2}\left(a - \frac{1}{2}\right)^2 + p_1(a-1)^2\right)\right] \end{aligned} \tag{5}$$

We can now write the first-order condition as follows:

$$\begin{aligned}\frac{\partial \mathbf{E}\left[\mathcal{M}(a;\theta) \mid \mathcal{I}\right]}{\partial a} &= 1 - \gamma\left(2p_0 a + 2p_{1/2}\left(a - \frac{1}{2}\right) + 2p_1(a-1)\right) \\ &= 1 - \gamma\left(a(2p_0 + 2p_{1/2} + 2p_1) - p_{1/2} - 2p_1\right) \\ &= 1 - \gamma\left(2a - p_{1/2} - 2p_1\right) \\ &= 0 \end{aligned} \tag{6}$$

Solving for the first-order condition, we can write:

$$a^* = \frac{p_{1/2} + 2p_1}{2} = \mathbf{E}[\theta \mid \mathcal{I}] \tag{7}$$

■

Proof of Corollary 1. The media firm's profit maximization problem is the following:

$$\begin{aligned}
\pi(x, a; \mathcal{I}) &= \max_{x,a} D(x) \cdot \mathbf{E}\left[\mathcal{M}(a; \theta) \mid \mathcal{I}\right] \\
&= \max_{x} D(x) \cdot \mathbf{E}\left[\mathcal{M}(a^*; \theta) \mid \mathcal{I}\right] \\
&= \max_{x} D(x) \cdot \mathbf{E}\left[1 - \gamma(a^* - \theta)^2 \mid \mathcal{I}\right] \\
&= \max_{x} D(x) \cdot \left(1 - \gamma\mathbf{E}\left[(a^* - \theta)^2 \mid \mathcal{I}\right]\right) \\
&= \max_{x} D(x) \cdot \left(1 - \gamma\mathbf{E}\left[(\mathbf{E}[\theta \mid \mathcal{I}] - \theta)^2 \mid \mathcal{I}\right]\right) \\
&= \max_{x} D(x) \cdot [1 - \gamma\mathbf{Var}(\theta|\mathcal{I})] \\
&= \max_{x} D(x) \cdot [1 - \gamma Privacy(\mathcal{I})]
\end{aligned}$$

■

Proof of Lemma 2. One can see that consumers whose type matches exactly the content's location will consumer for sure. So, we just need to determine whether other consumers prefer consuming the content to the outside option.

1. Mainstream content $x = 1/2$

   The demand from mainstream consumers is $1 - 2\lambda$.

   Consider a type 0 consumer. Her utility from consuming the content is $U(1/2, 0) = \epsilon$. She will consume the content if and only if $\epsilon > 0$, her utility from the outside option. The consumption probability is:

   $$P(\epsilon > 0) = 1 - F(0)$$

   Hence, the demand from type 0 consumers is $\lambda[1 - F(0)]$. By symmetry, the demand from type 1 consumers is also $\lambda[1 - F(0)]$.

   In sum, the total demand is $1 - 2\lambda + 2\lambda[1 - F(0)]$.

2. Niche content $x = 0$

   The demand from type 0 consumers is λ and the demand from type 1 consumers is 0.

Consider a type 1/2 consumer. Her utility from consuming the content is $U(0, 1/2) = \epsilon$. She will consume the content if and only if $\epsilon > 0$, her utility from the outside option. By the same argument in the previous case, one can see that the demand from type 1/2 consumers is $(1-2\lambda)[1-F(0)]$.

In sum, the total demand is $\lambda + (1-2\lambda)[1-F(0)]$.

3. Niche content $x = 1$

   It is symmetric to the previous $x = 0$ case.

■

Proof of Proposition 1. Under behavioral advertising, the match between the ad and the consumer type is always 1. Therefore, the monopoly chooses the content that leads to the highest demand. According to Lemma 1, mainstream content generates a higher demand than niche content does if and only if $1 - 2F(0)\lambda > 1 - F(0) + [2F(0) - 1]\lambda \Leftrightarrow [4F(0) - 1]\lambda > F(0)$. ■

Proof of Proposition 2. We prove a lemma first.

Lemma 3. *If $[4F(0) - 1]\lambda < F(0)$ and $\lambda > 1/3$, the ex-post privacy of mainstream content is higher than the ex-post privacy of niche content.*

Proof of Lemma 3. Due to the symmetry between two types of niche contents, we only need to compare the ex-post privacy of mainstream content and type 0 content.

1. Mainstream content $x = 1/2$

   By symmetry, $\mathbf{E}[\theta|x = 1/2] = 1/2$. So,

$$\begin{aligned}\mathbf{Var}[\theta|x = 1/2] =& \mathbf{E}[\theta - 1/2|x = 1/2]^2 \\ =& 2 \cdot \frac{\lambda[1 - F(0)]}{1 - 2\lambda F(0)} \cdot (1/2)^2 \\ =& \frac{1}{2} \cdot \frac{\lambda[1 - F(0)]}{1 - 2\lambda F(0)}\end{aligned}$$

2. Niche content $x = 0$

$$
\begin{aligned}
\mathbf{Var}[\theta|x=0] =& \mathbf{E}[\theta^2|x=0] - \mathbf{E}[\theta|x=0]^2 \\
=& \frac{[1-F(0)](1-2\lambda)}{\lambda + [1-F(0)](1-2\lambda)} \cdot (\frac{1}{2})^2 - \left\{ \frac{[1-F(0)](1-2\lambda)}{1-F(0)+[2F(0)-1]\lambda} \cdot \frac{1}{2} \right\}^2 \\
=& \frac{1}{4} \cdot \frac{[1-F(0)](1-2\lambda)}{\lambda + [1-F(0)](1-2\lambda)} - \frac{1}{4} \cdot \frac{[1-F(0)]^2(1-2\lambda)^2}{\{\lambda + [1-F(0)](1-2\lambda)\}^2} \\
=& \frac{1}{4} \cdot \frac{[1-F(0)](1-2\lambda)\lambda}{\{\lambda + [1-F(0)](1-2\lambda)\}^2}
\end{aligned}
$$

$$
\begin{aligned}
& \mathbf{Var}[\theta|x=0] < \mathbf{Var}[\theta|x=1/2] \\
\Leftrightarrow & \frac{1}{4} \cdot \frac{[1-F(0)](1-2\lambda)\lambda}{\{\lambda + [1-F(0)](1-2\lambda)\}^2} < \frac{1}{2} \cdot \frac{\lambda[1-F(0)]}{1-2\lambda F(0)} \\
\Leftrightarrow & (1-2\lambda)[1-2F(0)\lambda] < 2\{\lambda + [1-F(0)](1-2\lambda)\}^2 \\
\Leftrightarrow & \lambda + \lambda(1-2\lambda)F(0) + 2\lambda(1-2\lambda)[1-F(0)] + \lambda^2 + [1-F(0)]^2(1-2\lambda)^2 > \frac{1}{2}
\end{aligned}
$$

The LHS is greater than $\lambda+\lambda(1-2\lambda)F(0)+\lambda(1-2\lambda)[1-F(0)]+\lambda^2 = \lambda(2-\lambda)$. Since $\lambda \in (1/3, 1/2)$ and $\lambda(2-\lambda)$ increases in λ for $\lambda < 1$, we have the LHS $> \lambda(2-\lambda) > 1/3(2-1/3) = 5/9 > 1/2$. Therefore, $\mathbf{Var}[\theta|x = 0] < \mathbf{Var}[\theta|x = 1/2]$ always holds. By definition, the ex-post privacy of mainstream content is higher than the ex-post privacy of niche content. ■

Consider contextual advertising. By symmetry, we only need to compare the firm's expected profits from niche content and mainstream content.

$$
\begin{aligned}
& \pi^c(0) > \pi^c(1/2) \\
\Leftrightarrow\ & D(0) \cdot [1 - \gamma \mathbf{Var}(\theta|x=0)] > D(1/2) \cdot [1 - \gamma \mathbf{Var}(\theta|x=1/2)] \\
\Leftrightarrow\ & \gamma > \frac{\frac{D(1/2)}{D(0)} - 1}{\mathbf{Var}(\theta|x=1/2)\frac{D(1/2)}{D(0)} - \mathbf{Var}(\theta|x=0)}
\end{aligned}
$$

Denote $\frac{\frac{D(1/2)}{D(0)}-1}{\mathbf{Var}(\theta|x=1/2)\frac{D(1/2)}{D(0)}-\mathbf{Var}(\theta|x=0)}$ by γ^m. Since $D(1/2) > D(0)$ and $\mathbf{Var}[\theta|x = 0] < \mathbf{Var}[\theta|x = 1/2]$, we have $\gamma^m > 0$. So, the monopoly chooses niche content if $\gamma > \gamma^m$ and mainstream content if $\gamma < \gamma^m$ under contextual advertising. ■

Proof of Proposition 3. We first compute the consumer welfare for each content choice.

1. Mainstream content $x = 1/2$

   The consumer welfare is:

$$\begin{aligned}&(1-2\lambda)[1/2+\mathbf{E}[\epsilon]]+2[1-F(0)]\lambda\mathbf{E}[\epsilon|\epsilon>0]\\=&1/2-\lambda+2[1-F(0)]\lambda\mathbf{E}[\epsilon|\epsilon>0]\end{aligned}$$

2. Niche content $x = 0$

   The consumer welfare is:

$$\begin{aligned}&\lambda[1/2+\mathbf{E}[\epsilon]]+[1-F(0)](1-2\lambda)\mathbf{E}[\epsilon|\epsilon>0]\\=&\lambda/2+[1-F(0)](1-2\lambda)\mathbf{E}[\epsilon|\epsilon>0]\end{aligned}$$

3. Niche content $x = 1$

   It is symmetric to the previous $x = 0$ case.

Niche content leads to lower consumer welfare if and only if

$$\begin{aligned}&\lambda/2+[1-F(0)](1-2\lambda)\mathbf{E}[\epsilon|\epsilon>0]<1/2-\lambda+2[1-F(0)]\lambda\mathbf{E}[\epsilon|\epsilon>0]\\\Leftrightarrow&\lambda<\hat{\lambda}=\frac{1-2[1-F(0)]\mathbf{E}[\epsilon|\epsilon>0]}{3-8[1-F(0)]\mathbf{E}[\epsilon|\epsilon>0]}\end{aligned}$$

In addition, $\hat{\lambda} > 1/3 \Leftrightarrow 2[1 - F(0)]\mathbf{E}[\epsilon|\epsilon > 0] > 0$, which holds for any distribution of ϵ. ■

Proof of Proposition 4. We first characterize demand given different content choices.

Lemma 4. *Suppose the duopoly's content choices are (0,1). Firm 1's total demand is* $\lambda+\frac{1-F(0)^2}{2}(1-2\lambda)$*. Among them,* λ *are type 0 consumers and* $\frac{1-F(0)^2}{2}(1-2\lambda)$ *are mainstream consumers. Firm 2's total demand is* $\lambda+\frac{1-F(0)^2}{2}(1-2\lambda)$*. Among them,* λ *are type 1 consumers and* $\frac{1-F(0)^2}{2}(1-2\lambda)$ *are mainstream consumers.*

Suppose the duopoly's content choices are (1/2,1/2). Each firm's total demand is $\frac{1-2\lambda}{2}+[1-F(0)^2]\lambda$*. Among them,* $\frac{1-2\lambda}{2}$ *are mainstream consumers,* $\frac{1-F(0)^2}{2}\lambda$ *are type 0 consumers, and* $\frac{1-F(0)^2}{2}\lambda$ *are type 1 consumers.*

Suppose the duopoly's content choices are (0,1/2). Firm 1's total demand is $\lambda[1-\int_{-1/2}^{0}F(\epsilon_0)f(\epsilon+1/2)d\epsilon]+(1-2\lambda)\int_{-1/2}^{0}F(\epsilon)f(\epsilon+1/2)d\epsilon$*. Among them,* $\lambda[1-\int_{-1/2}^{0}F(\epsilon)f(\epsilon+1/2)d\epsilon]$ *are type 0 consumers and* $(1-2\lambda)\int_{-1/2}^{0}F(\epsilon)f(\epsilon+1/2)d\epsilon$ *are mainstream consumers. Firm 2's total demand is* $\lambda\int_{-1/2}^{0}F(\epsilon)f(\epsilon+1/2)d\epsilon+(1-2\lambda)[1-\int_{-1/2}^{0}F(\epsilon)f(\epsilon+1/2)d\epsilon]+[1-F(0)]\lambda$*. Among them,* $\lambda\int_{-1/2}^{0}F(\epsilon)f(\epsilon+1/2)d\epsilon$ *are type 0 consumers,* $(1-2\lambda)[1-\int_{-1/2}^{0}F(\epsilon)f(\epsilon+1/2)d\epsilon]$ *are type 1/2 consumers, and* $[1-F(0)]\lambda$ *are type 1 consumers.*

Proof. There are three candidate equilibria: (0,1/2), (1/2,1/2), (0,1).

1. (0,1)

   The demands for type 0 and type 1 content are symmetric. So, we only need to examine type 0 content. Each consumer will choose type 0 content if and only if her utility from type 0 content is positive (higher than the outside option) and higher than her utility from type 1 content.

   One can see that all type 0 consumers will choose type 0 content. Now consider type 1/2 consumers. A mainstream consumer will consume one of the contents if and only if her utility from at least one content is positive. By symmetry, her overall probabilities of choosing type

0 and type 1 content are identical.

$$
\begin{aligned}
&P(\text{a mainstream consumer chooses type 0 content})\\
=&\ \frac{1}{2}\cdot P(\max\{\epsilon_0,\epsilon_1\}>0)\\
\overset{\textit{independence of } \epsilon_j}{=}&\ \frac{1}{2}\cdot[1-P(\epsilon_0\leq 0)P(\epsilon_1\leq 0)]\\
=&\ \frac{1}{2}\cdot[1-F(0)^2]
\end{aligned}
$$

In sum, the demand of type 0 content from type 0 consumers is λ and from type 1/2 consumers is $\frac{1-F(0)^2}{2}(1-2\lambda)$. The total demand of type 0 content is $\lambda+\frac{1-F(0)^2}{2}(1-2\lambda)$.

2. (1/2, 1/2)

   One can see that all mainstream consumers will consume one of the contents. The demands from type 0 and type 1 consumers are symmetric. Consider type 0 consumers. She will consume one of the contents if and only if her utility from at least one content is positive. By symmetry, her overall probabilities of choosing either mainstream content are identical.

$$
\begin{aligned}
&P(\text{a type 0 consumer chooses firm } i)\\
=&\ \frac{1}{2}\cdot P(\max\{\epsilon_{1/2},\epsilon_{1/2'}\}>0)\\
\overset{\textit{independence of } \epsilon_j}{=}&\ \frac{1}{2}\cdot[1-P(\epsilon_{1/2}\leq 0)P(\epsilon_{1/2'}\leq 0)]\\
=&\ \frac{1-F(0)^2}{2}
\end{aligned}
$$

   In sum, the demand of either firm from mainstream consumers is $\frac{1-2\lambda}{2}$, from type 0 consumers is $\frac{1-F(0)^2}{2}\lambda$, and from type 1 consumers is $\frac{1-F(0)^2}{2}\lambda$. The total demand of either firm is $\frac{1-2\lambda}{2}+[1-F(0)^2]\lambda$.

3. (0, 1/2)

   Consider first the demand for type 0 content. Since a type 0 consumer's utility from consuming type 0 content is always positive, we only need to compare her utility from type 0 and type

1/2 contents.

$$
\begin{aligned}
&P(\text{a type 0 consumer chooses type 0 content}) \\
&= P(1/2 + \epsilon_0 > 0 + \epsilon_{1/2}) \\
&= P(\epsilon_{1/2} < \epsilon_0 + 1/2) \\
&= 1 - F(0) + \int_{-1/2}^{0} \int_{-1/2}^{\epsilon_0 + 1/2} f(\epsilon_{1/2}) d\epsilon_{1/2} f(\epsilon_0) d\epsilon_0 \\
&= 1 - F(0) + \int_{-1/2}^{0} F(\epsilon_0 + 1/2) f(\epsilon_0) d\epsilon_0 \\
&\overset{\text{integral by parts}}{=} 1 - \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0
\end{aligned}
$$

A type 1/2 consumer will choose type 0 content if and only if her utility of consuming it is positive and higher than her utility of consuming mainstream content. Since her utility of consuming mainstream content is always positive, we only need the condition that her utility of consuming type 0 content is higher than her utility of consuming mainstream content.

$$
\begin{aligned}
&P(\text{a type 1/2 consumer chooses type 0 content}) \\
&= P(0 + \epsilon_0 > 1/2 + \epsilon_{1/2}) \\
&= P(\epsilon_0 > \epsilon_{1/2} + 1/2) \\
&\overset{\textit{symmetry}}{=} P(\epsilon_{1/2} > \epsilon_0 + 1/2) \\
&= 1 - P(\epsilon_{1/2} < \epsilon_0 + 1/2) \\
&\overset{\textit{previous case}}{=} 1 - [1 - \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0] \\
&= \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0
\end{aligned}
$$

Therefore, the demand of type 0 content from type 0 consumers is $\lambda[1 - \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0]$ and from type 1/2 consumers is $(1-2\lambda) \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0$. The total demand of of type 0 content is $\lambda[1 - \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0] + (1 - 2\lambda) \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0$.

Consider now the demand for type 1/2 content.

$$\begin{aligned}
&P(\text{a type 0 consumer chooses type 1/2 content})\\
&= 1 - P(\text{a type 0 consumer chooses type 0 content})\\
&= \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0
\end{aligned}$$

$$\begin{aligned}
&P(\text{a type 1/2 consumer chooses type 1/2 content})\\
&= 1 - P(\text{a type 1/2 consumer chooses type 0 content})\\
&= 1 - \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0
\end{aligned}$$

$$\begin{aligned}
&P(\text{a type 1 consumer chooses type 1/2 content})\\
&= P(U(1/2, 1) > 0)\\
&= P(\epsilon_{1/2} > 0)\\
&= 1 - F(0)
\end{aligned}$$

Therefore, the demand of type 1/2 content from type 0 consumers is $\lambda \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0$, from type 1/2 consumers is $(1-2\lambda)[1 - \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0]$, and from type 1 consumers is $[1 - F(0)]\lambda$. The total demand of type 1/2 content is $\lambda \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0 + (1-2\lambda)[1 - \int_{-1/2}^{0} F(\epsilon_0) f(\epsilon_0 + 1/2) d\epsilon_0] + [1 - F(0)]\lambda$.

■

We first simplify the demand in Lemma 4 when $\epsilon \sim U[-1/2, 1/2]$.

Corollary 3. *Suppose* $\epsilon \sim U[-1/2, 1/2]$.

Suppose the duopoly's content choices are (0,1). Firm 1's total demand is $3/8 + \lambda/4$. *Among them,* λ *are type 0 consumers and* $3/8 - 3\lambda/4$ *are mainstream consumers. Firm 2's total demand is* $3/8 + \lambda/4$. *Among them,* λ *are type 1 consumers and* $3/8 - 3\lambda/4$ *are mainstream consumers.*

Suppose the duopoly's content choices are (1/2,1/2). Each firm's total demand is $1/2-\lambda/4$*. Among them,* $1/2-\lambda$ *are mainstream consumers,* $3\lambda/8$ *are type 0 consumers, and* $3\lambda/8$ *are type 1 consumers.*

Suppose the duopoly's content choices are (0,1/2). Firm 1's total demand is $1/8+5\lambda/8$*. Among them,* $7\lambda/8$ *are type 0 consumers and* $(1-2\lambda)/8$ *are mainstream consumers. Firm 2's total demand is* $7/8-9\lambda/8$*. Among them,* $\lambda/8$ *are type 0 consumers,* $\frac{7}{8}(1-2\lambda)$ *are type 1/2 consumers, and* $\lambda/2$ *are type 1 consumers.*

Now consider the equilibrium strategy.

1. Behavioral advertising

   Under behavioral advertising, the advertising revenue per consumer is 1. So, each firm's profit equals its total demand. Suppose the equilibrium is (0,1/2). Firm 1 will deviate from $x=0$ to $x=1/2$ if and only if:

$$\begin{aligned} &1/2-\lambda/4 > 1/8+5\lambda/8 \\ \Leftrightarrow &\lambda < \lambda_2 = 3/7 \end{aligned} \tag{8}$$

   Firm 2 will deviate from $x=1/2$ to $x=1$ if and only if:

$$\begin{aligned} &3/8+\lambda/4 > 7/8-9\lambda/8 \\ \Leftrightarrow &\lambda > \lambda_1 = 4/11 \end{aligned} \tag{9}$$

   From the above conditions and noting that $\lambda_1 < \lambda_2$, one can see that (0,1/2) will never be an equilibrium.

   Now suppose the equilibrium is (1/2,1/2). Equation 8 implies that the firm will deviate to 0 or 1 if $\lambda > \lambda_2$. Similarly, if the equilibrium is (0,1), Equation 9 implies that the firm will deviate to 1/2 if $\lambda < \lambda_1$.

   In sum, the equilibrium is (1/2,1/2) if $\lambda < \lambda_1$, (1/2,1/2) or (0,1) if $\lambda_1 < \lambda < \lambda_2$, and (0,1) if $\lambda > \lambda_2$.

2. Contextual advertising

We first characterize the equilibrium ad choices and profits for each firm.

(a) (1/2,1/2)

By symmetry, $a = \mathbf{E}(\theta) = 1/2$. $\mathbf{Var}(\theta) = \mathbf{E}[\theta - \mathbf{E}(\theta)]^2 = \frac{3\lambda}{4(2-\lambda)}$.

$$\begin{aligned}\pi(1/2,1/2) =& D(1/2,1/2) \cdot [1 - \gamma \mathbf{Var}(\theta)] \\ =& \frac{2-\lambda}{4} - \frac{3\gamma\lambda}{16}\end{aligned} \tag{10}$$

(b) (0,1)

By symmetry, we only need to consider firm 1. Firm 2's ad choice will be $1 - a_1$, and firm 2's profit will be identical to firm 1's.

$a_1 = \mathbf{E}(\theta) = \frac{3/8-3\lambda/4}{\lambda/4+3/8} \cdot \frac{1}{2} = \frac{3(1-2\lambda)}{2(3+2\lambda)}$. $\mathbf{Var}(\theta) = \mathbf{E}[\theta - \mathbf{E}(\theta)]^2 = \frac{6\lambda(1-2\lambda)}{(3+2\lambda)^2}$.

$$\begin{aligned}\pi(0,1) =& D(0,1) \cdot [1 - \gamma \mathbf{Var}(\theta)] \\ =& \frac{2\lambda+3}{8} - \frac{3\gamma\lambda(1-2\lambda)}{4(2\lambda+3)}\end{aligned} \tag{11}$$

(c) (0,1/2)

Consider firm 1 first. $a_1 = \mathbf{E}_1(\theta) = \frac{1-2\lambda}{2(1+5\lambda)}$. $\mathbf{Var}_1(\theta) = \mathbf{E}_1[\theta - \mathbf{E}_1(\theta)]^2 = \frac{7\lambda(1-2\lambda)}{4(1+5\lambda)^2}$.

$$\begin{aligned}\pi_1(0,1/2) =& D_1(0,1/2) \cdot [1 - \gamma \mathbf{Var}_1(\theta)] \\ =& \frac{1+5\lambda}{8} - \frac{7\gamma\lambda(1-2\lambda)}{32(1+5\lambda)}\end{aligned} \tag{12}$$

Consider firm 2 then. $a_2 = \mathbf{E}_2(\theta) = \frac{7-6\lambda}{2(7-9\lambda)}$. $\mathbf{Var}_2(\theta) = \mathbf{E}_2[\theta - \mathbf{E}_2(\theta)]^2 = \frac{35\lambda-54\lambda^2}{4(7-9\lambda)^2}$.

$$\begin{aligned}\pi_2(0,1/2) =& D_2(0,1/2) \cdot [1 - \gamma \mathbf{Var}_2(\theta)] \\ =& \frac{7-9\lambda}{8} - \frac{\gamma(35\lambda-54\lambda^2)}{32(7-9\lambda)}\end{aligned} \tag{13}$$

Suppose the equilibrium is (0,1/2). Firm 1 will deviate from $x = 0$ to $x = 1/2$ if and only if:

$$\begin{aligned} &\pi(1/2, 1/2) > \pi_1(0, 1/2) \\ \Leftrightarrow &\gamma < \gamma^d = (\lambda_2 - \lambda)\frac{28(1 + 5\lambda)}{\lambda(44\lambda - 1)} \end{aligned} \tag{14}$$

An immediate implication is that this deviation never happens if $\lambda \geq \lambda_2$.

Firm 2 will deviate from $x = 1/2$ to $x = 1$ if and only if:[7]

$$\begin{aligned} &\pi(0, 1) > \pi_2(0, 1/2) \\ \Leftrightarrow &\gamma > \gamma^{d'} = (\lambda_1 - \lambda)\frac{44(2\lambda + 3)(7 - 9\lambda)}{\lambda(-540\lambda^2 + 460\lambda - 63)} \end{aligned} \tag{15}$$

An immediate implication is that deviation always happens if $\lambda \geq \lambda_1$.[8] Therefore, (0,1/2) may only be an equilibrium if $\lambda < \lambda_1$. Now suppose $\lambda < \lambda_1$. In this case, we have shown that firm 1 will deviate if $\gamma < \gamma^d$ and firm 2 will deviate if $\gamma > \gamma^{d'}$. Some calculation yields that there exists a unique $\lambda_0 \in (1/3, \lambda_1)$ such that $\gamma^d < (>)\gamma^{d'}$ if $\lambda < (>)\lambda_0$. Therefore, at least one deviation happens and (0,1/2) will not be an equilibrium if $\lambda > \lambda_0$. There is no deviation and (0,1/2) is an equilibrium if $\lambda < \lambda_0$.

Now suppose the equilibrium is (1/2,1/2). Equation 14 implies that the firm will deviate to 0 or 1 if $\gamma > \gamma^d$, which may hold if $\lambda < \lambda_2$ and always holds if $\lambda > \lambda_2$. Similarly, if the equilibrium is (0,1), Equation 15 implies that the firm will deviate to 1/2 if $\gamma < \gamma^{d'}$, which may hold if $\lambda < \lambda_1$ and never holds if $\lambda > \lambda_1$.

In sum, (0,1) is an equilibrium if $\lambda > \lambda_1$ or if $\lambda < \lambda_1$ & $\gamma > \gamma^{d'}$. (1/2,1/2) is an equilibrium if $\lambda < \lambda_2$ & $\gamma < \gamma^d$.

■

Proof of Proposition 5. We first compute the consumer welfare for each content choice.

[7] Firm 2 has a stronger incentive to deviate to $x = 1$ rather than $x = 0$ to soften competition. So, we only need to consider its deviation to $x = 1$.

[8] $\frac{44(2\lambda+3)(7-9\lambda)}{\lambda(-540\lambda^2+460\lambda-63)} > 0$ for $\lambda \in (1/3, 1/2)$.

1. (0,1)

   The consumer welfare from each firm is:

$$
\begin{aligned}
&\lambda(1/2+\mathbf{E}[\epsilon_0]) + (3/8 - 3\lambda/4)\mathbf{E}[\epsilon_0|\epsilon_0 > 0 \text{ and } \epsilon_0 > \epsilon_1]) \\
=&\lambda/2 + (3/8 - 3\lambda/4)\cdot 5/18 \\
=&\frac{5-\lambda}{18}
\end{aligned}
$$

2. (1/2,1/2)

   The consumer welfare from each firm is:

$$
\begin{aligned}
&\frac{1-2\lambda}{2}(1/2+\mathbf{E}[\epsilon_{1/2}]) + 2(\frac{1-F(0)^2}{2})\lambda\mathbf{E}[\epsilon_{1/2}|\epsilon_{1/2} > 0 \text{ and } \epsilon_{1/2} > \epsilon_{1/2'}]) \\
=&\frac{1-2\lambda}{4} + \frac{3\lambda}{4}\frac{5}{18} \\
=&\frac{6-7\lambda}{24}
\end{aligned}
$$

3. (0,1/2)

   The consumer welfare from firm 1 is:

$$
\begin{aligned}
&\frac{7\lambda}{8}(1/2+\mathbf{E}[\epsilon_0|\epsilon_0+1/2 > \epsilon_{1/2}]) + \frac{1-2\lambda}{8}\mathbf{E}[\epsilon_0|\epsilon_0 > 0 \text{ and } \epsilon_0 > 1/2+\epsilon_{1/2}] \\
=&\frac{7\lambda}{8}\cdot(\frac{1}{2}+\frac{1}{21}) + \frac{1-2\lambda}{8}\cdot\frac{1}{3} \\
=&\frac{2+19\lambda}{48}
\end{aligned}
$$

   The consumer welfare from firm 2 is:

$$
\begin{aligned}
&\frac{\lambda}{8}\mathbf{E}[\epsilon_{1/2}|\epsilon_{1/2} > 0 \text{ and } \epsilon_{1/2} > 1/2+\epsilon_0] + \frac{7(1-2\lambda)}{8}(1/2+\mathbf{E}[\epsilon_{1/2}|\epsilon_{1/2} > 1/2+\epsilon_0)+ \\
&\frac{\lambda}{2}\mathbf{E}[\epsilon_{1/2}|\epsilon_{1/2} > 0] \\
=&\frac{\lambda}{8}\cdot\frac{1}{3} + \frac{7(1-2\lambda)}{8}(1/2+1/3) + \frac{\lambda}{2}\cdot\frac{1}{4} \\
=&\frac{35-62\lambda}{48}
\end{aligned}
$$

The consumer's total welfare is $\frac{2+19\lambda}{48} + \frac{35-62\lambda}{48} = \frac{37-43\lambda}{48}$.

The total consumer welfare under (0,1) content choice is higher than that under (0,1/2) content choice if and only if $2 \cdot \frac{5-\lambda}{18} > \frac{37-43\lambda}{48} \Leftrightarrow \lambda > 31/113$, which always holds. The total consumer welfare under (0,1/2) content choice is higher than that under (1/2,1/2) content choice if and only if $\frac{37-43\lambda}{48} > 2 \cdot \frac{6-7\lambda}{24} \Leftrightarrow \lambda < 13/15$, which always holds.

■

Proof of Proposition 6. We consider two cases:

1. $v < 1/2$:

   In this case, the consumer may not consume the content even if the content location perfectly match the consumer type. In addition, the consumer never consumes the content if the content location is on the opposite end of the consumer's location.

   (a) Monopoly

   We first summarize the demand for a monopoly by choosing a particular content.

   Suppose the monopoly chooses niche content $x = 0$.[9] One can show that the demand from type 0 consumer is $\lambda(1/2 + v)$, from type 1/2 consumer is $(1-2\lambda)v$, and from type 1 consumer is 0. Hence, the total demand is $\lambda(1/2 + v) + (1-2\lambda)v$.

   Suppose the monopoly chooses mainstream content $x = 1/2$. One can show that the demand from type 0 consumer is λv, from type 1/2 consumer is $(1-2\lambda)(1/2+v)$, and from type 1 consumer is λv. Hence, the total demand is $(1-2\lambda)(1/2+v) + 2\lambda v$.

$$D(0) < D(1/2) \Leftrightarrow \lambda < \frac{1}{3-2v}$$

   Consequently, the monopoly chooses mainstream content under behavioral advertising if and only if $\lambda < 1/(3-2v)$.

   Now let us study the equilibrium under contextual advertising. We restirct the attention to the case where $1/3 < \lambda < 1/(3-2v)$ because niche content gives the firm higher

[9] The case where $x = 1$ is symmetric to this case. We only need to consider one of these cases.

demand and better inference if $\lambda \geq 1/(3-2v)$ whereas mainstream content dominates niche content if $\lambda \leq 1/3$.

Some calculations yield that $\mathbf{Var}(\theta|x=0) = \frac{1}{4}\frac{v(1-2\lambda)(\lambda v+\lambda/2)}{(\lambda/2-\lambda v+v)^2}$ and $\mathbf{Var}(\theta|x=1/2) = \frac{1}{2}\frac{\lambda v}{1/2+v-\lambda}$. Niche content gives the firm better inference if and only if $\mathbf{Var}(\theta|x=0) < \mathbf{Var}(\theta|x=1/2)$, which always hold for $\lambda \in (1/3, 1/(3-2v))$.

The monopoly prefers niche content to mainstream content if:

$$
\begin{aligned}
&\pi_{con}(x=0) > \pi_{con}(x=1/2) \\
\Leftrightarrow &D(0)[1-\gamma\mathbf{Var}(\theta|x=0) > D(1/2)[1-\gamma\mathbf{Var}(\theta|x=1/2)] \\
\Leftrightarrow &\gamma > \frac{16[1+\lambda(2v-3)]}{\lambda v[2\lambda^3(1-2v)^2(1+2v)+4\lambda v(-1+2v+8v^2)+\lambda^2(-1+10v+4v^2-40v^3)-4(-4+v^2+2v^3)]}
\end{aligned}
$$

Denote the threshold as γ^m. One can show that γ^m decreases in λ and increases in v.[10]

Now consider consumer welfare. Denote the consumer welfare given content choice x as $CW(x)$. We have:

$$
\begin{aligned}
CW(1/2) =&(1-2\lambda)\mathbf{E}[v+\epsilon|v+\epsilon>0]P(v+\epsilon>0)+2\lambda\mathbf{E}[v-1/2+\epsilon|v-1/2+\epsilon>0]P(v-1/2+\epsilon> \\
=&\frac{(1-2\lambda)(v+1/2)^2}{2}+\lambda v^2 \\
CW(0) =&\lambda\mathbf{E}[v+\epsilon|v+\epsilon>0]P(v+\epsilon>0)+(1-2\lambda)\mathbf{E}[v-1/2+\epsilon|v-1/2+\epsilon>0]P(v-1/2+\epsilon>0 \\
=&\frac{\lambda(v+1/2)^2}{2}+\frac{(1-2\lambda)v^2}{2}
\end{aligned}
$$

Niche content leads to lower consumer welfare if: $CW(0) < Cw(1/2) \Leftrightarrow \lambda < \hat{\lambda} := \frac{v+1/4}{-v^2+3v+1/4}$. One can see that $\hat{\lambda} > 1/3$.

(b) Duopoly

We first compute the demand given different content choices.

i. (0,1)

The demand for type 0 content:

From $\theta = 0$ consumers: $\lambda P(v+\epsilon>0) = \lambda(v+1/2)$.

[10] Observe that γ^m can be negative for small v. In that case, we let $\gamma^m = 0$ for convenience.

From $\theta = 1/2$ consumers: $\frac{1}{2}(1-2\lambda)P(\max\{v-1/2+\epsilon_0, v-1/2+\epsilon_1\} > 0) = \frac{(1-2\lambda)(2v-v^2)}{2}$.

From $\theta = 1$ consumers: 0.

Total demand: $\lambda(v+1/2) + \frac{(1-2\lambda)(2v-v^2)}{2}$.

The demand for type 1 content is symmetric.

ii. (1/2, 1/2)

The demand for type 1/2 content:

From $\theta = 0/1$ consumers: $\frac{1}{2}\lambda P(\max\{v-1/2+\epsilon_{1/2}, v-1/2+\epsilon_{1/2'}\} > 0) = \frac{\lambda(2v-v^2)}{2}$.

From $\theta = 1/2$ consumers: $\frac{1}{2}(1-2\lambda)P(\max\{v+\epsilon_{1/2}, v+\epsilon_{1/2'}\} > 0) = \frac{(1-2\lambda)[1-(1/2-v)^2]}{2}$.

Total demand: $\lambda(2v-v^2) + \frac{(1-2\lambda)[1-(1/2-v)^2]}{2}$.

iii. (0,1/2)

The demand for type 0 content:

From $\theta = 0$ consumers: $\lambda P(v+\epsilon_0 > 0 \textit{ and } v+\epsilon_0 > v-1/2+\epsilon_{1/2}) = \lambda(v/2+5/8)$.

From $\theta = 1/2$ consumers: $\frac{1}{2}\lambda P(v-1/2+\epsilon_0 > 0 \textit{ and } v-1/2+\epsilon_0 > v+\epsilon_{1/2}) = \frac{(1-2\lambda)(v-v^2)}{2}$.

From $\theta = 1$ consumers: 0.

Total demand: $\lambda(v/2+5/8) + \frac{(1-2\lambda)(v-v^2)}{2}$.

The demand for type 1/2 content:

From $\theta = 0$ consumers: $\frac{\lambda(v-v^2)}{2}$.

From $\theta = 1/2$ consumers: $(1-2\lambda)(v/2+5/8)$.

From $\theta = 1$ consumers: $\lambda P(v-1/2+\epsilon_{1/2} > 0) = \lambda v$.

Total demand: $\frac{\lambda(v-v^2)}{2} + (1-2\lambda)(v/2+5/8) + \lambda v$.

Consider the equilibrium under behavioral advertising. Suppose (0,1/2) is an equilibrium. Firm 1 will deviate from $x = 0$ to $x = 1/2$ if and only if:

$$
\begin{aligned}
&\lambda(2v-v^2)+\frac{(1-2\lambda)[1-(1/2-v)^2]}{2}>\lambda(v/2+5/8)+\frac{(1-2\lambda)(v-v^2)}{2}\\
\Leftrightarrow&\lambda<\lambda_2:=\frac{3}{11-12v+8v^2}(<\frac{1}{3-2v})
\end{aligned}
$$

Firm 2 will deviate from $x=1/2$ to $x=1$ if and only if:

$$
\begin{aligned}
&\lambda(v+1/2)+\frac{(1-2\lambda)(2v-v^2)}{2}>\frac{\lambda(v-v^2)}{2}+(1-2\lambda)(v/2+5/8)+\lambda v\\
\Leftrightarrow&\lambda>\lambda_1:=\frac{5-4v+4v^2}{14-12v+12v^2}
\end{aligned}
$$

One can see that λ_1 and λ_2 increases in v. One can also show that there exists $\underline{v}<7/25$ such that $\lambda_1<\lambda_2$ when $\underline{v}<v<1/2$. Therefore, (0,1/2) will never be an equilibrium if $\underline{v}<v<1/2$. We assume that $\underline{v}<v<1/2$ in subsequent analyses.

Now suppose the equilibrium is (1/2,1/2). One can see that the firm will deviate to 0 or 1 if $\lambda>\lambda_2$. Similarly, if the equilibrium is (0,1), the firm will deviate to 1/2 if $\lambda<\lambda_1$.

In sum, the equilibrium is (1/2,1/2) if $\lambda<\lambda_1$, (1/2,1/2) or (0,1) if $\lambda_1<\lambda<\lambda_2$, and (0,1) if $\lambda>\lambda_2$.

Now consider the equilibrium under contextual advertising.

We first characterize the equilibrium ad choices and profits for each firm.

A. (1/2,1/2)

By symmetry, $a=\mathbf{E}(\theta)=1/2$. $\mathbf{Var}(\theta)=\frac{\lambda v(2-v)}{3+4v-4v^2+\lambda(4v^2-6)}$.

$$
\begin{aligned}
\pi(1/2,1/2)=&D(1/2,1/2)\cdot[1-\gamma\mathbf{Var}(\theta)]\\
=&\lambda(2v-v^2)+\frac{(1-2\lambda)[1-(1/2-v)^2]}{2}-\frac{\gamma\lambda(2v-v^2)}{4}
\end{aligned}
$$

B. (0,1)

By symmetry, we only need to consider firm 1. Firm 2's ad choice will be $1 - a_1$, and firm 2's profit will be identical to firm 1's.

$\mathbf{Var}(\theta) = \frac{1}{4}\frac{(1-2\lambda)(2v-v^2)/2}{(v+1/2)\lambda+(1-2\lambda)(2v-v^2)/2}\frac{(v+1/2)\lambda}{(v+1/2)\lambda+(1-2\lambda)(2v-v^2)/2}$.

$$\begin{aligned}\pi(0,1) =& D(0,1)\cdot[1-\gamma\mathbf{Var}(\theta)]\\ =&(v+1/2)\lambda+(1-2\lambda)(2v-v^2)/2-\gamma\frac{1}{4}\frac{\frac{1}{2}(1-2\lambda)(2v-v^2)(v+1/2)\lambda}{(v+1/2)\lambda+(1-2\lambda)(2v-v^2)/2}\end{aligned}$$

C. (0,1/2)

Consider firm 1 first. $\mathbf{Var}_1(\theta) = \frac{1}{4}\frac{(1-2\lambda)(v-v^2)/2}{(v/2+5/8)\lambda+(1-2\lambda)(v-v^2)/2}\frac{(v/2+5/8)\lambda}{(v/2+5/8)\lambda+(1-2\lambda)(v-v^2)/2}$.

$$\begin{aligned}\pi_1(0,1/2) =& D_1(0,1/2)\cdot[1-\gamma\mathbf{Var}_1(\theta)]\\ =&(\frac{v}{2}+\frac{5}{8})\lambda+\frac{(1-2\lambda)(v-v^2)}{2}-\gamma\frac{1}{4}\frac{\frac{1}{2}(1-2\lambda)(v-v^2)(v/2+5/8)\lambda}{(v/2+5/8)\lambda+(1-2\lambda)(v-v^2)/2}\end{aligned}$$

Consider firm 2 then. $\mathbf{Var}_2(\theta) = \frac{(1-2\lambda(v/2+5/8)}{D_2(0,1/2)}\frac{1}{4}+\frac{\lambda v}{D_2(0,1/2)}\cdot 1-[\frac{(1-2\lambda(v/2+5/8)}{D_2(0,1/2)}\frac{1}{2}+\frac{\lambda v}{D_2(0,1/2)}\cdot 1]^2$.

$$\begin{aligned}&\pi_2(0,1/2)\\ =&D_2(0,1/2)\cdot[1-\gamma\mathbf{Var}_2(\theta)]\\ =&\frac{1}{8}[5+4v-2\lambda(5-2v+2v^2)][1+\frac{\gamma\lambda v[-15-7v+4v^2+6\lambda(5-3v+4v^2)]}{[5+4v-2\lambda(5-2v+2v^2)]^2}]\end{aligned}$$

Suppose the equilibrium is (0,1/2). One can show that there exists a γ^d such that firm 1 will deviate from $x=0$ to $x=1/2$ if and only if:

$$\begin{aligned}&\pi(1/2,1/2) > \pi_1(0,1/2)\\ \Leftrightarrow&\gamma < \gamma^d \end{aligned}\tag{16}$$

Moreover, γ^d increases in v for $v \in (0,1/2)$ and $\lambda \in (1/3,1/2)$. Also, given that $1/3 < \lambda < 1/2$ and $0 < v < 1/2$, we have $\gamma^d > 0 \Leftrightarrow (3-\sqrt{5})/4 < v < 1/2$ and $1/3 < \lambda < \lambda_2$. An immediate implication is that this deviation never happens if

$\lambda \geq \lambda_2$.

One can show that there exists a $\gamma^{d'}$ such that firm 2 will deviate from $x = 1/2$ to $x = 1$ if and only if:[11]

$$\begin{aligned} &\pi(0,1) > \pi_2(0,1/2) \\ \Leftrightarrow &\gamma > \gamma^{d'} \end{aligned} \tag{17}$$

Moreover, given that $1/3 < \lambda < 1/2$ and $0 < v < 1/2$, we have $\gamma^{d'} \leq 0 \Leftrightarrow \lambda_1 \leq \lambda < 1/2$. An immediate implication is that deviation always happens if $\lambda \geq \lambda_1$. Therefore, (0,1/2) may only be an equilibrium if $\lambda < \lambda_1$. Now suppose $\lambda < \lambda_1$. In this case, we have shown that firm 1 will deviate if $\gamma < \gamma^d$ and firm 2 will deviate if $\gamma > \gamma^{d'}$.

Now suppose the equilibrium is (1/2,1/2). One can see that the firm will deviate to 0 or 1 if $\gamma > \gamma^d$, which may hold if $\lambda < \lambda_2$ and always holds if $\lambda > \lambda_2$. Similarly, if the equilibrium is (0,1), the firm will deviate to 1/2 if $\gamma < \gamma^{d'}$, which may hold if $\lambda < \lambda_1$ and never holds if $\lambda > \lambda_1$.

Since $\gamma^m > \max\{\gamma^d, \gamma^{d'}\}$, $\forall v \in (\underline{v}, 1/2), \lambda \in (1/3, 1/2)$, defining $\gamma^{d''}$ by $\min\{\gamma^d, \gamma^{d'}\}\mathbf{1}_{[\lambda<\lambda_1]} + \gamma^d \mathbf{1}_{[\lambda_1 \leq \lambda < \lambda_2]}$, we have that the equilibrium under contextual advertising is more polarized than that under behavioral advertising if $\gamma > \gamma^{d''}$ and $1/3 < \lambda < \lambda_2$. Moreover, $\gamma^{d''} < \gamma^m$ for all $\lambda \in (1/3, \lambda_2)$.

2. $v \geq 1/2$:

In this case, the consumer always derives a non-negative utility from consuming the content if the content location perfectly match the consumer type. In addition, the consumer may consume the content even if the content location is on the opposite end of the consumer's location. The analyses are similar to the previous case.

■

[11] Firm 2 has a stronger incentive to deviate to $x = 1$ rather than $x = 0$ to soften competition. So, we only need to consider its deviation to $x = 1$.